PROCEEDINGS

The HKBU 9th Computer Science Postgraduate Research Symposium

January 20, 2009

PG Day 2009

Department of Computer Science
Hong Kong Baptist University

The 9th HKBU-CSD Postgraduate Research Symposium (PG Day) Program

<table>
<tr><th colspan="3">January 20 Tuesday, 2009</th></tr>
<tr><th>Time</th><th colspan="2">Sessions</th></tr>
<tr><td>09:00-09:20</td><td colspan="2">On-site registration</td></tr>
<tr><td>09:20-09:30</td><td colspan="2">Welcome: Prof. LIU Jiming, Head of Computer Science Department (FSC501)</td></tr>
<tr><td rowspan="2">09:30-12:00</td><td>Session A: (Chair: Chang LIU) (FSC501)
Pattern Recognition</td><td>Session B: (Chair: Chun Fan WONG) (FSC701B)
Information System & Data Mining</td></tr>
<tr><td>• Face Template Protection: A Binary Linear Combination Representation
Yicheng FENG
• A SemiBoosted Co-Training Algorithm For Human Action Recognition
Chang LIU
• Image Analysis By An Improved Empirical Mode Decomposition Method
Dan ZHANG
• Recognition Of 3D Graphical Models By Using Shape Similarity
Yuesheng HE
• Automatic Lip Localization Under Changing Illumination Conditions
Meng LI</td><td>• Hiding Emerging Patterns By Using Local Recoding Generalization
Wai Kit CHENG
• Improving Patient Journey With Reduced Length Of Stay By Using A Multi-agent Approach
Chung Ho CHOI
• A Knowledge-based Approach For Histogram-distances Image Retrieval
Chun Fan WONG
• Index Convergence Behaviour For Collaborative Semantic Indexing Of Multimedia Data Objects
Wing Sze CHAN</td></tr>
<tr><td>12:00-14:00</td><td colspan="2">Noon Break</td></tr>
<tr><td>14:00-15:00</td><td colspan="2">Keynote Talk: Prof. Qiang Yang, Hong Kong University of Science and Technology (FSC 501)
(Chair: Prof. Jiming Liu)
Enjoy Life through Research</td></tr>
<tr><td>15:00-15:30</td><td colspan="2">Tea Break</td></tr>
<tr><td rowspan="2">15:30-18:00</td><td>Session C: (Chair: Yu LI) (FSC501)
Database, Networking & Optimization</td><td>Session D: (Chair: Victor CHENG) (FSC701B)
Information System & Data Mining</td></tr>
<tr><td>• L-BFGS And Delayed Dynamical Systems Approach For Unconstrained Optimization
Xiaohui XIE
• Transaction Support On Flash Devices
Sai Tung ON
• Exploiting Fast Random Reads For Flash-based Joins
Yu LI
• Lightweight Emulation To Study BitTorrent-Like Systems
Xiaowei CHEN
• Speeding Up Homomorphic Hashing Using GPUs
Kaiyong ZHAO</td><td>• Kernel Learning For Local Learning Based Clustering
Hong ZENG
• Analysis Of Recall Characteristics On Image Search Engine
Xiaoling WANG
• HMM-LDA Feature Extraction For Mining Of Product Ownership Of Online Forum Participants
Tianjie ZHAN
• Opinion Mining: A Survey Of The State-of-the-Art And Possible Extensions
Kwan Wai LEUNG</td></tr>
<tr><td>19:00</td><td colspan="2">Best Paper & Best Presentation Awards Announcement via Email</td></tr>
</table>

TABLE OF CONTENTS

Session A: Pattern Recognition

Session B: Information System & Data mining

Session C: Database, Networking & Optimization

Session D: Information System & Data mining

Face Template Protection: A Binary Linear Combination Representation*

Yicheng FENG

Abstract

This paper addresses the security issues of the face biometric templates stored in a database. In order to improve the security level of the stored face templates, cryptographic techniques are commonly employed. Since most of such techniques require a binary input, thresholding is usually employed to binarize the real valued face features. While binary templates are obtained, thresholding usually leads to loss of some useful template information. And in turn, the recognition performance is also affected. In order to overcome this limitation, this paper proposes a new approach to represent a real valued face template using a binary vector whose elements represent the weights given to a set of basis feature vectors such that the weighted linear combination approximates the original template. To estimate an optimal set of weights and the basis vectors, a new optimization method is developed. The proposed method is evaluated on three public domain databases, namely FERET, CMU-PIE and FRGC. Experimental results show that, in comparison to the existing thresholding-based methods, the proposed method improves the recognition accuracy by around 7% on the FERET and CMU-PIE databases and around 5% on the FRGC database at an FAR of 1%.

1 Introduction

Biometrics is a reliable, robust and convenient way for person authentication [8, 9, 5]. With the success of the biometrics research in the last two decades, several large scale recognition systems have been successfully deployed. With the growing use of biometrics, there is a rising concern about the security and privacy of the stored biometric templates (which refer to a set of features extracted from raw biometric data) stored in a database or a smartcard. Recent studies [10] show that simple attacks on a biometric system, such as hill climbing, are able to recover the raw biometric data from a stolen biometric template. Moreover, the attacker may be able to make use of the stolen template to access the system or cross-match across databases. A comprehensive analysis of eight types of attacks [5] on a biometric system has been reported. Therefore, biometric template security [8, 9, 5, 17] has been an important issue in deploying a biometric system.

In order to overcome the security and privacy problems [5, 6, 8], a number of biometric template protection algorithms have been reported in the last few years. These methods can be broadly categorized into two approaches, namely biometric cryptosystem approach and transformation-based approach. The basic idea of both the approaches is that instead of storing the original biometric template, the transformed/encrypted template is stored. In case the transformed/encrypted biometric template is stolen or lost, it is computationally hard to reconstruct the biometric template and the original raw biometric data from the transformed/encrypted template. Generally speaking, biometric cryptosystems offer better security than transformation-based approach because cryptographic technique is employed in the last step in the cryptosystems approach. It is computationally hard to reconstruct the original template from the encrypted template stored in the database. However, most (if not all) of the protection algorithms in biometric cryptosystem approach require a binary template for encryption. That means, the input template has to be converted into a binary template before applying the encryption because most of the biometric templates are not represented in binary form, but are real valued. In order to satisfy the input requirements, thresholding is a typically employed in existing algorithms [11, 12, 13, 14, 15, 16]. While the binary template can be obtained, there are two major limitations of thresholding technique. First, some useful and discriminative information in the original (real valued) template will be lost after thresholding leading to degraded matching accuracy [11, 12]. Second, the selection of optimal thresholds still remains unsolved.

In view of the limitations on existing thresholding-based algorithms, this paper adopt the well-known idea that a face can be represented by a linear combination of other faces [2]. If we can find a good approximation to the original template as a linear combination of certain basis feature vectors, the information loss will be minimized and therefore, the discriminability of the original template can be preserved. In turn, a real valued face template can be represented by

*This paper has been submitted to CVPR2009.

the weights (coefficients) corresponding to the basis vectors. Different from existing linear combination schemes, in this paper, the weights for each basis is binary. To our knowledge, there is no binary linear combination scheme available and this paper proposes a new algorithm for that. The rest of this paper is organized as follows. Section 2 gives a brief review of the existing binarization schemes. Our proposed algorithm is then reported in Section 3. Experimental results and analysis are given in Sections 4 and 5. Finally, Section 6 gives the conclusion.

2 Review on Existing Binarization Schemes

A comprehensive survey on biometric template protection has been reported in [8, 7]. This section mainly reviews the existing binarization schemes in biometric template protection. This is always a two-step architecture as illustrated in Figure 1. A binarization process transforms the original feature vectors into binary strings, and then the cryptographic method is employed to encrypt the binary template in order to improve the security level. In practice, some transformation is usually applied before the binarization step in order to provide cancelability and enhance discriminability at the same time [18]. This however, is not the focus of this paper.

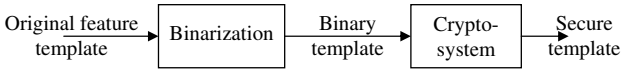


Figure 1. The binarization-combined cryptosystem architecture.

The binarization scheme was first proposed by Monrose $et\ al.$ [11, 12], who described a cryptographic key generation scheme from biometrics. The biometric data is transformed into a binary string called "feature descriptor" which has relatively small intra-class variation and large between-class variation. This binary string is generated by a thresholding technique based on mean and standard deviation of the biometric data.

Goh and Ngo [14] proposed a biohashing scheme. In their scheme, the original templates are first transformed using random mapping. Each element of the transformed template is thresholded to either 0 or 1, thus converting the template into binary form. Different versions of Biohashing algorithms have been proposed in the last few years. Teoh $et\ al.$ proposed a two-factor authentication algorithm [13] for fingerprint and a random multispace quantization (RMQ) algorithm [16] for face biometric. In RMQ algorithm, a user specific random transformation matrix is adopted to enhance the discriminability of the templates. Again, in all their schemes, thresholding technique is adopted.

Feng et al. [18, 19, 20] proposed a class-distribution-preserving transform (CDP transform) for binarization. Distinguishing points are determined. The distance between each distinguishing point and the face template is calculated and thresholded. With optimal positioning of the distinguishing points, the transform optimizes the discriminability of the binary strings. Thus the discriminability-preserving ability of this scheme is justified. But note that it is still a thresholding-based approach.

3 Proposed Binary Linear Combination Representation

From the reviews in Section 2, it is noted that a simple thresholding technique is employed in most existing template protection algorithms. It is important to point out that the binarization step significantly affects the overall biometric system performance, but has received only little attention. No matter how good the thresholding algorithm is, some information is bound to be lost. In order to avoid the thresholding process while we could obtain binary template representation, this paper proposes to represent the (real valued) face template by a linear combination of certain basis features with binary weights. The binary weights are then used to represent of the original face template.

Linear combination has been adopted in face recognition community for more than two decades. However, the problem in this paper is different from the traditional linear combination problem where weights are real values, and the optimal weights can be determined using a least square method. To the best of our knowledge, there is no standard method in determining the linear combination with binary weights. Therefore, a new method is proposed in this paper.

The problem of the binary linear combination is defined as follows. Given an original face template v, we would like to find an approximated template v_A which is a linear combination of basis $\beta_1, \beta_2 \ldots \beta_k$ with binary coefficients $b_1, b_2 \ldots b_k$ such that the error between v_A and v is minimum. Mathematically, it is written as,

$$v_A = \sum_{i=1}^{k} b_i \beta_i.$$

$$w = \underset{b_1, b_2 \cdots b_k}{\operatorname{argmin}} ||v_A - v||.$$

where w is the binary (template) representation of the original face template v.

3.1 Finding Optimal Basis and Weights for the Reference Template

In finding the optimal basis, we hope that the distance between any two face templates is preserved in the binary

representation space. Suppose v_1 and v_2 are two face templates, and v_{1A} and v_{2A} are their approximations with binary representations (weights) w_1 and w_2 respectively. Let $\beta_1, \beta_2 \ldots \beta_k$ are orthogonal columns of matrix $B_{l\times k}$. Then

$$||v_1 - v_2|| \approx ||v_{1A} - v_{2A}|| = ||B(w_1 - w_2)|| = m||w_1 - w_2||$$

where

$$B^T B = m^2 I. \tag{1}$$

It can be seen that if we want to preserve the distance ranking, the basis have to be orthogonal and m is a normalization constant.

In this paper, we assume that each class has its own set of basis, but same value of m should be used. A representative face template from each class (which could be an average of all input face template of the same individual) is given. Suppose there are c classes and the representatives of the various classes in the training set are $\{v_1, v_2, v_3 \ldots v_c\}$ respectively. Assume that m is given and fixed. For the i^{th} class, the optimal basis B_i and binary template w_i for representing v_i can be determined by minimizing the approximated error E_i which is given as follows,

$$\begin{aligned} E_i &= ||v_{iA} - v_i||^2 \\ &= ||v_{iA}||^2 - 2v_i^T v_{iA} + ||v_i||^2 \\ &= ||B_i w_i||^2 - 2v_i^T B_i w_i + ||v_i||^2 \\ &= m^2||w_i||^2 - 2v_i^T B_i w_i + ||v_i||^2 \end{aligned} \tag{2}$$

For a fixed w_i and $||v_i||^2$ is a positive constant, minimizing E_i is equivalent to maximizing the term $v_i^T B_i w_i$ which is then given as follows,

$$v_i^T \cdot (B_i w_i) = ||v_i|| \cdot ||B_i w_i|| \cos\theta, \tag{3}$$

where θ is the angle between v_i and $B_i w_i$. Obviously, Equation (3) achieves its maximum if and only if $\cos\theta = 1$, that is,

$$B_i w_i = x v_i, \quad (\forall x \in \mathbb{R}^+). \tag{4}$$

Substitute Equation (1) into Equation (4), we have

$$x||v_i|| = ||B_i w_i|| = m||w_i||.$$

$$x = \frac{m||w_i||}{||v_i||}.$$

Therefore,

$$B_i \frac{w_i}{||w_i||} = m\frac{v_i}{||v_i||}; \tag{5}$$

Solving Equations (1) and (5), $B_i(w_i)$ can be calculated. It is important to point out that B_i is a function of w_i.

Substituting Equation (5) into Equation (2), the error term becomes

$$E_i = ||B_i w_i - v_i||^2 = (m||w_i|| - ||v_i||)^2. \tag{6}$$

Since we want to minimize E_i, $||w_i||$ should be as close to $||v_i||/m$ as possible. Notice $||w_i||^2$ is an integer in interval $[0, k]$. Therefore, we set,

$$||w_i|| = \min(\sqrt{round(||v_i||^2/m^2)}, k); \tag{7}$$

where the function $round()$ outputs the integer nearest to the input value. The binary strings w_i satisfying Equation (7) will be the optimal weights (binary representation) for the v_i. Once w_i is determined, B_i can be calculated as follows.

Let w_{i0} and v_{i0} be the normalized vectors of w_i and v_i respectively. We construct two matrices $P_{n\times k}$ and $Q_{k\times k}$ such that

$$P e_1 = v_{i0} \tag{8}$$

and,

$$Q e_1 = w_{i0} \tag{9}$$

and,

$$P^T P = Q^T Q = I, \tag{10}$$

where n denotes the length of v_i and e_1 denotes the unit vector $(1, 0, 0 \ldots 0)^T$ with length k. With the constructed P and Q, set

$$B_i = mPQ^T, \tag{11}$$

The constructed B_i satisfies Equations (1) and (5).

Here, P and Q in Equations (8) and (9) can be constructed using linear independent bases with the first basis vector equal to v_{i0} and w_{i0} respectively. After that, the orthonormal criterion in Equation (10) can be satisfied using the Gram-Schmidt Orthonormalization method.

The next step we need to do is to determine the normalization constant m which is given in Equation (1). The ranges of $||w||$ and $||v_A||$ are $[0, \sqrt{k}]$ and $[0, m\sqrt{k}]$ respectively. If we want $v_A \approx v$, the range of $||v||$ should be close to $[0, m\sqrt{k}]$. However, the value of $||v||$ is database dependent and there is no guarantee that it will be close to $[0, m\sqrt{k}]$. Therefore m is used to normalize v to the range $[0, \sqrt{k}]$ after transformation. m is calculated as

$$m = \sqrt{\text{mean}(||v_i||^2)/(k/2)}. \tag{12}$$

With this value, the optimal binary strings $\{w_1, w_2 \ldots w_c\}$ will have roughly 50% of bits as "0" and the rest as "1". (Notice that $||w_i||^2$ represents the number of "1"-bits in the binary string w_i). This provides high entropy of the binary strings.

3.2 Finding Binary Representation of the Query Templates

When a query face template v_i' is presented, the corresponding basis matrix B_i' is then extracted from database.

The binary representation w_i' of the query can then be calculated as:

$$w_i' = \underset{w}{\operatorname{argmin}} \parallel v_i' - B_i' \cdot w \parallel . \tag{13}$$

Here

$$\begin{aligned} & ||v_i' - B_i' \cdot w_i'||^2 \\ = \;& ||v_i'||^2 - 2v_i'^T \cdot B_i' \cdot w_i' + w_i'^T \cdot B_i'^T \cdot B_i' \cdot w_i' \\ = \;& ||v_i'||^2 - 2v_i'^T \cdot B_i' \cdot w_i' + m^2 \parallel w_i' \parallel^2 \end{aligned} \tag{14}$$

Let $v_i'^T \cdot B_i' = [\gamma_1, \gamma_1 \ldots \gamma_k]$, $w_i' = [b_1', b_2' \ldots b_k']^T$. Substitute these into Equation (14), it becomes

$$\begin{aligned} & \parallel v_i' - B_i' \cdot w_i' \parallel^2 \\ = \;& ||v_i'||^2 - 2\sum_{i=1}^{k} b_i'\gamma_i + m^2 \sum_{i=1}^{k} b_i'^2 \\ = \;& ||v_i'||^2 + m^2 \sum_{i=1}^{k} (b_i' - \frac{2\gamma_i}{m^2}) b_i'. \end{aligned} \tag{15}$$

Since b_i' is binary, $\parallel v_i' - B_i' \cdot w_i' \parallel^2$ will be minimum when:

$$b_i' = \begin{cases} 1 & \text{if} \quad \gamma_i > m^2/2; \\ 0 & \text{if} \quad \gamma_i \leq m^2/2. \end{cases} \tag{16}$$

The binary representation of v_i i.e. $w_i = [b_1', b_2' \ldots b_k']^T$ is thus calculated as mentioned above.

3.3 Procedure to Implement the Binary Linear Combination Scheme

The whole procedure to implement the binary linear combination (BLC) algorithm is described as follows:

In enrollment:

a) The representative templates v_i for each user are extracted from the enrolled templates;

b) Determine m and the binary representations w_i by Equation (12) and (7);

c) The bases B_i are computed with determined w_i and m;

d) w_i is encrypted and stored in database with B_i.

In authentication:

a) When a query v_i' is presented, the corresponding bases B_i' is released from the database;

b) w_i' is computed from B_i' and v_i';

c) v_i' and the stored v_i are matched for a decision.

4 Experimental results

This section reports the performance of our proposed binarization algorithm using binary linear combination representation. Three public domain face databases, namely CMU PIE, FERET and FRGC are selected to evaluate our algorithm. The Fisherface [1] algorithm is employed to generate the original face templates. The detailed parameters settings are shown in Table 1, where n_c is the number of individuals in the database, n_p denotes the total number of images from each individual, n_t is the number of images used for training from each individual and k is the number of basis vectors used in our proposed binary linear combination (BLC) algorithm. Moreover, the random multispace quantization (RMQ) [16] algorithm which employs thresholding technique, is also used for comparison. It is important to note that full RMQ algorithm with random projection and thresholding steps, is implemented for comparison.

Table 1. The experiment settings

Database	n_c	n_p	n_t	k
CMU PIE	68	105	10	50
FERET	250	4	2	200
FRGC	350	40	5	200

Figures 2-4 show the experimental results for the Fisherface method (labeled as original), RMQ method (labeled as RMQ) as well as our proposed method (labeled as BLC) on CMU-PIE, FERET and FRGC databases respectively. In all three databases, it can be seen that the proposed method outperforms the RMQ method as well as the original (Fisherface) method. The results also show that the proposed method not only preserves, but also improves the original template discriminability. Following the popular setting, we fix the FAR at 0.01 and compare the accuracy. The results are shown in Table 2. It can be seen that, in comparison to the thresholding-based method, the recognition accuracy of the proposed method is increased by around 7% on CMU-PIE and FERET databases and around 5% on FRGC database. We also measure the equal error rate (EER) and the results are shown in Table 3. In all three databases, our proposed method gives the lowest EER.

In addition to the ROC, we also perform the analysis using histograms. The histograms of the genuine and impostor matching score distributions for CMU-PIE, FERET and FRGC databases are plotted in Figures 5-7 respectively. The Figures (a), (b) and (c) show the distributions corresponding to original Fisherface algorithm, RMQ scheme and our proposed binary linear combination algorithm respectively. To evaluate the histograms, we follow the method reported in [16] and use the decidability index d introduced by Daug-

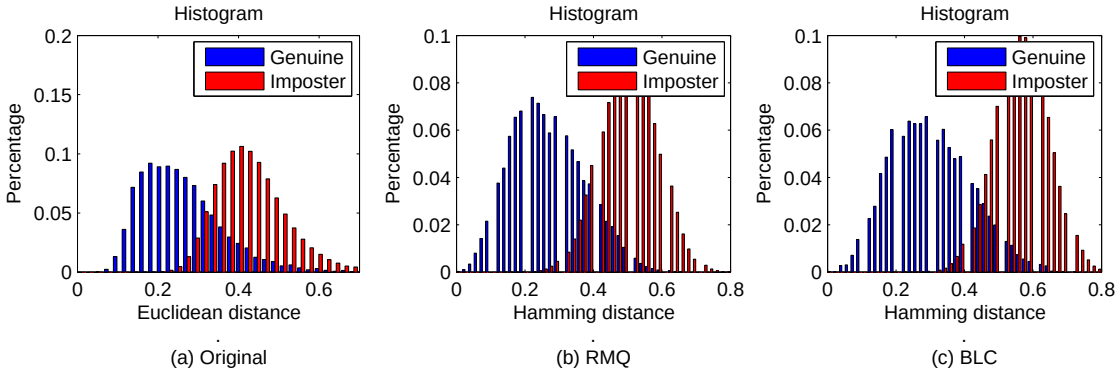


Figure 5. The genuine users and imposters distribution in CMU PIE database.

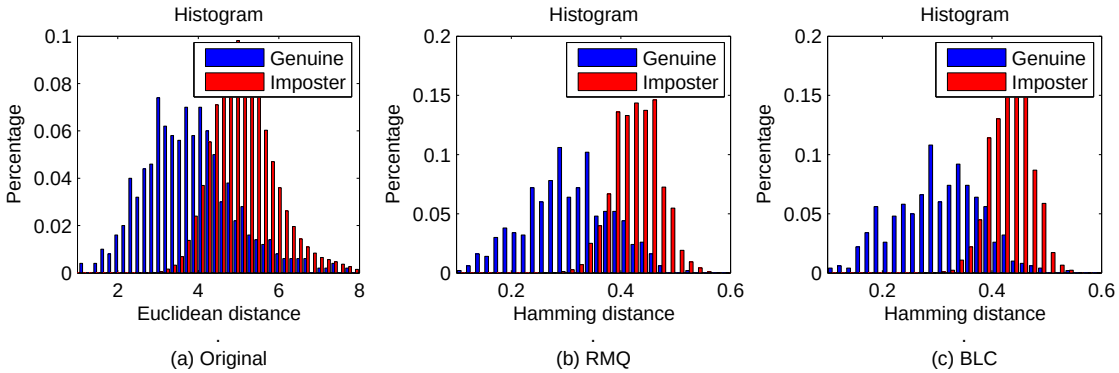


Figure 6. The genuine users and imposters distribution in FERET database.

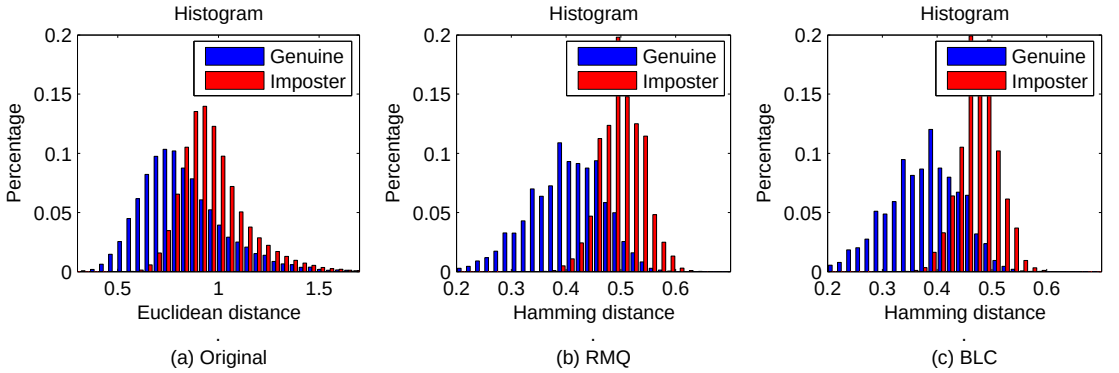


Figure 7. The genuine users and imposters distribution in FRGC database.

Table 2. The GARs(%) of the experiments with fixed FAR=0.01

Database	Original	RMQ	BLC
CMU PIE	59.26	66.17	73.28
FERET	45.47	62.25	69.24
FRGC	26.28	51.45	56.31

Table 3. The EERs(%) of the experiments

Database	Original	RMQ	BLC
CMU PIE	17.32	12.00	10.13
FERET	21.66	16.20	11.81
FRGC	31.75	17.56	16.83

Table 4. Evaluation of The Histograms

Database	Scheme	μ_g	μ_i	σ_g^2	σ_i^2	d
CMU PIE	Original	0.2610	0.4302	0.0108	0.0082	**1.74**
	RMQ	0.0115	0.0068	0.2699	0.5029	**2.43**
	BLC	0.3063	0.5662	0.0136	0.0063	**2.60**
FERET	Original	3.7857	5.1216	1.2489	0.6068	**1.39**
	RMQ	0.0061	0.0017	0.3045	0.4267	**1.95**
	BLC	0.3011	0.4334	0.0063	0.0013	**2.14**
FRGC	Original	0.8247	0.9753	0.0443	0.0287	**0.79**
	RMQ	0.0045	0.0015	0.4016	0.5003	**1.79**
	BLC	0.3768	0.4745	0.0044	0.0013	**1.85**

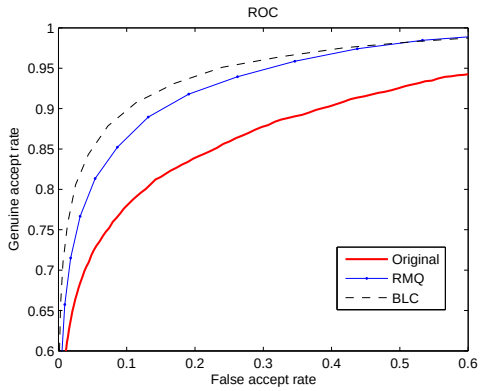


Figure 2. The experimental results of the CMU PIE database.

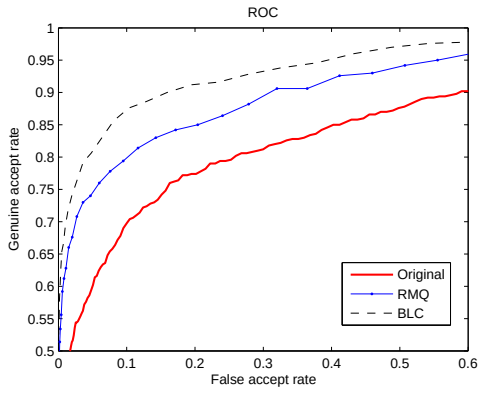


Figure 3. The experimental results of the FERET database.

man [4] as follows,

$$d = \frac{\mu_g - \mu_i}{\sqrt{\frac{1}{2}(\sigma_g^2 + \sigma_i^2)}}, \tag{17}$$

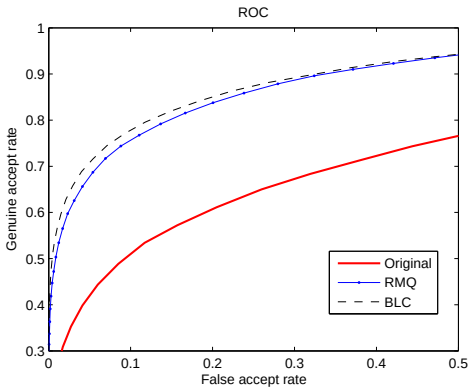


Figure 4. The experimental results of the FRGC database.

where μ_g and μ_i are the respective means of the genuine and imposter distributions, σ_g^2 and σ_i^2 are the respective variances. Note that d shows how well the genuine distribution is separated from the imposter distribution. The larger the value of d, the better the performance will be. The results on the Fisherface, RMQ and our proposed BLC method are shown in Table 4. It can be seen that our proposed method gives the best separation.

5 Analysis

A biometric template protection scheme should satisfy three requirements: discriminability, security and cancelability. That is, the protection scheme preserves the accuracy of the original system and the template stored in database should be secure and cancelable.

The accuracy performance of our proposed algorithm has been reported and discussed in Section 4. Experimental results shows that our algorithm gives very good performance.

The biometric system security relies on the security of the transformed binary template. Since binary template

stored in the database is encrypted, it is computationally hard to recover the original face template (eg. the MD5 [3] hashing provides a security level of 2^{128}). Attackers can only apply a brute-force attack to guess the original binary template bit by bit, resulting in a high security level of 2^k. However, since we select m such that the number of bits '1' and '0' are balanced in the binary templates w_i, attackers may assume that the stored binary templates have half of the bits to be '1' and half of the bits '0'. This will reduce their guessing times to $C_k^{[k/2]}$ ($[k/2]$ means the nearest integer to $k/2$). But still, the BLC algorithm gets a high security level when k is comparatively large (e. g. $k = 200$ the security level is about 2^{195}, only 5 bits lost in security level). Another thing should be mentioned is that the bases matrices B_i are stored in database without protection. As a result, they are exposed to attackers. This will offer attackers a relationship between v_i and w_i (referring to Equation (5)). However, since attackers know neither v_i nor w_i, the exposed B_i provides no useful information to attackers. It will not affect the security of our scheme.

The cancelability of our algorithm is also high because the reference binary templates are randomly generated with fixed number of bits '1' as shown in Section 3.1. If the binary template is compromised, it can be cancelled and a new reference template can be generated for replacement. So our algorithm has high cancelability capability.

6 Conclusion

This paper has proposed and reported a new method to generate a binary face template from a real valued face template. A new binary linear combination method is proposed to represent a real valued face template and the binary weights of the basis are then used as the binary representation. Since there is no standard method in determining the binary linear combination, a new method is proposed in order to minimize the approximation error. Three public domain available face databases have been used to evaluate the proposed method. The experimental results show that the proposed method outperforms the existing thresholding-based template protection method.

References

[1] P N Belhumeur, J P Hespanha and D J Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Trans. on PAMI*, 19(7), pp. 711-720, 1997.

[2] S Z Li, "Face Recognition Based on Nearest Linear Combinations," *Computer Vision and Pattern Recognition*, pp. 839-844,1998.

[3] R L Rivest, "The MD5 Message-Digest Algorithm," *RFC1321, Network Working Group, MIT Laboratory for Computer Science and RSA Data Security, Inc.*, 1992.

[4] J Daugman, "The Importance of Being Random: Statistical Principles of Iris Recognition," *Pattern Recognition*, Vol. 36, No. 2, pp. 279-291, 2003.

[5] N Ratha, J Connell and R Bolle, "Enhancing security and privacy in biometric-based authentication systems," *IBM Systems Journal*, Vol. 40. No. 3, pp. 614 - 634, 2001.

[6] S Prabhakar, S Pankanti and A K Jain, "Biometric Recognition: Security and Privacy Concerns," *IEEE Security and Privacy Magazine*, Vol. 1, No. 2, pp. 33-42, March-April 2003.

[7] A K Jain, K Nandakumar and A Nagar, "Biometric template security," *EURASIP Journal on Advances in Signal Processing*, Vol. 8, 2008.

[8] U Uludag, S Pankanti, S Prabhakar and A K Jain, "Biometric cryptosystems: issues and challenges," *Proceedings of the IEEE*, vol. 92, no. 6, pp. 948-960, 2004.

[9] A K Jain, A Ross and S Pankanti, "Biometrics: A Tool for Information Security", *IEEE Transactions on Information Forensics and Security*, Vol. 1, No. 2, pp. 125-143, 2006.

[10] A Alder, "Images can be regenerated from quantized biometric match score data", *Proceedings of Canadian conference of Electrical and Computer Engineering*, pp. 469-472, 2004.

[11] F Monrose, M K Reiter and S Wetzel, "Password Hardening Based on Key Stroke Dynamics," *Proc. ACM Conf. Computer and Comm. Security*, pp. 73-82, 1999.

[12] F Monrose, M Reiter, Q Li and S Wetzel, "Cryptographic Key Generation from Voice," *Proc. IEEE Symp. Security and Privacy*, pp.202-213, May 2001.

[13] A Teoh, D Ngo and A Goh, "Biohashing: Two Factor Authentication Featuring Fingerprint Data and Tokenised Random Number," *Pattern Recognition*, vol. 37, no. 11, pp. 2245-2255, Nov. 2004.

[14] A Goh and D C L Ngo, "Computation of cryptographic keys from face biometrics", *in Proc. 7th IFIP TC6/TC11 Conf. Commun. Multimedia Security*, vol. 22, pp. 1-13, 2003.

[15] D Ngo, A Teoh and A Goh, "Biometric Hash: High-Confidence Face Recognition", *IEEE transactions on circuits and systems for video technology*, vol. 16, no. 6, 2006.

[16] A Teoh, A Goh and D. Ngo, "Random Multispace Quantization as an Analytic Mechanism for BioHashing of Biometric and Random Identity Inputs," *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 1892-1901, Dec. 2006.

[17] N Ratha, S Chikkerur, J Connell and R Bolle, "Generating Cancelable Fingerprint Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.29, no.4, pp. 561-572, 2007.

[18] Y C Feng, P C Yuen and A K Jain, "A Hybrid Approach for Face Template Protection," *Proceedings of SPIE Defense and Security Symposium*, 2008.

[19] Y C Feng and P C Yuen, "Class-Distribution Preserving Transform for Face Biometric Data Security," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 141-144, 2007.

[20] Y C Feng and P C Yuen, "Selection of Distinguish Points for Class Distribution Preserving Transform for Biometric Template Protection," *Proceedings of IEEE International Conference on Biometrics (ICB)*, pp. 636-645, 2007.

A SemiBoosted Co-Training Algorithm For Human Action Recognition*

Chang LIU

Abstract

This paper proposes a SemiBoosted Co-Training method for human action recognition. Two confidence measures namely inter-view confidence and intra-view confidence are proposed and estimated from co-training and self-training perspectives, and are dynamically fused into one semi-supervised learning process. For co-training, two discriminative views from temporal and spatial information of the video, namely action saliency view and action eigen-projection view, are proposed for the training process. For self-training, a multi-class SemiBoost algorithm is introduced in order that the performances of classifiers in different views are effectively boosted in each iteration. Given a small set of labeled videos and a large set of unlabeled videos, the proposed semi-supervised learning algorithm trains a classifier by dynamically incorporating a set of unlabeled data into the labeled data set, the performance of the classifier will be improved in each iteration. The KTH and Weizmann human action database are used to evaluate our proposed algorithm, average recognition accuracy of 99.2% *and* 91.3% *are obtained respectively.*

1. Introduction

[1] Human action recognition has been receiving increasing attentions from researchers in computer vision community. The aim of human action recognition is to recognize human actions from videos so that the system could understand the scene, and make further classification or semantic description of the scene [18] [10] [31] [17] [30]. The results can be applied to many applications such as visual surveillance, human-computer interfaces, video summarization, content based video retrieval etc. Human action recognition is a challenging research area because the dynamic human body motions have almost unlimited underlying representations, there are also difficulties from perspective distortions, different viewpoints and illumination variations. To recognize human actions, an action model [26] [19] is often required. In order to train a good action model, a large amount of labeled data should be needed so that there are sufficient sample features for training [15] [28] [7], and achieve a strong generalization ability. However, labeled videos are often very costly to obtain because they require much of human efforts, while the unlabeled videos can be easily obtained from public surveillance cameras. In this case, how to fully use the labeled data and more importantly, how to employ the large amount of unlabeled data to boost the performance of the overall system is a crucial problem.

Semi-supervised learning is currently a hot topic that utilizes the labeled data as well as a large amount of unlabeled data to learn the hypothesis [34]. It shows great advantage in automatically exploiting huge amount of information from the unlabeled data and boost the generalization ability of the trained system. For extracting specific information from the unlabeled data, a number of semi-supervised learning methods are proposed. Graph-based methods [1] [27] [2] construct graphs in which the nodes represent labeled and unlabeled samples, and weighted edges represent the similarity among the samples. These kinds of methods often have label smoothness assumptions and they are said to be transductive in nature. SemiBoost [16] is a semi-supervised improvement graph-based algorithm. By using a boosting framework, SemiBoost has the ability to improve the performance of any given supervised classifier in the presence of unlabeled samples. The SemiBoost has already been successfully employed in image analysis applications [13], while only two-class classification problem is considered.

Co-training [4] is a multi-view semi-supervised learning algorithm. In co-training, two learners are trained from the same set of examples, with each learner using an ideally independent set of features for each example. During the co-training process, each learner iteratively labels several unlabeled examples that show highest confidence value from its point of view, and these newly labeled examples are added to the labeled training set of the other learner. Therefore, the labeled training set of both learners will be simultaneously augmented. By assuming the two views are conditionally independent and each view is sufficient for learning a classifier, it has been shown that any weak hypothesis can be boosted from the unlabeled data [5]. However, the problem is that the confidence of a newly labeled data can only be measured on one learner while ignoring the confidence of the other learner. Therefore, the performance of the over-

[1]* This paper has been submitted to IEEE International Conference on Computer Vision and Pattern Recognition 2009.

all system is largely depended on the selection of the two learners.

This paper proposes a SemiBoosted Co-Training algorithm for human action recognition. In order to automatically label new samples and utilize them as labeled training data, both co-trained the inter-view confidence measured from other views and the self-trained intra-view confidence measured from self view on unlabeled data are used, so that different views measure their confidence on unlabeled data in an interacted manner. Two discriminative views from temporal and spatial information of the video, namely action saliency view and action eigen-projection view, are proposed for the training process. Simultaneously, each view provides confidence information of the unlabeled data to the other view, while the other view decides the labeled data from both the inter-view confidence and intra-view confidence by a Multi-class SemiBoost process. Therefore, the decisions from both views are effectively boosted in each iteration with new unlabeled data adding into the labeled data set for training, and the performance of the classifiers are improved cooperatively.

2. Related Work

In literature, there are many existing research works on human action recognition, two good survey papers were reported [10] [18]. In this section, we mainly focus on existing methods which are related to our work in this paper.

For representing human actions, space-time based approach becomes more and more popular recently because it does not require segmentation or tracking of the human. Shechtman et al. [25] have recently proposed a spatio-temporal patch correlation based method for human activity recognition. Small spatio-temporal reference volumes are correlated against the entire video sequences in the target volume. The overall peak correlation values shows the matched activities. Yilmaz et al [33] extracted differential geometry features from the 3D contour of the action volume. The 3D contour is then projected to a 2D surface. The projection on the time axis forms the new spatio-temporal volume. Then the human moving speed, moving direction and human shape can be extracted from the volume. Activity recognition can then be performed. Gorelick et al [8] [3] utilized the properties of the poisson equation solution to analyze the spatio-temporal volume. Three dimensional space-time shapes are generated from the silhouettes of the spatio-temporal volume. The space-time salient features are then extracted from the space-time shape. It shows that these features are very useful for activity recognition. Niebles et al [20] proposed a mixture hierarchical model for human activity recognition based on spatial and spatio-temporal features. They showed that static shape features can improve the recognition performance when using the spatio-temporal features.

Another approach represent human actions by spatial-temporal interest points. Laptev [12] proposed a space-time interest point detector. It can find local salient pixels in space-time volume where the pixel values have significant local variations in both spatial and temporal domain, the local saliency maxima is detected based on the Harris operator. However, this method detect a small number of stable interest points which may not sufficient to characterize complex events. Dollar et al. [6] used separable linear filters in the spatio-temporal volume and detected interest points. A number of descriptors are then proposed for each interest point. Oikonomopoulos et al. [22] proposed to expend spatial salient region detector to spatio-temporal volume. Two sets of spatio-temporal salient points are detected and compared by chamfer distance. The result is shown to be promising.

Alternatively, some researchers do not work on learning an action model from the labeled action data, but directly learning from unlabeled action dataset in a unsupervised manner. This is because that labeled videos are often very costly to obtain as they require much of human efforts, while unlabeled videos are much easier to obtain. Niebles et al. [21] quantize local space-time features and represent human actions as a bag of spatio-temporal words by extracting space-time interest points from unlabeled action videos. Then probabilistic latent semantic analysis model and latent dirichlet allocation model are used to classify these action videos. There are also other researchers develop algorithms to directly extract discriminative action features from unlabeled data [32] [29]. However, there are very few semi-supervised learning methods for human action analysis, which can fully use both the labeled and unlabeled data. Guan et al. [9] proposed an En-Co-training method to make use of the unlabeled action videos. They use different classifiers namely decision tree, Naive Bayes and k-nearest neighbors for the three views. It showns that the learning performance can be improved by utilizing the unlabeled data. But they did not consider the independence of different views, the comparative experimental results with the state of the art methods on publicly dataset are not reported.

3. SemiBoosted Co-Training for Human Action Recognition

This paper proposes a new semi-supervised learning framework for human action recognition. Both labeled and unlabeled data are used for training the system, while the labeled data set is augmented in each iteration from the inter-view confidence and intra-view confidence. The confidence is decided by the two classifiers trained in the two views. The block diagram of our proposed framework is shown in Figure 1. It is an online learning process, in each iteration, simultaneously, the existing labeled data will be used for

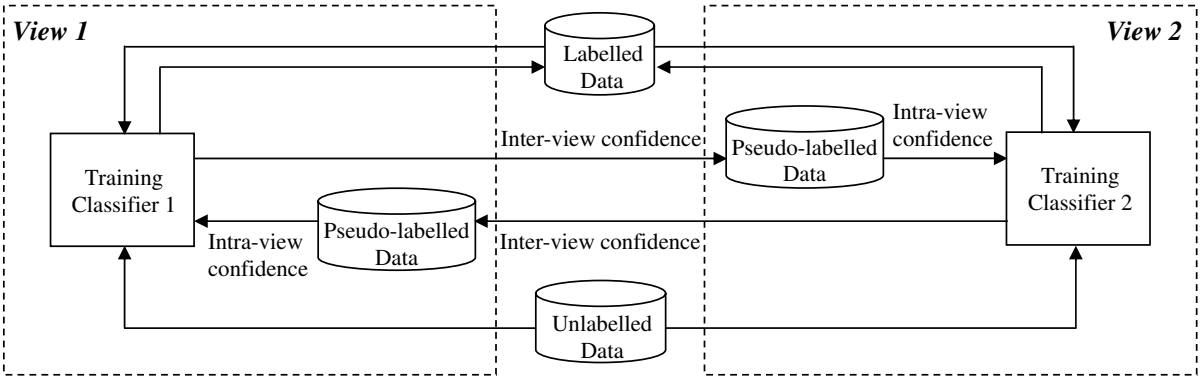


Figure 1. Block diagram of SemiBoosted Co-Training.

training two classifiers from two views. For each classifier, it calculates its confidence on the unlabeled data, and select a set of pseudo-labeled data which have high confidence (inter-view confidence). The other classifier will then decide a set of labeled data from the pseudo-labeled data from its own confidence (intra-view confidence), and add the newly labeled data into the labeled data set. After the labeled dataset is augmented, two new classifiers will be trained from the labeled dataset, and this iterative process continues.

3.1. Data Representation

To represent video data from two views as shown in Figure 1, a logical way is to consider the video data as a spatio-temporal volume [11] [23], and extract the spatial and temporal information separately from the video. It is reasonable to consider the spatial information and temporal information are conditionally independent, and these two kinds of information have the properties to be co-trained in one framework. In this paper, we adopt the Information Saliency Map (ISM) [14] to represent the video data. The ISM was built from the video while each entry of ISM reflects the saliency of the corresponding pixel in that video frame. Considering the current frame Im_0 which can be divided into $h \times w$ smaller patches $\{Im_{1,1}, Im_{1,2}, ..., Im_{h,w}\}$, the ISM for Im_0 can be obtained by the Spatio-Temporal ISM model [14]

$$I_{r,s} = -\log_2\left(\frac{P(X|V)P(x_0|X)+P(X'|V)P(x_0|X')}{P(X|V)[1-P(x_0|X)]+P(X'|V)[1-P(x_0|X')]}\right) \quad (1)$$

where x_0 is the $< m \times 1 >$ vector form of $Im_{i,j}$, the temporal vector set $X = \{x_0, x_1, ..., x_{N-1}\}$ is constructed from concatenating N temporal patches located at $Im_{r,s}$, which is called a sub-volume. The spatial vector set $X' = \{x'_0, x'_1, ..., x'_{N'-1}\}$ is constructed by the patch x_0 in Im and its $N'-1$ spatial neighborhoods. V is chosen as the spatio-temporal cube that contains X and X', where $\{X, X'\} \subset V$. $P(x_0|X)$ is the conditional probability of x_0 given X. A Gaussian kernel is then used to estimate the conditional probability and Eq.(1) becomes Eq.(2) and can be solved.

$$\hat{f}(y) = \frac{1}{(2\pi)^{q/2}N}\sum_{i=0}^{N-1}[(D_{KL}(\hat{f}||\hat{f}_i))^{-q/2} \cdot \exp(-\tfrac{1}{2}(y-y_i)^T(D_{KL}(\hat{f}||\hat{f}_i))^{-1}I)(y-y_i))] \quad (2)$$

where kullback-Leibler divergence D_{KL} is used to measure the similarities between density functions. After the Information Saliency Map is calculated, an information saliency curve can be generated as shown in Figure 2. The images show five key frames of a person walking and their corresponding ISM, the information saliency curve is obtained from object overall saliency value along time axis. Three primitive actions are then detected from the curve where each primitive action is represented by a complete saliency changing period, and the primitive actions are noted as Salient Action Unit (SAU) [14]. Each curve in one SAU are represented as action information saliency vector set $V_S = \{v_s^1, v_s^2, ..., v_s^m\}$. The SAU is then used for projecting the original video to the eigen space, where we get another vector set named human action eigen-projection vector set $V_E = \{v_e^1, v_e^2, ..., v_e^n\}$. It can be seen that V_S and V_E represent the temporal and spatial view of the video data respectively. They are considered to be conditionally independent, and will be used as the raw data for training the semi-supervised action recognition system.

3.2. Inter-view Confidence

We propose to fully utilize the unlabeled data for training the human action recognition system. Inspired from the co-training method, two different kinds of feature sets are extracted from the original data, namely action saliency vector set V_S and action eigen-projection set V_E. These two sets are used to train two different classifiers corresponding to two different views. Two classifiers are then integrated in one system and interacted with each other as shown in Figure 1.

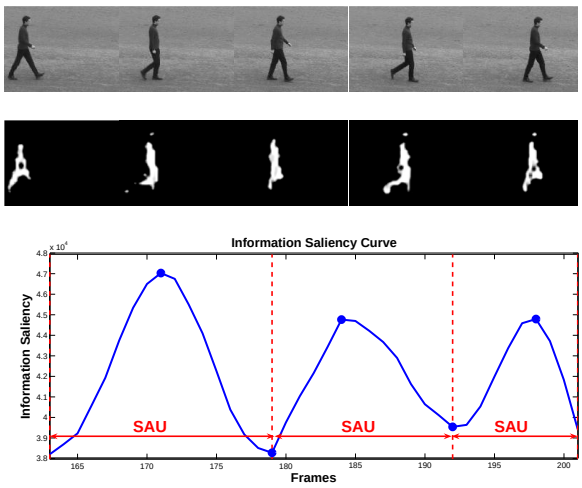


Figure 2. Information saliency curve generated from "*person*01_*walking_d*1.*avi*" in KTH human action database [24], images in the first row are from the original video, the second row shows the ISM for each frame. An information saliency curve is shown in the third row, five solid points represent the five information saliency values of the five frames. Notice that three Salient Action Units are generated from this curve.

We consider that the information obtained from temporal context and spatial context in videos are conditionally independent to each other. The action saliency vector set $V_S = \{v_s^1, v_s^2, ..., v_s^m\}$ is generated from the information saliency curve, which measures pixel densities in time axes. The action eigen-projection vector set $V_E = \{v_e^1, v_e^2, ..., v_e^n\}$ measures spatial information in the eigen space. Therefore, we consider that V_S and V_E are conditionally independent to each other. And it has been shown in [14] that these vectors are useful for training an action recognition system. So these two feature sets can be used as two views for co-training.

Co-training method can boost the performance of individual classifier while augmenting the labeled data set, but the problem is that the confidence of a newly labeled data can only be measured on one learner while ignoring the confidence of the other learner. Therefore, if only the co-trained inter-view confidence is used, the performance of the overall system is largely dependent on the selection of the two learners. As the labeling process does not has a re-validation step, the newly labeled data will then be used as labeled training data. Therefore, the confidence of the unlabeled data in both views should be measured before the labeling process. To solve this problem, we further introduce self-trained intra-view confidence into the whole semi-supervised learning framework, where the confidence of both learners are incorporated. As shown in Figure 1, the unlabeled data with high confidence in one learner will not be immediately given a label, but a pseudo-label instead. The other view decides whether the data should be labeled by further considering the intra-view confidence measure.

3.3. Intra-view Confidence

To fully incorporate the confidence of both views, we propose to introduce two intra-view confidences by a self-training process together with the inter-view confidence measured by co-training. In practice, the learner h_S trained from action saliency vector set $V_S = \{v_s^1, v_s^2, ..., v_s^m\}$ decides a set of unlabeled data that h_S has the highest confidence, then these unlabeled data are given pseudo-labels and augment the pseudo-label data set for the other view. The learner h_E is then updated by a self-training process from the labeled data and pseudo-labeled data. Some pseudo-labeled data which have the highest intra-view confidence are given their labels and added into the labeled data set. The rest of the data from the pseudo-label data set are added back to the unlabeled data set. Simultaneously, the learner h_S also labels a set of data from the pseudo-labeled data set into the labeled dataset.

For self-training, we employ the SemiBoost algorithm [16] and extend it for multi-class classification task. SemiBoost is a semi-supervised improvement algorithm which is able to boost any supervised learning algorithm in each iteration given both labeled data and unlabeled data for training. With the assumption that the unlabeled data having high similarity often share the same labels. Because the original SemiBoost algorithm can be only applied in two-class classification problems, it is necessary to extend this algorithm to multi-class case, so that it can be applied for human action recognition applications . The multi-class objective function that measures the overall inconsistency between data is shown in Eq.(3), by minimizing this objective function, the inconsistency between all the training samples will be minimized

$$F(\mathbf{x}, \mathbf{y}, S) = F_l(\mathbf{x}, \mathbf{y}, S) + F_u(\mathbf{x^u}, \mathbf{y^u}, S) \tag{3}$$

where x represents the mixture of labeled and unlabeled data set, y is the mixture label set, x^u is the unlabeled data set, y^u is the label set for these unlabeled data, S is the similarity matrix. F_l measures the overall inconsistency between all the labeled data and unlabeled data, F_u measures the overall inconsistency among unlabeled data.

Because the space is limited, we solely show the important steps of the whole derivation. By employing the harmonic function approach to measure data similarity, the objective function Eq.(4) can be obtained,

$$F(\mathbf{x}, \mathbf{y}, S) = \frac{1}{n^l n^u} \sum_{i=1}^{n^l} \sum_{j=1}^{n^u} S_{i,j} \exp(-2y_i^l y_j^u) + \frac{1}{n^u n^u} \sum_{i=1}^{n^u} \sum_{j=1}^{n^u} S_{i,j} \exp(y_i^u - y_j^u) \tag{4}$$

where n_l is the total number of labeled data, and n_u is the total number of unlabeled data.

By minimizing the objective function Eq.(4). The optimal class label for the unlabeled data and optimal sampling weight can be achieved in each iteration during the boosting process. Considering the multi-class problem can be divided into a set of two-class classification problems. From this aspect, each binary decision $[\![y_i^u = \ell]\!]$ has two labels of $\{1,-1\}$, and its label can be estimated from the Adaboost classifier as $\sum_{\ell \in L}(H_i(x_i,\ell)+\alpha h_i(x_i,\ell))$, where L is the whole label set, then we obtain

$$\begin{aligned} F(\mathbf{x},\mathbf{y},S) &= \frac{1}{n^l n^u}\sum_{i=1}^{n^l}\sum_{j=1}^{n^u}\sum_{\ell\in L} S_{i,j}\exp(-2y_i^l(H_j(x_j,\ell)+\alpha h_j(x_j,\ell))) \\ &+ \frac{1}{n^u n^u}\sum_{i=1}^{n^u}\sum_{j=1}^{n^u}\sum_{\ell_1\in L}\sum_{\ell_2\in L} S_{i,j}\exp(H_i(x_i,\ell_1)+\alpha h_i(x_i,\ell_1)) \\ &\cdot \exp(-H_j(x_j,\ell_2)-\alpha h_j(x_j,\ell_2)) \end{aligned} \tag{5}$$

To minimize the objective function Eq.(5), we find its upper bound as

$$\begin{aligned} F(\mathbf{x},\mathbf{y},S) &\le \frac{1}{n^l n^u}\sum_{i=1}^{n^l}\sum_{j=1}^{n^u}\sum_{\ell\in L} S_{i,j}\exp(-2y_i^l H_j(x_j,\ell))\exp(-2y_i^l\alpha h_j(x_j,\ell)) \\ &+ \frac{1}{n^u n^u}\sum_{i=1}^{n^u}\sum_{j=1}^{n^u}\sum_{\ell_1\in L}\sum_{\ell_2\in L} S_{i,j}\exp(H_i(x_i,\ell_1)-H_j(x_j,\ell_2)) \\ &\cdot \exp(\alpha h_i(x_i,\ell_1)\delta(y_i,\ell_1)) \end{aligned} \tag{6}$$

We set the upper bound in Eq.(6) as $\overline{F_1}$, it can be further represented as Eq.(7)

$$\begin{aligned} \overline{F_1} &= \frac{1}{n^l n^u}\sum_{i=1}^{n^l}\sum_{j=1}^{n^u}\sum_{\ell_1\in L}\sum_{\ell_2\in L} S_{i,j}\exp(-2H_j(x_j,\ell_1)-2\alpha h_j(x_j,\ell_1))\delta(y_i,\ell_2) \\ &+ \frac{1}{n^u n^u}\sum_{i=1}^{n^u}\sum_{j=1}^{n^u}\sum_{\ell_1\in L}\sum_{\ell_2\in L} S_{i,j}\exp(H_i(x_i,\ell_1)-H_j(x_j,\ell_2)) \\ &\cdot \exp(\alpha h_i(x_i,\ell_1)\delta(y_i,\ell_1)) \end{aligned} \tag{7}$$

It can be seen that this upper bound of total inconsistency also consist two parts: $F_l(\mathbf{x},\mathbf{y},S)$ and $F_u(\mathbf{x^u},\mathbf{y^u},S)$, we combine these two parts into one expression and further divide the expression into $\overline{F_1}$ and $p_{i,k}$ as Eq.(8) is shown

$$\begin{aligned} \overline{F_1} &= \frac{1}{n^u}\sum_{i=1}^{n^u}\sum_{\ell_1\in L}\sum_{k\in L}\exp(\alpha h_i(x_i,\ell_1))\delta(y_i,k)p_{i,k} \\ p_{i,k} &= \frac{1}{n^l}\sum_{j=1}^{n^l} S_{i,j}\exp(H_i(x_i,\ell_1))\delta(y_j,k)+ \\ &\frac{1}{n^u}\sum_{j=1}^{n^u}\sum_{\ell_2\in L} S_{i,j}\exp(H_i(x_i,\ell_1)-H_j(x_j,\ell_2)) \end{aligned} \tag{8}$$

where p_i^k can be interpreted as the confidence in classifying the unlabeled data x_i to class k. Considering $n_u \gg n_l$, we find the upper bound of $\overline{F_1}$ as

$$\begin{aligned} \overline{F_1} \le \overline{F_2} = &\frac{1}{n^u}\sum_{i=1}^{n^u}\sum_{\ell_1\in L}\sum_{k\in L} p_{i,k}(\alpha h_i(x_i,\ell_1)\delta(y_i,k)-1)- \\ &\frac{1}{n^u}\sum_{i=1}^{n^u}\sum_{\ell_1\in L}\sum_{k_1\in L}\sum_{k_2\in L} 2\alpha h_i(x_i,\ell_1)(p_{i,k_1}-p_{i,k_2}) \end{aligned} \tag{9}$$

The upper bound $\overline{F_2}$ is composed of two terms, of which the first is independent of h_i. Therefore, to minimize the objective function at each iteration, the optimal class label for y_i^u and its weight w_i should be

$$\widehat{y_i^u} = \underset{k_1,k_2\in L}{argmax_i}(p_{i,k_1}-p_{i,k_2}) \tag{10}$$

$$\widehat{w_i} = |p_{i,k_1}-p_{i,k_2}| \tag{11}$$

The detailed SemiBoosted Co-Training algorithm is shown in Algorithm 1. It shows that after the y_i^u and w_i are obtained, in each iteration, a set of unlabeled data will be labeled and added into the labeled data set, and classifiers will be updated and improved

Algorithm 1 SemiBoosted Co-Training Algorithm

- Given labeled video example set $\{L\}$, unlabeled video example set $\{U\}$
- For the labeled example set $\{L\}$, calculate the two views of vector set: action saliency vector set $X_S^l = \{x_s^1,...,x_s^{n^l}\}$ and action eigen-projection vector set $X_E^l = \{x_e^1,...,x_e^{n^l}\}$.
- Create a sub data set U' by choosing u examples from U randomly

Do for k=1,2,...,K

- ▶ For unlabeled example set $\{U'\}$, calculate the two views of vector set: $U_S^u = \{u_s^1,...,u_s^{n^u}\}$, $U_E^u = \{u_e^1,...,u_e^{n^u}\}$
- ▶ Use X_S^l to train a classifier h_S, use h_S to label example set U_S^u from U'
- ▶ Use X_E^l to train a classifier h_E, use h_E to label example set U_E^u from U'
- ▶ Compute the pairwise similarity matrix $S_{i,j}^S$ between any two examples in X_S^l and U_S^u, and $S_{i,j}^E$ between any two examples in X_E^l and U_E^u

Do for t=1,2,...,T

- ◇ Compute $p_{i,k}$ for every example in $\{X_S^l, U_S^u\}$ and in $\{X_E^l, U_E^u\}$ using Eq.(8).
- ◇ Compute class label $\underset{k_1,k_2\in L}{argmax_i}(p_{i,k_1}-p_{i,k_2})$ for each example, sample new example sets U_S^l and U_E^l from U_S^u and U_E^u respectively by Eq.(11)
- ◇ Use $\{X_S^l, U_S^l\}$ and $\{X_E^l, U_E^l\}$ to train Adaboost.M2 classifiers h_t^S and h_t^E respectively, compute their weights α_t^S and α_t^E
- ◇ Update the strong classifier $H^S = \sum_t \alpha_t h_t^S$, $H^E = \sum_t \alpha_t h_t^E$

- ▶ Add all the $\aleph$ labeled examples to L
- ▶ Randomly choose $\aleph$ examples from U to replenish U'

- Obtain final classifier $H \leftarrow H^S + H^E$

4. Experimental Results

We perform extensive experiments to evaluate our proposed SemiBoosted Co-Training method on publicly avail-

able datasets. Details on the experiments and comparative results are given below.

4.1. Dataset and Experimental Settings

We evaluate our proposed method with two human action databases, namely Weizmann [8] human action database and KTH [24] human action database. Weizmann database contains 90 low-resolution (180×144) video sequences from nine people, each performing 10 natural actions: 'run', 'walk', 'skip', 'jumping jack', 'jump forward on two legs', 'jump in place on two legs', 'galloping sideways', 'wave one hand', 'wave two hands' and 'bend'. All the videos are captured from a fixed viewpoint. KTH database contains 600 low-resolution (160×120) video sequences from 25 people, each performing 6 natural actions: 'boxing', 'handclapping', 'handwaving', 'jogging', 'running' and 'walking'. Each action is performed under 4 different conditions: outdoors, outdoors with scale variations, outdoors with different clothes and indoors. Each video sequence contains one person repeatedly performing one action. This dataset has challenges of scale changes, action frequency changes and illumination changes.

Due to the periodic nature of the actions, we obtained 412 primitive actions from Weizmann dataset and 9572 primitive actions from KTH dataset. These primitive actions are the SAUs that represent the repeated actions. For calculating the ISM, we use temporal window size=20 and patch size=4. Then we get the action saliency vector set V_S and the action eigen-projection vector set V_E. Both the saliency vectors and the eigen-projection vectors are normalized before training. A multi-class Adaboost.m2 classifier is employed as the basic supervised learner for co-training. The termination for the iterative training is controlled by setting training error rate threshold for the learner.

4.2. Results and Analysis

We test our proposed SemiBoosted Co-Training algorithm on the KTH dataset with different unlabeled data sizes are used. For the whole primitive action dataset that consists 9572 samples, we randomly select 20% of samples as the testing data. For the remaining 80 % of samples, 5 % are randomly selected as the labeled training data, the remaining 75 % are the unlabeled training data. So there is no overlapping between the training data and the testing data. The results for using 10 %, 20 %, 40 % and 75 % of the total dataset as unlabeled dataset for training are shown in Figure 3. It can be clearly seen from the figure that with more unlabeled training data, the performance will be better. If all the 75 % of the total data are used as unlabeled data for training, we obtain the 92.7% accuracy rate at false alarm rate 5%. Because the size of Weizmann dataset is relatively small, the difference of choosing different ratio of unlabeled data is not obvious. We did not perform this experiment on

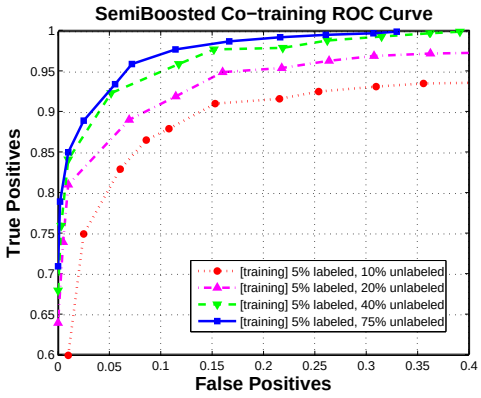


Figure 3. Comparative results of our proposed SemiBoosted Co-Training for using 10%, 20%, 40% and 75% of the total dataset as unlabeled data for training

Runs	CT	View_S	View_E	SBCT
1	0.51±0.076	0.51±0.048	0.56±0.021	0.54±0.028
2	0.57±0.050	0.64±0.032	0.69±0.013	0.67±0.016
3	0.60±0.048	0.74±0.033	0.77±0.017	0.76±0.023
4	0.67±0.036	0.80±0.026	0.82±0.019	0.81±0.022
5	0.70±0.059	0.78±0.021	0.83±0.007	0.81±0.011
6	0.71±0.041	0.81±0.038	0.84±0.020	0.82±0.027
7	0.72±0.035	0.82±0.017	0.86±0.005	0.84±0.009
8	0.73±0.042	0.84±0.026	0.85±0.022	0.84±0.024
9	0.73±0.051	0.83±0.029	0.88±0.014	0.86±0.021
10	0.74±0.037	0.85±0.023	0.89±0.014	0.87±0.017

Table 1. Results for the first 10 iterations of training process, Co-Training (CT), action saliency view classifier ($View_S$), action eigen-projection view classifier ($View_E$), SemiBoosted Co-Training (SBCT)

Weizmann dataset, but we have comparative results for this dataset later.

During the Semi-supervised learning process, the two classifiers are boosted by their intra-view confidence. We perform experiments to test the classifiers in both views individually, the results for the first 10 iterations are shown in Table 1. The first column shows results for co-training method, where the intra-view confidence is not considered. It can be seen that the performance of the two classifiers improved with more iterations. The final classifier is combined by these two classifiers by majority voting. From the overall performance, it shows that our proposed SemiBoosted Co-Training method outperforms the original co-training method. This shows that the incorporation of inter-view confidence and intra-view confidence improves the accuracy of labeling new data.

To compare our proposed method with supervised learning method, we use the same testing dataset (20 % ran-

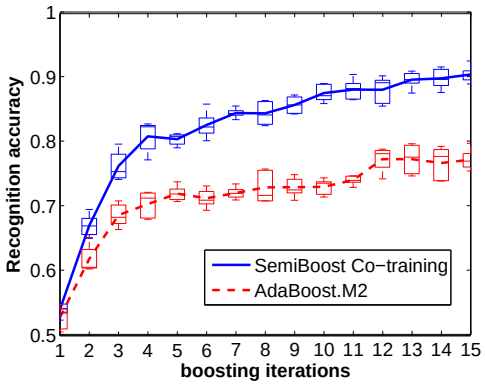
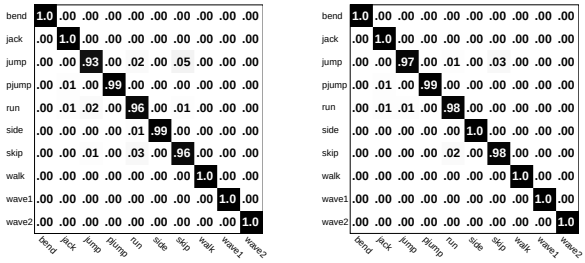
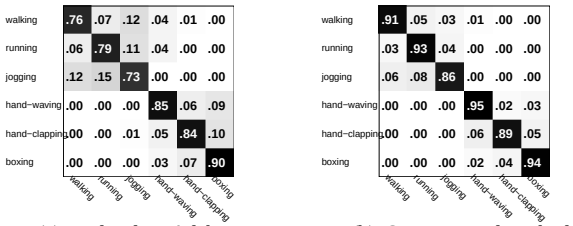


Figure 4. Comparison results for SemiBoosted Co-Training and Multi-class Adaboost [14]. 20% data are randomly selected for testing. For Multi-class Adaboost, the other 80% data are used for training, for SemiBoosted Co-Training, 5% of the remaining data are used as labeled training data, the other 75% are used as unlabeled training data



(a) Multi-class Adaboost (b) Our proposed method

Figure 5. Confusion matrix results on Weizmann dataset (a)Multi-class Adaboost[14], average accuracy:98.3% and (b) Our proposed SemiBoosted Co-Training method, average accuracy:99.2%



(a) Multi-class Adaboost (b) Our proposed method

Figure 6. Confusion matrix results on KTH dataset (a)Multi-class Adaboost[14], average accuracy:81.2%, and (b) Our proposed SemiBoosted Co-Training method average accuracy:91.3%

dom selected from the total data) for supervised learning, the other 80 % are all used for labeled training data. We use leave-one-out cross-validation scheme, where 24 subjects are used for training, the remaining one for testing, the procedure is repeated for 10 permutations, and the results are averaged. We compare our method with Multi-class Adaboost algorithm [14] because it is a Adaboost classifier learned in a supervised manner. The comparison results are shown in Figure 4. Each point on the boxplot shows the lowest accuracy, lower quartile, median, upper quartile, and largest accuracy, while the curve shows the average observation of accuracy. As can be seen, our proposed method outperforms the supervised learning algorithm of Multi-class Adaboost.

We also analyze our proposed method by calculating accuracy for each action class on Weizmann and KTH databases. The confusion matrix results are shown in Figure 5 and Figure 6. Comparing with Multi-class Adaboost algorithm, we have obtain an average accuracy of 99.2% against 98.3% on Weizmann dataset, and 91.3% against 81.2% on KTH dataset. Particularly, our proposed algorithm is more discriminative to classify ”jogging”, ”walking” and ”running” actions.

5. Conclusions and Future Works

This paper proposes a SemiBoosted Co-Training method for human action recognition. Both the co-trained inter-view confidence and the self-trained intra-view confidence are estimated and incorporated into one interacted semi-supervised learning process. Two separated views namely action saliency view and action eigen-projection view are extracted directly from the video data, and two classifiers are trained from the two views and updated in each iteration. Furthermore, we have proposed a multi-class SemiBoost algorithm to boost the performance of each classifier. We have tested our proposed method on publicly used human action databases, the results are encouraging. We believe that the framework of proposed semi-supervised algorithm can be also very helpful for other applications where manual labels are not costly.

Our future work will be concentrated on theoretical analysis of the convergence of the objective function and generalization error, and exploring more effective representations for action saliency.

References

[1] M. Belkin, P. Niyogi, and V. Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *The Journal of Machine Learning Research*, 7:2399–2434, 2006.

[2] Y. Bengio, O. Dellalleau, and N. L. Roux. Label propagation and quadratic criterion. *O. Chapelle, B. Schlkopf and A. Zien (Eds.), Semi-supervised learning. MIT Press.*, pages 193–216, 2006.

[3] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri. Actions as space-time shapes. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1395–1402, 2005.

[4] A. Blum and T. Mitchell. Combining labeled and unlabeled data with co-training. *Annual conference on computational learning theory*, pages 92–100, 1998.

[5] S. Dasgupta, M. L. Littman, and D. Mcallester. Pac generalization bounds for co-training. *Advances in Neural Information Processing Systems*, pages 375–382, 2002.

[6] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie. Behavior recognition via sparse spatio-temporal features. *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pages 65–72, 2005.

[7] A. Fathi and G. Mori. Action recognition by learning mid-level motion features. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[8] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri. Actions as space-time shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2247–2053, 2007.

[9] D. Guan, W. Yuan, Y.-K. Lee, A. Gavrilov, and S. Lee. Activity recognition based on semi-supervised learning. *International Conference on Embedded and Real-Time Computing Systems and Applications*, pages 469–475, 2007.

[10] W. Hu, T. Tan, L. Wang, and S. Maybank. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on SMC*, 34(3):334–352, 2004.

[11] K. Jia and D.-Y. Yeung. Human action recognition using local spatio-temporal discriminant embedding. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[12] I. Laptev. On space-time interest points. *International Journal on Computer Vision*, 64(2-3):107–123, 2005.

[13] C. Leistner, H. Grabner, and H. Bischof. Semi-supervised boosting using visual similarity learning. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[14] C. Liu and P. C. Yuen. Boosting eigenactions: A new algorithm for human action categorization. *IEEE International Conference on Automatic Face and Gesture Recognition*, 2008.

[15] J. Liu, S. Ali, and M. Shah. Recognizing human actions using multiple features. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[16] P. K. Mallapragada, R. Jin, A. K. Jain, and Y. Liu. Semiboost: Boosting for semi-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008.

[17] K. Mikolajczyk and U. Hirofumi. Action recognition with motion-appearance vocabulary forest. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[18] T. B. Moeslund, A. Hilton, and V. Kruger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2):90–126, 2006.

[19] P. Natarajan and R. Nevatia. View and scale invariant action recognition using multiview shape-flow models. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[20] J. C. Niebles and F.-F. Li. A hierarchical model of shape and appearance for human action classification. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.

[21] J. C. Niebles, H. Wang, and L. Fei-Fei. Unsupervised learning of human action categories using spatial-temporal words. *International Journal of Computer Vision*, pages 299–318, 2008.

[22] A. Oikonomopoulos, I. Patras, and M. Pantic. Spatiotemporal salient points for visual recognition of human actions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 36(3):710–719, 2005.

[23] M. Rodriguez, J. Ahmed, and M. Shah. Action mach: A spatio-temporal maximum average correlation height filter for action recognition. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[24] C. Schuldt, I. Laptev, and B. Caputo. Recognizing human actions: a local svm approach. *Proceedings of International Conference on Pattern Recognition*, pages 32–36, 2004.

[25] E. Shechtman and M. Irani. Space-time behavior based correlation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):2045–2056, 2007.

[26] Q. Shi, L. Wang, L. Cheng, and A. Smola. Discriminative human action segmentation and recognition using semi-markov model. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[27] V. Sindhwani, P. Niyogi, and M. Belkin. Beyond the point cloud: from transductive to semi-supervised learning. *international conference on Machine learning*, pages 824–831, 2005.

[28] R. Souvenir and J. Babbs. Learning the viewpoint manifold for action recognition. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[29] Y. Wang, H. Jiang, M. S. Drew, Z.-N. Li, and G. Mori. Unsupervised discovery of action classes. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1654–1661, 2006.

[30] D. Weinland and E. Boyer. Action recognition using exemplar-based embedding. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[31] T. Xiang and S. Gong. Video behavior profiling for anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):893–908, 2008.

[32] S. Yang, L. Goncalves, and P. Perona. Unsupervised learning of human motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7):814–827, 2003.

[33] A. Yilmaz and M. Shah. Actions sketch: a novel action representation. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 984–989, 2005.

[34] X. Zhu. Semi-supervised learning literature survey. *Computer Sciences Technical Report 1530, University of Wisconsin-Madison*, 2008.

Image Analysis By An Improved Empirical Mode Decomposition Algorithm

Dan ZHANG

Abstract

An alternative Empirical Mode Decomposition (EMD) method has been proposed in this paper. EMD is an adaptive method to decompose non-linear and multi-component signals. The essence of EMD is to decompose a signal into intrinsic mode functions obtained in the sifting process. One of the key process is to detect the local mean. The original EMD uses envelope mean interpolated by cubic spline, however it is sensitive to extrema and always inaccurate, which often leads poor decomposition. We proposed a new approach to measure the local means in this paper. Compared to the traditional EMD, the alternative EMD is easy to implemented, faster and adaptive.

1 Introduction

Signal analysis is an important part in both research and practical applications. It aims to find hidden information and structures in data. Though Fourier spectral analysis and wavelet transform have provided some general methods for analyzing signals and data, they are still weak at non-stationary and nonlinear data process.

EMD is an empirically based data-analysis method proposed by Huang [1, 2, 3]. Because its basis of expansion is adaptive, it can product physically meaningful representations of data from nonlinear and non-stationary processes. Due to the data-driven advantages and efficiency in non-stationary and nonlinear data process, EMD has been more and more widely applied in signal analysis including ocean waves, rogue water waves, sound analysis, earthquake time records [2] as well as image analysis such as texture analysis [8], image compression [10] and so forth [12, 11, 9, 14, 13, 15, 16].

The essence of EMD is to decompose a signal into a set of Intrinsic Mode Functions (IMFs) by a sifting process, in which one key step is to detect the mean of the data. The original EMD adopted the average of upper and lower envelopes interpolated by cubic spline. However, it is sensitive to extrema and the interpolation is always difficult and inaccurate which often result in poor decomposition. Moreover, though as powerful as EMD is in many applications, a mathematical foundation is virtually nonexistent. Many fundamental mathematical issues such as the convergence of the sifting algorithm, the orthogonality of IMFs and others can't be proved. Building a mathematical foundation remains a big challenge in the study of EMD.

A lot of researchers tried to study toward this direction, many improved methods and alternative approaches has been proposed. In [20], the cubic splines were replaced by B-splines, which gives an alternative way for EMD but again this modification does not resolve those mathematical issues. In [22], the author finds the local mean using support vector regression machines, which is insensitive to the sampling frequency and can eliminate mode mixing in small amplitude sine waves intermittence. In [21], the sawtooth function has been constructed by connecting the successive extrema of the original data function with straight line segment. Sharif [18, 19] adopted order statistics filter as surface interpolation method, and then follow a smoothing process. The window size using for order statistics filters is specially determined. In recent work, Lixin Shen [24] proposed to use a more general polynomial of m degree to represent the local mean instead of the cubic spline. The undetermined problem converts to a quadratic programming optimal problem to solve the coefficients. However, these methods still can't solve the problems perfectly.

In this paper, a new alternative approach has been proposed. The mean of envelopes has been replaced by a certain "moving average" obtained through a low pass filter. The low pass filters are completely adaptive because they are data dependent. The window size for the filters are determined by the local extrema. This approach is easy implemented, faster and efficient in decomposition. The simulation results demonstrate its efficiency.

The paper is organized as follows. Section 2 presents a overview of the EMD sifting process. Section 3 presents the details of the proposed alternative EMD algorithm. The simulation results are demonstrated in Section 4. Finally, a conclusion and future work outlook are presented in Section 5.

2 EMD Overview

2.1 Steps of EMD

EMD is a signal analysis technique for adaptive representation of non-stationary signals as sum of a set of IMFs. It captures information about local trends in the signal by measuring oscillations, which can be quantized by a local high frequency or a local low frequency, corresponding to finest detail and coarsest content. Here we briefly review the sifting process of EMD. Four main steps are contained, S1, S2, S3 and S4 are abbreviation for Step 1 to Step 4. Given a signal x(t),

S1. Identify all the local minima and maxima of the input signals $x(t)$;

S2. Interpolate between all minima and maxima to yield two corresponding envelopes $E_{max}(t)$ and $E_{min}(t)$. Calculate the mean envelope $m(t) = (E_{max}(t) + E_{min}(t))/2$;

S3. Compute the residue $h(t) = x(t) - m(t)$. If it is less than the threshold predefined then it becomes the first IMF, go to Step 4. Otherwise, repeat Step 1 and Step 2 using the residue $h(t)$, until the latest residue meets the threshold and turns to be an IMF;

S4. Input the residue r(t) to the loop from Step 1 to Step 3 to get the next remained IMFs until it can not be decomposed further.

2.2 Issues related to EMD

One of the main unresolved questions mentioned earlier is the convergence of $h(t)$ in general. Theoretically, the judgement should accord two conditions [1]. First, the number of extrema and the number of zero-crossing must be at most differ by one. Second, the mean envelops obtained by the maximum envelop and the minimum envelop must equal to zero. However, it is difficult to achieve these two conditions strictly in the sense of realistic implementation. Even though in practice we stop the iteration once some stopping criterion is met (e.g. define a small threshold), it is still important to know whether such criterion will ever be met. Although there is no mathematical proof for the convergence, there have been no examples in which the sifting algorithm fails to stop. Because of the convergence in practical sense, the value of $h(t)$ was degressive. Thus we define the stop criterion in this way that can control the number of IMFs: $h_{current}(t) <= (1-\alpha)h_{up-iter}(t)$, where α is a constant in $[0, 1)$. The larger the α, the more IMFs, the smaller the α, the faster the stopping.

An orthogonality index (OI), denoted as O, has been proposed in [1] which is defined as follows:

$$O = \sum_{t=1}^{T}(\sum_{j=1}^{n+1}\sum_{k=1}^{n+1})C_j(t)C_k(t)/x^2(t).$$

For the case of 2-dimensional image, the extended formula:

$$O = \sum_{x=1}^{M}\sum_{y=1}^{N}(\sum_{j=1}^{n+1}\sum_{k=1}^{n+1}\frac{C_j(x,y)C_k(x,y)}{\sum_C^2(x,y)}.$$

A low value of OI indicates a good decomposition in terms of local orthogonality among the IMFs.

3 New Algorithm

Here we propose an alternative algorithm for EMD. Instead of using the envelopes generated by splines we use a low pass filter to generate a "moving average" to replace the mean of the envelopes. The essence of the sifting algorithm remains.

3.1 Moving average filters

The moving average is the most common filter in digital signal processing. It operates by averaging a number of points from the input signal to produce each point in the output signal, it is written:

$$y[i] = \frac{1}{M}\sum_{j=0}^{M-1}x[i+j],$$

where $x[]$ is the input signal, $y[]$ is the output signal, and M is the number of points used in the moving average. It is actually a convolution using a simple filter $[a_i]_{i=1}^{M}, a_i = \frac{1}{M}$, and $[A_{i,j}]_{i=1,j=1}^{M,N}, A_{i,j} = \frac{1}{M\times N}$ for the 2-dimensional case.

3.2 Determining window size for average filters

3.2.1 Detection of local extrema

Detection of local extrema means finding the local maxima and minima points from the given data. No matter for 1D signal or 2D array, neighboring window method is employed to find local maxima and local minima points. The data point/pixel is considered as a local maximum (minimum) if its value is strictly higher (lower) than all of its neighbors.

3.2.2 Determining window size for average filters

We illustrated 1-dimensional case and 2-dimensional case separately.

- 1-dimensional case:
  For each extrema map, the distance between the two neighborhood local maxima (minima, extrema, zero-crossing) has been calculated called as adjacent maxima (minima, extrema, zero-crossing) distance vector Adj_max ($Adj\_min, Adj\_ext, Adj\_zer$). Four types of window size:
  - Window-size I: $\max(Adj\_max)$;
  - Window-size II: $\max(Adj\_min)$;
  - Window-size III: $\max(Adj\_zer)$;
  - Window-size IV: $\max(Adj\_ext)$.
- 2-dimensional case:
  The window size for average filters is determined based on the maxima and minima maps obtained from a source image. For each local maximum (minimum) point, the Euclidean distance to the nearest local maximum (minimum) point is calculated, denoted as adjacent maxima (minimum) distance array Adj_max (Adj_min).
  - Window-size I: $\max(Adj\_max)$;
  - Window-size II: $\max(Adj\_min)$;

4 Simulation Results

In all our numerical experiments we determine the window size in each decomposition with Window-size I. Unless otherwise specified we use $\alpha = 0.5$ for our stopping criterion.

4.1 Signal analysis

We test our EMD on a couple of very standard test examples, where the test functions are combinations of two sinusoidal functions with well separated frequencies.

The first signal is given by Fig.1 top as

$$f(t) = sin(t) + sin(4t).$$

The proposed algorithm easily separates out the two components as the IMFs, which are shown as the middle and bottom plots in Fig.1.

Another similar test function is

$$f(t) = sin(t) + \frac{1}{10} sin(20t),$$

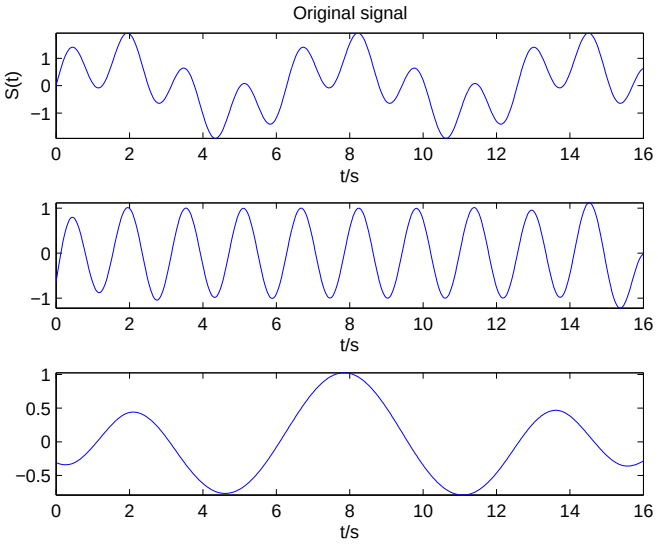


Figure 1. Decomposition of $f(t) = sin(t) + sin(4t)$.

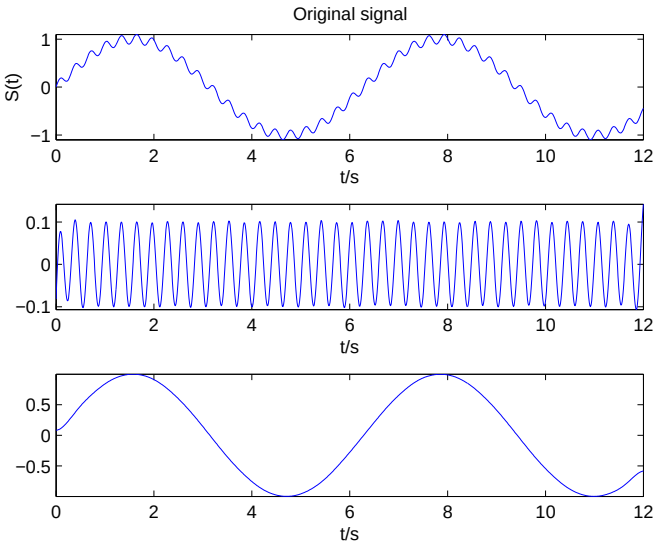


Figure 2. Decomposition of $f(t) = sin(t) + \frac{1}{10} sin(20t)$.

shown by the top figure of Fig.2. Again we easily separate the two components, as shown by the middle and right figures of Fig.2.

Fig.3 showed the decomposition results of

$$f(t) = sin(20\pi t)+4sin(40\pi t)sin(\frac{\pi}{5}t)+sin(10\pi t)+10t.$$

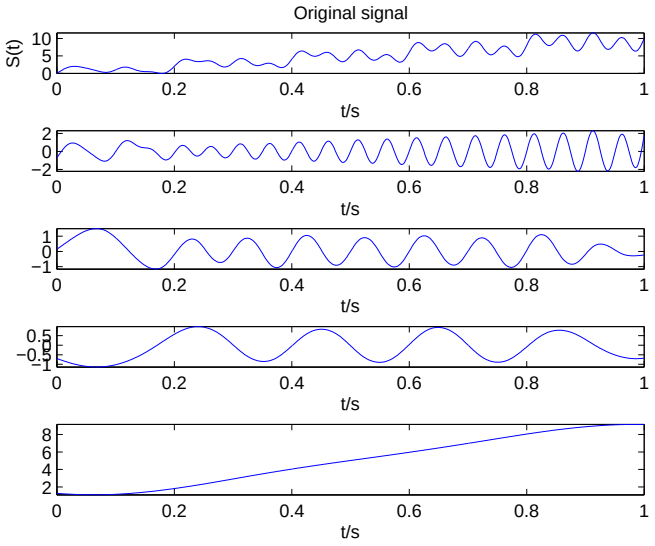


Figure 3. Decomposition of $f(t) = sin(20\pi t) + 4sin(40\pi t)sin(\frac{\pi}{5}t) + sin(10\pi t) + 10t$.

Table 1. Orthogonality Index (OI) and Consuming Time of the three signals

Signals	Orthogonality Index (OI)	Consuming Time
I	0.0047	1.00
II	0.0008	1.01
III	0.0007	1.36

4.2 Image analysis

We have evaluated the proposed EMD approach on an image of 256×256. Table.2 shows the OI and decomposition time of three different methods. Fig.4, Fig.5 and Fig.6 showed the IMF components obtained by the three methods respectively. Generally, a lower OI represents a better decomposition. Our approach performed well both in decomposition quality and time consuming.

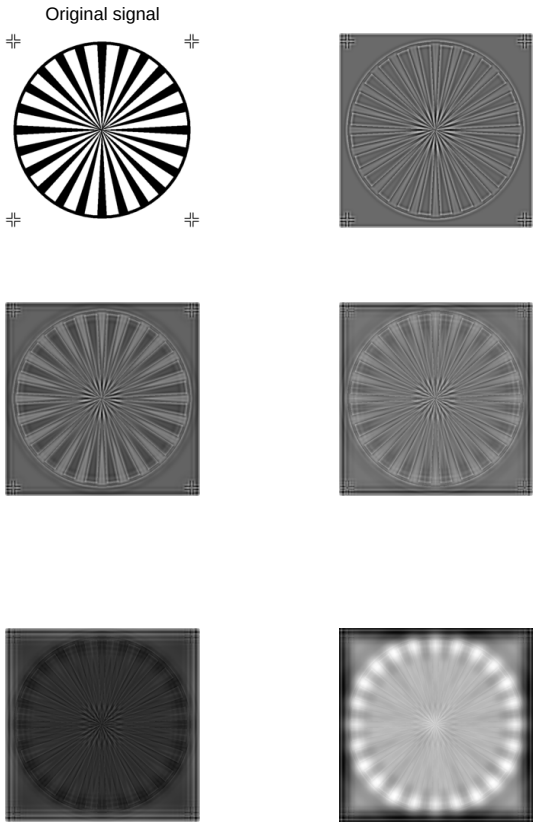


Figure 4. 5 components by our proposed method.

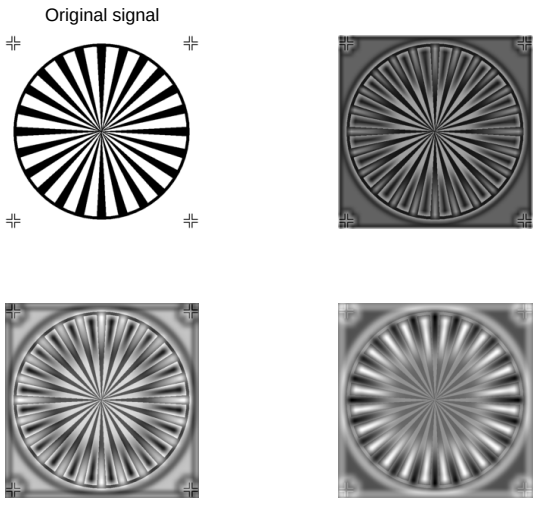


Figure 5. 3 components by Sharif method[18].

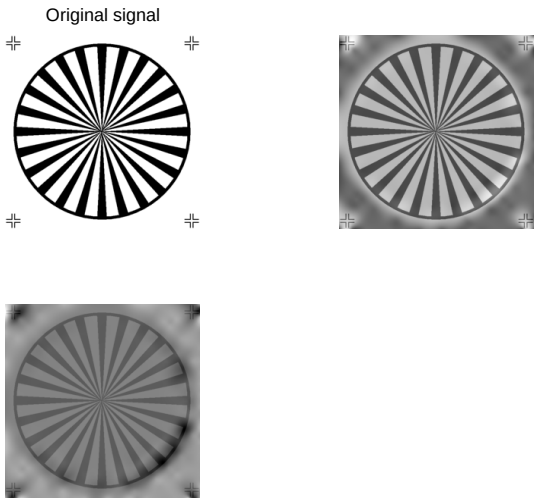


Figure 6. 2 components by envelope method[9].

Table 2. Orthogonality Index (OI) and Consuming Time of the three methods

Methods	Orthogonality Index (OI)	Consuming Time
Sharif [18]	0.0015	11.09
Envelope [9]	0.0013	108.98
Our method	8.0914e-004	8.84

5 Conclusions

We proposed an alternative Empirical Mode Decomposition (EMD) approach in this paper. One key process is to detect the local mean of the data. The original EMD uses envelope mean interpolated by cubic spline, however it is sensitive to extrema and always inaccurate that often leads poor decomposition. The proposed new approach measures the local means using a low-pass filter. It is easy to implemented, and the simulation results demonstrate that it is much faster and more accurate the original EMD. What's more, it is not difficult to evaluate the convergence of this approach in rigorous mathematical view while the original EMD can't, which is also our future work.

References

[1] N. E. Huang, Z. Shen, S. R. Long, et al.. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A, vol. 454, no. 1971, pp. 903C995, 1998.*

[2] HILBERT-HUANG TRANSFORM AND ITS APPLICATIONS. *Book in Interdisciplinary Mathematical Sciences, Vol. 5, edited by N. E. Huang, and Samuel S P Shen, 2005.*

[3] N. E. Huang, M. L. C. Wu, S. R. Long, et al.. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society A, vol. 459, no. 2037, pp. 2317C2345, 2003.*

[4] S. R. Long. Applications of HHT in image analysis. in Hilbert-Huang Transform and Its Applications, N. E. Huang and S. S. P. Shen, Eds., World Scientific, River Edge, NJ, USA, 2005.

[5] Harishwaran Hariharan, Andrei Gribok, Besma Abidi, and Mongi Abidi. Multi-modal Face Image Fusion using Empirical Mode Decomposition. *The Biometrics Consortium Conference, Crystal City, VA, September 2005.*

[6] H. Hariharan, A. Koschan, B. Abidi, A. Gribok, and M.A. Abidi. Fusion of visible and infrared images using empirical mode decomposition to improve face recognition. *IEEE International Conference on Image Processing ICIP2006, Atlanta, GA, pp. 2049-2052, October 2006.*

[7] Bhagavatula, R., Marios Savvides, and M. Acoustics. Analyzing Facial Images using Empirical Mode Decomposition for Illumination Artifact Removal and Improved Face Recognition. *IEEE International Conference on Speech and Signal Processing, 2007 (ICASSP 2007). Vol. 1, Issue , 15-20 April 2007 pp. I 505-508.*

[8] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and Ph. Bunel. Image analysis by bidimensional empirical mode decomposition. *Image and Vision Computing Volume 21, Issue 12, Pages 1019-1026, November 2003.*

[9] Nunes J. C., Guyot S., and Delé*c*helle E. Texture analysis based on local analysis of the Bidimensional Empirical Mode Decomposition. In *Machine Vision and Applications 16, 3, pp. 0932-8092, 2005.*

[10] A. Linderhed. 2-D empirical mode decompositions in the spirit of image compression. *in Wavelet and Independent Component Analysis Applications IX, vol. 4738 of Proceedings of SPIE, pp. 1C8, Orlando, Fla, USA, April 2002.*

[11] A. Linderhed. Compression by image empirical mode decomposition. *IEEE International Conference on Image Processing, 2005 (ICIP 2005), Vol. 1, pp. I 553-6, 2005.*

[12] H. Hariharan, A. Gribok, M. Abidi, and A. Koschan. Image Fusion and Enhancement via Empirical Mode Decomposition. *Journal of Pattern Recognition Research, Vol. 1, No. 1, pp. 16-32, January 2006.*

[13] Sinclair, S. and Pegram, G. G. S. Empirical Mode Decomposition in 2-D space and time: a tool for space-time rainfall analysis and nowcasting. *Hydrol. Earth Syst. Sci. Discuss., 2, 289-318, 2005.*

[14] Jian Wan, Longtao Ren, and Chunhui Zhao. Image Feature Extraction Based on the Two-Dimensional Empirical Mode Decomposition. *2008 Congress on Image and Signal Processing, Vol. 1, pp. 627-631, 2008.*

[15] Jalil Taghia, Mohammad Ali Doostari and Jalal Taghia. An Image Watermarking Method Based on Bidimensional Empirical Mode Decomposition. *2008 Congress on Image and Signal Processing, Vol. 5, pp. 674-678, 2008.*

[16] Fauchereau, N., Sinclair, S., and Pegram, G. 2-D Empirical Mode Decomposition on the sphere, application to the spatial scales of surface temperature variations. *Hydrol. Earth Syst. Sci. Discuss., 5, 405-435, 2008.*

[17] C. Damerval, S. Meignen, and V. Perrier. A fast algorithm for bidimensional EMD. *IEEE Signal Processing Letters, vol. 12, no. 10, pp. 701C704, 2005.*

[18] Sharif M. A. Bhuiyan, Reza R. Adhami, and Jesmin F. Khan. Fast and Adaptive Bidimensional Empirical Mode Decomposition Using Order-Statistics Filter Based Envelope Estimation. *EURASIP Journal on Advances in Signal Processing, vol. 2008, Article ID 728356, 18 pages, 2008.*

[19] Sharif M. A. Bhuiyan, Reza R. Adhami, and Jesmin F. Khan. A novel approach of fast and adaptive bidimensional empirical mode decomposition. *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008 (ICASSP 2008), pp. 1313-1316, 2008.*

[20] Sherman Riemenschneider, Bao Liu, Yuesheng Xu and Norden E. Huang. B-spline based empirical mode decomposition. *Hilbert-Huang Transform and Its Applications, Book chapter 2, 2005.*

[21] Louis Yu Lu. Fast intrinsic mode decomposition of time series data with sawtooth transform. *Technical report, Nov 2007.*

[22] Yong-Ping Huang, Xue-Yao Li and Ru-Bo Zhang. A research on local mean in empirical mode decomposition. *Proceedings of the 7th international conference on Computational Science, Part III: ICCS 2007, Lecture Notes In Computer Science, Vol. 4489, pp. 125-128.*

[23] Luan Lin, Yang Wang, and Haomin Zhou. A new approach to empirical mode decomposition. *Preprint.*

[24] Lixin Shen. Local mean and empirical mode decomposition. *Report on The Second International Conference on the Advances of Hilbert-Huang Transform and Its Applications, Guangzhou, Dec 2008. Preprint.*

Recognition of 3D Graphical Models by Using Shape Similarity

Yuesheng HE

Abstract

Recognizing the 3D models in the graphical environment is a fundamental problem with applications in computer graphics, virtual reality, especially the intelligent virtual human for Humanoid Animation. A challenging aspect of this problem is to find a suitable shape feature that can be used to compared quickly, while still discriminating between similar and dissimilar shapes. We propose a method of recognizing shape features for surface-based 3D shape models based on their shape similarity. The the features of shape of 3D models are computed by first converting an input surface based model into an oriented point set model and then computing joint 2D histogram of distance and Shape Distributions. Then, Support Vector Machines (SVM) are used to classify the features of the models. By the classification the models can be given semantic meanings in the 3D environment.

1 Introduction

If the 3D models can be recognized by the system, then they can be given topological and/or semantic meaning and easy to be used and stored, for instance, interacted with 3D human-like animation object(virtual Human) [17][11]. Moreover,proliferation of 3D models on the Internet and in in-house databases prompted development of the technology for effective recognition of three-dimensional (3D) models.

A 3D model could be described by its textual annotation by using a conventional text-based search engine. This approach wouldnt work in many of the application scenarios for the 3D shape model, however. The annotations added by human beings depend on different applications and other factors. However,it is extremely difficult to describe by words a shape that is not in a well known shape or semantic category. It is thus necessary to develop content-based recognition systems for 3D models that are based on the features intrinsic to the 3D models, one of the most important of which is shape.

In the study of shape similarity recognition of 3D models, first step is to extract robust, concise, yet expressive shape features, and on the development of similarity (or, dissimilarity) comparison methods that conform well to the human notion of shape similarity[1][5].

In developing the shape features for 3D models, we first have to decide which class of 3D shape representation we are targeting. A 3D shape may be defined by using any of a number of shape representations, many of which are not mutually compatible. Some of the shape representations are mathematically well founded, allowing for computations of such well-defined properties as volume, surface curvature, or surface (or volume) topology. Unfortunately, since most 3D file formats (VRML, 3D Studio, etc.) have been designed for visualization, they contain only geometric and appearance attributes, and usually lack semantic information to be recognized. A great part of shape representations are less nicer. For example, a polygon-soup model is a topologically disconnected collection of independent polygons and/or polygonal meshes[1][2]. Neither volume nor surface curvature can be computed for the model.

The second step is to classify the features of shapes been extracted from the 3D models. Because the shapes seldom have any topology or solid model information; they rarely are manifold; and most are not even self-consistent, it is important to separate them on their features.

In this paper, we use method for computing 3D shape signatures and dissimilarity measures for arbitrary objects described by possibly degenerate 3D polygonal models. The object is to represent the signature of an 3D model on measuring global geometric properties of the object.Thus,we used a pair of methods to represent features of shapes - 1-dimensional descriptor (e.g. Osada, et al.)[1] and 2- dimensional descriptor (e.g. mutual Absolute-Angle Distance histogram(AAD))[2].

Intuitively A given a set of feature points which belongs to either one of two classes can be linearly separated by SVM with the hyperplane leaving the largest possible fraction of points of the same class on the same side. Since the model features are vectors representing different shapes.Then we classify the models by using SVM RBF and polynomial kernel .

The paper is organized as follows. In the next section, we introduce the ways to extract features of shapes. The classification method is described in Section 3, and the method

and results for the experimental evaluation of our algorithm are presented in Section 4. We conclude the paper in Section 5.

2 Features of Shapes

A method for shape similarity comparison of 3D models can be classified by the shape representation it is targeting. Some of the shape comparison algorithms assume well-defined shape representation, that are, 3D solid represented by using voxels, boundary representation, or constructive solid geometry . Others assume topologically well-defined 2-manifold surfaces . However these methods cant be used to compare polygon soup models. In this section, we review shape similarity comparison methods for not-so-well-defined shape representations, especially those for polygon-soup models. Another possible classification is by the method used to achieve invariance of the shape comparison method to a class of geometrical transformations[3].

Osada et al. proposed what they call shape distributions. Osadas shape distributions, a set of shape features, have the advantage of being invariant, without pose normalization, to similarity transformations. Moreover, they are designed to be applicable to a not-so-well-defined mesh-based model, i.e., a polygon soup defining a non-solid object consisting of non-manifold surfaces, multiple connected components, and such degenerate surfaces as zero-area polygons. D1 is one of the distributions.

We choose this distribution, because it is simple and computational efficient.

The 2-dimensional feature descriptor uses vectors to represent shapes. For instance, the mutual Absolute-Angle Distance histogram (AAD) shape feature is a 2D histogram. This method considers not only the distance, but the the orientation of vectors as well.

The process of obtaining the features is:

1. Calculating the items of features (e.g. L1-norm, L2-norm distance and inner product);

2. Normalization by a certain criteria (e.g. By maximum,by average or by median);

3. Making histogram.

2.1 1-Dimensional Descriptor

The first issue of the method is to select a function whose distribution provides a good signature for the shape of a 3D polygonal model. Ideally, the distribution should be invariant under similarity transformations, and it should be insensitive to noise, cracks, tessellation, and insertion/removal of small polygons. In general, any function could be sampled to form a shape distribution, including ones that incorporate domain-specific knowledge, visibility information (e.g., the distance between random but mutually visible points), and/or surface attributes (e.g., color, texture coordinates, normals and curvature). However, as our interesting is on the topology information of the models,we use the features based on geometric measurements (e.g., angles, distances, areas, and volumes).

Shape features should be independent of the representation, topology, or application domain of the sampled 3D models. As a result, the shape similarity method can be applied equally well to databases with 3D models stored as polygon soup, meshes, constructive solid geometry, voxels, or any other geometric representation as long as a suitable shape function can be computed from each representation.

The Osada, et al.[1] is the set of shape distributions which has been used to represent the geometric properties.

The D2 definition is:

- ***D2***: Measures the distance between two random points on the surface.

Figure 1 shows two original 3D models which are going to be made the feature descriptors.

We use the L2-norm to measure the distance between the points of the surface:

$$d = \sqrt{(p_i - p_j)^2}$$

The following step is normalizing the result by by maximum distance. Thus the results are between 0 and 1.

Then the D2 shape histogram has been made. Figure 2 shows the different feature histograms according to the models.

To achieve geometrical transformation invariance, some methods employ pose normalization.

2.2 2-Dimensional Descriptor

The 2-Dimensional Descriptor is 2D histogram of distances and angles formed by pairs of oriented points that are generated on the surfaces of the given 3D shape model.

We use the AAD [2] to form the histogram.

The AAD method is based on AD[2]. The AD shape feature measures, for each pair of points p_1 and p_2 , the 3D Euclidian distance $d = \sqrt{(p_1 - p_2)^2}$ between the points and the inner product $a =< n_1, n_2 >$ of the orientation vectors n_1 and n_2 of the points. The AD shape feature described above is sensitive to the sign of the orientation vector of the point set model. If the models to be compared have a consistent surface orientation, e.g. a consistent traversal order of the vertices among polygons, the AD shape feature performs well. If, however, the database contains models

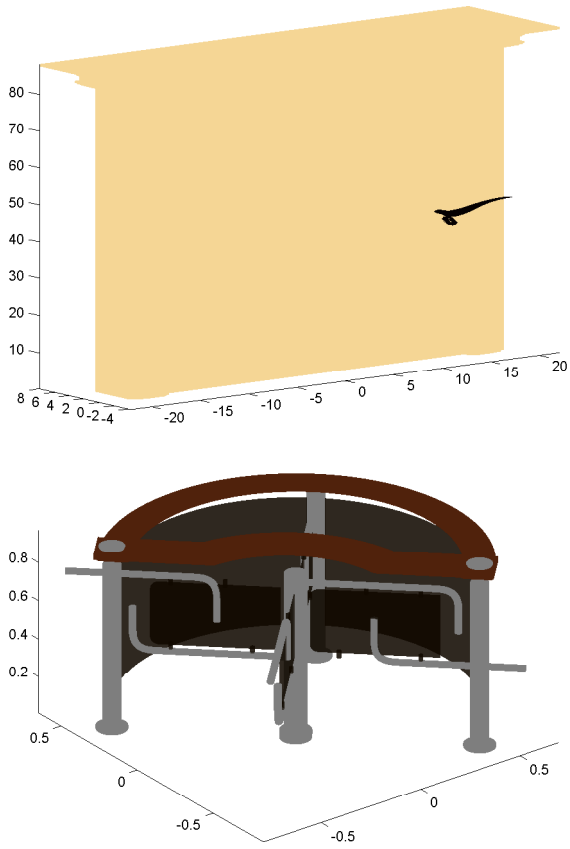


Figure 1. Two Different 3D Models

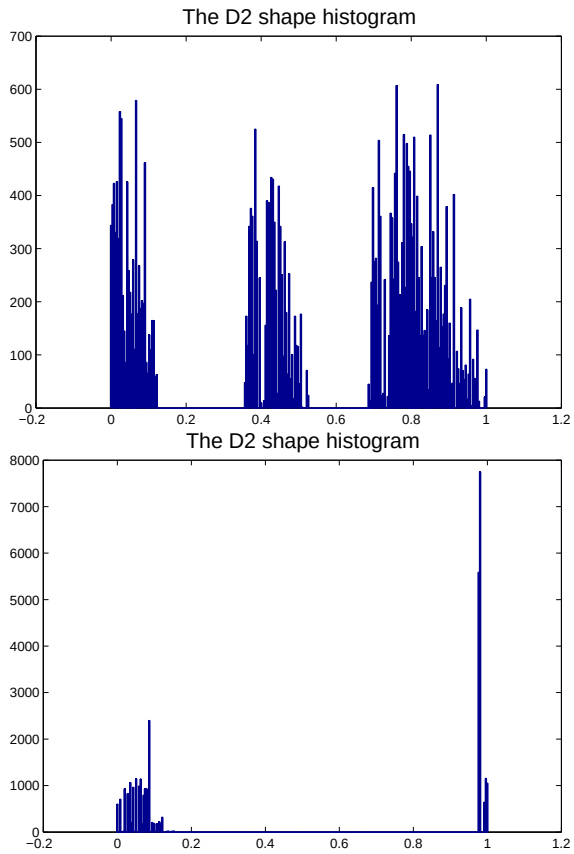


Figure 2. The 3D Models' D2 Shape Histograms

having surfaces that are inconsistently oriented, AD shape feature suffers.

One of the important requirement for a 3D shape similarity comparison method is invariance of the method to a required class of geometrical transformations. Most of the time, an invariance to similarity transformation, that is, a combination of translation, rotation, and uniform scaling, is required for a 3D shape similarity comparison.

The mutual Absolute Angle and Distance (AAD) is computed similarly to the AD, except that the AAD ignores the sign of the inner product. This makes the AAD a more robust shape feature than the AD for the models having unoriented or inconsistently oriented surface orientations.

Its definition is:

- ***AAD***: 2D histogram of distances and angles formed by pairs of oriented points that are generated on the surfaces of the given 3D shape model.

The AAD has properties:

- orientation insensitive shape feature;
- normalization prior to applying a pose orientation sensitive shape feature.

Models of Figure 1 generated the different shape descriptors of AAD Figure 3. The mutual Absolute Angle and Distance (AAD) histogram is computed similarly to the AD, except that the AAD ignores the sign of the inner product. This makes the AAD a more robust shape feature than the AD for the models having unoriented or inconsistently oriented surface orientations.

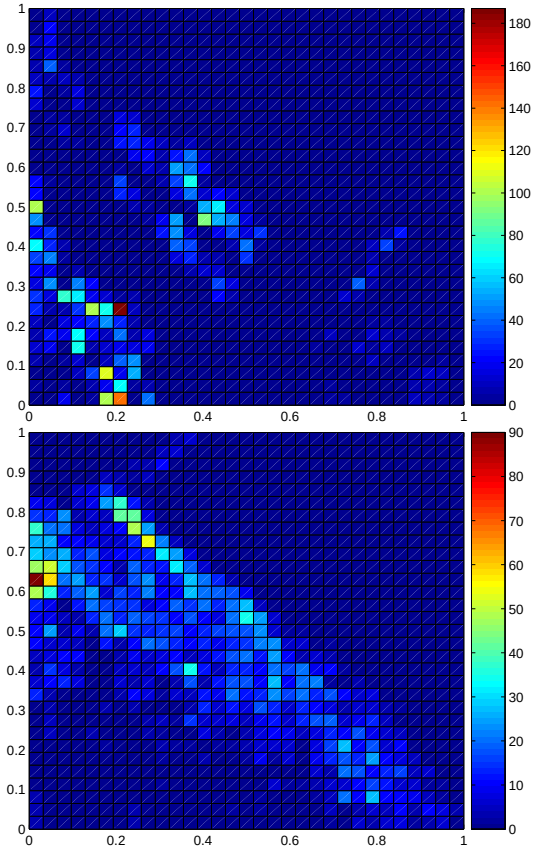


Figure 3. The 3D Models' AAD Shape Histograms

3 Classification and Recognition

The SVM [9] methodology comes form the application of statistical learning theory to separating hyperplanes for binary classification problems. The central idea of SVM is to adjust a discriminating function so that it makes optimal use of the separability information of boundary cases. Given a set of cases which belong to one of two classes, training a linear SVM consists in searching for the hyperplane that leaves the largest number of cases of the same class on the same side, while maximizing the distance of both classes from the hyperplane. If the training set is linearly separable, then a separating hyperplane, defined by a normal w and a bias b, will satisfy the inequalities:

$$y_i(w \cdot x_i + b \geq 1 \forall i \in \{1 \cdots N\} \quad (1)$$

where $x_i \in \Re^d$is a case of the training set $i = (1, \cdots, N)$, and d being the dimension of the input space, and $y_i \in \{-1, 1\}$ is the corresponding class.Since such a

distance is $\frac{1}{\|w\|}$, finding the optimal hyperplane is equivalent to minimizing $\|w\|^2$ under constraints (1).

The dual problem is:

$$\max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j < x_i x_j > \quad (2)$$

under the constraint:

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \quad \text{where} \quad 0 \leq \alpha_i \leq C \ \ \forall i \in \{1 \cdots N\} \quad (3)$$

The SVM approach can be extended to non-linear decision surfaces through a non-linear function Φ which maps the original feature space $\Re^d$ a higher dimensional space H. Since the only operation needed on H is the inner product, if we have a kernel function k [10]:

$$k(x', x") = \Phi(x') \cdot \Phi(x") \quad (4)$$

The objective function becomes:

$$\max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (5)$$

The RBF kernel:

$$k(x', x") = exp(-\frac{\|x' - x''\|^2}{\delta^2}) \quad (6)$$

where δ (a positive real) are parameters of the kernel.

The Poly kernel:

$$k(x', x") = (x' \dot{x}" + 1)^p \quad (7)$$

We applied different SVM-based classification strategies and feature sets to semantically classify regions in graphical 3 D models.

During training procedure, we compute a large amount of surface Points features with various combinations from parameter space. We apply different kernels to select the most discriminating features. SVM kernel parameters are determined by observer the result of the procedure. All feature settings and parameters determined on the training-set are then applied for classification and recognition tasks on the test-set. Due to the histogram representation of the features, the usage of a Histogram-Intersection kernel implying a L2-norm distance measure is straight forward and empirically delivered the best results.

4 Experimental Result

We used the 3D studio models from World Wide Web (randomly search and download) to construct the training and testing sets. The examples are shows in Figure 4

Figure 4. Examples of 3 D Models

Then, we built the 3D shape features for every models.

One issue we must be concerned with is sampling density. The more samples we take, the more accurately and precisely we can reconstruct the shape distribution. On the other hand, the time to sample a shape distribution is linearly proportional to the number of samples, so there is an accuracy/time tradeoff in the choice of N.

We compared the performance of the AAD and the D2 shape features by using the models. The parameters used for this experiment are as follows:

1. *D2*: The number of points per model $P_n = 256$. Distance is computed by using the L2 norm. The normalization is performed by using the maximum-based method.

2. *AAD*: The number of points per model $P_n = 128$.Distance is computed by using the L2 norm-based method The normalization is performed by the maximum-based method.

Their shape histograms are constructed as follows:

1. *D2*: Bins numbers $B_d = 256$;

2. *AAD*: Bins numbers $B = D \times I$ (distance D = 32, inner product I = 32).

We further investigated the robustness of our method by testing it with different polygon tessellations of two 3D shapes. Thus, we test the different positions in the environment of the models and rotation of the models. Those did not influence the basic property of every feature descriptors.

To classify and recognize the certain features, we used the SVM. For instance, the object of the work is to recognize more chair-like models from others. We first trained the machine on the training set, then used it to classify the testing set.

Although the results are quite better than those of random guessing, they are still far from satisfactory. Moreover, the 2-Dimensional feature always performance better than

$Kernels$	$D2shape$	$AADshape$
Gaussian	52%	55%
Gaussianslow	53%	56%
Multiquadric	52%	54%
Poly	64%	82%

Table 1. Result of Classification

1-Dimensional one. However, the Ploy Kernel achieved more accurate rate of recognition than others. The Table 1 shows the result.

5 Conclusion and Discussion

In this paper, we proposed and evaluated a method of classification and recognition of shape features for shape similarity search of 3D models. The shape features, which have been represented by 1-Dimensional and 2-Dimensional descriptors, are robust against topological and geometrical irregularities and degeneracies, which make them applicable to 3D Studio format and other so called polygon soup models. They are also invariant to similarity transformation, a quality valuable in classifying 3D shape models.

According to the experiments, though the AAD have computational cost somewhat higher (have to compute inner product) than the D2, they significantly outperformed D2 in our classification experiments by SVM. Although a further comparison has not been made, the 2-Dimensional descriptor might have the performance better than that of the 1-Dimensional methods, such as the Table 1 shows. However, the computational costs of former one is higher than the later one. Thus, both the distance and angular information be useful for recognizing 3D shapes. As a future work, we would like to improve our shape feature, for example by adding some form of multi-resolution approach to matching 3D shapes. We also would like to explore a hybrid shape feature that combines, possibly adaptively, shape features having different characteristics.

Depend on the shape features, the SVM has been used to train the system to classify different shapes. According to the experimental result, the poly kernel perform well in the task. In feature, we would like to explore different kind of kernels and parameters of them.

References

[1] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin, *ACM Transactions on Graphics*, Vol. 21, No. 4, October 2002, Page 807—832.

[2] Ryutarou Ohbuchi, Takahiro Minamitani, Tsuyoshi Takei, *International Journal of Computer Applications in Technology*,Vol. 23, No. 2/3/4, 2005, Page 70—85.

[3] J. Fehr, H. Burkhardt, Harmonic Shape Histograms for 3D Shape Classification and Retrieval, *MVA2007 IAPR Conference on Machine Vision Applications*, May 16-18, 2007, Tokyo, JAPAN.

[4] Ding-Yun Chen, Xiao-Pei Tian, Yu-Te Shen and Ming Ouhyoung, Visual Similarity Based 3D Model Retrieval,*EUROGRAPHICS 2003 / P. Brunet and D. Fellner (Guest Editors)* Volume 22 (2003), Number 3,Page 223—232.

[5] FRED L. BOOKSTEIN, Principal Warps: Thin-Plate Splines and the Decomposition of Deformations, *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*. Vol.II., NO. 6, JUNE 1989, page 576—585.

[6] Edward Grabczewski, John Cosmas, Peter Van Santen Damian Green, Takebumi Itagaki, Fred Weimer,3D MURALE: Multimedia Database System Architecture,*Association for Computing Machinery*, 2002

[7] Michael M. Bronstein and Ron Kimmel, Topology-invariant similarity of nonrigid shapes Alexander M. Bronstein, June 10, 2008.

[8] Ryutarou Ohbuchi, Akihiro Yamamoto and Jun Kobayash, Learning Semantic Categories for 3D Model Retrieval, *9th ACM SIGMM International Workshop on Multimedia Information Retrieval (ACM MIR 2007)*.

[9] Vladimir Naumovich Vapnik,*The Nature of Statistical Learning Theory*.

[10] Dmitry Zelenko, DMITRY Chinatsu, Aone CHINATSU, Anthony Richardella, Kernel Methods for Relation Extraction, *Journal of Machine Learning Research*, FEB 2003, Page 1083—1106.

[11] T. Conde, D. Thalmann, An Integrated Perception for Autonomous Virtual Agents: Active and Predictive Perception, Computer Animation and Virtual Worlds, Volume 17, Issue 3-4, John Wiley, 2006

[12] Moccozet L, Thalmann N. M., Dirichlet Free-Form Deformation and their Application to Hand Simulation[J], *Proceedings Computer Animation97*, IEEE Computer Society, 1997, Page 93—102.

[13] Catherine Zanbaka1, Amy Ulinski, Paula Goolkasian, Larry F. Hodges, Social Responses to Virtual Humans: Implications for Future Interface Design, *CHI 2007 Proceedings*, Page 1561-1570.

[14] Edward M. Sims, Reusable, lifelike virtual humans for mentoring and role-playing, *Computers & Education*, 49 2007, Page 75-92.

[15] Lucio Ieronutti , Luca Chittaro, Employing virtual humans for education and training in X3D/VRML worlds, *Computers & Education*, Page 93-109.

[16] Alain Rakotomamonjy,Variable Selection Using SVM-based Criteria,*Journal of Machine Learning Research*, March 2003, Page 1357—1370

[17] Weixiong Zhang, Randall W. Hill, Jr., A Template-Based and Pattern-Driven Approach to Situation Awareness and Assessment in Virtual Humans, *Agents*, 2000, Page 116—123.

[18] Z.M. , D. Reidsma, A. Nijholt, Human Computing, Virtual Humans and Artificial Imperfection, *ICMI'06*, Page 179—184.

Automatic Lip Localization under Changing Illumination Conditions

Meng LI

Abstract

This paper addresses mainly the problem of lip localization for the purpose of lip-reading under complex situations. We propose a novel approach to automatic localization of the minimum enclosing rectangle of lip based on the gray level information of the mouth image. This approach consists of two phases: estimation of the crucial points of the ROI (region of interesting) in horizontal, and normal direction. For the former one, a loop refinement is designed, in which some traditional transformations in gray level space are employed to ensure that the result of estimation converged quickly and exactly. For the latter one, based on the result of the former phase, the normal crucial points are located via filter. Experimental result shows the accuracy and robustness of the proposal approach.

1 Introduction

Motivated by human ability to lip read, the useful information on speech content can be obtained through analyzing the subtle cue conveyed by lip movement of speakers [9]. The intimate relation between the audio and visual sensory modality in human recognition can be demonstrated with audio-visual illusions such as the "McGurk effect" [8]. It suggests that speech perception is multimodal involving information from more than one sensory modality. In 1984, the first automatic lip-reading system was presented by Petajan [3, 4]. From then on, lip-reading has received considerable attention from the community because of its potential attractive applications in information security, speech recognition, secret communication, and so forth [17, 6].

In lip-reading, one key issue is the lip localization, i.e. how to obtain the accurate position of lip or mouth from image. Paper [5] demonstrates that the error rate of this AVSR system in studio environment, i.e. ideal light condition without shadow, is 37.3%. In contrast, the visual-only word error rate will reach 76.2% when an Automatic Visual Speech Recognition (AVSR) system is used in real world. It can be seen that the degraded performance is mainly caused by the imprecise localization of lip under "changing illumination condition". The term "changing illumination condition" means real world like environment in which the illumination may come from different directions. The difference in performance implies that the accuracy of lip localization is one of the most important factors that determine the recognition rate. So that, in this paper, we will therefore concentrate on the lip localization only.

Thus far, several methods have been proposed to enhance the performance of lip localization for AVSR system. For instance, paper [18] presents an approach that employs the hue-filter to distinguish lip and surrounding skin region. The papers in [7, 10, 16, 12] utilize the information of red component and saturation to localize the lip region. Also, paper [13] utilizes a gradient based Canny edge detector to locate the mouth corner. In [2], the input image is projected into YUV color coordinate system and the accumulations of V value in each row and column of the image are utilized to estimate the crucial points (i.e. the top, bottom and two mouth corners) of a lip.

In general, the methods stated above make the lip localization in a studio environment. Under a more challenging environment, e.g. some parts of mouth are covered by shadow, the boundary between lip and surrounding skin region, especially the area near mouth corners, cannot be distinguished precisely. Under the circumstances, those methods may not locate the ROI accurately, thus degrading the subsequent recognition accuracy. To circumvent the shadow effect, some approaches need to make the landmarks around mouth, e.g. see [14, 15]. Nevertheless, to the best of our knowledge, the lip localization under the challenging environment has not been well solved yet in the literature.

In this paper, we focus on lip localization under changing illumination conditions. The geometric lip features of interest are four outer lip crucial points along the horizontal and vertical directions. We propose an approach to localize the minimum enclosing rectangle of lip automatically based upon the gray-level image. This approach utilizes the mean filter and the image transformations to circumvent the noise caused by shadow. Subsequently, the crucial points of the ROI are estimated via analyzing the curve of gray-level values along with the vertical and horizontal midline of mouth, respectively. Experiments have shown the promising re-

sult of the proposed approach in comparison with a existing method.

2 The Proposed Lip Localization Method

2.1 Horizontal crucial points localization

Image of the mouth region are acquired by a video system which captures a 100×100 window at 20 frames per second and 24 bit pixel resolution.

The images captured by the camera are comprised of RGB values. We project these RGB value into gray level space according to the equation 1.

$$Y = 0.299R + 0.587G + 0.114B \tag{1}$$

The example of a frame is shown in Figure 1.

Figure 1. Lip area in an example frame

In order to enhance the contrast between lip and surrounding skip region, we adjust the histogram of the image and make it equalized for the first step. Then we make an accumulation of gray level value for each row of the image, which shown in Figure 2. The slopes of the curve contain the information about the boundaries between the lips and the surrounding skin region. The minimum value on the curve retained as the row position of mouth corner points or the nearby position, the row can be named as horizontal midline of mouth. The midline is shown in Figure 3.

The curve of gray level values along with the horizontal midline is saved in vector G. Building a sub-vector G_s by a segment of G which between the first maximum from left and the first maximum from right. Using the following equations to make the curve smooth and save it into a new vector which named C. The G and C are shown in Figure 4.

$$C_l^{(i)} = \begin{cases} G_s^{(i)} & (C_l^{(i-1)} > G_s^{(i)}) \\ C_l^{(i-1)} & (C_l^{(i-1)} \leq G_s^{(i)}) \end{cases} \quad i = 1, 2, \ldots, n \tag{2}$$

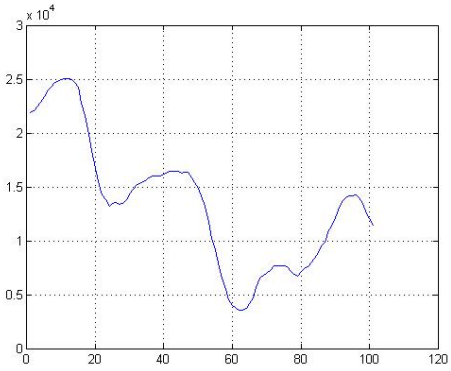


Figure 2. Accumulation of gray level value for each row

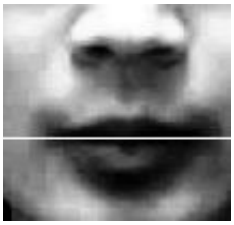


Figure 3. The midline calculated by the accumulation of gray level value for each row

$$C_r^{(i)} = \begin{cases} G_s^{(i)} & (C_r^{(i+1)} > G_s^{(i)}) \\ C_r^{(i+1)} & (C_l^{(i+1)} \leq G_s^{(i)}) \end{cases} \quad i = n-1, n-2, \ldots, 1 \tag{3}$$

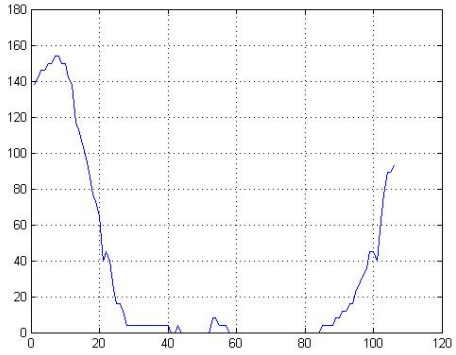
$$C = C_l + C_r \tag{4}$$

where C_l and C_r are assistant vectors, $C_l^{(i)}$ is the ith element in vector C_l, n is the dimension of the vector G. The initial values of the two vectors are shown in equation 5.

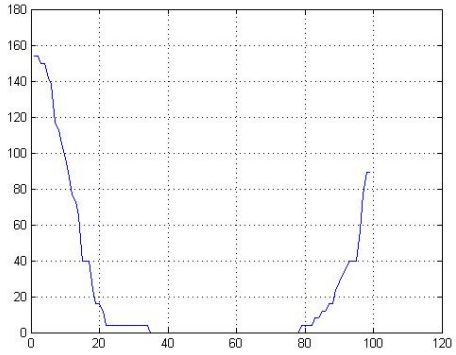
$$\begin{aligned} C_l^{(1)} &= G^{(1)} \\ C_r^{(n)} &= G^{(n)} \end{aligned} \tag{5}$$

Set the minimum of the most left and most right value in C as threshold. Elements in C less than the threshold build a new vector C' . Accordingly, the average can be calculated by the equation 6.

$$c_{avg} = \frac{\sum_{i=1}^{m} C'^{(i)}}{m} \tag{6}$$



(a)



(b)

Figure 4. Curve of original (a) and adjusted (b) gray-level value along with the horizontal midline of mouth.

The equation 7 is employed to adjust the contrast of image.

$$I_{out} = \begin{cases} 255 & (1.5c_{avg} < I_{in} < 1) \\ \frac{500}{c_{avg}} - 500 & (0 < I_{in} \leq 1.5c_{avg}) \end{cases} \quad (7)$$

where I_{in} is the input gray level value, and the I_{out} is the output.

For the adjusted image, a 11×1 searching block is performed along with the midline, the positions of the most left and right non-all white block are marked as the column of mouth corner candidates. The procedure is performed through an iterative process in the steps above, repeated until the position of mouth corner candidates no longer changed or the image turned into binary. Figure 5 illustrate the change of image. Figure 6 shows the estimate of the horizontal crucial points, i.e. the mouth corners.

Figure 5. The terminal image in the extraction procedure

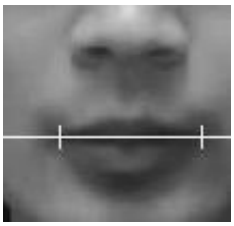


Figure 6. The estimate of crucial points in horizontal direction

2.2 Normal crucial points localization

As shown in equation 8 and 9 , a 33 mask is employed to perform mean filter in the initial image.

$$M = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix} \quad (8)$$

$$I^{(i+1)} = I^{(i)} * M \quad (9)$$

where the $I^{(i)}$ is the result of ith time filter. The times filter performed is determined by the equation 10.

$$\delta_i = dist(I^{(i+1)}, I^{(i)}) \quad (10)$$

where δ_i is the Euclidean distance between $I^{(i)}$ and $I^{(i+1)}$. The procedure should be ceased once δ_i less than a given threshold, and the $I^{(i+1)}$ can be marked as I_f. The example of $I_f = I^{(100)}$ is shown in Figure 7.

Due to the position of left and right mouth corners have been estimated in section 2.1, we can utilize them to calculate the center of mouth easily. For each $I^{(i)}$, a gray value vector $G^{(i)}_{mu}$ is built by the segment from the center point to the top of image along with the normal direction respectively. Then the vector ΔG_{acc} is calculated by the equation 11. The corresponding curve is shown in Figure 8.

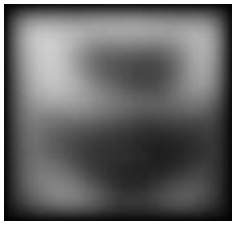


Figure 7. The image performed 100 times mean filter

$$\Delta G_{acc} = \sum_{i=1}^{n} (|G_{mu}^{(0)} - G_{mu}^{(i)}|) \qquad (11)$$

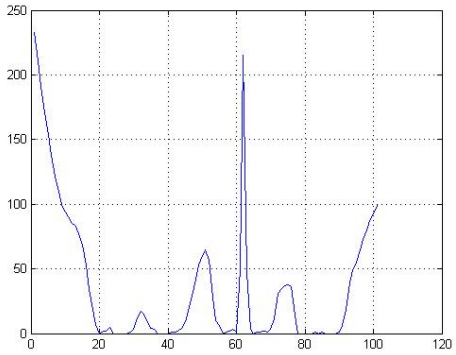


Figure 8. The curve of vector ΔG_{acc}

The point correspond to extreme value of maximum (except boundary value) is retained as the row position of upper bound of mouth.

Then the subtracted image between $I^{(0)}$ and I_f can be calculated. For observing conveniently, an image inversing transformation is employed. The processing result is shown in Figure 9.

Figure 9. The image produced by subtraction and inversion.

We get the gray level value along with the normal direction pass the middle point of mouth to the bottom of the image. The curve is shown in Figure 10.

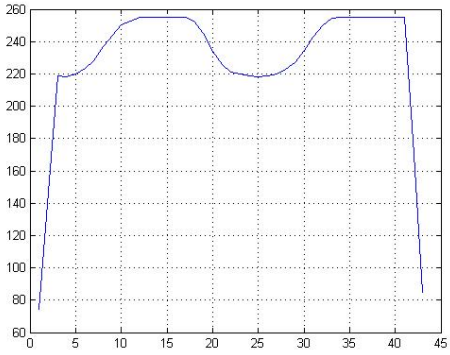


Figure 10. The gray level values from center point to the bottom of the image along with the normal direction

The point perform extreme value of minimum (except boundary value) is retained as the row position of lower bound of mouth. Hence, the normal crucial points are localized which is shown in Figure 11.

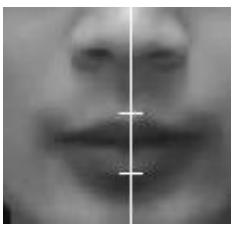


Figure 11. The estimate of crucial points in normal direction

Then we can get the minimum enclosing rectangle as shown in Figure 12.

3 Experimental result

Since most of the existing databases for lip-reading are not available for public a number of research groups develop their own visual-speech database [11]. Under the circumstances, we established a database for our experiments. The database consisted of six speakers (3 males and 3 females). Each speaker uttered ten isolated digits from zero to nine in Mandarin Chinese in several lighting condition (involve the shadow condition).

The lip localization approach is tested on some representative mouth region images from the database introduced

Figure 12. The minimum enclosing rectangle of mouth

above and some others images from other database. An existing approach is performed using the same testing set for comparing. In the following figures, the left column is the original gray images, the middle column is the processed results of employ our approach, the right column is the processed results of employ the approach proposed in [1].

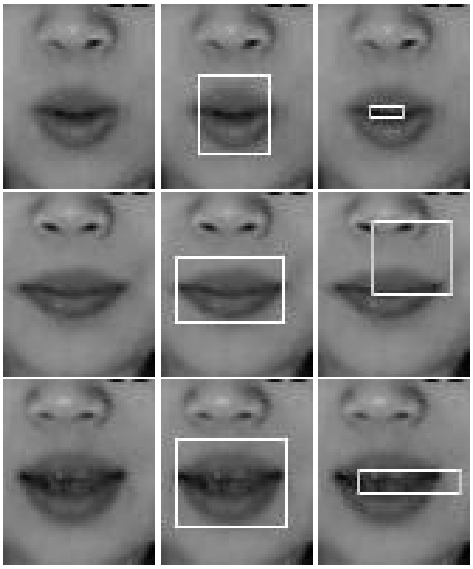


Figure 13. Results of the lip localization in shadow condition

4 Conclusion

In this paper, a lip localization approach has been proposed. Such a method provides an effective way to localize the four crucial points from the lip image in complex situation. The proposed approach has been empirically investigated by different speakers. The experiments have shown the promising results.

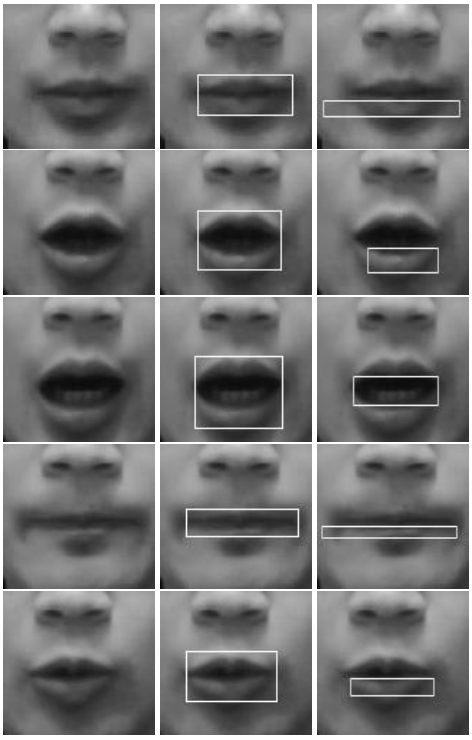


Figure 14. Results of the lip localization in high brightness image

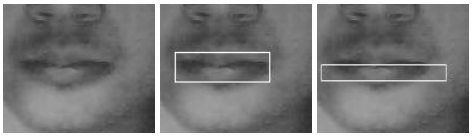


Figure 15. Results of the lip localization in low resolution and low contrast

Figure 16. Results of the lip localization with mustache

References

[1] A.R.Baig, R.Seguier, and G.Vaucher. Image sequence analysis using a spatio-temporal coding for automatic lipreading. In *Proc. of the 10th International Conference on Image Analysis and Processing*, Venice, Italy.

[2] A.R.Baig, R.Séguier, and G.Vaucher. Image sequence analysis using a spatio-temporal coding for automatic lipreading. In *Proc. IEEE International Conference on Image Analysis and Processing*, pages 544–549, Venice, Italy, 1999.

[3] E.D.Petajan. *Automatic lipreading to enhance speech recognition*. PhD thesis, University of Illinois, 1984.

[4] E.D.Petajan. Automatic lipreading to enhance speech recognition. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 40–47, 1985.

[5] G.Potamianos and C.Neti. Audio-visual speech recognition in challenging environments. In *Proc. European Conference on Speech Communication and Technology*, pages 1293–1296, Geneva, Switzerland, 2003.

[6] G.Potamianos, C.Neti, J.Luettin, and I.Matthews. Audio-visual automatic speech recognition: An overview. In G.Bailly, E.Vatikiotis-Bateson, and P.Perrier, editors, *Issues in Visual and Audio-Visual Speech Processing*. MIT Press, 2004.

[7] G.Potamianos, H.P.Graf, and E.Cosatto. An image transform approach for hmm based automatic lipreading. In *Proc. IEEE International Conference on Image Processing*, pages 173–177, Seattle,WA, 1998.

[8] H.McGurk and J.McDonald. Hearing lips and seeing voices. *Nature*, 264:746–748, 1976.

[9] J.Bulwer. *Philocopus, or the Deaf and Dumbe Mans Friend*. Humphrey and Moseley, 1648.

[10] A. W. C. Liew, S. H. Leung, and W. H. Lau. Segmentation of color lip images by spatial fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, 11(4):542–549, 2003.

[11] L.Liang, Y.Luo, F.Huang, and A.V.Nefian. A multi-stream audio-video large-vocabulary mandarin chinese speech database. In *Proc. IEEE International Conference on Multimedia and Expo*, pages 1787–1790, 2004.

[12] Nakata and Ando. Lipreading method using color extraction method and eigenspace technique. *Systems and Computers in Japan*, 35(3):12–23, 2004.

[13] H. Ouyang and T. Lee. A new lip feature representation method for video-based bimodal authentication. In *Proceedings of the 2005 NICTA-HCSNet Multimodal User Interaction Workshop*, volume 57, pages 33–37, Sydney, Australia.

[14] S.Basu. A three-dimensional model of human lip motion. Master's thesis, Massachusetts Institute of Technology, 1997.

[15] S.Basu, N.Oliver, and A.Pentland. 3d modeling and tracking of human lip motion. In *Proc. IEEE International Conference on Computer Vision*, pages 337–343, Bombay, India, 1998.

[16] S.Werda, W.Mahdi, and A.B.Hamadou. Colour and geometric based model for lip localisation: Application for lipreading system. In *Proc. IEEE International Conference on Image Analysis and Processing*, pages 9–14, Modena, Italy, 2007.

[17] T.Chen and R.R.Rao. Audio-visual integration in multimodal communication. In *Proceedings of the IEEE*, pages 837–851, 1998.

[18] T.Coianis, L.Torresani, and B.Capril. 2d deformable models for visual speech analysis. In *NATO Advanced Study Institute: Speechreading by Man and Machine*, pages 391–398. Springer Verlag, 1995.

Hiding Emerging Patterns by using Local Recoding Generalization

Wai Kit CHENG

Abstract

Due to the increase in privacy-awareness of the public, the privacy-preserving data publishing (PPDP) has recently been a hot research topic. In this paper, we propose a new kind of privacy protection issue in PPDP, namely emerging patterns hiding. Emerging patterns refer to the itemsets exist in pair of datasets where their support has a significant difference in the two datasets. We adopt local recoding generalization for sanitization, which has been extensively studied in the context of k-anonymity but not in the context of hiding itemsets. One major contribution of this paper is to bring the generalization technique into the context of hiding emerging patterns. Also, we how one can preserve the frequent itemsets during the hiding process. Based on an existing distortion-based generalization quality metric, we propose a modified version of distortion that suits our problem setting.

1 Introduction

Data has known to be an invaluable asset to business. For all reasons that business data may be published to public, there are cases where data, among other things, is forced to publish for analysis, e.g., the investigation of the sales of certain "minibonds" products in certain banks in Hong Kong in 2008 [1]. Before data is released, the data owners may perfer to hide sensitive information, if possible. For example, there may be patterns whose supports are significant in a subset of data but not significant in another subset of data. Such patterns have known to be *emerging patterns*. For instance, certain saving plans may often be signed up by a certain group of bank customers but not popular among the overall customers. Such patterns can be directly related to a bank's marketing strategies and are therefore desirable to remain private. Certainly, such sensitive information may also appear in data from other industry sectors, including insurance, medical research and biology research [5, 18, 19, 20].

Published data is subjected to data analysis. In light of this, there has been a stream of work on *privacy-preserving data publishing (PPDP)* [4, 12, 17, 21, 22]. The objective of PPDP has been to protect different kinds of sensitive information being revealed by data mining.

In this paper, we study a particular case of privacy-preserving data mining where the sensitive information to be protected is emerging patterns in a dataset and the data analysis considered is one of the classical data mining technique, namely frequent itemset mining. More specifically, given two transactional datasets D_1 and D_2, a threshold of growth rate ρ and a threshold of support σ, we want to determine a transformation T on D_1 and D_2 such that the emerging patterns with a growth rate higher than ρ are eliminated and the distortion of ρ-frequent itemsets is minimized. (The details of the problem formulation shall be given in Section 4.)

Sanitization techniques, e.g., generalization [15, 4, 12, 23, 32], randomization [2, 7, 9], injection of unknowns [28], have been recently proposed to hide sensitive information. However, the sanitized data must not lose so much information such that further data analysis on the published data becomes meaningless. For instance, different kinds of data mining techniques, such as mining of frequent itemsets, classification and clustering, may be applied on the sanitized data. A sanitization is obviously unacceptable if a data mining technique would either (i) reveal sensitive information or (ii) obtain a distorted result from the sanitized data. In other words, the problem of sanitization is to protect sensitive information while minimizing the distortion of data mining results from a given data. The two competing objectives make the sanitization problem technically intriguing.

There have been a number of existing work on hiding frequent itemsets. For example, adding unknowns and removing items are widely adopted techniques for sanitizing data. However, either fake information generation or incomplete dataset are an intrinsic drawbacks of these techniques. In many real applications, e.g. data from hospitals, the completeness and truthfulness of the data are important.

In this paper, we adopt the generalization approach for sanitizing data. Generalization has been extensively studied and used in the context of achieving k-anonymity. Generalization is basically an items grouping process according to a given generalization hierarchy. Since the generalization hierarchy is always published with the data, data recipients

can correctly interpret every generalized values occurred in the sanitized data. As a result, the truthfulness and completeness of the sanitized dataset can be "blurred" but not lost. This works focuses on exploring the possibilities of adopting the generalization technique in the context of hiding itemsets.

Hiding emerging patterns is a more technically challenging task than hiding frequent itemsets. The reason the apriori anti-monotone property of frequent itemsets does not hold for emerging patterns. Subsequently, the total number of emerging patterns is often significantly larger than that of frequent itemsets in a dataset. Thus, the search space of emerging patterns cannot be pruned as effective as frequent itemsets. To the best of our knowledge, there have not been work on adopting generalization techniques to hide emerging patterns in a dataset.

The rest of this paper is organized as follows: In section 2, we give a brief summary on the field of PPDP. In Section 3, we provide background information on emerging patterns and generalization. In Section 4, we formulate our research problem. In Section 5, we propose a new metric for measuring the generalization quality. In Section 6, we describe our research plans on this study.

2 Related Work

Recently, there are works focused on different areas of PPDP. In the following, we give a brief summary on several popular PPDP techniques.

2.1 k-Anonymity

In [22], it showed that even if explicit identifiers had been removed from the data before publishing, personal identity can also be revealed by linking the published data with other external information. As such, the k-anonymity model was proposed to address this issue. A published data is said to be k-anonymized if at least k individuals are linked with a particular record in the published data even the data is cross-referenced with other external information. Various methods had been proposed for achieving k-anonymity, such as generalization [21, 32, 14] and suppression [21]. These methods sanitize the data without targeting any specific mining tasks on the sanitized data but use generic quality metric for guiding their sanitization process.

2.2 Privacy-Preserving Classification

In this kind of privacy notion[4, 13], they also target on achieving k-anonymity, but they used classification accuracy to measure the information loss instead of generic quality metrics. They made use of the class labels associated with the data to guide the sanitization process, as such, the difference in classification error before and after the sanitization process can be minimized.

2.3 Privacy-Preserving Clustering

This privacy notion was proposed in [11]. The major difference with the Privacy-Preserving Classification is no class labels can be use to guide the sanitization process. Therefore, they first extract the cluster structure from the raw data and encode it in the form of class labels. During the sanitization process, they try to preserve such class label maintain the cluster structure.

2.4 Frequent Itemset Hiding

Apart from k-anonymity, another privacy notion - Frequent Itemset Hiding, had been studied in [25, 29, 24]. In their studies, user need to first specify a subset of frequent itemsets, namely sensitive frequent itemsets, that they do not want to disclose publicly. The main objective of their work is to sanitize the data such that the sensitive frequent itemsets are absent in the sanitized data and retain as much non-sensitive frequent itemsets as possible.

3 Preliminaries

3.1 Emerging Patterns

Emerging patterns (EPs) [5] is a kind of special pattern exists in datasets and it was discovered in 1999. EPs are defined as itemsets whose supports increase significantly from one dataset/class to another. In other words, EPs are itemsets whose growth rates (i.e. the ratio of the support of an itemset in one dataset/class to that in the others) are larger than a given threshold.

Definition 1 Given two datasets, namely D_1 and D_2, the *growth rate* of an itemset *X*, denoted by *GR(X)*, from D_1 to D_2 is defined as *GR(X)* =

$$\begin{cases} 0 & \text{,if } Supp_{D1} = 0 \text{ and } Supp_{D2} = 0 \\ \infty & \text{,if } Supp_{D1} = 0 \text{ and } Supp_{D2} > 0 \\ \frac{Supp_{D2}(X)}{Supp_{D1}(X)} & \text{,otherwise} \end{cases}$$

Definition 2 For a growth rate threshold ρ, an itemset X is said to be a ρ-emerging pattern (ρ-EP) from D1 to D2 if GR(X) $>= \rho$.

Since EPs are describing some distinctive features from one class to the others, many works [16, 31, 10, 6] have been focused on how to use EPs for classification. The results of

many of these works have shown that the accuracy of the EP-based classifiers are better than of the traditional classification models, such as C5.0, Nave Bayes, CAEP. Due to the high classification power, EP-based classifiers have been successfully used for predicting the likelihood of diseases such as acute lymphoblastic leukemia [18] and discovering knowledge in gene expression data [19, 20].

Example 1 Table 1 shows a small, hypothetical dataset taken from [20] containing gene expression data, which records expression levels of genes under specific experimental conditions. There are 6 tissues samples in total: 3 normal and 3 cancerous tissues. Each tissue sample is described by the 4 gene expressions (namely, gene_1, gene_2, gene_3 and gene_4).

We call gene_j@[l, r] an item, meaning the values of expression of gene_j is limited inclusively between l and r. Inspecting Table 1, we find the following interesting patterns.

- The pattern {gene_1@[0.3, 0.5], gene_4@[0.41, 0.82]} has a frequency of 0
- The pattern {gene_2@[1.1, 1.3], gene_3@[-0.83, -0.7]} appear three times in the sub-dataset with normal cells but only once with cancerous cells.

These patterns represent a group of gene expressions that have certain ranges of expression levels frequently in one type of tissue but less frequently in another. Therefore, they are excellent discriminators to distinguish the normal and cancer cells.

ID	Cell type	gene_1	gene_2	gene_3	gene_4
1	Normal	0.1	1.2	-0.7	3.25
2	Normal	0.2	1.1	-0.83	4.37
3	Normal	0.6	1.3	-0.75	5.21
4	Cancerous	0.4	1.4	-1.21	0.41
5	Cancerous	0.5	1.1	-0.78	0.75
6	Cancerous	0.3	1	-0.32	0.82

Table 1. A simple gene expression dataset

3.1.1 Emerging Pattern Mining

Mining EPs efficiently is a challenging problem because of two reasons. First, the total number of EPs present in large datasets is very huge. In the worst case, the total number of EPs growth exponentially with respect to the total number of attributes. Thus, it is almost computationally infeasible to enumerate all the EPs exhaustively. Second, the Apriori anti-monotone property of frequent itemset - every subset of a frequent itemset must also be frequent, does not hold for EP. As such, we cannot adopt the search space pruning strategy in frequent itemset mining.

There are several approaches developed in the literature to address the EP mining issue:

Border-based Approach
In [5], EPs are represented by borders and border differential operation is used for discovering EP. Given D_1, D_2 and growth rate threshold ρ, we first fix the minimum support threshold σ_1 for D_1. Then we can use a border-discovery algorithm such as Max-Miner [26] to obtain the large border (i.e. the maximal frequent itemsets) for D_1. For D_2, we use the border-discovery algorithm to find the large border by using the minimum support threshold $\sigma_2 = \sigma_1 \cdot \rho$. After both large borders for D_1 and D_2 are found, the border of EPs can be obtained by using the border differential operation.

Constraint-based Approach
ConsEPMiner [33] utilizes two major types of constraint to prune the search space of EPs effectively: External constraints are user-given minimums on support, growth rate, and growth-rate improvement to confine the resulting EP set. Inherent constraints, including same subset support, top growth rate, and same origin are derived from the properties of EPs and datasets for pruning the search space and saving computation.

Jumping emerging patterns (JEPs)
JEPs is a special kind of EP which have infinite growth rate. In [3], fast algorithms for mining JEP were proposed and its performance is typically around 5 times faster than the border-based approach. It used tree construction approach to target the likely distribution of JEPs.

3.2 Generalization

In privacy-preserving data mining, generalization is one of the most common techniques for sanitizing datasets in order to achieve anonymization - to prevent the inference of personal identity from released data. The key idea of generalization is to modify the original values in dataset into more general values such that more tuples will share the same set of attribute values. Thus, anonymization can be achieved.

When comparing with other sanitizing techniques, such as removing data, randomization and using unknowns to replace some data in datasets, generalization has several advantages.

First, the sanitized dataset by generalization is semantically consistent with the original dataset. Unlike other techniques, the information conveyed by a generalized dataset

is always representing the truth and having no missing values or records. Second, since the generalization hierarchy is published with the sanitized dataset, the generalized values in the published dataset can always be correctly interpreted by the recipient of the dataset.

There are three major types of generalization have been studied and adopted in the literature of achieving k-anonymity, namely single-dimensional global recoding, multidimensional global recoding and local recoding.

Single-dimensional global recoding
Single-dimensional global recoding was studied in [4, 12, 17, 27, 21, 30]. It performs generalization at domain-levels. It changes a value in a single domain to another globally. Under this kind of generalization, if we decide to generalize a particular value, all tuples contain this value will be generalized. As a result, the original datasets are very often to be over-generalized.

Multidimensional global recoding
Multidimensional global recoding was studied in [15] .It performs generalization at cell levels. After this kind of generalization has performed both original and generalized values may co-exist in the generalized dataset. However, it has a constraint on *equivalence class*. A set of tuples are said to be in a *equivalence class* with respect to a set of attributes $T = \{t_1, t_2, \ldots, t_m\}$ if they all contain the same attribute values for T. In multidimensional global recoding, all equivalence classes in the original dataset should not be generalized into two or more difference equivalence classes in the generalized dataset. Since it allows partial generalization on the attribute values, over-generalization can be avoided.

Local recoding
Local recoding was studied in [8, 23, 32, 14]. It also performs generalization at cell levels. In fact, it is the same as the multidimensional global recoding except it relaxes the constraint on equivalence class. In this paper, we mainly focus on this kind of generalization.

3.2.1 Metrics for generalization quality

In many of the previous works, different kinds of quality metric have been proposed for measuring the quality of generalization. Most of them are utility-based metrics (i.e. the metrics are associated with specific mining tasks), such as classification accuracy. However, we normally do not know what kind of mining tasks will be performed by data recipient on the released dataset. As such, general-purpose quality metrics should be used instead as a fairer measurement on the generalization quality. Following are several general-purpose metrics for generalization quality proposed by recent works:

Discernability metric (DM)
DM was proposed by Bayardo et al. [4]. It measures the generalization quality by the size of equivalence classes. Tuples in a dataset are said to be in the same equivalence class with respect to a set of attributes $X_1, X_2, \ldots X_d$ if they all have the same attribute values for $X_1, X_2, \ldots X_d$. The Discernability metric is defined as follow:

$$DM = \sum_{EquivClasses\ \mathrm{E}} |E|^2$$

where $|E|$ is the size of equivalence class

Clearly, the smaller the size of equivalence classes, the better the generalization quality is. Therefore, if DM is used as a metric, the objective during generalization is to minimize the discernability.

Normalized average equivalence class (CAVG)
CAVG was proposed by LeFevre et al. [8]. It is also a metric based on the size of equivalence classes. But instead of simply calculate the sum of square of the equivalence classes size, it calculate the average class size. CAVG is defined as follow:

$$CAVG = \left(\frac{\text{Total no. of records}}{\text{Total no. of equivlance classes}} \right) / k$$

Similar to DM, the objective during generalization is to minimize the CAVG.

Normalized certainty penalty (NCP)
NCP was proposed by Xu et al. [32]. It is a metric for measuring the degree of generalization of each attribute values. The intuition is to calculate the ratio of the total number of values which it generalized with to the total number of different attribute values. In the case of categorical attributes, NCP for a single attribute value is defined as follow:

$$NCP(p) = \begin{cases} 0 & \text{,if } |u_p| = 1 \\ |u_p|/|I| & \text{,otherwise} \end{cases}$$

,where u_p is the immediate antecedent of p in the generalization hierarchy, $|u_p|$ is the number of leave nodes under $|u_p|$ and $|I|$ is the total number of leave nodes for the entire generalization hierarchy. Then, the NCP for the whole dataset D is defined as:

$$NCP(D) = \frac{\sum_{p \epsilon I} C_p \cdot NCP(p)}{\sum_{p \epsilon I} C_p}$$

where C_p is the total number of occurrences of item p in the dataset.

Distortion
Distortion [14] measures the distortion caused by generalization which takes the hierarchical structure of attribute into consideration. It first defines a metric measuring the distance between different levels in an attribute hierarchy, called Weighted Hierarchical Distance (WHD). The WHD between level p and level q is defined as:

$$WHD(p,q) = \frac{\sum_{j=q+1}^{p} w_{j,j-1}}{\sum_{j=2}^{h} w_{j,j-1}}$$

where h is the height of the domain hierarchy and $w_{j,j-1}$ is the weight between domain level j and $j-1$. The distortion of a generalization which generalizes tuple $t = \{v_1, v_2, \ldots, v_m\}$ to $t' = \{v'_1, v'_2, \ldots, v'_m\}$ is defined as:

$$Distortion(t,t') = \sum_{j=1}^{m} WHD(level(v_j), level(v'_j))$$

where level(v_j) be the domain level of v_j in an attribute hierarchy. Finally, the distortion for the whole dataset is defined as:

$$Distortion(D,D') = \sum_{i=1}^{|D|} Distortion(t_i, t'_i)$$

Inconsistency
In multidimensional global recoding and local recoding generalization, the attribute values in the same attribute may belong to different domains. In other word, both original and generalized attribute values may co-exist in the same attribute. This maybe one of the obstacles for performing analysis on the published data in many real world applications. Inconsistency [14] is a metric for measuring the domain diversity of attribute. Inconsistency for a single attribute i is defined as:

$$inconsist_i = (1 - max_j(p_{ij}))$$

where p_{ij} is the fraction of values in domain level j of attribute i over all values in attribute i. Then, the inconsistency for the whole dataset D is defined as:

$$inconsist_D = max_i(inconsist_i)$$

4 Problem Formulation

Given a dataset D, the support of an itemset X, denoted as $Supp(X)$, is the percentage of tuples in D that contain X. For a support threshold σ, X is said to be σ-$frequent$ if $Supp_D(X) >= \sigma$.

Given D, σ, and ρ , our goal is to adopt local recoding generalization to transform D to D' such that no EPs can be mined from D' while the distortion on σ-$frequent\ itemsets$ is minimized (i.e. to preserve as much σ-$frequent\ itemsets$ as possible).

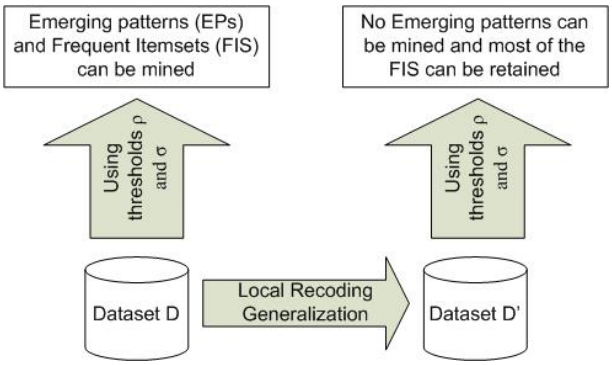


Figure 1. A graphical illustration of the problem definition

4.1 Local Recoding Generalization

In global recoding generalization, if we decide to generalize two or more attribute values together, all tuples which contain either one of the values will be affected. As a result, the original dataset will most likely to be over-generalized. Thus, the σ-$frequent\ itemsets$ in D will suffer from a significant level of distortion.

In order to strike for a better balance between hiding EP and preserving σ-$frequent\ itemsets$ in D, local recoding generalization is a more suitable approach for the transformation. In local recoding generalization, partial generalization on attribute values is allowed. For example, C and D are the attribute values to be generalized to a new values CD. After local recoding generalization, the values C, D and CD may co-exist in the transformed dataset D'. That is, only a portion of the attribute values C and D are generalized.

In the following, we give an illustrative example to use local recoding generalization to hide EP and preserve σ-$frequent\ itemsets$:

Given that $\sigma = 40\%$ and $\rho = 3$, we now particularly consider the EP = $\{A\ C_1\}$ and the σ-$frequent\ itemset$ = $\{C_1\ F\}$ exist in the original dataset (Table 2).

If we want to hide the EP = $\{A\ C_1\}$, one of the possible way is to perform generalization on attribute 2 according to the hierarchy in figure 2.

If we adopt global recoding generalization approach for hiding the EP = $\{A\ C_1\}$, all the attribute values C_1, C_2, D_1 and D_2 in the dataset are all generalized to CD. The

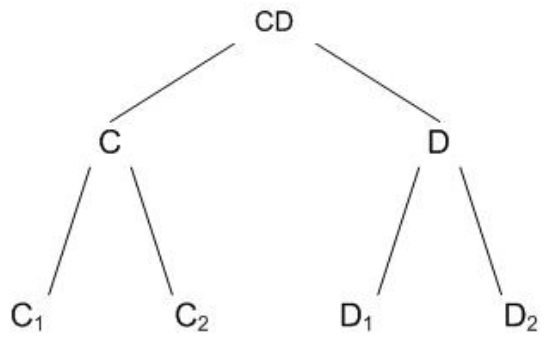


Figure 2. Generalization hierarchy for attribute 2

σ-$frequent\ itemset$ = {C$_1$ F} disappeared because the dataset is over-generalized (Table 3).

If we adopt local recoding generalization approach for hiding the EP = {A C$_1$}, only a portion of C$_1$, C$_2$, D$_1$ and D$_2$ are generalized to CD. As a result, the EP = {A C$_1$} is removed while the σ-$frequent\ itemset$ = {C$_1$ F} is retained. (Table 4)

Class C_1			Class C_2		
Attr. 1	Attr. 2	Attr. 3	Attr. 1	Attr. 2	Attr. 3
B	C$_1$	F	B	C$_1$	F
A	C$_1$	F	B	C$_2$	F
A	C$_1$	F	B	C$_1$	F
B	C$_2$	F	B	C$_1$	F
B	D$_1$	E	A	D$_2$	E

Table 2. Original Dataset

Class C_1			Class C_2		
Attr. 1	Attr. 2	Attr. 3	Attr. 1	Attr. 2	Attr. 3
B	CD	F	B	CD	F
A	CD	F	B	CD	F
A	CD	F	B	CD	F
B	CD	F	B	CD	F
B	CD	E	A	CD	E

Table 3. Transformed dataset by adopting global recoding generalization

5 Metrics for the quality of generalization

In this section, we discuss several metrics needed for measuring the quality of the transformed dataset and guide the decision on choosing each candidate generalization.

Class C_1			Class C_2		
Attr. 1	Attr. 2	Attr. 3	Attr. 1	Attr. 2	Attr. 3
B	C$_1$	F	B	C$_1$	F
A	CD	F	B	C$_2$	F
A	CD	F	B	C$_1$	F
B	C$_2$	F	B	C$_1$	F
B	D$_1$	E	A	CD	E

Table 4. Transformed dataset by adopting local recoding generalization

5.1 Measure the distortion on σ-frequent itemsets

Since generalization is a process of grouping existing attribute values to some new attribute values, some *σ-frequent itemsets* may disappear because of some of their attribute values no longer exist after the generalization. On the other hand, some *σ-frequent itemsets* consist of generalized attribute values maybe newly generated.

For example, consider the global generalization which transformed the dataset from Table 1 to Table 2. After the generalization, the *σ-frequent itemsets* {C F} which exists in the dataset disappeared and another *σ-frequent itemsets* {CD F} is newly generated.

Therefore, we need a metric for measuring the distance between the *σ-frequent itemsets* in D and D'.

First, we will present our proposed metric in the case of global recoding generalization. After that, we will show how the metric can be adapted to the case of local recoding generalization.

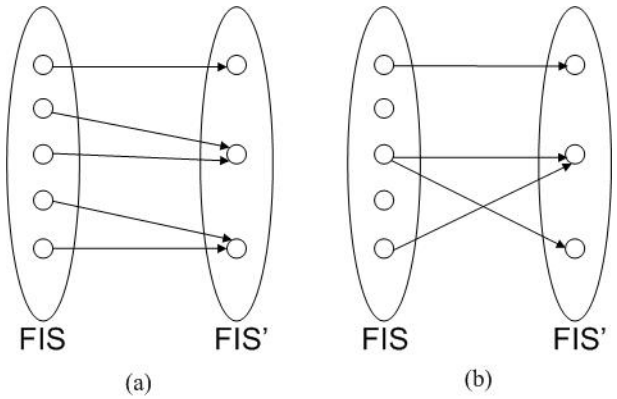


Figure 3. The relationship between FIS and FIS' in (a)global recoding (b)local recoding

5.1.1 Global Recoding

In global recoding, no *σ-frequent itemsets* (FIS) will be disappeared after the process of generalization but they may appear in a generalized form (Figure 2(a)). In other word, for a particular *σ-frequent itemsets* in D, there always exists a corresponding *σ-frequent itemsets*(FIS') in D' in either the original or generalized form.

Inspired by the distortion metric proposed in [14], we propose a metric for measuring the *generalization distance* (GD) between the original and generalized form of a tuple which has taken both the hierarchical distance and distortion in every single generalization into consideration.

Here, we define a metric called *value distance* (VD) for measuring the distance between the original and generalized form of a single attribute value. We will then use VD as a building block for the definition of generalization distance (GD).

Definition 3 (Generalization Distortion (GenDist)) Consider a generalization G which generalizes a set of attribute values $\{\mathbf{c}\}$ to a single generalized value p. The generalization distortion of G is defined as:

$$GenDist(G) = |c|,$$

where $|c|$ is the number of leave nodes under p in the generalization hierarchy.

Definition 4 (Value Distance (VD)) Let h be the height of the hierarchy tree of the attribute. Level h is the most general level and level 0 is the most specific level in the hierarchy tree. Consider an attribute value v at level p is generalized to v' at level q. Let G_n be the generalization which generalizes the attribute from level n-1 to n, where $0 < n \leq h$. The valude distance between v and v' is defined as:

$$VD(v, v') = \sum_{i=p}^{q} \frac{i \cdot GenDist(G_i)}{h}$$

Definition 5 (Tuple Distance (TD)) Consider a tuple $T = t_1, t_2, \ldots, t_n$ is generalized to $T' = t'_1, t'_2, \ldots, t'_n$ for $n > 0$. The tuple distance between T and T' is defined as:

$$TD(T, T') = \sum_{i=1}^{n} VD(t_i, t'_i)$$

Definition 6 (Generalization Distance (GD)) Let FIS = $\{f_1, f_2 \ldots f_n\}$ be the set of *σ-frequent itemsets* in D and FIS' = $\{f'_1, f'_2 \ldots f'_n\}$ be the set of *σ-frequent itemsets* in D', where $0 < m \leq n$. In the case of global recoding generalization, for every f_n, there must be a corresponding f'_m is either the original or generalized form of f_n, we denote f'_m as G(f_n).

The generalization distance between FIS and FIS' is defined as:

$$GD(FIS, FIS') = \sum_{i=1}^{n} TD(f_i, G(f_i))$$

For example, the generalization distance incurred in table 3 can be calculated as follow:

$$\begin{aligned} GD(FIS, FIS') &= TD(\{C_1, F\}, \{CD, F\}) \\ &= VD(C_1, CD) \\ &= \sum_{i=0}^{2} \frac{i \cdot GenDist(G_i)}{h} \\ &= \frac{1 \times 2}{2} + \frac{2 \times 4}{2} \\ &= 5 \end{aligned}$$

5.1.2 Local Recoding

In the case of using local recoding generalization, two problems for calculating the generalization distance (GD) will be occurred. (Figure 2(b))

(1) An itemset in FIS having no correspondence in FIS' Since local recoding generalization allows attribute values to be partially generalized, it allows only a portion of the tuples which contains a particular itemset in FIS to be generalized and the remaining portion be retained in the original form. In the case of the support of both portions are smaller than σ, the particular itemset in FIS may find no corresponding itemsets in FIS' neither in original nor generalized form.

Solution: The generalization hierarchy is published with D'. Thus, everyone can laways replace attribute values in D' with more general values according to the hierarchy.

As such, if users (i.e. the one who receive D' and the generalization hierarchy) generalize all the attribute values to the most general form exists in each attribute of D', they may discover more *σ-frequent itemsets*, in generalized form, which are hidden before.

Therefore, if no correspondence can be found in FIS' for a particular itemset, f_x, in FIS, we may first create a tuple, f_{max}, by generalizing each attribute values in FIS to the most general form exists in D'. Then we can calculate the generalization distance (GD) by using the tuple distance between f_x, in FIS.

(2) An itemset in FIS having more than one itemsets exist in FIS' Another side effect of local recoding generalization is that there may exist more than one correspondence in FIS' for a particular itemset, f_x, in FIS.

Solution: In this case, we can use the corresponding itemset in FIS' which has minimum tuple distance (TD) with f_x to calculate the generalization distance (GD)

After taken the above two problems in local recoding generalization into consideration, the definition of generalization distance (GD) can be refined as:

$$GD(FIS, FIS') = \sum_{i=1}^{n} TD_i$$

where TD_i =

$$\begin{cases} TD(f_i, G(f_i)) & \text{,if no. correspondence in FIS' = 1} \\ TD(f_i, f_{max}) & \text{,if no. correspondence in FIS' = 0} \\ min(TD(f_i, G(f_i)) & \text{,if no. correspondence in FIS' } > 1 \end{cases}$$

For example, the generalization distance incurred in table 3 is zero. It is because the only frequent itemset $\{C_1,F\}$ is preserved. This is also the major benefit of local recoding generalization - the distortion on the data can be minimized.

5.2 Measure the efficiency for each generalization

In order to choose the best candidate generalization in each step, we need a metric for measuring the efficiency for each generalization. After such a metric is defined, then we may choose the candidate generalization which yields the highest efficiency in each step.

The efficiency for a generalization G is defined as:

$$Efficiency(G) = \frac{\textit{Total Growth rate reduction}}{GD(FIS, FIS')}$$

By defining the efficiency as above, the higher the efficiency of a generalization means it can remove a larger portion of EP and only sacrifice a smaller portion of *σ-frequent itemsets*.

After this metric is defined, we may use it as an objective function to develop a heuristic for performing generalization.

6 Future Work

In the near future, we will use the generalization efficiency as an objective function to develop heuristic for choosing candidate generalizations at each step to hide emerging patterns.

In addition, some optimization techniques may be adopted for pruning the set of candidate generalizations. As such, we need not to take all possible generalizations into consideration and thus the runtime of the heuristic can be shorten.

We will evaluate the performance of the proposed heuristic by performing frequent itemsets mining on both the original and the sanitized dataset. By compare the mining results, we can know how much frequent itemsets is actually preserved during sanitization.

References

[1] Hong kong legislative council resolution under the legislative council (powers and privileges) ordinance (cap. 382) moved by ir dr. hong raymond ho chung-tai at the legislative council meeting of wednesday, 12 november 2008 wording of motion available at http://www.legco.gov.hk/yr08-09/english/legcorpt/legcoreso1113-e.pdf, 2008.

[2] D. Agrawal and C.Aggarwal. On the design and quantification of privacy preserving data mining algorithms. In *Proc. of the 20th ACM SIGACT-IGMOD-IGART Symposium on Principles of Database Systems*, Santa Barbara, California, USA, May 32-23, 2001.

[3] J. Bailey, T. Manoukian, and K. Ramamohanarao. Fast algorithms for mining emerging patterns. In *Proc. of the 6th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'02)*, Helsinki, Finland, 2002.

[4] R. Bayardo and R. Agrawal. Data privacy through optimal k-anonymization. In *Proc. 21st Int'l Conf. Data Eng. (ICDE'05)*, pages 217–228, 2005.

[5] G. Dong and J. Li. Efficient mining of emerging patterns:discovering trends and differences. In *Proc. KDD'99*, volume 3rd. ed., pages 43–52, San Diego, CA, USA, 1999.

[6] G. Dong, X. Zhang, and L. Wong. Caep: Classification by aggregating emerging patterns. In *Proc. of the 2nd Int'l Conference on Discovery Science(DS'99)*, pages 30–42, Tokyo, Japan, 1999(December).

[7] W. Du and Z. Zhan. Using randomized response techniques for privacy-preserving data mining. In *Proc. of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, DC, USA, August 24-27, 2003.

[8] Y. Du, T. Xia, Y. Tao, D. Zhang, and F. Zhu. On multidimensional k-anonymity with local recoding generalization. In *Proc. 23 rd Int'l Conf. Data Eng. (ICDE'07)*, pages 1422–1424, 2007.

[9] A. Evfimievski, R. Strikant, R. Agrawal, and J. Gehrke. Privacy preserving mining of association rules. In *Proc. of 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton, Alberta, Canada, 2002.

[10] H. Fan and K. Ramamohanarao. A bayesian approach to use emerging patterns for classification. In *Proc. 14th Australasian Database Conference (ADC2003)*, pages 39–48, 2003.

[11] B. Fung, K. Wang, L. Wang, and M. Debbabi. A framework for privacy-preserving cluster analysis. In *Proc. of the 2008 IEEE International Conference on Intelligence and Security Informatics (ISI)*, page 4651, Taipei, Taiwan, 2008.

[12] B. Fung, K. Wang, and P. Yu. Top-down specialization for information and privacy preservation. In *Proc. 21st Int'l Conf. Data Eng. (ICDE'05)*, pages 205–216, 2005.

[13] B. Fung, K. Wang, and P. Yu. Anonymizing classification data for privacy preservation. In *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, volume 10, no.5, pages 711–725, May, 2007.

[14] J.Li, R. Wong, A. Fu, and J. Pei. Anonymization by local recoding in data with attribute hierarchical taxonomies. In *Proc. IEEE Transactions on Knowledge and Data Engineering*, pages 1181–1194, 2008.

[15] K.LeFevre, D. Dewitt, and R.Ramakrishnan. Mondrian multidimensional k-anonymity. In *Proc. 22nd Int'l Conf. Data Eng.(ICDE'06*, page 25, 2006.

[16] H. K.Ramamohanarao. Pattern based classifiers. In *World Wide Web*, pages 71–83, 2007.

[17] K. LeFevre, D. Dewitt, and R. Ramakrishnan. Incognito:efficient full-domain k-anonymity. In *Proc. 24th ACM SIGMOD '05*, pages 49–60, 2005.

[18] J. Li, H. Liu, J. Downing, A.-J. Yeoh, and L. Wong. Simple rules underlying gene expression profiles of more than six subtypes of acute lymphoblastic leukemia (all) patients. In *Bioinformatics*, volume 19(1), pages 71–78, 2003.

[19] J. Li, H. Liu, S.-k. Ng, and L. Wong. Discovery of significant fules for classifying cancer ddiagnosis data. In *Bioinformatics*, volume 19(Suppl. 2), pages ii93–ii102, 2003.

[20] J. Li and L. Wong. Identifying good diagnostic gene groups from gene expression profiles using the concept of emerging patterns. In *Bioinformatics*, pages 725–734, 2002.

[21] L.Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. In *Int'l J. Uncertainty, Fuzziness and Knowledge Based Systems*, pages 571–588, 2002.

[22] L.Sweeney. k-anonymity: a model for protecting privacy. In *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, pages 557–570, 2002.

[23] A. Meyerson and R. Williams. On the complexity of optimal k-anonymity. In *Proc. 23rd ACM SIGMOD-SIGACT-SIGART Symp. Principles of Database Systems (PODS '04)*, pages 223–228, 2004.

[24] G. Moustakides and V. Verykios. A maxmin approach for hiding frequent itemsets. In *Proc. of Data & Knowledge Engineering*, volume 65, issue 1, pages 75–79, 2008.

[25] S. Oliveira and O.R.Zaiane. Privacy preserving frequent itemset mining. In *Proc. of the IEEE international conference on Privacy, security and data mining*, volume 14, pages 43–54, Maebashi City, Japan, 2002.

[26] B. J. R.J. Efficiently mining long patterns from databases. In *Proc. of the 1998 ACM-SIGMOD Int'l Conference Management of Data*, pages 85–93, Seatle, WA, 1998.

[27] P. Samarati. Protecting respondents' identities in microdata release. In *IEEE Trans. Knowledge and Data Eng.*, pages 1010–1027, Nov./Dec. 2001.

[28] Y. Saygin, V. Verykios, and C. Clifton. Using unknowns to prevent discovery of association rules. In *SIGMOD Record*, volume 30(4), pages 45–54.

[29] X. Sun and P. Yu. A border-based approach for hiding sensitive frequent itemsets. In *Proc. of the Fifth IEEE International Conference on Data Mining*, pages 426–433, 2005.

[30] K. Wang, P. Yu, and S. Chakraborty. A data mining solution to privacy protection. In *Proc. Fourth IEEE Int'l Conf. Data Mining (ICDM'04)*, pages 249–256, 2004.

[31] Z. Wang, H. Fan, and K. Ramamohanarao. Exploiting maximal emerging patterns for classification. In *Proc. 17th Australian Joint Conf. on Artificial Intelligence*, pages 1062–1068, Cairns, Queensland, Australia, 2004(December).

[32] J. Xu, W. Wang, J. Pei, X. Wang, B. Shi, and A. Fu. Utility-based anonymization using local recoding. In *Proc. 12th ACMSIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD'06)*, pages 785–790, 2006.

[33] X. Zhang, G. Dong, and K. Ramamohanarao. Exploring constraints to efficiently mine emerging patterns from large high-dimensional datasets. In *Proc. 6th ACM SIGKDDD Int'l Conf. Knowledge Discovery and Data Mining (KDD'00)*, pages 310–314, Boston, 2000.

Improving patient journey with reduced length of stay by using a multi-agent approach

Chung Ho CHOI

Abstract

Delivering healthcare services in high quality is one of the most challenging and critical tasks for healthcare providers. For the sake of minimizing dissatisfactions, healthcare managers are now struggling to provide an improved patient journey for their clients. In doing so, it is thought that the most unambiguous way is to reduce the length of stay (LoS) of a patient journey. Hence, an effective patient scheduling scheme is needed. Although there are quite a lot of different approaches for scheduling, an multi-agent approach is proposed in view of having a distributed and dynamic nature in hospitals. Currently, in Hong Kong, a lengthy length of stay is not uncommon as patient schedulings in hospitals are not being done in an efficient manner. In particular, patients in Hong Kong are being scheduled without a sophisticated mechanism such that they have to endure long waiting times for treatments which are discouraging and life-threatening. Hence, in this paper, we will address how a multi-agent approach could be used for patient scheduling such that a reduced length of stay is achieved.

1. Introduction

With the constraint of having limited resources, it is difficult for healthcare providers to deliver healthcare services that meet the increasing expectations from patients. In spite of facing such kind of difficulty, it is imperative for healthcare managers to raise the standard for healthcare services such that it meets the regulations enacted by the government and the needs of different patients. However, it can sometimes be subjective in defining improvements for healthcare services. For instance, the improvement of providing a better environment in hospital settings can easily be biased as patients possess different values in evaluating things. Thus, in order to improve healthcare services in an objective fashion, the most obvious and direct way is to reduce the length of stay (LoS) of a patient journey. In particular, patient flows should be improved such that undesired delays can be minimized if not entirely eliminated. More importantly, it is known that reduced delays in patient journeys can dramatically improve medical outcomes and patient satisfactions [1]. Hence, it is wise for the healthcare providers to figure out an effective scheme for patient scheduling in order to reduce unnecessary delays.

In fact, patient scheduling can be widely applied in healthcare management. Particularly, an effective scheduling scheme can be deployed for optimizing healthcare operations such that there is a better match between the supply and demand of medical resources. For instance, with the aid of a proper scheduling scheme, hospitals can conduct admission planning [2, 11], patient mix optimization [3], etc such that medical resources can be utilized in the most efficient manner. However, although these kinds of scheduling can boost the efficiency of resources utilization, they are not utterly patient-focused and may not guarantee a conspicuous reduction of delays in patient journeys. In particular, such inability of achieving a considerable reduction of delays is mainly caused by the fact that delays can still occur with better resource utilization if there is an increasing number of patients being admitted. Therefore, with the goal of reducing delays or length of stays for patients, one should focus on scheduling schemes that consider the whole picture of patient journey.

It is known that hospitals have a decentralized and dynamic structure [4, 5]. In particular, patients need to be routed to different operation units within a hospital or even different hospitals for different treatments. With such insight, the approach chosen for patient scheduling should be able to deal with the unique hospital structure and thus should be dynamic as well. Hence, with the strength of dealing with dynamic environments and the ability of overseeing the entire patient journey, an multi-agent approach is proposed for effective patient scheduling.

In order to limit our scope, we decide to focus on scheduling for cancer patients in Hong Kong. The reason for choosing cancer patients is that neoplasms (i.e. cancers) have the highest mortality rate among different diseases in this city [18]. Currently, in Hong Kong, there are seven cancer centers [19] which are located at different districts (i.e. Hong Kong Island, Kowloon, and New Territories).

The problem is that each cancer center operates on its own and does not have information about the others. Thus, in receiving treatments, patients are frequently being assigned to one particular center, even though it is a busy one. In other words, patients in a particular cancer center may have to wait long for treatments despite the same could be provided earlier by other centers which are however unknown to the scheduler.

Hence, in this paper, we are going to demonstrate how a multi-agent approach can be used for patient scheduling such that cancer patients in Hong Kong can receive treatments in different cancer centers with a reduction of waiting times. And the paper is organized as follows: in section 2, we will first discuss what is patient journey and the causes and consequences of having delays in it. Thereafter, in section 3, we will discuss how patient scheduling could be done by using a multi-agent approach. In section 4, we will formulate the scheduling problem for cancer patients in Hong Kong. Lastly, in section 5, we will present our conclusion and future work.

2. Patient Journey

Patient journey is generally defined as the process of how patients proceed through the care delivery system [6]. While patient journey is the pathway in which every patient must go through, it is also believed that how the patients feel about the quality of healthcare services depend heavily on their experiences in patient journeys. Hence, patient journey deserved to receive well attentions from healthcare providers.

Generally speaking, a patient journey includes five main stages, namely clinical assessment, investigation, clinical decision, admission, treatment, and discharge [2, 6, 7]. However, some variations may exist due to some specific circumstances. For instance, a follow up (or a revisit) after discharge would be required if there exists a deterioration.

There are lots of factors contributing to patients' satisfactions amid patient journey. Particularly, the following eight factors have been identified as the most important by patients [7]:

- Fast access to health advices
- Effective treatments delivered by staff whom the patients can trust
- Patients' involvement in decisions and their preferences being respected
- Clear, comprehensible information and support to regain or increase independence
- Physically comfort, clean, and safe environment
- Emotional support and alleviation of anxiety
- Involvement of family and friends
- Continuity of care and smooth transitions

But in reality, lots of the above factors are missing and patients often feel confused, disappointed, and frustrated [8]. In such condition, patients are said to have bad emotional experiences.

2.1. Managing patient's emotional experience

Patient's emotional experience is how a patient feels about his/her experience of using medical service [8]. Although it may often be neglected, a good emotional experience is critical in patient journey as it has a close relationship with effective treatment. In particular, it is found that a bad emotional experience will result in increased psychosocial morbidity [9]. Moreover, some patients said that their feelings are linked to their situations and medical conditions [8].

Hence, in order to deliver healthcare service that meets the needs of patients, one should address patient's emotional experience properly. In doing so, it is thought that the most effective and unbiased way is to reduce the length of stay (LoS) for patients as it could be done in the most objective manner. With such idea, it is important to eliminate undesired delays such that the length of stay could be kept as minimal.

2.2. Causes of delays in patient journey

One of the key causes of having delays amid patient journey is handoffs, where patient care or information is handed from one individual to another [6]. For instance, a patient may be routed through different hospital units for different operations and treatments. In such routing, delays are common as the schedules in different units are seldom considered as a whole.

A high degree of uncertainty also yields delays in patient journey. As treatment pathways and the arrivals of patients are stochastic [4, 10, 11], it is difficult for healthcare providers to come up with an optimal schedule that caters the will of each patient. Hence, undesired delays seem to be a natural consequence for patient journey.

Another possible cause of having delays is the existence of time lags in obtaining necessary investigations. Some diseases may need to be investigated for several months. For instance, a suspected case of bowel cancer may take about four months to investigate [12]. Interestingly, besides investigations, patients' preferences may also cause delays in patient journey [9]. And we do think that such kind of delay is mainly caused by a lack of supportive guidance or information.

While it is not uncommon, shortage of resources possesses lots of challenges for healthcare providers. Particularly, competitions for scarce resources between hospital units may cause bottlenecks and hence delays amid patient journey.

2.3. Consequences of having delays in patient journey

While some studies pointed out that the impacts of having delays in patient journey up to several months are minimal [13], some others do not. For instance, it was found that the survival rates for colon cancer and female breast cancer in the first six months after diagnosis in England and Wales are lower than the rest of Western Europe. And it is believed that such differences may be related to late presentations or delays in treatments for British patients [14].

Moreover, it was found that long waiting time may cause psychosocial stress for patients [9]. The situation is even worse when there is inadequate information available for patients. And we believe that waiting time should be shortened as much as possible such that patients can receive treatments at the most appropriate time with minimal adverse impacts.

With the existence of undesired delays, patients' conditions may deteriorate rapidly and they may require urgent medical treatments. And such kind of urgency is thought to induce an adverse effect on the original waiting list [1]. For instance, a patient admitted urgently for a surgery may occupy an operating room and a team of medical staff which are originally reserved for another patient. Under such circumstances, the patient who has been initially scheduled for a surgery has to wait for another available timeslot. Hence, a chain reaction with increasing delays would occur if undesired delays cannot be curbed at the very beginning.

2.4. Reducing delays in patient journey

With the goal of improving patient journey, it is imperative for healthcare providers to take the initiatives to minimize undesired delays. In doing so, the most common practice is to increase the resource utilization such that the corresponding patient throughput can be maximized. However, it is found that an increased resource utilization can lead to bottlenecks which in turn inhibit patient flows [11]. To illustrate, consider there is a group of patients who have to undergo an operation shortly after a X-ray examination. Although the number of patients being served will increase with better utilization of X-ray machines, patient flows could still be inhibited if the operation rooms are already operated in their full capacities. Hence, solely increase the resource utilization is not an ideal strategy for reducing delays in patient journey.

In fact, while delays can be minimized through careful forecasting, process improvement, and information management [1], it is thought that the most effective and efficient way for reducing delays amid patient journey is to devise a proper scheduling scheme.

Generally, patient schedulings can be classified into two categories: resource-focused scheduling and patient-focused scheduling.

2.4.1. Resource-focused scheduling

Resource-focused scheduling emphasizes the improvement of resource utilizations. It makes the assumption that patients' waiting times could be reduced with better resource utilizations. In addition, instead of taking the whole patient journey into account, resource-focused scheduling aims at improving the efficiency of a particular medical resource or unit. Two common examples are admission planning [2, 11] and patient mix optimization [3] which both aim at improving the resource utilization by finding a better match between the supply and demand of medical resources.

However, while it is thought that resource-focused scheduling may help to reduce delays in some extents, there is no guarantee. For instance, though the patient throughput may increase with better resource utilization, patients could still have to wait long if there is an increasing number of patients being admitted. Hence, with such deficiency of resource-focused scheduling, it would be wise to pay much attention to patient-focused scheduling if reducing delays is considered as the ultimate goal.

2.4.2. Patient-focused scheduling

Unlike resource-focused scheduling, patient-focused scheduling emphasizes the reduction of waiting times for patients and takes the whole patient journey into account. It aims at finding an improved if not optimal schedule such that patients can enjoy themselves with fewer undesired delays.

With the goal of reducing delays in a decentralized and dynamic hospital environment, a non-localized and dynamic approach is needed for patient-focused scheduling. Hence, while it is capable of responding to the decentralized hospital environment [5], an agent-based approach is being seen as an effective solution because it is also capable of overseeing the dynamic patient journey. On the other hand, though operations research techniques are effective for centralized optimization problems [15], they do not suit well for the dynamic nature of patient journey [4, 5, 11]. Thus, in the followings, we will focus on how a multi-agent approach could be used for patient-focused scheduling.

3. Patient scheduling - a multi-agent approach

For a multi-agent approach, every coordinated object is modeled as an autonomous agent. Under such a multi-agent environment, agents are distributed, self-interested and have their own goals or constraints [16]. Furthermore, while agents are acting autonomously on behalf of their owners, they would interact with each other such that a better overall solution can be achieved [4]. In fact, it is the unique characteristic of agents and their willingness to interact with others that provide the foundation for efficient patient scheduling.

With the goal of reducing delays in patient journey, two types of agents are typically defined by using a multi-agent approach: 1) patient agent and 2) resource agent.

3.1. Patient agent

A patient agent is used to act on behalf of a single patient and contains information about its represented patient. Although the types of information contained in a patient agent could be uniquely designated according to the design of a scheduling algorithm, the most common ones are:

- Treatment operations needed by the represented patient
- Temporal constraints between treatment operations
- Patient preferences
- Utility cost of the represented patient

A. Treatment operations

In order for patient scheduling to take place, a patient agent must have the information about the treatment pathway of its represented patient. In fact, each patient would have an unique treatment plan (which is determined by medical staff) that states all the needed treatments in his/her patient journey. Hence, all the necessary treatment operations (e.g. X-ray, blood test) should be known in advance by a patient agent.

B. Temporal constraints between treatment operations

The temporal constraints between treatment operations of a patient should also be known by a patient agent. The reason for imposing temporal constraints in a patient agent is to ensure all the necessary treatment operations are performed in the correct sequence according to a doctor's decision or a treatment protocol.

C. Patient preferences

A patient agent may also carry information about patient preferences. Particularly, patient preferences specify what a patient wants to achieve. For instance, a patient may have the preference of receiving treatments in a nearby hospital. Although patient preferences (soft constraints) may also have their substantial influences on patient scheduling, they are considered as less critical than the temporal constraints between treatment operations (hard constraints) in terms of constructing a schedule with minimal delays.

D. Utility cost

In patient scheduling, with the goal of minimizing delays within patient journey, each patient agent will try to acquire an earlier timeslot for its represented patient. However, while resources are scarce, not all patient agents can get their desired timeslots. Thus, in order to determine who can get a particular timeslot, a utility cost is kept by each patient agent.

Utility cost is a mathematical representation of a person's well-being [17]. It allows the comparison between agents in a multi-agent environment. By incorporating an utility cost in each patient agent, the priorities in acquiring a particular timeslot can be objectively identified. Often, patient agent with the highest utility cost has the top priority in reserving a timeslot for its represented patient.

Although the design of an utility cost is unique to the design of a scheduling algorithm, it should be something that reflects the medical situation of a patient in patient scheduling. In addition, utility cost should not be based upon monetary values in hospitals in contrast to commercial domains [4].

In devising an utility cost, Paulussen *et al.* [4, 11] introduced the notion of health state which is a continuous function over time. Health state is defined as a measure of a patient's health development during his or her hospital stay. It depreciates over time until the patient has received proper treatments. In fact, although health state may reflect a patient's medical situation well, it is difficult to quantify such utility cost for a patient [5]. Hence, while an utility cost should be able to reflect the medical situation of a patient, it is also important to consider the practicality in formulating it.

3.2. Resource agent

A resource agent is used to represent one unique medical resource such as operation room, X-ray machine, and medical staff. Like patient agents, resource agents contain information about their represented medical resources. The followings are the most common information carried by a resource agent:

- The service that can be provided by the represented

medical resource

- The operating time of the represented medical resource
- The number of patients that can be served in a particular time interval by the represented medical resource

A. Service provided by the represented medical resource

In order to act on behalf of its represented medical resource, a resource agent should specify the service that can be provided by its represented medical resource. For instance, a resource agent representing a X-ray machine would specify *X-ray examination* as the service provided by that machine.

B. Operating time of the represented medical resource

A resource agent may also specify the operating times of their represented medical resources. For instance, a resource agent representing an operation room may specify *8:00 am to 5:00 pm* as the operating hours of that room.

C. Number of patients being served

In order to model the capacities of medical resources, resource agents may also specify the number of patients that can be served by their represented medical resources in a particular time interval. For instance, a resource agent representing a medical staff may specify *10 patients per day* as the number of patients that can be served daily.

3.3. Mechanism of multi-agent patient scheduling

Unlike operations research, a multi-agent approach works in a decentralized and dynamic manner. While agents are self-interested and acting on behalf of their owners, they have to interact with each other so as to achieve a better overall performance. Thus, a mechanism or protocol is needed for the collaboration between agents. In fact, there are two common collaboration mechanisms in distributing tasks under a multi-agent environment: 1) market mechanism and 2) contract net protocol [16].

3.3.1. Market mechanism

Under market mechanism, agents act like the buyers and sellers in a physical market and they would exchange things with others if a deal is both accepted by two involving parties. The underlying principle of market mechanism is to facilitate collaboration between participants with low communication needs [4]. Particularly, while it is assumed that the price of a product is the only concern for both the buyers and sellers, only prices are communicated between market participants.

In patient scheduling, a patient agent will try to acquire the earliest possible resource timeslot for its represented patient. However, the desired timeslot may sometimes be already occupied by another patient agent. In such case, the former one (the buyer) will try to exchange its timeslot with the latter one (the seller). In reaching a deal, the two involving patient agents will ensure their states (i.e. utility cost) would not be worsen after the exchange. And it is so called a Pareto optimal solution [4, 5, 17].

3.3.2. Contract net protocol

In contract net protocol, agents are trying to acquire their desires by submitting bids in an auction. In particular, there are mainly three phases, namely announcement phase, bidding phase, and awarding phase [10, 16].

In the context of patient scheduling, resource agents first announce their available timeslots to those interested patient agents during the announcement phase. Thereafter, during the bidding phase, interested patient agents will submit their bids based on their corresponding utility costs. Once the resource agents have received bids from patient agents, they will award the timeslots to those with the highest bid (or utility cost).

3.3.3. Comparisons of market mechanism and contract net protocol

For market mechanism, tasks are matched to agents by generalized agreement [16]. In particular, a deal for exchange is only accepted if it is both agreed by the buyer agent and the seller agent (or the generalized agreement is satisfied). The generalized agreement is known in advance by each agent before a negotiation takes place. In each negotiation, only two parties (i.e. buyer agent and seller agent) are involved. For contract net protocol, tasks are assigned to the agent with the highest bid. Particularly, agents try to acquire their desires by bidding instead of negotiating with others. In such bidding, more than two parties could be involved.

In terms of efficiency, while it is thought that contract net protocol is communication-intensive [16], market mechanism would perform better as the amount of information transmitted (i.e. the network traffic) in each interaction is lower compared to contract net protocol.

Table 1 summarizes the comparisons of market mechanism and contract net protocol.

4. Problem formulation

So far, we have briefly discussed some fundamental ideas of multi-agent patient scheduling. In this section, we

Table 1. Comparisons of market mechanism and contract net protocol

	Market mechanism	Contract net protocol
Criteria for task assignments	Generalized agreement is satisfied	Tasks are assigned to agent with the highest bid
Means of interaction between agents	Negotiation	Bidding
Number of agents involved in each interaction	Two	More than two
Network traffic	Lower	Higher

are going to formulate the real problem of patient scheduling in hospital settings. As mentioned in Section 1, we will focus on scheduling for cancer patients in Hong Kong so as to limit our scope.

In Hong Kong, there are seven cancer centers. We denote Z as the set of cancer centers with $Z = \{\text{HKE, HKW, KE, KC, KW, NTE, NTW}\}$, where HKE = Hong Kong Island East, HKW = Hong Kong Island West, KE = Kowloon East, KC = Kowloon Center, KW = Kowloon West, NTE = New Territories East, NTW = New Territories West. Currently, these cancer centers operate on their owns and do not have information about the others. Hence, cancer patients in Hong Kong are frequently being scheduled to receive treatments in a particular cancer center though the same could be provided earlier by other centers which are however unknown to the scheduler. Figure 1 shows the distribution of the seven cancer centers in Hong Kong.

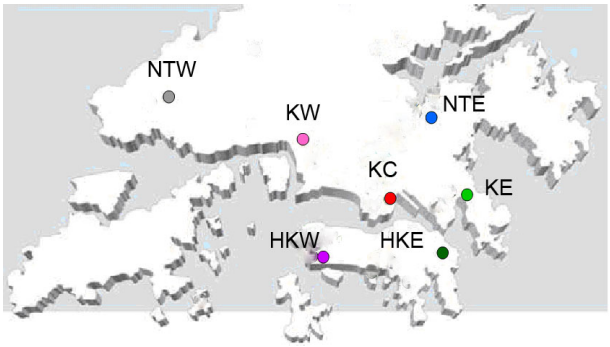


Figure 1. Distribution of the seven cancer centers in Hong Kong

From the data given by the Hospital Authority in Hong Kong, there were 7250 cancer patients admitted in 2007. Figure 2 shows the allocation of these 7250 cancer patients in the seven cancer centers in 2007.

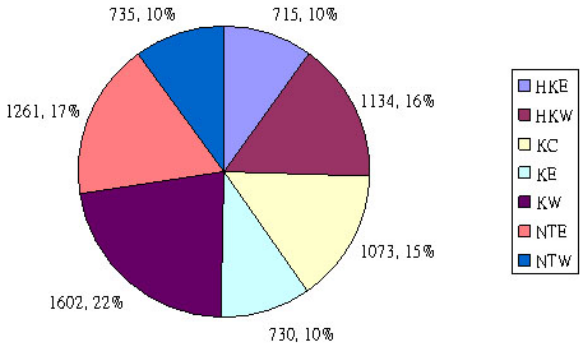


Figure 2. Allocation of 7250 cancer patients in the seven cancer centers in Hong Kong (2007)

For any patient admitted to a hospital, he or she is first to be diagnosed by a doctor. Once the diagnosis has been completed, the doctor will specify a number of actions to be scheduled and a treatment plan will be established for the patient. In particular, a treatment plan depicts all the necessary treatment operations in a patient journey [5]. We denote the set of treatment operations for our domain by $T = \{\text{radiotherapy planning, radiotherapy, surgery, chemotherapy}\}$.

In our work, we define length of stay (LoS) as the duration between the date of the last treatment operation in a treatment plan and the date of a patient's admission. Figure 3 shows the average LoS (in days) for cancer patients in the seven different cancer centers in 2007.

While it is found that cancer patients in Hong Kong are not uniformly allocated (see Figure 2), it is also found that the average LoS for cancer patients varies among the seven cancer centers (with a maximum variation of about 2 weeks as observed in Figure 3). In fact, such variations should be minimized such that cancer patients in Hong Kong can receive treatments sooner rather than later. Hence, a sophisticated mechanism is needed for patient scheduling in Hong Kong.

In receiving treatments, patients have to appoint medical resources. We let α be the set of medical resources for our domain with $\alpha = \{\text{radiotherapy planning unit, radiotherapy unit, operation unit, chemotherapy unit}\}$. With the assumptions that one treatment operation is performed in one medical unit and one medical unit is regarded as one type of

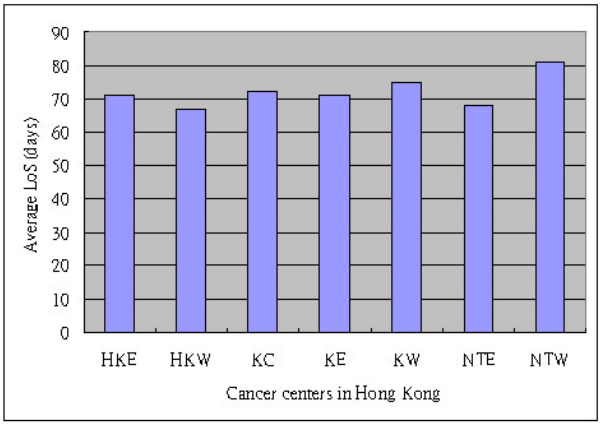


Figure 3. Average LoS (in days) for cancer patients in the seven different cancer centers in Hong Kong (2007)

resource, we have an one-to-one mapping between α and T such that we get $M = \{$(radiotherapy planning unit $\Rightarrow$ radiotherapy planning), (radiotherapy unit $\Rightarrow$ radiotherapy), (operation unit $\Rightarrow$ surgery), (chemotherapy unit $\Rightarrow$ chemotherapy)$\}$, with $(H \Rightarrow J)$ being understood as "H is responsible for conducting J".

4.1. Resource modeling

A resource agent is used to manage a specific type of medical resource. Here, we denote R_{ab} as a resource agent representing medical resource a in cancer center b with $a \in \alpha$; $b \in Z$.

Also, each resource agent has access to all the available timeslots that can be provided by its corresponding medical resource. It is assumed that one particular timeslot could only be occupied by one patient. Hence, for each resource agent R_{ab}, we let:

C_{ab} be the maximum number of patients (i.e. capacity) that can be served daily with resource a in cancer center b;

V^{ab}_{xyg} be the x^{th} available timeslot to which a patient can be assigned by R_{ab} on date y; $x = 1, 2, \ldots, C_{ab}$; $g =$ "a patient's identifier" of whom has occupied the timeslot;

L_m be the maximum duration (i.e. maximum LoS) in which the cancer patients being scheduled have to wait;

Then, we can define $E_{ab}(L_m)$ as the set of timeslots that can be appointed from R_{ab} within the maximum duration L_m days, given as:

$$E_{ab}(L_m) = \{V^{ab}_{xyg}\}, \quad (1)$$

such that $x \leq C_{ab}$; $y \leq L_m$.

4.2. Patient modeling

A patient agent is used to represent one identifiable cancer patient and is denoted as P_i with $i = 1, 2, \ldots, N_p$; $N_p =$ number of cancer patients being scheduled.

In a patient journey, it is common that some of the treatment operations have to be performed in prior to another. For instance, a patient has to undergo a radiotherapy planning before radiotherapy such that the latter could be precisely performed. In other words, the treatment operations needed by a patient should be ordered and satisfy certain temporal constraints. Hence, each patient agent P_i maintains an ordered set S_i which carries all the necessary treatment operations as a sequence, denoted as:

$$S_i = \{O_1, O_2, ..., O_{N_i}\}, \quad (2)$$

where N_i is the number of treatment operations needed by P_i; $O_k \in T$ for $k = 1, 2, \ldots, N_i$; O_{k-1} precedes O_k.

To have O_k scheduled, P_i needs to check with the resource agents R_{ab} where $(a \Rightarrow O_k) \in M$; $a \in \alpha$; $b \in Z$, with the assumption that there is no restriction in patient referrals between cancer centers.

4.3. Scheduling algorithm

In devising a scheduling algorithm for patient scheduling, both Paulussen *et al.* [4] and Vermeulen *et al.* [5] proposed a two phases algorithm by using a multi-agent approach. The two phases are known as initial assignment and rescheduling. During the initial assignment phase, each patient agent was first being assigned an initial schedule. Thereafter, patients agents would interact with each other so as to improve their initial schedules during the rescheduling phase. Particularly, during the rescheduling phase, patient agents would try to improve their initial schedules by selecting the date of a particular treatment operation for exchange.

Although the work of Vermeulen *et al.* [5] could be quite effective in reducing the length of stays in patient journeys, it does not account for the temporal constraints between treatment operations. Moreover, as they claim that the time of the last appointment effectively determines a patient's waiting time, they only select the last treatment operation for rescheduling. In fact, while we agree with their claim, we do think that undesired waiting times could be reduced in a more effective manner if treatment operations other than the last one are also involved for rescheduling. Particularly, it is found that a greater maximum waiting times could be reduced by involving all treatment operations for rescheduling. Figure 4(a) shows the maximum waiting times that could be reduced by only selecting the last treatment operation for rescheduling and Figure 4(b)

shows the maximum waiting times that could be reduced by selecting all the treatment operations for rescheduling.

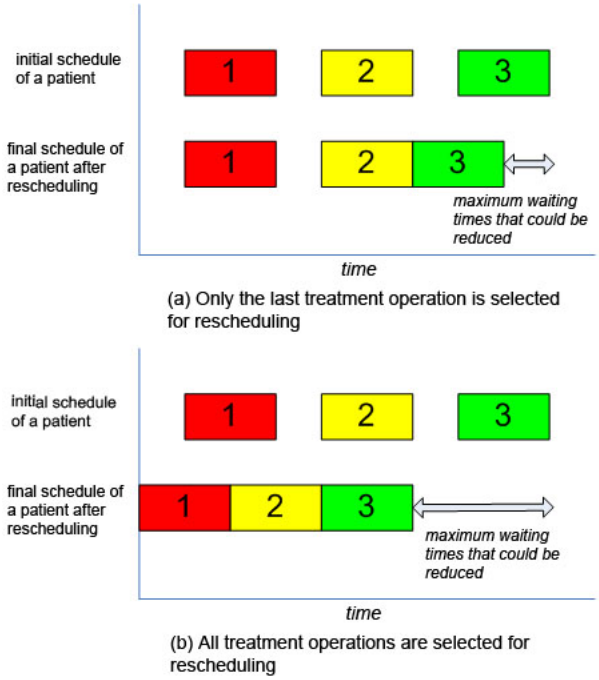


Figure 4. The maximum waiting times that could be reduced by (a) only selecting the last treatment operation for rescheduling; (b) selecting all treatment operations for rescheduling

On the other hand, though the rescheduling scheme proposed by Paulussen *et al.* [4] involves all the treatment operations, it also does not take the temporal constraints between treatment operations into account.

In the scheduling algorithm that we propose, while there are also an initial assignment phase and a rescheduling phase, we will select all the treatment operations for rescheduling so as to maximize the waiting times that could be reduced. More importantly, our scheduling algorithm will also cater the temporal constraints between treatment operations which were not considered in [4] and [5].

4.3.1. Initial assignment

During the initial assignment phase, each patient agent is given an initial schedule based on the admission order of its represented patient. Patients who arrive earlier would give their corresponding agents a higher priority in consuming medical resources. In addition to specifying all the timeslots in which treatment operations will take place, an initial schedule also tells how the resources are committed.

With the assumption that medical staff will schedule their patients professionally, it is fair to assume that the timeslots within an initial schedule would not conflict each other (i.e. no two timeslots of a patient are on the same date). Also, we assume that the initial schedule of a patient agent respects the treatment pathway (which is determined by medical staff) of its represented patient.

In particular, an initial schedule of patient agent P_i with length of stay L_i can be defined as $VP_i(L_i)$:

$$VP_i(L_i) = \{V^{a_1 b_1}_{x_1 y_1 P_i}, V^{a_2 b_2}_{x_2 y_2 P_i}, ..., V^{a_{N_i} b_{N_i}}_{x_{N_i} y_{N_i} P_i}\}, \quad (3)$$

where $(a_k \Rightarrow O_k) \in M$; $a_k \in \alpha$; $b_k \in Z$; $x_k \leq C_{a_k b_k}$; $y_{k-1} < y_k < L_i$.

4.3.2. Rescheduling

Since its construction is solely based on the admission order of a patient, an initial schedule is not an ideal one and may possess lots of undesired delays. Therefore, patient agents will undergo the rescheduling phase such that a reduced waiting times could be achieved for their represented patients. In doing so, a patient agent P_i will try to improve its initial schedule $VP_i(L_i)$ by exchanging timeslots with another patient agent, with the assumption that no patient agent is worse off in terms of the overall schedule (i.e. y_{N_i} corresponding to $V^{a_{N_i} b_{N_i}}_{x_{N_i} y_{N_i} P_i}$ does not worse off). Such exchange will iterate for different patient agents until no further improvement could be achieved.

For the sake of identifying which patient agent should initiate an exchange in each iteration, an utility cost U_i is maintained by each patient agent P_i. And as we want to minimize the waiting times of patients, the utility cost should be defined accordingly. Hence, we let:

A_i be the admission date of whom the patient agent P_i is representing;

Then we can define U_i as the time (in days) between the date of the last timeslot in the initial schedule (i.e. y_{N_i}) and the admission date of whom the patient agent P_i is representing:

$$U_i = y_{N_i} - A_i \quad (4)$$

In order to minimize the adverse effects of having long waiting times, patient agent with the highest utility cost (i.e. the longest waiting time) is chosen to initiate a negotiation in each iteration. Here, we let:

W be the set of utility costs with $W = \{U_1 , U_2 , ... , U_N\}$, where N is the number of cancer patients being scheduled.

Obviously, in each iteration, patient agent P_i with the longest waiting time will have the maximum utility cost Q:

$$Q = \max\{W\} \tag{5}$$

Once the patient agent P_i with the maximum utility cost Q (we call it as "initiating patient agent", P_I) has been identified, it will initiate an exchange for timeslots with another patient agent (we call it as "target patient agent", P_t). Particularly, the initiating patient agent P_I will first select the 1^{st} timeslot of its initial schedule (i.e. $V^{a_1 b_1}_{x_1 y_1 P_I}$) for exchange. Since then, it will continue to select the next timeslot for exchange until the last timeslot of its initial schedule (i.e. $V^{a_k b_k}_{x_k y_k P_I}$ for $k = 2, 3, \ldots, N_I$).

While it is assumed that the 1^{st} treatment operation (i.e. O_1 in S_i) could start on any particular date since a patient's admission A_i, it is also assumed that O_1 cannot be performed on the same date as A_i (i.e. the date for performing O_1 does not equal a patient's admission date, $y_1 \neq A_i$) because a treatment plan would only be established once the patient has been admitted to the hospital. Hence, for each initiating patient agent P_I, we let:

y_{O_1} be a suitable date on which the 1^{st} treatment operation (i.e. O_1 in S_I) can be duly performed;

Then, in order to find an earlier date for the 1^{st} treatment operation O_1 (or the 1^{st} timeslot in the initial schedule of P_I, i.e. $V^{a_1 b_1}_{x_1 y_1 P_I}$), y_{O_1} should satisfy:

$$A_I < y_{O_1} < y_1, \tag{6}$$

where A_I is the admission date of P_I; y_1 is the date of the 1^{st} timeslot of P_I's initial schedule;

Once the smallest value of y_{O_1} (i.e. the following day of P_I's admission, $A_I + 1$ *day*) is obtained, the initiating patient agent P_I would contact the corresponding resource agent to see if there is an alternative timeslot for exchange. In particular, as it is assumed that patients are willing to go across different cancer centers for a reduction of waiting times, the initiating patient agent P_I should contact all the corresponding resource agents one by one across different cancer centers (i.e. $R_{a_1 b}$, $b \in Z$; $(a_1 \Rightarrow O_1) \in M$). And in order to maximize the chance of receiving the 1^{st} treatment operation O_1 in the same center as which is planned in the initial schedule, P_I will first contact $R_{a_1 b_1}$ as b_1 is the cancer center which is responsible for the 1^{st} timeslot in P_I's initial schedule.

By contacting with $R_{a_1 b_1}$, if there are timeslots available for exchanges on date y_{O_1}, P_I will get all the timeslots on that date (i.e. $V^{a_1 b_1}_{x y_{O_1} g}$ for $x = 1, 2, \ldots, C_{a_1 b_1}$). In addition, by the variable g in (1), all the owners of these timeslots are also identified such that the initiating patient agent P_I could know who to initiate an exchange for timeslots.

Once the initiating patient agent P_I get all the timeslots as well as their owners on date y_{O_1}, it will first propose an exchange to the one (i.e. the target patient agent, P_t) who currently owns the 1^{st} timeslot on that date (i.e. $x = 1$ for $V^{a_1 b_1}_{x y_{O_1} g}$). In order to accept an exchange, the target patient agent P_t should ensure the followings:

- The date of its last treatment operation does not worse off after an exchange (i.e. y_{N_t} corresponding to $V^{a_{N_t} b_{N_t}}_{x_{N_t} y_{N_t} P_t}$ does not worse off)
- The temporal constraints in (2) would not be violated after an exchange (i.e. treatment operations are performed in the correct order)

In fact, while it is quite easy and straightforward for P_t to determine whether or not the last treatment operation has been worse off, it could be tricky in determining whether the temporal constraints between treatment operations have been violated. The reason is that treatment operations are performed with durations and they may overlap with each other after an exchange though the temporal constraints have not been violated. In order to deal with that, we let:

$\overline{D}_{O_m O_n}$ be the average duration between the start of treatment operation O_m and the start of treatment operation O_n where $O_m, O_n \in T$; O_m precedes O_n;

In addition, while $\overline{D}_{O_m O_n}$ can be easily obtained from the set of patients' initial schedules, we assume that all treatment operation n could be performed duly after treatment operation m with a duration $\overline{D}_{O_m O_n}$ so as to reduce complexity.

Then, given that the date of the last treatment operation does not worse off (i.e. y_{N_t} corresponding to $V^{a_{N_t} b_{N_t}}_{x_{N_t} y_{N_t} P_t}$ remains constant), an exchange for timeslots would only be accepted by the target patient agent P_t if the followings are satisfied:

$$y_{t_{k-1}} + \overline{D}_{O_{k-1} O_k} < y_{I_k} < y_{t_{k+1}} - \overline{D}_{O_k O_{k+1}}, \tag{7}$$

where $y_{t_{k-1}}$, $y_{t_{k+1}}$ are the dates on which the (k-1)th treatment operation (i.e. O_{k-1}) and the (k+1)th treatment operation (i.e. O_{k+1}) are scheduled for P_t respectively; y_{I_k} is the date for the k^{th} treatment operation (i.e. O_k) proposed by P_I; $O_{k-1}, O_k, O_{k+1} \in S_t$;

During the search for an alternative timeslot for its 1^{st} timeslot, the initiating patient agent P_I could receive its 1^{st} treatment operation O_I in an earlier fashion if the target patient agent P_t accepts the exchange (i.e. y_{N_t} does not worse off and (7) has not been violated after the proposed exchange). However, if the proposed exchange has been turned down (i.e. either (7) is violated or y_{Nt} is worse off), the target patient agent P_t will inform the initiating patient agent P_I that the proposal for exchange has been rejected. Since then, P_I will continue to find an alternative timeslot for its 1^{st} treatment operation O_1 (or the 1^{st} timeslot of

its initial schedule, i.e $V^{a_1 b_1}_{x_1 y_1 P_I}$) by contacting the one who owns the next timeslot on that date (i.e. $V^{a_1 b_1}_{x y_{O_1} g}$ for $x = 2, 3, \ldots, C_{a_1 b_1}$).

As time past, if an alternative timeslot still cannot be found in center b_1 on date y_{O_1}, the initiating patient agent P_I would then first contact the agents in other centers (i.e. $R_{a_1 b}$ for $b \neq b_1$; $b \in Z$) one by one for timeslots on that date. As the centers contacted are no longer the same as which is planned in the initial schedule, the order of contact is considered as unimportant and thus these centers could be contacted in random. Thereafter, if there is no an alternative timeslot found again, the value of y_{O_1} has to be added by 1 (i.e. $A_I + 2$ *day*) and the above process would continue as long as (6) is satisfied.

Once there is an alternative timeslot found for the 1^{st} timeslot or (6) is no longer satisfied, the initiating patient agent P_I would then continue the process by finding an alternative timeslot for its next treatment operation (i.e. $O_k \in S_I$ for $k = 2, 3, \ldots, N_I$) or the next timeslot of its initial schedule (i.e. $V^{a_k b_k}_{x_k y_k P_I}$ for $k = 2, 3, \ldots, N_I$). Again, as we have to assure that the new timeslots acquired by P_I would not overlap with each other, (7) is modified and the followings have to be satisfied:

$$y_{I_{k-1}} + \overline{D}_{O_{k-1} O_k} < y_{t_k} < y_{I_{k+1}} - \overline{D}_{O_k O_{k+1}}, \quad (8)$$

where $y_{I_{k-1}}$, $y_{I_{k+1}}$ are the dates on which the (k-1)th treatment operation (i.e. O_{k-1}) and the (k+1)th treatment operation (i.e. O_{k+1}) are scheduled for P_I respectively; y_{t_k} is the date for the k^{th} treatment operation (i.e. O_k) acquired from P_t; O_{k-1}, O_k, $O_{k+1} \in S_I$;

In addition, in order to get an earlier timeslot from P_t after an exchange, y_{t_k} should satisfy:

$$y_{t_k} < y_k, \quad (9)$$

where y_k is the date for the k^{th} timeslot of P_I's initial schedule.

Once there is an alternative timeslot found for the last timeslot of P_I's initial schedule (i.e. $V^{a_{N_I} b_{N_I}}_{x_{N_I} y_{N_I} P_I}$) or there is no longer such a timeslot found that satisfies both (8) and (9), an iteration ends for the initiating patient agent P_I and the one with the next highest utility cost (or waiting times) will be selected to initiate another exchange for timeslots.

In each iteration, if there is an alternative timeslot found for the last timeslot of P_I's initial schedule (i.e. $V^{a_{N_I} b_{N_I}}_{x_{N_I} y_{N_I} P_I}$), the new last timeslot of P_I would become $V^{a_{N_I} b_{N_I}}_{x_{N_I} y_{t_N} P_I}$. Clearly, by (9), as y_{t_N} is smaller than y_{N_I}, a reduction of length of stay (LoS) or waiting times is achieved.

5. Conclusion and future work

In this paper, we have discussed how a multi-agent approach can be adopted for patient scheduling. Particularly, we have pointed out some fundamental concepts of patient journey and the importance of reducing waiting times for patients.

While we have categorized patient schedulings into resource-focused scheduling and patient-focused scheduling, we focus on the latter one as our ultimate goal is to minimize undesired delays for patients.

In devising an algorithm for patient-focused scheduling, we decided to base our work on a multi-agent approach as it caters well the dynamic and decentralized nature of hospital settings. We have also formulated the real problem of patient scheduling by limiting our scope to cancer patients in Hong Kong.

In the future, we are going to apply our model to the real hospital settings in Hong Kong by using the data given by the Hospital Authority. Particularly, we will compare the performances in terms of the amount of length of stay that could be reduced by varying the scheduling strategy (i.e. involve all treatment operations for rescheduling VS involve only the last treatment operation for rescheduling).

References

[1] R. W. Hall. *Patient Flow: Reducing Delay in Healthcare Delivery*. Springer Science+Business Media, LLC, Springer Street, New York, NY 10013, USA, 2006.

[2] J. Vissers and R. Beech. *Health operations management : patient flow logistics in health care*. Routledge, 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN, 2005.

[3] J. Vissers, J. Bekkers, and I. Adan. Patient mix optimization in tactical cardiothoracic surgery planning: a case study. *IMA Journal of Management Mathematics*, 16, 2005.

[4] T. O. Paulussen, N. R. Jennings, K. S. Decker, and A. Heinzl. Distributed patient scheduling in hospitals. *In: 18th International Joint Conference on Artificial Intelligence*, 2003.

[5] Ivan Vermeulen, Sander Bohte, Koye Somefun, and Han La Poutré. Improving patient activity schedules by multi-agent pareto appointment exchanging. *Service Oriented Computing and Applications*, 1(3):185–196, 2007.

[6] NHS Scotland. Understanding the Patient Journey - Process Mapping. 2006.

[7] Patient Liaison Group. Improving Your Elective Patient's Journey. 2007.

[8] Department of Health, UK. Now I feel tall: What a patient-led NHS feels like. 2005.

[9] M. Simunovic, A. Gagliardi, D. McCready, A. Coates, M. Levine, and D. DePetrillo. A snapshot of waiting times for cancer surgery provided by surgeons affiliated with regional cancer centres in Ontario. *CMAJ*, 165(4):421–425, August 2001.

[10] Torsten Paulussen, Anja Zoller, Franz Rothlauf, Armin Heinzl, Lars Braubach, Alexander Pokahr and Winfried Lamersdorf. Agent-based patient scheduling in hospitals. 2006.

[11] A. K. Hutzschenreuter, P. A. N. Bosman, I. Blonk-Altena, J. van Aarle, and H. L. Poutré. Agent-based patient admission scheduling in hospitals. In *AAMAS '08: Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems*, pages 45–52, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems.

[12] T. J. Cantor. Waiting times for patients with cancer - waiting lists are putting patients' lives in jeopardy. *BMJ*, 321(7255):236, July 2000.

[13] M. Samur, H. S. Bozcuk, G. Dalmaz, E. Karaveli, F. G. Koseoglu, T. Colak, and E. Pestereli. Treatment delay in breast cancer; does it really have an impact on prognosis? *Turkish Journal of Cancer*, 32(4):138–147, 2002.

[14] P. Spurgeon, F. Barwell, and D. Kerr. Waiting times for cancer patients in England after general practitioners' referrals: retrospective national survey. *BMJ*, 320:838–839, 2000.

[15] J. Patrick and M. Puterman. Reducing wait times through operations research: optimizing the use of surge capacity. *Healthc Q*, 11(3):77–83, 2008.

[16] G. Weiss. *Multiagent systems : a modern approach to distributed artificial intelligence*. Cambridge, Mass. : MIT Press, 1999.

[17] H. Czap and M. Becker. Multi-agent systems and microeconomic theory: A negotiation approach to solve scheduling problems in high dynamic environments. In *HICSS '03: Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track 3*, page 83.2, Washington, DC, USA, 2003. IEEE Computer Society.

[18] Hospital Authority Statistical Report (2006-2007). http://www.ha.org.hk/visitor/ha_visitor_index.asp?Parent_ID=652&Content_ID=136136.

[19] Hospital Authority. HA Clinical Practice Guideline for Management of Lung Cancer. 2007.

A Knowledge-based Approach for Histogram-distances Image Retrieval

Chun Fan WONG

Abstract

Search and retrieval models are vital to image management, and are increasingly receiving attention with the growing use of image libraries and the explosion of digital media on the Web. To fulfill these requirements, we develop an automated annotation model based on image capture metadata in conjunction with image processing techniques. In this paper, we propose an extension of image indexing models which utilizes knowledge-based expansion and contextual feature-based index expansion with adaptive segmentation of HSV color space. Our system is evaluated quantitatively using more than 100,000 web images and around 1,000,000 tags. Experimental results indicate that this approach is able to deliver highly superior performance.

1 Introduction and related work

With the rapid increase of the volume of digital image collections, image retrieval has been become one of the most research area. As the number of images available in online repositories is growing dramatically, exploring the frontier between image and language is an interesting and challenging task.

Most important elements in a retrieval system are the features used to express an image and image features can be divided between two main categories: concept-based image retrieval and content-based image retrieval. The former focuses on retrieval by objects and high-level concepts, while the latter focuses on the low-level visual features of the image.

Content-based image retrieval, the problem of searching large image repositories according to image content, has been the subject of a significant amount of research in the last decade. These methods aim at accessing the knowledge embedded in images by extract low-level visual features and capturing image similarity by relying on some specific characteristic of images. Typically, these models are based on color, texture and shape [4, 6, 9, 10, 13, 22, 25, 30, 33]. Some studies [8, 11, 15, 17, 23, 29] of image segmentation using image partitions, sign detection, region segmentation techniques while some studies rely on compute general similarity between images based on statistical image properties [1–3, 18, 21, 26, 27]. Researches on semantic retrieval of the image database [9, 10, 13, 14, 22, 30, 33] combined with a region based image decomposition is used, which aims to extract semantic properties of images based on spatial distribution of color and texture properties. The advantage of content-based image retrieval methods is it do not incur any indexing cost as they can be extracted by automatic algorithms.

Concept-based image retrieval is to create a set of metadata to describe the images content, namely, concept indexing. Some studies [11, 20] include users in a search loop with a relevance feedback mechanism to adapt the search parameters based on user feedback. Some researches [5] focuses on implicit image indexing which involve an implicit and, in consequence, augments the original indexes with additional concepts that are related to the query. With the advent of Semantic Web technology, knowledge is playing a key role as the core element of knowledge representation architecture. Some effort [7, 12, 19, 24, 28] has been made for image retrieval using Semantic Web techniques.

We propose an integrated framework for image retrieval based on generative modelling approaches. In [31,32], a semantic indexing technique named Automatic Semantic Annotation (ASA) approach is developed which is based on the use of image parametric dimensions and metadata. Using decision trees and rule induction, a rule-based approach to formulate explicit annotations for images fully automatically is developed, so that, semantic query such as ”night scene of Arch of Triumph in Paris in winter” can be answered and indexed purely by machine. In this paper, we propose an extension of such image indexing models by using knowledge-based expansion and contextual feature-based index expansion. Experimental evidence on more than 100,000 web images and over 990,000 tags shows that semantically meaningful retrieval are inferred and it is able to deliver highly competent performance.

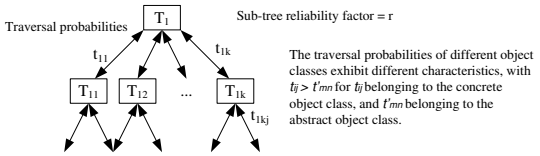


Figure 1. Hierarchical expansion

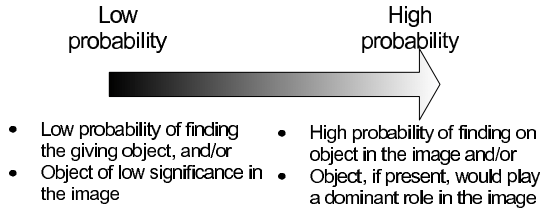


Figure 2. Measurement of Image Indexing

2 Knowledge-based Framework for Image Retrieval

2.1 Knowledge-based expansion

Our approach provides operations to perform image retrieval with knowledge-based expansion enabled. It aims in introducing knowledge-based expansion into the image retrieval problem and using the sub-objects as surrogate terms for general queries is to improve the precision since, in certain applications, the presence of particular objects in an image often implies the occurrence of other objects. The application of such inferences will allow the concept of an image to be automatically expanded.

Aggregation hierarchical expansion is a particularly useful technique, which relates to the aggregation hierarchy of sub-objects that constitute an object. These can be classified into two categories, concrete and abstract hierarchical expansion. Concrete hierarchical expansion is the relevant objects are well-defined (Fig. 1). For example, an concept "wedding" expanded to bridge, groom, flower, wedding cake. Abstract hierarchical expansion is the objects are not concretely defined. Although "conflict" is not a definite visual object, it contains certain common characteristics.

2.2 Contextual feature-based Expansion

In order to perform direct extraction of high-level semantic content automatically, we establish associations between low-level features with high-level concepts, and such associations take the following forms. In the contextual feature-based expansion, the presence of certain low-level features F may suggest a finite number of m object possibilities. Sometimes, a combination of basic features may be used to infer the presence of high-level concepts for inclusion in the semantic index.

The presence of certain basic features alone may not be sufficient to infer the presence of specific objects, but such features if augmented by additional information may lead to meaningful inferences. When a particularly context is known, a concept may be indexed more precisely. Such contextual information will typically be provided through knowledge-based expansion, which may lead to the creation of a new index term, or a revision of the score of an existing index term. An iterative feedback loop will be risen where the determination of new objects will lead to new meaningful feature-object combinations, where further objects may be determined.

2.3 Measurement of Image Indexing

The quality assessment of the machine-inferred boundaries between parts of the depicted scenes is based on the precision. Thus, our system is measured quantitatively (given in a later section) in order to compute the effectiveness of our approach. The reliability of a given annotation will given rise to a numerical measure, which signifies how good the annotation is. For annotations with a low measure, this would mean that the annotation is not very occurrence, or in extreme cases, what is being annotated is absent from an image. A high annotation measure indicates that the chance of finding the corresponding object or content in the given image is high. In addition, apart from measuring the likelihood of whether something is present or not, it can be used to indicate the importance of an object in the image. For example, an object which is very prominently present in the foreground of an image would have a much large value than an object of small size in the remote background. Hence the annotation measure is used to signify two aspects; the likelihood of finding the object in the given image and the prominence of the object in the given image (Fig. 2).

2.4 HSV Color Space

A color space is defined as a model for representing color in terms of intensity values. Technology based on color tone is most widely used due to compactness of calculation and information expression. Histogram is mainly used for color tone information and, currently, histograms are an important tool for the retrieval of images and video from visual databases. Methods using color tone is robust with respect to object movement, rotation and to changes like distortion within an image and may be implemented easily. Color is defined by the three characteristics: hue, saturation and value. In Fig. 3, HSV is an expression of color tones that

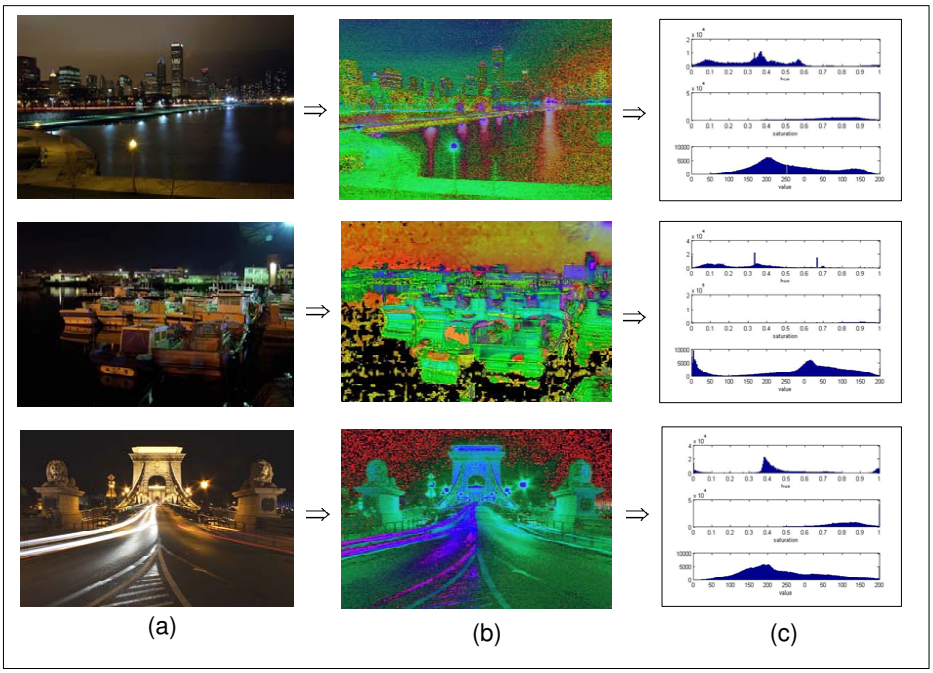


Figure 4. Examples of conversion of color space conversion histogram. Column (a) shows the original images in RGB color space. Column (b) are image in HSV space and column (c) show their HSV color histogram.

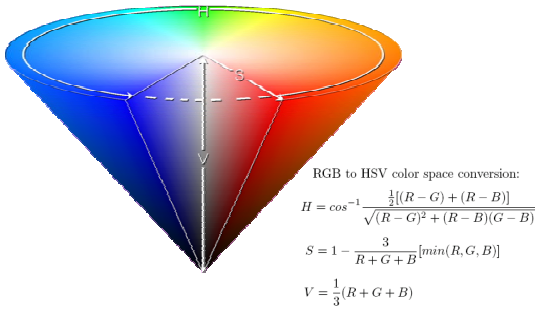


Figure 3. HSV Color space

can be sensed by humans using these characteristics.

An image histogram gives to the probability mass function of the image intensities. This is extended for color images to capture the joint probabilities of the intensities of the three color channels. Histogram search characterizes an image by its color distribution, or histogram but the drawback of a global histogram representation is that information about object location, shape, and texture is discarded.

Many histogram distances have been used to define the similarity of two color histogram representations. The Bhattacharya Distance (BD), Chi-squared Distance (CD) and Euclidean Distance (ED) are also used and, from [16, 25], the definition are listed in Equation 1, 2 and 3 respectively.

$$d_{Bhattacharya}(p,q) = \sqrt{1 - \sum_{u=1}^{n} \sqrt{q_u p_u(m)}} \quad (1)$$

$$d_{x^2} = \sum_{u=1}^{n} \frac{(q_u - p_u(m))^2}{(q_u + p_u(m))} \quad (2)$$

$$d_{euclidean} = \sqrt{\sum_{u=1}^{n} (p_u - q_u)^2} \quad (3)$$

where m is the center of the image region, n is the number of bins in the distribution, and q_u and p_u are the weighted histograms of the model and candidate respectively.

3 Experimental Evaluation

The main purpose in introducing knowledge-based expansion into the image retrieval problem and using the sub-objects as surrogate terms for general queries is to improve the precision in the image sets. In this paper, we

mainly focus on the knowledge-based expansion and contextual feature-based index expansion, specially focus on histogram search which characterizes an image by its color distribution. The knowledge concept are organized and used to build the basic content index within a relational database where it is designed for maximum query effectiveness by distributing the semantic elements across different relations. A further concept is built on top of these relations to support rapid discovery.

Our system is evaluated quantitatively in order to compute the effectiveness of our approach. The quality assessment of the machine-inferred boundaries between parts of the depicted scenes is based on the precision. A set of standard evaluation queries are used for experimentation. Comparison is made between base-level indexing and the expanded level indexing, and the widely accepted measures of retrieval performance of precision and recall are used to assess system performance. To numerically assess the accuracy and effectiveness of our annotation approach, we have retrieved 103,521 sets of images with 991,074 associated tags from flickr.com which are a popular photo sharing web site and online community platform offering a fairly comprehensive web-service API that allows developers to create applications that can perform almost any function on images. In our evaluation, we decide that a relevant image must include a representation of the category in such a manner that a human should be able to immediately associate it with the assessed concept.

3.1 Results

In [31, 32] by using decision trees and rule induction, a rule-based approach to formulate explicit annotations for images fully automatically has been developed. In relation to image acquisition, many images may be broken down to few basic scenes, such as nature and wildlife, portrait, landscape and sports. In the case of aggregation hierarchical expansion, we decided to test our system using the aggregation hierarchy of basic categories ”night scenes” and extend the image hierarchy to find a sub-scene ”night scene of downtown”, ”downtown” can be expanded to ”business district”, ”commercial district”, ”city center” and ”city district”, while ”city district” can be expanded to ”road”, ”building”, ”architecture”, ”highway” and ”hotel”.

To extend the original approach, firstly, we annotate night scenes based on the prior rule-based approach to extract 422 out of 103,527 images. We also gather 1108 tags associated with those images and totally 417 unique terms are formed. We list the top 217 out of 417 unique terms list in Fig. 5. We present the results of the evaluation in Fig. 6.

To establish associations between low-level features with high-level concepts, associating basic features with semantic concepts may be applied to arbitrary images for inclusion

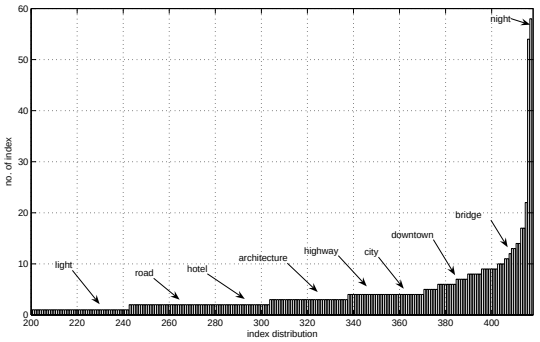


Figure 5. Index distribution associated with night scene images

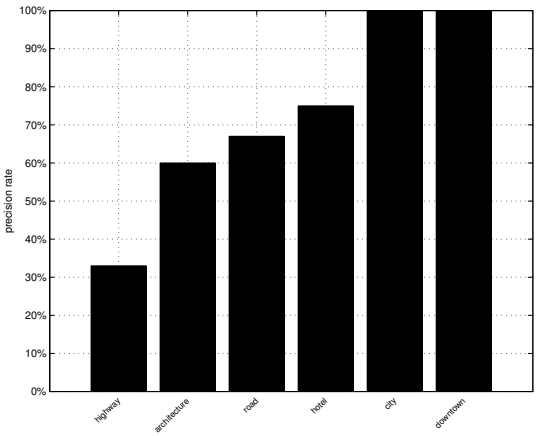


Figure 6. Experimental results on aggregation hierarchical expansion

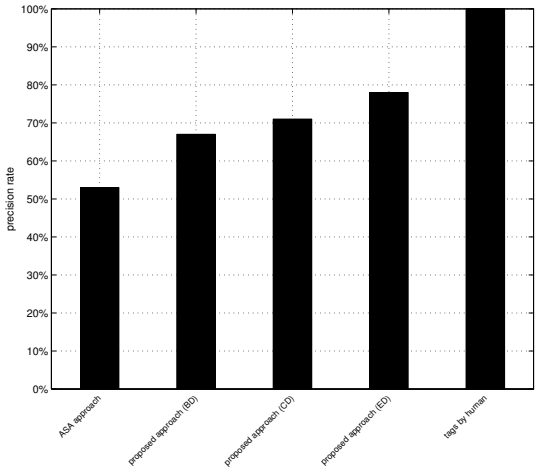


Figure 7. Experimental results on contextual feature-based index expansion

in the semantic index. Methods using color tone is robust with respect to object movement, rotation and to changes like distortion within an image and may be implemented easily. Histogram is mainly used for color tone information, particularly in the areas of feature detection and feature extraction, to refer to algorithms described in Equation 1, 2 and 3. Here, we adapt color tone histogram algorithms [16,25] to extract high-level concepts from low-level features.

We performed experiments to show the good optimality and convergence of our approach. Firstly, we randomly select one "downtown" image from the test set and carried out evaluation by comparing the original Automatic Semantic Annotation (ASA) approach) with our approach which combines the original ASA approach using adaptive annotation of HSV color space and distance algorithms (see Fig. 8) and the use of human tags.

In Fig 7, experimental results show that tags by human deliver excellent precision rate with 100% precision but this tagging approach relies heavily on human involvement. Significantly better results can be obtained by ASA Approach combining HSV color space histogram distances where the precision rate grows to 67.2% (BD), 71.4% (CD) and 77.9% (ED). Obviously, compared to annotation without the contextual feature-based index expansion enabled, the performance is around 52.8%. From the joint application of these, we can formulate semantic annotations for specific image fully automatically and index images purely by machine without any human involvement.

4 Conclusion

In this paper, a knowledge-based framework of image retrieval together with contextual feature-based expansion is combined. Our method combines the advantages of original ASA approach and contextual feature-based expansion while preserving the necessary image and knowledge coherence. Our system is evaluated quantitatively, and experimental results indicate that this approach is able to deliver highly competent performance.

References

[1] G. Amato and C. Meghini. Combining features for image retrieval by concept lattice querying and navigation. *International Conference on Image Analysis and Processing Workshops*, 0:107–112, 2007.

[2] J. Amores, N. Sebe, and P. Radeva. Context-based object-class recognition and retrieval by generalised correlograms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(10):1818–1833, October 2007.

[3] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios. Boostmap: A method for efficient approximate similarity rankings. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 02:268–275, 2004.

[4] I. Azzam, A. G. Charlapally, C. H. C. Leung, and J. F. Horwood. Content-based image indexing and retrieval with xml representations. *Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong*, pages 181–185, 2004.

[5] I. A. Azzam, C. H. C. Leung, and J. F. Horwood. Implicit concept-based image indexing and retrieval. In *Proceedings of the IEEE International Conference on Multi-media Modeling*, pages 354–359, Brisbane, Australia, January 2004.

[6] I. A. Azzam, C. H. C. Leung, and J. F. Horwood. A fuzzy expert system for concept-based image indexing and retrieval. *International Conference on Multimedia Modeling Conference*, 0:452–457, 2005.

[7] K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, and M. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3:1107–1135, 2003.

[8] Y. Chen, J. Z. Wang, and R. Krovetz. Content-based image retrieval by clustering. In *MIR'03: Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval*, pages 193–200, New York, NY, USA, 2003. ACM.

[9] J. S. Cho and J. Choi. Contour-based partial object recognition using symmetry in image databases. In *SAC'05: Proceedings of the 2005 ACM symposium on Applied computing*, pages 1190–1194, New York, NY, USA, 2005. ACM Press.

[10] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215, 2007.

[11] R. Datta, J. Li, and J. Z. Wang. Content-based image retrieval: approaches and trends of the new age. In *MIR'05:*

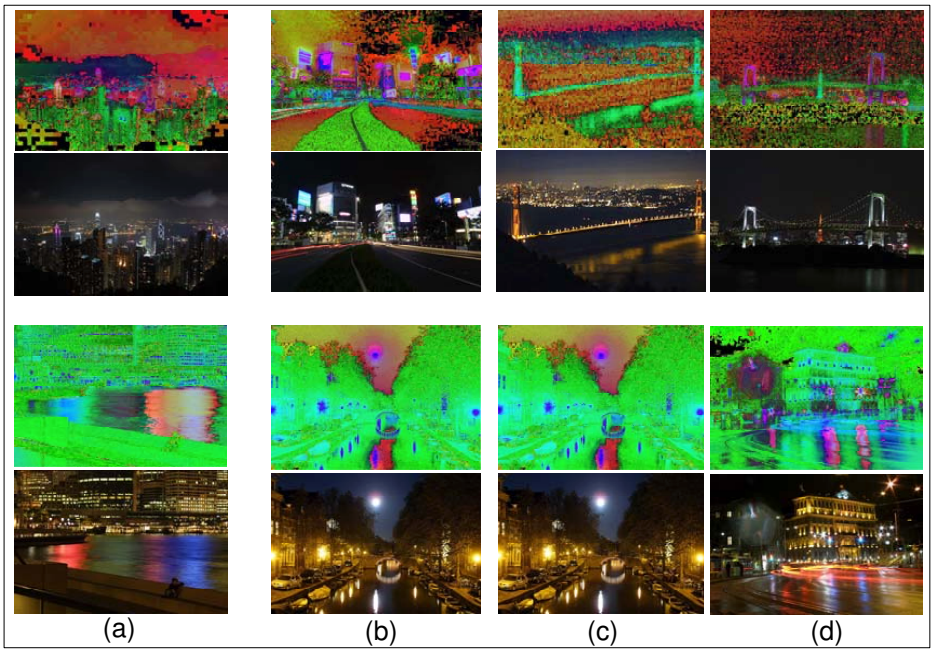


Figure 8. Example of experimental results of nearest histogram distance in RGB and HSV color space (a) query images. (b) Bhattacharya Distance(BD) (c) Chi-squared Distance(CD) (d) Euclidean Distance(ED)

Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval, pages 253–262, New York, NY, USA, 2005. ACM.

[12] L. Fan and B. Li. A hybrid model of image retrieval based on ontology technology and probabilistic ranking. *Web Intelligence*, 0:477–480, 2006.

[13] J. Gausemeier, J. Fruend, C. Matysczok, B. Bruederlin, and D. Beier. Development of a real time image based object recognition method for mobile ar-devices. In *International conference on Computer graphics, virtual Reality, visualisation and interaction in Africa*, pages 133–139, New York, NY, USA, 2003. ACM Press.

[14] K. Hornsby. Retrieving event-based semantics from images. *International Symposium on Multimedia Software Engineering*, 00:529–536, 2004.

[15] M. Jian, J. Dong, and R. Tang. Combining color, texture and region with objects of user's interest for content-based image retrieval. *International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 01:764–769, 2007.

[16] A. Kale, K. Kwan, and C. Jaynes. Epipolar contrained user pushbutton selection in projected interfaces. *Computer Vision and Pattern Recognition Workshop*, 10:156–164, 2004.

[17] R. Krishnapuram, S. Medasani, S. H. Jung, Y. S. Choi, and R. Balasubramaniam. Content-based image retrieval based on a fuzzy approach. *IEEE Transactions on Knowledge and Data Engineering*, 16(10):1185–1199, 2004.

[18] J. Li and J. Z. Wang. Automatic linguistic indexing of pictures by a statistical modeling approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1075–1088, 2003.

[19] H. Lieberman and H. Liu. Adaptive linking between text and photos using common sense reasoning. *Adaptive Hypermedia and Adaptive Web-Based Systems, Second International Conference, AH 2002, Malaga, Spain*, pages 2–11, May 2002.

[20] D. Liu and T. Chen. Content-free image retrieval using bayesian product rule. *IEEE International Conference on Multimedia and Expo*, 0:89–92, 2006.

[21] A. P. Natsev, A. Haubold, J. Tešić, L. Xie, and R. Yan. Semantic concept-based query expansion and re-ranking for multimedia retrieval. In *International conference on Multimedia*, pages 991–1000, New York, NY, USA, 2007. ACM.

[22] R. Pawlicki, I. Kókai, J. Finger, R. Smith, and T. Vetter. Navigating in a shape space of registered models. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1552–1559, 2007.

[23] A. Perina, M. Cristani, and V. Murino. Natural scenes categorization by hierarchical extraction of typicality patterns. *International Conference on Image Analysis and Processing*, pages 801–806, 2007.

[24] A. Popescu, G. Grefenstette, and P. A. Moellic. Using semantic commonsense resources in image retrieval. In *International Workshop on Semantic Media Adaptation and Personalization*, pages 31–36, Washington, DC, USA, 2006. IEEE Computer Society.

[25] M. Riaz, G. Kang, Y. Kim, S. Pan, and J. Park. Efficient image retrieval using adaptive segmentation of hsv color space.

Computational Science and its Applications, International Conference, 0:491–496, 2008.

[26] K. Stevenson and C. H. C. Leung. Comparative evaluation of web image search engines for multimedia applications. *IEEE International Conference on Multimedia and Expo*, 0:4 pp., 2005.

[27] Y. Sun, S. Shimada, and M. Morimoto. Visual pattern discovery using web images. *MIR'06: Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, pages 127–136, 2006.

[28] A. M. Tam and C. H. C. Leung. Semantic content retrieval and structured annotation: Beyond keywords. *ISO/IEC JTC1/SC29/WG11 MPEG00/M5738, Noordwijkerhout, Netherlands*, March 2000.

[29] N. Vasconcelos. From pixels to semantic spaces: Advances in content-based image retrieval. *IEEE Computer*, 40(7):20–26, 2007.

[30] J. Vogel, A. Schwaninger, C. Wallraven, and H. H. Bülthoff. Categorization of natural scenes: Local versus global information and the role of color. *ACM Transactions on Applied Perception*, 4(3):19, 2007.

[31] R. C. F. Wong and C. H. C. Leung. Automatic semantic annotation of real-world web images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11):1933–1944, November 2008.

[32] R. C. F. Wong and C. H. C. Leung. Knowledge-based expansion for image indexing. In *International Computer Symposium*, volume 1, pages 161–165, November 2008.

[33] T. Zöller and J. M. Buhmann. Robust image segmentation using resampling and shape constraints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1147–1164, July 2007.

Index Convergence Behaviour for Collaborative Semantic Indexing of Multimedia Data Objects

Wing Sze CHAN

Abstract

While multimedia technology is growing dramatically, the searching of multimedia data objects becomes an important activity. However, this kind of search activity is far more challenging than the searching of the text-based documents. Here, we present an innovative approach that enables the semantic search of the multimedia objects by the discovery and meaningful indexing of their semantic concepts. By analyzing the users' search queries, relevance feedback and selection patterns, semantic concepts can be discovered and migrated through an index hierarchy. Through the growth and evolution of the index hierarchy, the semantic index may be dynamically constructed, validated, and built-up. Index convergence behaviour and modelling are also discussed.

1 Introduction

A huge amount of multimedia data objects, in various forms and formats, exist and grow explosively in our daily life. It is estimated that by 2010, over a billion digital images will be created each day [5]. The multimedia data object retrieval problem becomes important and necessary. Multimedia information search is far more difficult than searching text-based documents since the content of text-based documents can be extracted automatically while the content of multimedia objects cannot be automatically determined [13, 14].

Research in image retrieval has been divided into two main categories: "concept-based" image retrieval, and "content-based" image retrieval [1, 3, 4, 6, 11, 15, 16, 19, 22]. The former focuses on higher-level human perception using words to retrieve images (e.g. title, keywords, captions), while the latter focuses on the visual features of the image (e.g. size, colour, texture). In an effective "concept-based" multimedia retrieval system, efficient and meaningful indexing is necessary [8,9]. Due to current technological limitations, it is impossible to extract the semantic content of multimedia data objects automatically [18, 21, 23]. Meanwhile, the discovery and insertion of new indexing terms are always costly and time-consuming. Therefore, novel indexing mechanisms are required to support their search and retrieval.

2 The Indexing Approach & Hierarchical Evolution

Our approach concentrates on the indexing of semantic contents of multimedia objects and will exclude metadata from consideration, since indexing by metadata is relatively straightforward and less meaningful than semantic contents as perceived by humans [14]. This indexing approach enables user to search multimedia objects conceptually by the semantic visual features.

2.1 Index Structure & Hierarchy

We consider a set of multimedia data objects $\{O_j\}$, such as images, video, or music, where their semantic characteristics and contents cannot be extracted automatically. Each O_j has an index set I_j, which consists of a number of elements $\{e_{j1}, e_{j2}, ..., e_{jM_j}\}$. Each index element e is a triple, which is composed of an index term ID t_{jk}, a corresponding index score s_{jk}, and an object ID O_{j-ID}. The index scores reflect the significance of an index term to the object; the higher the index score, the more important is the index term to the object. In other words, the lower the index score means the less important is the index term to the object. Fig. 1 shows a clear view of the decomposition structure of the multimedia data objects and index elements.

The index hierarchy refers to the collective index sets I of all the objects O_j in the database [14]. Fig. 2 shows that the index set I is partitioned into N levels L_1, L_2, ..., L_N by partitioning the score value s_{jk} with a set of parameters P_1, P_2, ..., P_N. For a given index term with score x, the index term will be placed in level L_i if $P_i \leq x < P_{i+1}$, where $i = 1, ..., N-1$. Otherwise, it would be placed in level N if $P_N \leq x$. In this index hierarchy, the higher the level, the more important it is.

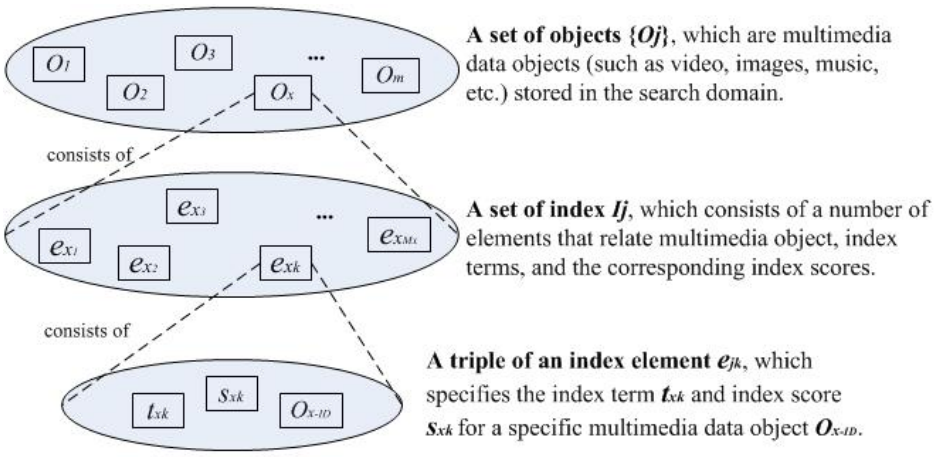


Figure 1. Composition of Multimedia Data Objects and Index Elements

2.2 Minimal Indexing & Index Growth

In order to be discovered by users, each multimedia object should be minimally indexed initially. When an object O_j is minimally indexed, it means that O_j has only a single index term T where T consists of a single word only. Through successive usage of the system, the index set of an object would be grow, such that an object which is minimally indexed with term $T = T_1$ may become indexed with multiple terms $T = T_1, T_2, ..., T_n$. Consider an object O_J, which is minimally indexed with an index term T_1. Consider a user input query $Q(T_1, T_2)$, the system would return an answer vector V_{ans} which consists a set of objects that is indexed with T_1 or T_2. When the user select O_J in V_{ans}, T_2 would be added to O_J at the low level of the index hierarchy. If many queries that contain T_2 also select O_J continuously, the index score of T_2 of O_J would be increased and promoted to the high level of the index hierarchy. Consequently, T_2 of O_J would be properly indexed. Meanwhile, T_1 of O_J may drop to the lower level of the index hierarchy since it would be affected by the user relevance feedback.

2.3 Index Score Update Influenced by Relevance Feedback

The index scores are directly affected by the user relevance feedback, either positive or negative. By the continuous use of the system, our system can collect and analyze both explicit and implicit relevance feedback from users. When the system receives a positive feedback from user, the index score(s) that relate to the search terms of the query would be increased. Similarly, the index score(s) would be decreased when the a negative feedback is received. Those positive and negative feedbacks can be the relevance feedback that are collected directly or indirectly from users.

Considering an example of a user input search query $Q(T_1, T_2)$ that consists of two search terms T_1 and T_2, suppose there are k multimedia objects $O_1, O_2, ..., O_k$ returned in the answer vector V_{ans} disregarding the object rankings. When the user provides a positive feedback or selects the desired object O_x in the answer vector, the index scores of T_1 and T_2 of O_x would be increased by a predefined value Δ_+. In contrast, when the user provides a negative feedback on O_x, the index scores of T_1 and T_2 of O_x would be decreased by a predefined value Δ_-. Moreover, when a user do not select any object in the answer vector, the index scores of T_1 and T_2 for all objects in the answer vector $(O_1, O_2, ..., O_k)$ would also be decreased.

3 User Relevance Feedback

Relevance feedback (RF) is a classical information retrieval (IR) technique where users relay their agreement with the system's evaluation of document relevance back to the system, which then uses this information to provide a revised list of search results [20]. It allows user to mark relevant (positive feedback) or irrelevant (negative feedback) to the object(s) of the result list by their relevance judgments. The user relevance feedback collected would be useful for refining the index scores, such that it helps tuning the index hierarchy to fit user preferences.

Our model collects both explicit and implicit relevance feedback from the user community. The explicit feedback refers to the relevance, indicating the relevance of the object retrieved for a query, and is collected directly from user judgements. Our model enables users to indicate relevance explicitly using a binary relevance system. Binary relevance feedback indicates that a multimedia data object is either

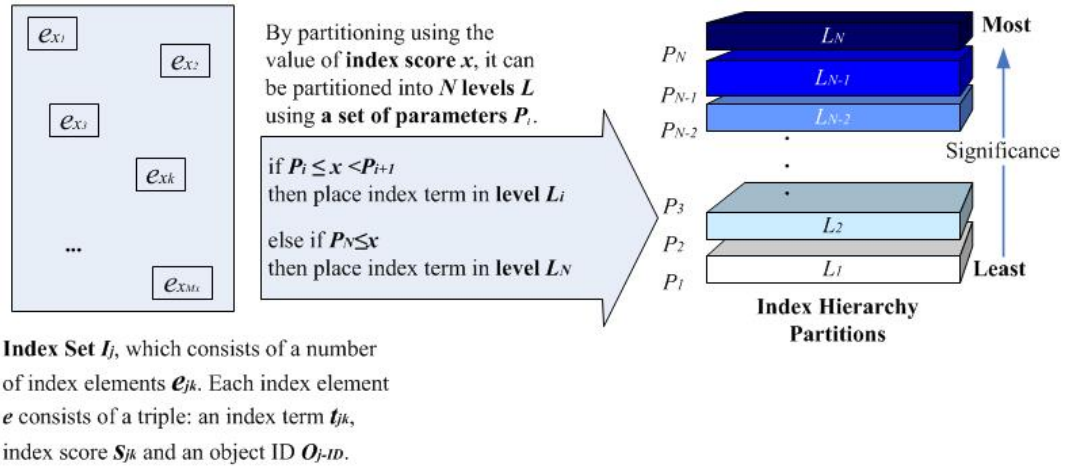


Figure 2. Index Hierarchy

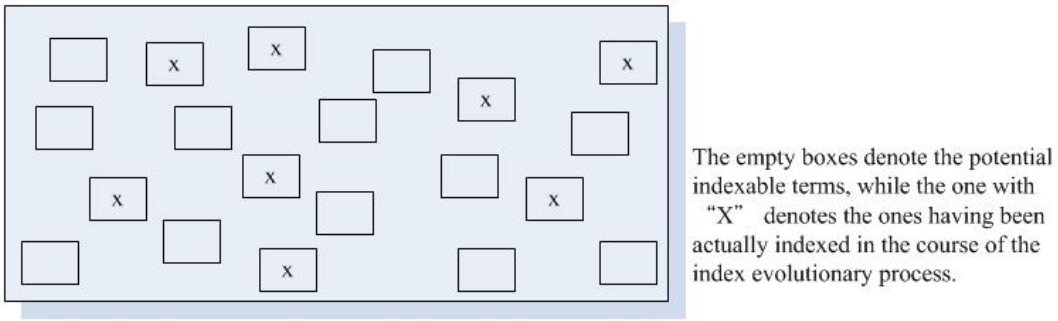


Figure 3. Potential Indexable Terms

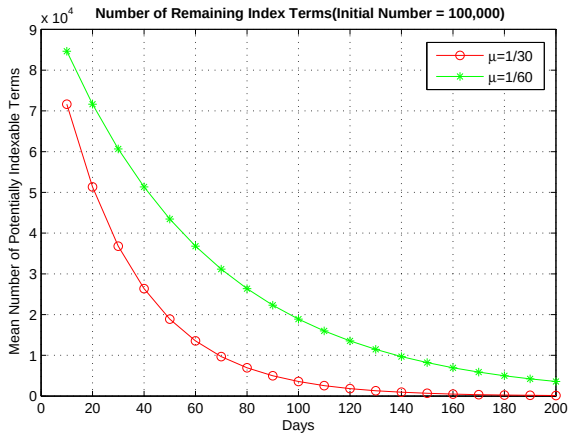


Figure 4. Number of Remaining Index Terms

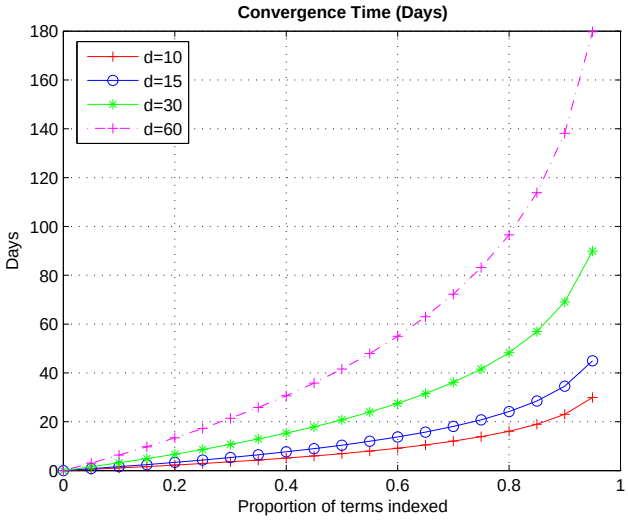


Figure 5. Convergence Time

relevant or irrelevant for a specific query. Once a user submit a query, our system will return a list of query results to the user. In order to maintain the spirit of Web 2.0 [2, 7, 24] collaborative users involvement, our system allows users to provide their relevance feedback for the multimedia objects of the query results. Their feedback can be either positive or negative.

Although the idea of exploiting user's feedback to rate relevance seems promising, it is not easy to convince a community of users to spend their time to explicitly rate objects. Therefore, our model also collects implicit relevance feedback from them. The implicit feedback is inferred from user behaviour and their history, such as noting which object(s) that users do and do not select for viewing, the duration of time spent in viewing an object. All such information can be collected automatically and would reflect user satisfaction and expectation of the query result. When users click on an object in the answer vector, we can infer that the selected object may be relevant to the user query. Our system will treat it as a kind of positive feedback from the user implicitly. On the contrary, when users do not select any object in the answer vector, we can infer that they may think that the objects in the answer vector are irrelevant to their input query or they are not interested in those objects. Our system will treat it as a kind of negative feedback from the user implicitly.

4 Modelling Index Convergence Behaviour

Since the measurement of the relevance of an index term to an object is based on the related index score, the index scores of the system are expected to evolve to an ideal situation. We assume that each index score for an index term of a specific object would have a hidden ideal score value S_H. When the actual index score reaches S_H, this index can be considered as having convergent. By the continuous usage of the system, the indexes would be convergent.

4.1 Index Convergence Behaviour

In theory, indexes would evolve to an ideal state by the index continuous convergence processes. Consider there are J objects in the search space, each of the objects are indexed with m initial index terms, and M be the number of the maximal index terms.

Let N_t be the state of the system which signifies the number of terms remaining to be indexed. N_t is a random variable that changes over time. Let the process starts at $t = 0$. Thus initially, we have

$$N_0 = J(M - m). \quad (1)$$

As time goes on, N_t will gradually decrease. N_t will decrement by 1 whenever a potential indexable term is being indexed. We assume that the random indexing pattern for a given term follows a Poisson process [12] with indexing rate μ, where in a small time interval Δh, a potential indexable term has a probability of $\approx \mu\Delta t$ of being actually indexed. The rate is dependent on the usage and indexing frequency of objects in the collection. Thus, over time, each potential indexable term is gradually being deleted as they are become indexed. Fig. 3 shows this situation where the empty boxes signify the potential indexable terms, while the ones with "X" signify the ones having been actually indexed in the course of this evolutionary process.

From the property of the Poisson distribution [10,17], the probability that a potential indexable term remaining unin-

dexed at time t is $e^{-\mu t}$. Therefore, we obtain the following binomial distribution [17] for N_t

$$Prob[N_t = k] = \binom{N_0}{k} (1 - e^{-\mu t})^{N_0 - k} e^{-\mu t k}, \quad (2)$$

which gives

$$E(N_t) = N_0 e^{-\mu t}, \quad (3)$$

$$Var(N_t) = N_0(1 - e^{-\mu t})e^{-\mu t}, \quad (4)$$

Adopting a time unit of days, $\frac{1}{\mu}$ can be taken as the average time elapsed to install the index term. For example, if $\mu = 0.1$, this means that the average time to install the index term is 10 days. Fig. 4 plots the number of remaining index terms over time for $N_0 = 100,000$, and $\mu = \frac{1}{30}, \frac{1}{60}$. We see that the number of indexable terms drops quickly at first, then do so slowly as time goes on. As $t \to \infty$, we see from equation (3) that the collection tends to be fully indexed with $E(N_t) \to 0$, irrespective of the initial number of potential indexable terms. Also, from equation (4), $Var(N_t) \to 0$ as $t \to \infty$, which indicates that the effect of stochastic fluctuation would be small; this implies that, over a long period of time, the process may be viewed as a deterministic one.

From equation (3), we can determine the time T_p, on average, when a certain proportion of p of the potential indexable terms have been indexed; i.e. letting

$$p = \frac{(N_0 - N_0 e^{-\mu T_p})}{N_0}, \quad (5)$$

we obtain

$$T_p = \frac{1}{\mu} \ln(\frac{1}{1-p}). \quad (6)$$

Replacing μ by $d = \frac{1}{\mu}$ in the above gives

$$T_p = d \ln(\frac{1}{1-p}). \quad (7)$$

Fig. 5 shows the convergence behaviour for $d = 10, 15, 30, 60$. In this model, each potential index term behaves independently of other index terms. When there are many terms remaining to be indexed, the collective indexing rate tends to be high, and this collective rate will decline as fewer and fewer terms are available to be indexed; this is evident from Figure 3, where the curves rise much more steeply as $p \to 1$.We observe that in order to complete the indexing of 95% of the terms, it takes approximately three times the amount of time for indexing an individual term. Indeed taking $p = 0.95$, we have $\ln(\frac{1}{1-p}) = 2.99$.

5 Experiments on Index Convergence

We performed experiments to examine the convergence behaviour of the index hierarchy. In order to evaluate the convergence behaviour of the model, we simulated the search processes including query submission from user, searching the relevant multimedia data objects, ranking the results, and collecting user relevance feedback. The goal of the series of experiments is to investigate the convergence behaviour of the index hierarchy.

In our simulation model, we assume the arrival of user queries follow poisson distribution. We tested the runs in the same environment with the same initial settings, such as number of queries (i.e., 50,000), number of initial index terms per object (i.e., 3), score increment / decrement after reeving user relevance feedback, indexing threshold (i.e., 50%) etc. By varying some variables, such as maximum number of index terms per object t_{max} and number of multimedia data objects O in the search domain, we test its effect on the convergence behaviour.

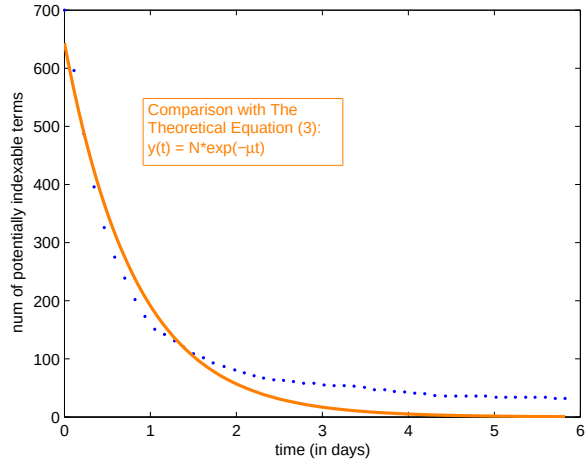
5.1 Decay Behaviour of Number of Remaining Index Terms

In the first series of runs, we performed the queries with 100 data objects and $d = 80$. We collected the number of potentially indexable terms N_t for every 1,000 queries. Initially, each data object O is indexed with 3 index terms. In each query, the number of search terms is fixed as 2 and the number of objects returned in search results are fixed as 10. We tested it with different number of maximum number of indexable terms t_{max}. Fig. 6 (a) (i.e., $t_{max} = 10$ and $N_0 = 700$) and (b) (i.e., $t_{max} = 30$ and $N_0 = 1700$) show the points that are the number of potentially indexable terms N_t collected from the runs. Then, we fit the data points exponentially with the equation (3). Our results show clearly that the number of potentially indexable terms 'decay' exponentially through the time.

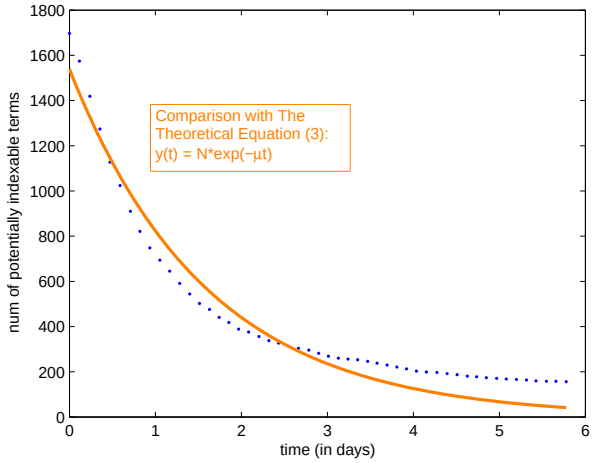
Then, we performed the queries with a different number of initial data objects O in the search domain and $d = 0.1$ to test the scalability of our index model. We also collected the number of potentially indexable terms N_t for every 1,000 queries with 3 initial index terms and 10 for the number of objects in search results. Fig. 6(c) (i.e., $O = 500$ and $N_0 = 3500$) and (d) (i.e., $O = 1000$ and $N_0 = 7000$) show the points that are the number of potentially indexable terms N_t collected from the runs and its best fitting curve exponentially with the equation (3). Our results also show that the number of potentially indexable terms 'decay' exponentially over time, although the number of data objects O increases.

5.2 Index Convergence

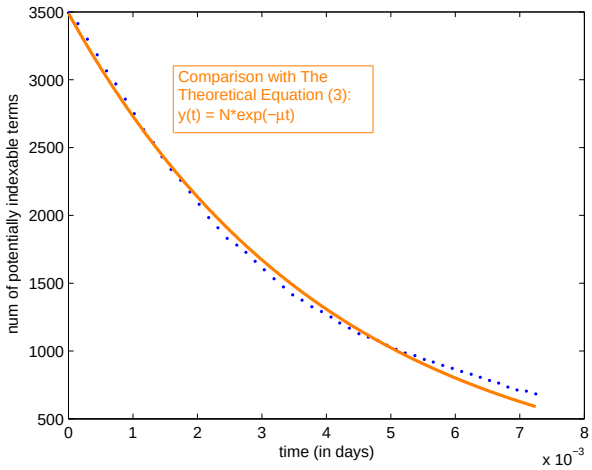
We further investigate the index convergence of our indexing model by another set of runs. We performed the runs with different values for the duration variable d. We col-



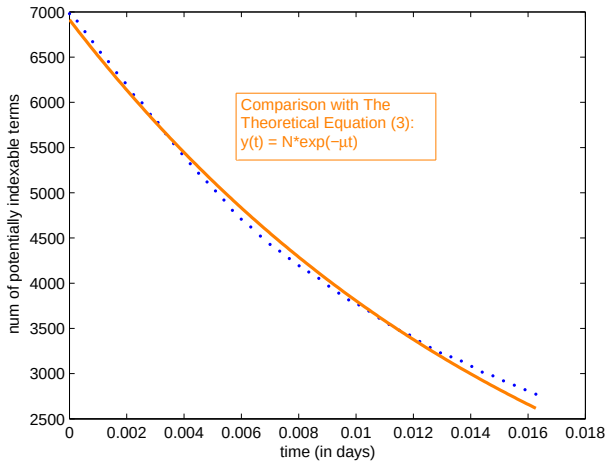
(a) t_{max}=10 and N_0=700



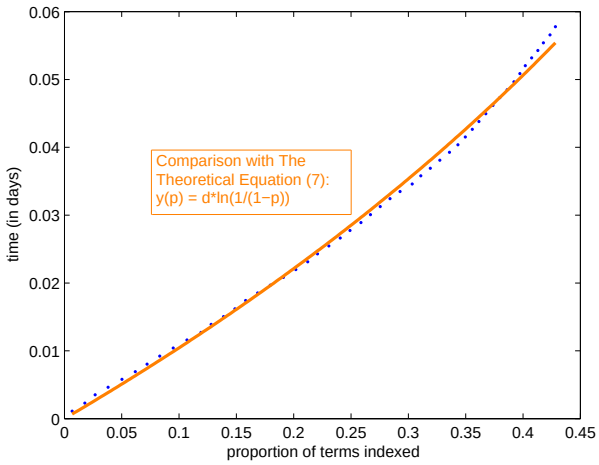
(b) t_{max}=20 and N_0=1700



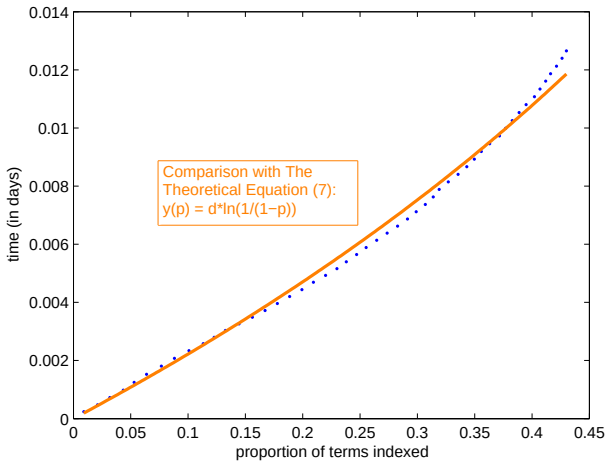
(c) O=500 and N_0=3500



(d) O=1000 and N_0=7000



(e) d=10



(f) d=50

Figure 6. Test Results.

lected N_t and calculated the proportion of terms indexed p for every 1,000 queries. Fig. 6 (e) (i.e., $d = 10$) and (f) (i.e., $d = 50$) show the points that are the p collected throughout the run and its best fitting curve with equation (7).

6 Conclusions & Future Works

We presented a collaborative indexing approach for enabling multimedia retrieval within a large collection of multimedia data objects. Our indexing approach helps to discover multimedia resources systematically by keeping track of the user query behaviour. By analyzing the search information, the user relevance feedback helps the index hierarchy to evolve towards to users' desired preferences. Thus, user satisfaction would be maximized. Our experimental result shows that the index converge successfully after successive use. In the future, we will further focus deeply on examining the index convergence behaviours.

References

[1] I. A. Azzam, C. H. C. Leung, and J. F. Horwood. Implicit concept-based image indexing and retrieval. In Y.-P. P. Chen, editor, *Proceedings of the 10th International Multimedia Modeling Conference (MMM 2004), 5-7 January 2004, Brisbane, Australia*, page 354. IEEE Computer Society, 2004.

[2] S. Boll. Multitube–where web 2.0 and multimedia could meet. *IEEE MultiMedia*, 14(1):9–13, 2007.

[3] L. Finkelstein, E. Gabrilovich, Y. Matias, E. Rivlin, Z. Solan, G. Wolfman, and E. Ruppin. Placing search in context: the concept revisited. *ACM Transactions on Information Systems*, 20(1):116–131, 2002.

[4] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1):83–105, 2003.

[5] J. F. Gantz, D. Reinesel, C. Chute, W. Schlichting, J. McArthur, S. Minton, I. Xheneti, A. Toncheva, and A. Manfrediz. The expanding digital universe: a forecast of worldwide information growth through 2010. *IDC White Paper*, March 2007.

[6] T. Gevers and A. W. M. Smeulders. Image search engines - an overview. 2004.

[7] E. Giannakidou, I. Kompatsiaris, and A. Vakali. Semsoc: Semantic, social and content-based clustering in multimedia collaborative tagging systems. In *ICSC '08: Proceedings of the 2008 IEEE International Conference on Semantic Computing*, pages 128–135, Washington, DC, USA, 2008. IEEE Computer Society.

[8] J. Go'mez and J. L. Vicedo. Next-generation multimedia database retrieval. *IEEE MultiMedia*, 14(3):106–107, 2007.

[9] G. Goth. Multimedia search: Ready or not? *IEEE Distributed Systems Online*, 5(7), 2004.

[10] F. A. Haight. *Handbook of the Poisson Distribution*. Wiley, 1967.

[11] R. Hawarth and H. Buxton. Conceptual-description from monitoring and watching image sequences. 18, 2000.

[12] J. F. C. Kingman. *Poisson processes*. Oxford University Press, 1993.

[13] C. Leung, J. Liu, W. S. Chan, and A. Milani. An architectural paradigm for collaborative semantic indexing of multimedia data objects. In *VISUAL '08: Proceedings of the 10th International Conference on Visual Information Systems*, Salerno, Italy, 2008. (To Appear).

[14] C. H. C. Leung and J. Liu. Multimedia data mining and searching through dynamic index evolution. In *VISUAL '07: Proceedings of the 9th International Conference on Visual Information Systems*, pages 298–309, Shanghai, China, 2007.

[15] H. Mu'ller, W. Mu'ller, D. M. Squire, S. Marchand-Maillet, and T. Pun. Performance evaluation in content-based image retrieval: Overview and proposals. 22(5), 2001.

[16] P. Over, C. H. C. Leung, H. H.-S. Ip, and M. Grubinger. Multimedia retrieval benchmarks. *IEEE MultiMedia*, 11(2):80–84, 2004.

[17] A. Sleeper. *Six Sigma Distribution Modeling*. McGraw-Hill, 2006.

[18] C. G. M. Snoek, M. Worring, J. C. van Gemert, J.-M. Geusebroek, and A. W. M. Smeulders. The challenge problem for automated detection of 101 semantic concepts in multimedia. In *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, pages 421–430, New York, NY, USA, 2006. ACM.

[19] A. M. Tam and C. H. C. Leung. Structured natural-language descriptions for semantic content retrieval of visual materials. *J. Am. Soc. Inf. Sci. Technol.*, 52(11):930–937, 2001.

[20] V. Vinay, K. Wood, N. Milic-Frayling, and I. J. Cox. Comparing relevance feedback algorithms for web search. In *WWW '05: Special interest tracks and posters of the 14th international conference on World Wide Web*, pages 1052–1053, New York, NY, USA, 2005. ACM.

[21] X. Y. Wei and C. W. Ngo. Fusing semantics, observability, reliability and diversity of concept detectors for video search. In *MM '08: Proceedings of the 16th ACM international conference on Multimedia*, pages 81–90, New York, NY, USA, 2008. ACM.

[22] R. C. F. Wong and C. H. C. Leung. Automatic semantic annotation of real world web images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008. (To Appear).

[23] B. Yang and A. R. Hurson. Ad hoc image retrieval using hierarchical semantic-based index. In *AINA '05: Proceedings of the 19th International Conference on Advanced Information Networking and Applications*, pages 629–634, Washington, DC, USA, 2005. IEEE Computer Society.

[24] Y. Yesilada and S. Harper. Web 2.0 and the semantic web: hindrance or opportunity?: W4a – international cross-disciplinary conference on web accessibility 2007. *SIGACCESS Access. Comput.*, (90):19–31, 2008.

L-BFGS and Delayed Dynamical Systems Approach for Unconstrained Optimization

Xiaohui XIE

Abstract

The dynamical (or ode) systems approach for optimization problems has existed for two decades. The main feature of this approach is that a continuous path starting from the initial point can be generated and eventually the path will converge to the solution. This approach is quite different from conventional optimization methods where a sequence of points, or a discrete path, is generated. An advantage of the dynanmical systems approach is that it is more suitable for large scale problems. Common examples of this approach are o.d.e.'s based on the steepest descent direction or the Newton direction. In this research we apply the L-BFGS scheme to the o.d.e. model, hopefully to improve on the rate of convergence over the steepest descent direction, but not to suffer from the large amount of computational work in the Newton direction.

1. Problem background and introduction

This paper studies computational methods for a local or the global minimizer of an unconstrained optimization problem. Optimization problems are classified into:

(a) Unconstrained Problem:

$$\min_{x\in R^n} f(x) \qquad f: R^n \to R^1 \qquad \text{(UP)}$$

(b) Equality Constrained Problem:

$$\min_{x\in R^n} f(x)$$
$$h(x) = 0 \qquad h: R^n \to R^p$$

(c) Inequality Constrained Problem:

$$\min_{x\in R^n} f(x)$$
$$g(x) \le 0 \qquad g: R^n \to R^m$$

(d) General Constrained Problem:

$$\min_{x\in R^n} f(x)$$
$$g(x) \le 0$$
$$h(x) = 0$$

The motivation of unconstrained methods is to generate a sequence of points $\{x_k\}$ (x_0 given) such that **(i)** $f(x_k) > f(x_{k+1})$; (**ii)** $\{x_k\}$ is convergent, and **(iii)** the limit point of the sequence is a stationary point of (UP). Different methods advance from x_k to x_{k+1} differently. Well-known methods include the steepest descent method, Newton's method and quasi-Newton method. A common theme behind all these methods is to find a direction $p \in R^n$ so that there exists an $\bar{\varepsilon} > 0$ such that

$$f(x + \varepsilon p) < f(x) \qquad \forall \varepsilon \in (0, \bar{\varepsilon}).$$

This direction is called a descent direction of $f(x)$ at x. Once we have found a descent direction, we may go along this direction to approach one more step toward the optimum solution.

The following paragraphs summarize the advantages and disadvantages of these methods.

1.1. Steepest descent method

Using directional derivative in multivariable calculus, it is clear that for (UP), p is a descent direction at $x \Leftrightarrow \nabla f(x)^T p < 0$. Hence $p = -\nabla f(x)$, or equivalently, $p = -\nabla f(x)/\|\nabla f(x)\|_2$ is obviously a descent direction for $f(x)$. This direction is called the steepest descent direction.

Method of Steepest Descent: At each iteration k: find the lowest point of f in the direction $-\nabla f(x_k)$ from x_k, i.e., find λ_k that solves

$$\min_{\lambda>0} f(x_k - \lambda \nabla f(x_k)).$$

Then $x_{k+1} = x_k - \lambda_k \nabla f(x_k)$.

Unfortunately, the steepest descent method converges only linearly, and sometimes very slowly linearly. In fact, if $\{x_k\}$, which is generated by the steepest descent method, converges to a local minimizer x^* where

$\nabla^2 f(x^*)$ is positive definite (p.d.), and $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of $\nabla^2 f(x^*)$, then one can show that $\{x_k\}$ satisfies

$$\lim_{k\to\infty} \sup \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \le c, c = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}},$$

in a particular weighted l_2 norm, and that the bound on c is tight for some starting x_0. This property indicates that the steepest descent method is q-linearly convergent. When $\lambda_{\max}$ and $\lambda_{\min}$ are far apart, then *c* is close to 1, and the convergence will be slow.

1.2. Newton's method

At point x_k, if $\nabla^2 f(x_k)$ is p.d., the function $f(x)$ can be approximated by a quadratic function based on the Taylor expansion:

$$f(x) \cong f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k) \quad (1)$$

Then the minimizer of (1) is given by

$$\nabla f(x) = 0$$
$$\Rightarrow \nabla f(x_k) + \nabla^2 f(x_k)(x - x_k) = 0$$
$$\Rightarrow x = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k),$$

where $-\left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$ is called Newton's direction. Then we define

$$x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k),$$

and the resulting method of computing $\{x_k\}$ is called the Newton's method.

Newton's method

Given x_0, compute

$$x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k), \quad k \leftarrow k+1.$$

A key requirement for Newton's method is the p.d. of $\nabla^2 f(x_k)$.

Descent directions guarantee that $f(x)$ can be further reduced and therefore they form the basis of some global methods. However, in some real applications, if the starting point is far away from the optimal solution, or the Hessian is not positive definite, Newton's direction is not adopted.

Although Newton's method converges very fast ($\{x_k\}$ converges to x^* q-quadratically), the Hessian matrix is difficult to compute. So we would like to find more feasible methods with (A) no second-order information, i.e., no Hessian; and (B) fast convergence. A rule of thumb is that first-order information normally gives slow (linear) convergence, while second-order information normally gives fast (quadratic) convergence. Let us discuss several practical considerations. In general, the convergence is quadratic: the error is essentially squared at each step (that is, the number of accurate digits doubles in each step). There are some caveats, however. Firstly, Newton's method requires that the derivative be calculated directly. (If the derivative is approximated by the slope of a line through two points on the function, the secant method results; this can be more efficient depending on how one measures computational effort.) Secondly, if the initial value is too far from the true zero, Newton's method can fail to converge. Because of this, most practical implementations of Newton's method put an upper limit on the number of iterations and perhaps on the size of the iterates. Thirdly, if the root being sought has multiplicity greater than one, the convergence rate is merely linear (errors reduced by a constant factor at each step) unless special procedures are taken.

Finding the inverse of the Hessian is an expensive operation. Therefore the descent direction $-\left[\nabla^2 f(x)\right]^{-1} \nabla f(x)$ is often solved approximately (but to great accuracy) using methods such as the conjugate gradient method. There also exist various quasi-Newton methods, where an approximation for the Hessian is used instead.

Table1 compares the advantages and disadvantages of the steepest descent method and the Newton's method.

Table1. Comparison of the methods

	Advantage	Disadvantage
Steepest descent method	Simple and inexpensive, guarantees descent	Slow convergence
Newton's method	Very fast convergence if applicable	Expensive, second-order information matrix inversion

1.3. Quasi-Newton method — BFGS

Instead of using the Hessian matrix, the quasi-Newton methods approximate it.

Quasi-Newton methods are based on Newton's method to find the stationary point of $f(x)$, where the gradient $\nabla f(x)$ is 0. In Quasi-Newton methods the Hessian matrices of second derivatives of $f(x)$ do not need to be computed. The Hessian is updated by analyzing successive gradient vectors instead. Quasi-Newton methods are a generalization of the secant method to find the root of the first derivative for multidimensional problems. In multi-dimensions the secant equation is under-determined, and quasi-Newton methods differ in how they constrain the solution, typically by adding a simple low-rank update to the current estimate of the Hessian.

In quasi-Newton methods, the inverse of the Hessian matrix is approximated in each iteration by a p.d. matrix, say H_k, where k is the iteration index. Thus, the k th iteration has the following basic structure:

(a) set $p_k = -H_k g_k, \left(g_k = \nabla f\left(x_k\right)\right)$

(b) line search along p_k giving $x_{k+1} = x_k + \lambda_k p_k$,

(c) update H_k giving H_{k+1}

The initial matrix H_0 can be any positive definite symmetric matrix, although in the absence of any better estimate, the choice $H_0 = I$ often suffices. Potential advantages of the method are:

(1) only first-order information is required;

(2) H_k being symmetric and p.d. implies the descent property; and

(3) $\mathrm{O}\left(n^2\right)$ multiplications per iteration.

The most important quasi-Newton formula was suggested by Broyden, Fletcher, Goldfarb, and Shanno independently in 1970, and is subsequently known as the BFGS formula. It is used to solve an unconstrained nonlinear optimization problem.

$$H_{k+1}^{BFGS} = H_k + \left(1 + \frac{y_k{}^T H_k y_k}{s_k{}^T y_k}\right)\frac{s_k s_k{}^T}{s_k{}^T y_k} - \left(\frac{s_k y_k{}^T H_k + H_k y_k s_k{}^T}{s_k{}^T y_k}\right) \qquad (2)$$

where $s_k = x_{k+1} - x_k$,

$y_k = \nabla f\left(x_{k+1}\right) - \nabla f\left(x_k\right) = g_{k+1} - g_k$.

We have the following theorem.

THEOREM 1 If H_k^{BFGS} is a p.d. matrix, and $s_k{}^T y_k > 0$, then H_{k+1}^{BFGS} in (2) is also positive definite.

Proof: For any $z \neq 0$, it is sufficient to prove that

$$z^T\left[H_k + \left(1 + \frac{y_k{}^T H_k y_k}{s_k{}^T y_k}\right)\frac{s_k s_k{}^T}{s_k{}^T y_k} - \left(\frac{s_k y_k{}^T H_k + H_k y_k s_k{}^T}{s_k{}^T y_k}\right)\right] > 0$$

In the rest of the proof, the subscript k will be omitted.

Since H is p.d., we can write $H = LL^T$, and let $a = L^T z$ and $b = L^T y$, then

$$\begin{aligned}
z^T H_{k+1} z &= z^T\left[H_k + \left(1 + \frac{y_k{}^T H_k y_k}{s_k{}^T y_k}\right)\frac{s_k s_k{}^T}{s_k{}^T y_k} - \left(\frac{s_k y_k{}^T H_k + H_k y_k s_k{}^T}{s_k{}^T y_k}\right)\right] z \\
&= z^T\left[H + \left(1 + \frac{y^T H y}{s^T y}\right)\frac{s s^T}{s^T y} - \left(\frac{s y^T H + H y s^T}{s^T y}\right)\right] z \\
&= a^T a + \frac{z^T b^T b s s^T z}{\left(s^T y\right)^2} + \frac{\left(z^T s\right)^2}{s^T y} - \frac{z^T s b^T a}{s^T y} - \frac{a^T b s^T z}{s^T y} \\
&= \left(a - \frac{z^T s b}{s^T y}\right)^T\left(a - \frac{z^T s b}{s^T y}\right) + \frac{\left(z^T s\right)^2}{s^T y} \\
&= \left\|a - \frac{z^T s b}{s^T y}\right\|^2 + \frac{\left(z^T s\right)^2}{s^T y} \\
&\geq 0
\end{aligned}$$

If the norm above equals zero, i.e., $\left\|a - \frac{z^T s b}{s^T y}\right\| = 0$, then we have $a = \frac{z^T s b}{s^T y}$,

or $L^T z = \dfrac{z^T s L^T y}{s^T y}$, which means that $y \propto z$.

However, since $s^T y > 0$,

$$\frac{\left(z^T s\right)^2}{s^T y} > 0$$

as $y \propto z$. Thus the theorem is proved. □

1.4. Limited-Memory Quasi-Newton Methods — L-BFGS

Limited-memory quasi-Newton methods are useful for solving large problems whose Hessian matrices cannot be computed at a reasonable cost or are not sparse. These methods maintain simple and compact approximations of Hessian matrices: instead of storing fully dense $n \times n$ approximations, they save only a few vectors of length n that represent the approximations implicitly. Despite these modest storage requirements, they often yield an acceptable rate of convergence. Various limited-memory methods have been proposed; we focus mainly on an algorithm known as L-BFGS, which, as its name suggests, is based on the BFGS updating formula. The main idea of this method is to use curvature information from only the most recent iterations to construct the Hessian approximation. Curvature information from earlier iterations, which is less likely to be relevant to the actual behavior of the Hessian at the current iteration, is discarded in the interest of saving storage.

As we have discussed in section 1.3, each step of the BFGS method has the form

$$x_{k+1} = x_k - \alpha_k H_k \nabla f_k ,$$

where α_k is the step length and H_k is updated at every iteration by means of the formula

$$H_{k+1} = V_k{}^T H_k V_k + \rho_k s_k s_k{}^T , \qquad (3)$$

where

$$\rho_k = \frac{1}{y_k{}^T s_k}, \qquad V_k = I - \rho_k y_k s_k{}^T \qquad (4)$$

and

$$s_k = x_{k+1} - x_k , \quad y_k = \nabla f_{k+1} - \nabla f_k . \qquad (5)$$

Since the inverse Hessian approximation H_k will generally be dense, the cost of storing and manipulating it is prohibitive when the number of variables is large. To circumvent this problem, we store a modified version of H_k implicitly, by storing a certain number (say, *m*) of the vector pairs $\{s_i, y_i\}$ used in the formulas (3)-(5). The product $H_k \nabla f_k$ can be obtained by performing a sequence of inner products and vector summations involving ∇f_k and the pairs $\{s_i, y_i\}$. After the new iterate is computed, the oldest vector pair in the set of pairs $\{s_i, y_i\}$ is replaced by the new pair $\{s_k, y_k\}$ obtained from the current step (5).

We now describe the updating process in a little more detail. At iteration k, the current iterate is x_k and the set of vector pairs is given by $\{s_i, y_i\}$ for $i = k-m, \ldots k-1$. We first choose some initial Hessian approximation H_0 (in contrast to the standard BFGS iteration, this initial approximation is allowed to vary from iteration to iteration) and find by repeated application of the formula (3) that the L-BFGS approximation H_{k+1} satisfies the following formula [39]:

In general, we have for $k+1 \le m$ the usual BFGS updated

$$\begin{aligned} H_{k+1} = {} & V_k^T V_{k-1}^T \cdots V_0^T H_0 V_0 \cdots V_{k-1} V_k \\ & + V_k^T \cdots V_1^T \rho_0 s_0 s_0^T V_1 \cdots V_k \\ & \vdots \\ & + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\ & + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\ & + \rho_k s_k s_k^T . \end{aligned} \qquad (6)$$

For $k+1 > m$ we have the update

$$\begin{aligned} H_{k+1} = {} & V_k^T V_{k-1}^T \cdots V_{k-m+1}^T H_0 V_{k-m+1} \cdots V_{k-1} V_k \\ & + V_k^T \cdots V_{k-m+2}^T \rho_{k-m+1} s_{k-m+1} s_{k-m+1}^T V_{k-m+2} \cdots V_k \\ & \vdots \\ & + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\ & + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\ & + \rho_k s_k s_k^T . \end{aligned} \qquad (7)$$

A method for choosing H_0 that has proved effective in practice is to set $H_0 = \gamma_k I$, where

$$\gamma_k = \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}} .$$

The strategy of keeping the *m* most recent correction pairs $\{s_i, y_i\}$ works well in practice; indeed no other strategy has yet proved to be consistently better. However, the main weakness of the L-BFGS method is that it converges slowly on ill-conditioned problems -- specifically, on problems where the Hessian matrix contains a wide distribution of eigenvalues. On certain applications, the nonlinear conjugate methods are competitive with limited-memory quasi-Newton methods.

Algorithm (L-BFGS)

Choose starting point x_0, integer $m > 0$;

$k \leftarrow 0$;

repeat

Choose H_0

Compute $p_k \leftarrow -H_k \nabla f_k$

Compute $x_{k+1} \leftarrow x_k + \alpha_k p_k$ where α_k is chosen to satisfy the Wolfe conditions;

if $k > m$

Discard the vector pair $\{s_{k-m}, y_{k-m}\}$ from storage;

Compute and save

$s_k \leftarrow x_{k+1} - x_k, y_k = \nabla f_{k+1} - \nabla f_k$;

$k \leftarrow k+1$;

until convergence.

2. Analysis for dynamical systems with time delay

2.1. Introduction of dynamical systems

For the easiness of reading, the (UP) problem is reproduced here:

$$\min_{x \in R^n} f(x) \qquad f: R^n \to R^1 \qquad (8)$$

It is very important that the optimization problem (8) itself is posted in the continuous form, i.e., x can be changed continuously. In the literature, the necessary and sufficient conditions of a local optimum are also presented in the continuous form. Furthermore, almost all the theoretical study for problem (8) is in the continuous form. However, it is very interesting to say that when it comes down to the numerical solution of (8), most of the conventional methods, such as the gradient/steepest descent method, Newton's method and quasi-Newton's method, are all addressed in the discrete form. This interesting situation is mainly due to the fact that the computer's computation can be only done discretely. However, is it possible to study both the optimization problem and the solution methods in its original form, i.e., continuous form? In this sense, we may use the dynamical system approach or neural network approach to solve the original optimization problem.

●Dynamical system approach. The essence of this approach is to convert problem (8) into a dynamical system or an ordinary differential equation (ode) so that the solution of problem (8) corresponds to a stable equilibrium point of this dynamical system.

●Neural network approach. The mathematical representation of neural network is an ordinary differential equation which is asymptotically stable at any isolated solution point. A companion of this neural network is an energy function which is a Lyapunov function. And as time evolves, the solution of the ode will converge to the optimum, and in this whole process, the energy function will decrease monotonically in time.

The following discussion reviews the research results in the dynamical system approach, and identifies the merits of this approach.

Consider the following simple dynamical system or ordinary differential equation

$$\frac{dx(t)}{dt} = p(x). \qquad (9)$$

We first state some classical result on the existence and uniqueness of the solution, and some stability definitions for the dynamical system (9) [66,72].

THEOREM 2. [72] Assume that $p(x)$ is a continuous function from R^n to R^n. Then for arbitrary $t_0 \geq 0$ and $x_0 \in R^n$ there exists a local solution $x(t)$ satisfying $x(t_0) = x_0, t \in [t_0, \tau)$ to (9) for some $\tau > t_0$. If furthermore $p(x)$ is locally Lipschitz continuous at x_0, then the solution is unique, and if $p(x)$ is Lipschitz continuous in R^n then τ can be extended to ∞.

DEFINITION 1. (Equilibrium point). A point $x^* \in R^n$ is called an equilibrium point of (9) if $p(x^*) = 0$.

DEFINITION 2. (Stability in the sense of Lyapunov). Let $x(t)$ be the solution of (9). An isolated equilibrium point x^* is Lyapunov stable if for any $x_0 = x(t_0)$ and any scalar $\varepsilon > 0$, there exists a $\delta > 0$ such that if $\|x(t_0) - x^*\| < \delta$, then $\|x(t) - x^*\| < \varepsilon$ for $t \geq t_0$.

DEFINITION 3. (Convergence). Let $x(t)$ be the solution of (9). An isolated equilibrium point x^* is convergent if there exists a $\delta > 0$ such that if $\|x(t_0) - x^*\| < \delta$, $x(t) \to x^*$ as $t \to \infty$.

A dynamical system or o.d.e. (9) arising from an optimization problem needs to have *p(x)* being a descent direction for the objective function *f(x)*. Some well-known versions are:

Dynamical system based on the steepest descent direction

$$\frac{dx(t)}{dt} = -\nabla f(x(t))$$

Dynamical system based on the Newton direction:

$$\frac{dx(t)}{dt} = -\left[\nabla^2 f(x(t))\right]^{-1} \nabla f(x(t)) \qquad (10)$$

As in the discrete optimization methods in the previous chapter, the steepest descent direction has a slow convergence rate – meaning that it takes a very "large" value of t to approach the equilibrium point. The Newton direction has a much faster convergence rate, but the amount of work in evaluating the Jacobian is much greater.

Some other dynamical systems in the literature are:

$$\frac{dx(t)}{dt} = s(t) \cdot p\left(x(t)\right), \qquad (11)$$

$$a(t) \cdot \frac{d^2x(t)}{dt^2} + b(t) \cdot B\left(x(t)\right) \cdot \frac{dx(t)}{dt} = p\left(x(t)\right), \qquad (12)$$

where $p(x)$ is a descent direction for $f(x)$, $B(x) \in R^{n \times n}$ is a positive definite matrix, $a(t)$ and $b(t)$ are scalar functions in t, and $s(t)$ is a positive scalar function in t and bounded above.

A major advantage of the dynamical systems approach is that very large problems can be solved [41]. No matter whether we use any of (9)-(12), existing o.d.e. methods are quite mature to tackle these problems. Solving systems with tens or hundreds of thousands of unknowns poses no problem to o.d.e. solvers. The problem size handled can be much larger than traditional methods described in Chapter 1, which are of order of magnitude in the thousands. The research in the o.d.e. approach is to find a "good" *p(x)* in (9) that balances the convergence rate and the amount of work.

The dynamical systems approach normally consists of the following three steps:

(a) to establish an ode system;

(b) to study the convergence of the solution $x(t)$ of the ode as $t \to \infty$; and

(c) to solve the ode system numerically.

The convergence study of $x(t)$ as $t \to \infty$ and the stability of the corresponding dynamical system have mostly been addressed on a case by case base. No standard theory and/or methodology are given. This phenomenon certainly limits the systematic study of the dynamical system approach and its application potential as well. Two papers are worth mentioning, one by Tanabe [61] which used the stability theory of the dynamical system to study the ode system, and the other one by Yamashita [71] which employed Lyapunov's direct method to study the ode system.

Even though the solutions of ode systems are continuous, yet the actual computation has to be done discretely. In all the dynamical systems (10)-(12), the numerical solutions were mainly solved by either discrete optimization methods or finite difference methods.

In summary, the main attractiveness of this approach is its simplicity and its originality in pursuing the continuous form. Furthermore, there is not any restriction on the form of the objective function *f(x)* in (8).

2.2. Delayed dynamical systems approach

As stated above, the steepest descent direction and the Newton direction of the dynamical systems approach both have their weakness. The main idea in this paper is to apply the theme of the L-BFGS algorithm in Chapter 1 to the dynamical systems approach, making it a bridge between the steepest descent direction and the Newton direction. The resulting dynamical system is a delayed o.d.e., thus we call it the delayed dynamical systems approach.

The delayed dynamical systems approach solves the delayed o.d.e.:

$$\frac{dx(t)}{dt} = -H(x(t), x(t-\tau_1(t)), ..., x(t-\tau_m(t)))\nabla f\left(x(t)\right), \qquad (13)$$

where H and $t-\tau_1(t)$, ..., $t-\tau_m(t)$ are to be defined below.

As the delayed o.d.e. (13) is numerically solved, we compute approximations $x_0, x_1, ..., x_k, x_{k+1}, ...$ to $x(t)$ at time points $t_o, t_1, ..., t_k, t_{k+1}, ...$ We define $H = H_k$ to be a different function in the interval $(t_{k-1}, t_k]$ iteratively by:

Given t_0, $x(t_0) = x_0$, and an initial H_0, for $t_0 <= t$, we define

$$H(x(t), x(t_0)) := H_1(x(t), x(t_0)) := V_0(t)^T H_0 V_0(t) + \rho_0(t)\, s_0(t)\, s_0(t)^T,$$

where

$$s_0(t) = x(t) - x_0,$$
$$y_0(t) = \nabla f(x(t)) - \nabla f(x_0),$$
$$\rho_0(t) = 1/(y_0(t)^T s_0(t)),$$
$$V_0(t) = I - \rho_0(t)\, y_0(t)\, s_0(t)^T,$$

in the R.H.S. of (13) and determine a stepsize $h_1 = (t_1 - t_0)$ to compute, using some numerical o.d.e. method, an approximation x_1 to $x(t_1)$ at t_1. Then for $t_1 <= t$, we define

$H(x(t), x(t_1), x(t_0))$

$:= H_2(x(t), x(t_1), x(t_0))$
$:= V_1(t)^T V_0(t_1)^T H_0 V_0(t_1) V_1(t)$
$+V_1(t)^T \rho_0(t_1) s_0(t_1) s_0^T(t_1) V_1(t)$
$+ \rho_1(t) s_1(t) s_1(t)^T$,

where

$s_1(t) = x(t) - x(t_1)$,
$y_1(t) = \nabla f(x(t)) - \nabla f(x(t_1))$,
$\rho_1(t) = 1/(y_1(t)^T s_1(t))$,
$V_1(t) = I - \rho_1(t) y_1(t) s_1(t)^T$,

in the R.H.S. of (13) and determine a stepsize $h_2 = (t_2 - t_1)$ to compute, using some numerical o.d.e. method, an approximation x_2 to $x(t_2)$ at t_2. Of course, computationally x_1 is used instead of $x(t_1)$. This process is repeated until we have accepted x_{m-1} at t_{m-1}. Then for $t_{m-1} <= t$, we use

$H(x(t), x(t_{m-1}), \ldots, x(t_1), x(t_0))$ (13A)
$:= H_m(x(t), x(t_{m-1}), \ldots, x(t_1), x(t_0))$
$:= V_{m-1}(t)^T V_{m-2}(t_{m-1})^T \ldots V_1(t_2)^T V_0(t_1)^T H_0$
$\cdot V_0(t_1) V_1(t_2) \ldots V_{m-2}(t_{m-1}) V_{m-1}(t)$
$+V_{m-1}(t)^T V_{m-2}(t_{m-1})^T \ldots V_1(t_2)^T \rho_0(t_1) s_0(t_1)$
$\cdot s_0(t_1)^T V_1(t_2) \ldots V_{m-2}(t_{m-1}) V_{m-1}(t)$
$+\ldots +V_{m-1}(t)^T \rho_{m-2}(t_{m-1}) s_{m-2}(t_{m-1}) s_{m-2}(t_{m-1})^T$
$V_{m-1}(t) + \rho_{m-1}(t) s_{m-1}(t) s_{m-1}(t)^T$,

where

$s_{m-1}(t) = x(t) - x(t_{m-1})$,
$y_{m-1}(t) = \nabla f(x(t)) - \nabla f(x(t_{m-1}))$,
$\rho_{m-1}(t) = 1/(y_{m-1}(t)^T s_{m-1}(t))$,
$V_{m-1}(t) = I - \rho_{m-1}(t) y_{m-1}(t) s_{m-1}(t)^T$.

to compute x_m at t_m. Beyond this point we save only m previous values of x. The definition of H is now, for $m <= k$,

for $t_k <= t$,

$H(x(t), x(t_k), \ldots, x(t_{k-m+2}), x(t_{k-m+1}))$ (13B)
$:= H_{k+1}(x(t), x(t_k), \ldots, x(t_{k-m+2}), x(t_{k-m+1}))$
$:= V_k(t)^T V_{k-1}(t_k)^T, \ldots, V_{k-m+2}(t_{k-m+3})^T$
$\cdot V_{k-m+1}(t_{k-m+2})^T H_0 V_{k-m+1}(t_{k-m+2})$
$\cdot V_{k-m+2}(t_{k-m+3}) \ldots V_{k-1}(t_k) V_k(t)$
$+V_k(t)^T V_{k-1}(t_k)^T ,,, V_{k-m+2}(t_{k-m+3})^T$
$\cdot \rho_{k-m+1}(t_{k-m+2}) s_{k-m+1}(t_{k-m+2}) s_{k-m+1}(t_{k-m+2})^T$
$\cdot V_{k-m+2}(t_{k-m+3}) \ldots V_{k-1}(t_k) V_k(t)$
$+\ldots +V_k(t)^T \rho_{k-1}(t_k) s_{k-1}(t_k) s_{k-1}(t_k)^T V_k(t)$
$+ \rho_k(t) s_k(t) s_k(t)^T$,

where

$s_k(t) = x(t) - x(t_k)$,
$y_k(t) = \nabla f(x(t)) - \nabla f(x(t_k))$,
$\rho_k(t) = 1/(y_k(t)^T s_k(t))$,
$V_k(t) = I - \rho_k(t) y_k(t) s_k(t)^T$.

It is obvious that the delayed o.d.e. (13) is a continuous version of the L-BFGS scheme. The $H=H_k$ in (13) attempts to approximate the inverse of the Jacobian in the Newton method. It is worth mentioning that the matrix H_k is never computed explicitly. We only need to compute the R.H.S. of (13), i.e., the product of H_k and a vector.

2.3. Main stages of this research

(A) Prove that the function H in (13) is positive definite. (Done and shown in Appendix *I*)
(B) Prove that H is Lipschitz continuous.
(C) Show that the solution to (13) is asymptotically stable.
(D) Show that (13) has a better rate of convergence than the dynamical system based on the steepest descent direction.
(E) Perform numerical testing.
(F) Apply this new optimization method to practical problems.

APPENDIX I: To show that *H* in (13) is positive definite.

Without loss of ambiguity, in the subsequent proof, we drop the t, t_o, t_1, ..., t_k, t_{k+1}, etc. in $s_k(t)$, $y_k(t)$, $V_k(t)$, and so on below.

PROPERTY 1. If H_0 is positive definite, the matrix H defined by (13) is positive definite (provided that $y_i^T s_i > 0$ for all i).

Proof: We prove the result by induction. From the above discussion we know that (13), the continuous analog of the L-BFGS formula, has two cases. Hence our proof needs to cater for each of them.

For the first case $k+1 > m$, note that when *m=1*

$$H_{k+1} = V_k^T H_0 V_k + \rho_k s_k s_k^T$$
$$= H_0 - \frac{H_0 y_k s_k^T}{y_k^T s_k} - \frac{s_k y_k^T H_0}{y_k^T s_k} + \frac{s_k y_k^T H_0 y_k s_k^T}{\left(y_k^T s_k\right)^2} + \frac{s_k s_k^T}{y_k^T s_k}$$

It is obvious that the proof of p.d. of this matrix is the same as that of Theorem 1 in section 1.3. Therefore, H_{k+1} is p.d. when *m=1*.

Now suppose they are true for $m = l$, we show that they are true for $m = l+1$.

When $m = l$, we have (denoting H_{k+1} by H^l_{k+1} to emphasize $m = l$)

$$
\begin{aligned}
H_{k+1}^{l} &= V_k^T V_{k-1}^T \cdots V_{k-l+1}^T H_0 V_{k-l+1} \cdots V_{k-1} V_k \\
&\quad + \{V_k^T \cdots V_{k-l+2}^T \rho_{k-l+1} s_{k-l+1} s_{k-l+1}^T \\
&\qquad \cdot V_{k-l+2} \cdots V_k \\
&\quad + V_k^T \cdots V_{k-l+3}^T \rho_{k-l+2} s_{k-l+2} s_{k-l+2}^T \\
&\qquad \cdot V_{k-l+3} \cdots V_k \\
&\quad \vdots \\
&\quad + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\
&\quad + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\
&\quad + \rho_k s_k s_k^T \}.
\end{aligned}
$$

being positive definite. (There are $l+1$ terms in H_{k+1}^{l}.)

If $m = l+1$, from (13B)

$$
\begin{aligned}
H_{k+1}^{l+1} &= V_k^T V_{k-1}^T \cdots V_{k-l}^T H_0 V_{k-l} \cdots V_{k-1} V_k \\
&\quad + V_k^T \cdots V_{k-l+1}^T \rho_{k-l} s_{k-l} s_{k-l}^T V_{k-l+1} \cdots V_k \\
&\quad + \{V_k^T \cdots V_{k-l+2}^T \rho_{k-l+1} s_{k-l+1} s_{k-l+1}^T \\
&\qquad \cdot V_{k-l+2} \cdots V_k \\
&\quad + V_k^T \cdots V_{k-l+3}^T \rho_{k-l+2} s_{k-l+2} s_{k-l+2}^T \\
&\qquad \cdot V_{k-l+3} \cdots V_k \\
&\quad \vdots \\
&\quad + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\
&\quad + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\
&\quad + \rho_k s_k s_k^T \}.
\end{aligned}
$$

(There are $l+2$ terms in H_{k+1}^{l+1}.)

Comparing these two equations we find that the terms in curly braces are the same, and let

$$
\begin{aligned}
(*) &= V_k^T \cdots V_{k-l+2}^T \rho_{k-l+1} s_{k-l+1} s_{k-l+1}^T \\
&\qquad \cdot V_{k-l+2} \cdots V_k + V_k^T \cdots V_{k-l+3}^T \\
&\qquad \cdot \rho_{k-l+2} s_{k-l+2} s_{k-l+2}^T V_{k-l+3} \cdots V_k \\
&\quad \vdots \\
&\quad + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\
&\quad + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\
&\quad + \rho_k s_k s_k^T
\end{aligned}
$$

Thus,

$$
\begin{aligned}
H_{k+1}^{l} &= V_k^T V_{k-1}^T \cdots V_{k-l+1}^T H_0 V_{k-l+1} \cdots V_{k-1} V_k \\
&\quad + (*)
\end{aligned}.
$$

$$
\begin{aligned}
H_{k+1}^{l+1} &= V_k^T V_{k-1}^T \cdots V_{k-l}^T H_0 V_{k-l} \cdots V_{k-1} V_k \\
&\quad + V_k^T \cdots V_{k-l+1}^T \rho_{k-l} s_{k-l} s_{k-l}^T V_{k-l+1} \cdots V_k + (*) \\
&= V_k^T V_{k-1}^T \cdots V_{k-l+1}^T \left(V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T \right) \\
&\quad V_{k-l+1} \cdots V_{k-1} V_k + (*).
\end{aligned}
$$

Since we have assumed that H_{k+1}^{l} is p.d., if we try to prove that H_{k+1}^{l+1} is also p.d., we should prove that $V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T$ and H_0 have the same property, i.e., $V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T$ is also p.d..

Now we move forward to prove that $V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T$ is p.d.

For any $z \neq 0$,

$$
\begin{aligned}
& z^T \left(V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T \right) z \\
&= z^T \left[H_0 - \frac{H_0 y_{k-l} s_{k-l}^T}{s_{k-l}^T y_{k-l}} - \frac{s_{k-l} y_{k-l}^T H_0}{s_{k-l}^T y_{k-l}} + 1 \right. . \\
&\qquad \left. + \frac{s_{k-l} y_{k-l}^T H_0 y_{k-l} s_{k-l}^T}{\left(s_{k-l}^T y_{k-l} \right)^2} + \frac{s_{k-l} s_{k-l}^T}{s_{k-l}^T y_{k-l}} \right] z
\end{aligned}
$$

Since H_0 is p.d., we can write $H_0 = LL^T$, and let $a = L^T z$ and $b = L^T y_{k-l}$. Then

$$
\begin{aligned}
& z^T \left(V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T \right) z \\
&= a^T a - \frac{a^T b s_{k-l}^T z}{s_{k-l}^T y_{k-l}} - \frac{z^T s_{k-l} b^T a}{s_{k-l}^T y_{k-l}} \\
&\qquad + \frac{z^T s_{k-l} b^T b s_{k-l}^T z}{\left(s_{k-l}^T y_{k-l} \right)^2} + \frac{\left(z^T s_{k-l} \right)^2}{s_{k-l}^T y_{k-l}} \\
&= a^T a - \frac{2 a^T b s_{k-l}^T z}{s_{k-l}^T y_{k-l}} + \frac{\left\| s_{k-l}^T z \right\|^2 \cdot \|b\|^2}{\left(s_{k-l}^T y_{k-l} \right)^2} \\
&\qquad + \frac{\left(z^T s_{k-l} \right)^2}{s_{k-l}^T y_{k-l}} \\
&= \left(a - \frac{z^T s_{k-l} b}{s_{k-l}^T y_{k-l}} \right)^T \left(a - \frac{z^T s_{k-l} b}{s_{k-l}^T y_{k-l}} \right) \\
&\qquad + \frac{\left(z^T s_{k-l} \right)^2}{s_{k-l}^T y_{k-l}} \\
&= \left\| a - \frac{z^T s_{k-l} b}{s_{k-l}^T y_{k-l}} \right\|^2 + \frac{\left(z^T s_{k-l} \right)^2}{s_{k-l}^T y_{k-l}} \\
&\geq 0
\end{aligned}
$$

If the norm above equals zero,

i.e., $\left\| a - \frac{z^T s_{k-l} b}{s_{k-l}{}^T y_{k-l}} \right\| = 0$,

then we have $a = \frac{z^T s_{k-l} b}{s_{k-l}{}^T y_{k-l}}$

or $a \propto b$, which means that $y_{k-l} \propto z$. However, since $s_{k-l}{}^T y_{k-l} > 0$,

$$\frac{\left(z^T s_{k-l}\right)^2}{s_{k-l}{}^T y_{k-l}} > 0$$

as $y_{k-l} \propto z$.

Thus we have proved that $V_{k-l}^T H_0 V_{k-l} + \rho_{k-l} s_{k-l} s_{k-l}^T$ is p.d.. As we have assumed that H_{k+1}^l is p.d. when H_0 is p.d., we conclude that H_{k+1}^{l+1} is p.d..

For the second case $k+1 \le m$,

$$\begin{aligned} H_{k+1} &= V_k^T V_{k-1}^T \cdots V_0^T H_0 V_0 \cdots V_{k-1} V_k \\ &\quad + V_k^T \cdots V_1^T \rho_0 s_0 s_0^T V_1 \cdots V_k \\ &\quad \vdots \\ &\quad + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\ &\quad + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\ &\quad + \rho_k s_k s_k^T . \end{aligned}$$

We also use induction to prove H_{k+1} is p.d..

Firstly, $H_1 = V_0{}^T H_0 V_0 + \rho_0 s_0 s_0^T$

from above, so it is clearly that H_1 is p.d..

Secondly, assume that H_k is p.d.

We are going to prove that H_{k+1} is also p.d. We have assumed

$$\begin{aligned} H_k &= V_{k-1}^T V_{k-2}^T \cdots V_0^T H_0 \\ &\quad \cdot V_0 \cdots V_{k-2} V_{k-1} \\ &\quad + V_{k-1}^T \cdots V_1^T \rho_0 s_0 s_0^T \\ &\quad \cdot V_1 \cdots V_{k-1} \\ &\quad + V_{k-1}^T \cdots V_2^T \rho_1 s_1 s_1^T \\ &\quad \cdot V_2 \cdots V_{k-1} \\ &\quad \vdots \\ &\quad + V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} \\ &\quad + \rho_{k-1} s_{k-1} s_{k-1}^T . \end{aligned}$$

is p.d.. Hence

$$\begin{aligned} H_{k+1} &= V_k^T V_{k-1}^T \cdots V_0^T H_0 V_0 \cdots V_{k-1} V_k \\ &\quad + V_k^T \cdots V_1^T \rho_0 s_0 s_0^T V_1 \cdots V_k \\ &\quad + V_k{}^T V_{k-1}^T \cdots V_2^T \rho_1 s_1 s_1^T V_2 \cdots V_{k-1} V_k \\ &\quad \vdots \\ &\quad + V_k^T V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} V_k \\ &\quad + V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\ &\quad + \rho_k s_k s_k^T \\ &= V_k^T (V_{k-1}^T V_{k-2}^T \cdots V_0^T H_0 \\ &\qquad \cdot V_0 \cdots V_{k-2} V_{k-1} \\ &\quad + V_{k-1}^T \cdots V_1^T \rho_0 s_0 s_0^T V_1 \cdots V_{k-1} \\ &\quad + V_{k-1}^T \cdots V_2^T \rho_1 s_1 s_1^T V_2 \cdots V_{k-1} \\ &\quad + \cdots + V_{k-1}^T \rho_{k-2} s_{k-2} s_{k-2}^T V_{k-1} \\ &\quad + \rho_{k-1} s_{k-1} s_{k-1}^T) V_k + \rho_k s_k s_k^T \\ &= V_k^T H_k V_k + \rho_k s_k s_k^T \end{aligned}$$

From the proof before, we have the following conclusion:

Consider the formula

$B = V_k^T A V_k + \rho_k s_k s_k^T \qquad k = 0,1,\cdots$

(The definition of V_k, ρ_k, s_k are the same as in the L-BFGS formula.) If we know *A* is a p.d. matrix, then *B* is also p.d..

Therefore, we have proved that H_{k+1} is p.d. when $k+1 \le m$.

So we have proved the property for both cases $k+1 \le m$ and $k+1 > m$. □

The proof of the property above is part of our work on the delayed dynamical systems approach for unconstrained optimization. The requirement that $y_i{}^T s_i > 0$ for all i in Property 1 is not a major issue, because we work with a continuous o.d.e. and the numerical method can always be restarted.

References:

[1] Aluffi-Pentini, F., Parisi, V. and Zirilli, F., A differential-equations algorithm for nonlinear equations, ACM Trans. on Math. Software 10 (3), 1984, 299–316.

[2] Aluffi-Pentini, F., Parisi, V. and Zirilli, F. Algorithm 617 DAFNE: A differential equations algorithm for nonlinear equations, ACM Trans. on Math. Software, 10 (3), 1984, 317–324.

[3] Anstreicher, K. M., Linear programming and the Newton barrier flow, Math. Prog. 41, 1988, 367–373.

[4] Arrow, K. J., Hurwicz, L. and Uzawa, H., Studies in Linear and Nonlinear Programming, Stanford University Press, Stanford, CA, 1958.
[5] Barron, A. R., Universal approximation bounds for superpositions of a sigmoidal function, IEEE Trans. Inform. Theory 39 (3), 1993, 930–945.
[6] Boggs, P. T., The solution of nonlinear systems of equations by A-stable integration techniques, SIAM J. Numer. Anal. 8 (4), 1971, 767–785.
[7] Botsaris, C. A. and Jacobson, D. H., A Newton-type curvilinear search method for optimization, JMAA, 54, 1976, 217–229.
[8] Botsaris, C. A., Differential gradient methods, JMAA 63, 1978, 177–198.
[9] Botsaris, C. A., A curvilinear optimization method based upon iterative estimation of the eigensystem of the Hessian matrix, JMAA 63, 1978, 396–411.
[10] Botsaris, C. A., A class of methods for unconstrained minimization based on stable numerical integration techniques, JMAA 63, 1978, 729–749.
[11] Bouzerdoum, A. and Pattison, T. R., Neural network for quadratic optimization with bound constrains, IEEE Trans. Neural Networks 4, 1993, 293–304.
[12] Branin, Jr. F. H.,, Widely convergent method for finding multiple solutions of simultaneous nonlinear equations, IBM Journal of Research and Development 16, 1972, 504–522.
[13] Brown, A. A. and Bartholomew-Biggs, M. C., Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations, JOTA 62 (2), 1989, 211–224.
[14] Brown, A. A. and Bartholomew-Biggs, M. C., ODE versus SQP methods for constrained optimization, JOTA 62 (3), 1989, 371–386.
[15] Chen, Y. H. and Fang, S. C., Solving convex programming problem with equality constraints by neural networks, Computers Math. Appl. 36, 1998, 41–68.
[16] Chu, M. T., On the continuous realization of iterative processes, SIAM Review 30 (3), 1988, 375–387.
[17] Cichocki, A. and Unbehauen, R., Neural Networks for Optimization and Signal Processing. Wiley, Chichester. 1993.
[18] Cichocki, A., Unbehauen, R., Weinzierl, K. and Holzel, R.,, A new neural network for solving linear programming problems, European J. Operational Res., 93, 1996, 244–256.
[19] Chua, L. O. and Lin, G. N., Nonlinear programming without computation, IEEE Trans. Circuits Syst., 31, 1984, 182–188.
[20] Cybenko, G., Approximation by superpositions of a sigmoidal function, Math. Control Signals Systems, 2, 1989, 303–314.
[21] Evtushenko, Yu. G. and Zhadan, V. G., A relaxation method for solving problems of non-linear programming, U.S.S.R. Comput. Math. Math. Phys. 17 (4), 1978, 73–87.
[22] Diener, I. and Schaback, R.,, An extended continuous Newton method, JOTA 67 (1), 1990,57–77.
[23] Glazos, M. P., Hui, S. and ˙ Zak, S.,, Sliding modes in solving convex programming problems, SIAM J. Control Optim. 36, 1998, 680–697.
[24] Goldstein, A. A., Convex programming in Hilbert space, Bulletin of American Mathematical Society 70, 1964, 709–710.
[25] Han, Q., Liao, L.-Z., Qi, H. and Qi, L., Stability analysis of gradient-based neural networks for optimization problems, J. Global Optim. 19 (4), 1962, 363–381.
[26] Hassan, N. and Rzymowski, W., An ordinary differential equation in nonlinear programming, Nonlinear Analysis, Theory, Method & Applications 15 (7), 1990, 597–599.
[27] Haykin, S. S., Neural Networks: A Comprehensive Foundation, Prentice-Hall, Englewood Cliffs, NJ, 1994.
[28] He, B. S., Solving a class of linear projection equations, Numerische Mathematik 68, 1994, 71–80.
[29] He, B. S., A class of projection and contraction methods for monotone variational inequalities, Applied Mathematics and Optimization 35, 1997, 69–76.
[30] He, B. S., Inexact implicit methods for monotone general variational inequalities, Mathematical Programming, 86 (1), 1999, 199–217.
[31] He, B. S. and Yang H., A neural network model for monotone linear asymmetric variational inequalities, IEEE Trans. Neural Networks, 11, 2000, 3–16.
[32] Hopfield, J. J.,, Neural networks and physical systems with emergent collective computational ability, Proc. Natl. Acad. Sci. USA, 79, 1982, 2554–2558.
[33] Hopfield, J. J., Neurons with graded response have collective computational properties like those of two-state neurons, Proc. Natl. Acad. Sci., 81, 1984, 3088–3092.
[34] Hopfield, J. J. and Tank, D. W., Neural computation of decisions in optimization problems, Biolog. Cybernetics, 52, 1985, 141–152.
[35] Hornik, K., Approximation capabilities of multilayer feedforward networks, Neural Networks, 4, 1991, 251–257.
[36] Hornik, K., Stinchcombe, M. and White, H., Multilayer feedforward networks are universal approximators, Neural Networks, 2, 1989, 359–366.
[37] Hou, Z.-G., Wu, C.-P. and Bao, P., A neural network for hierarchical optimization of nonlinear large-scale systems, International Journal of Systems Science 29 (2), 1998, 159–166.
[38] Incerti, S., Parisi, V. and Zirilli, F., A new method for solving nonlinear simultaneous equations, SIAM J. Numer. Anal. 16, 1979, 779–789.
[39] Jorge Nocedal, Updating Quasi-Newton Matrices With Limited Storage, Mathematics of Computation, Vol 35, No 151, July 1980, pp. 773-782.
[40] Kennedy, M. P. and Chua, L. O., Neural networks for nonlinear programming, IEEE Trans. Circuits Syst. 35, 1988, 554–562.
[41] Liao, L.-Z. and Qi, H., A neural network for the linear complementarity problem, Math. Comput. Modelling 29 (3), 1999, 9–18.
[42] L.Z. Liao, L.Q. Qi, and H.W. Tam, A gradient-based continuous method for large-scale optimization problems, *Journal of Global Optimization*, Vol 31 , Apr 2005, pp. 271-286.
[43] Liao, L.-Z., Qi, H. and Qi, L., Solving nonlinear complementarity problems with neural networks: a reformulation method approach, JCAM 131 (12), 2001, 343–359.
[44] Lillo, W. E., Loh, M. H., Hui S. and Zak, S. H., On solving constrained optimization problems with neural networks: a penalty method approach, IEEE Trans. Neural Networks, 4, 1993, 931–940.

[45] LI-ZHI LIAO, HOUDUO QI and LIQUN QI, (2004), Neurodynamical Optimization, Journal of Global Optimization 28: 175-195.
[46] Maa, C. Y. and Shanblatt, M. A., Linear and quadratic programming neural network analysis, IEEE Trans. Neural Networks 3, 1992, 580–594.
[47] Maa, C. Y. and Shanblatt, M. A., A two-phase optimization neural network, IEEE Trans. Neural Networks 3, 1992, 1003–1009.
[48] Mangasarian, O. L., Mathematical programming in neural networks, ORSA J. Comput. 5 (4), 1993, 349–360.
[49] More, J. J., Garbow, B. S. and Hillstrom, K. E., Testing unconstrained optimization software, ACM Trans. Math. Software 7 (1), 1981, 17–41.
[50] Novakovi´c, Z. R., Solving systems of non-linear equations using the Lyapunov direct method, Computers Math. Applic. 20 (12), 1990, 19–23.
[51] Pan, P.-Q., New ODE methods for equality constrained optimization (1) – equations, JCM 10 (1), 1992, 77–92.
[52] Pan, P.-Q., New ODE methods for equality constrained optimization (2) – algorithms, JCM 10 (2), 1992, 129–146.
[53] Polyak, B. T., Constrained minimization problems, USSR Computational Mathematics and Mathematical Physics 6, 1966, 1–50.
[54] Rodriguez-Vazquez, A., Dominguez-Castro, R., Rueda, A., Huertas J. L. and Sanchez-Sinencio, E., Nonlinear switch-capacitor 'neural' networks for optimization problems, IEEE Trans. Circuits Syst. 37, 1990, 384–398.
[55] Schaffler, S. and Warsitz, H., A trajectory-following method for unconstrained optimization, JOTA 67 (1), 1990, 133–140.
[56] Schnabel, R. B. and Eskow, E., A new modified Cholesky factorization, SIAM J. Sci. Stat. Comput. 11, 1990, 1136–1158.
[57] Slotine, J.-J. E. and Li, W., Applied Nonlinear Control, Prentice-Hall, Englewood Cliffs, NJ, 1991.
[58] Snyman, J. A., A new and dynamic method for unconstrained minimization, Appl. Math. Modelling 6, 1982, 449–462.
[59] Solodov, M. V. and Tseng, P., Modified projection-type methods for monotone variational inequalities, SIAM J. Control and Optimization 34, 1996, 1814–1830.
[60] Sudharsanan, S. and Sundareshan, M., Exponential stability and a systematic synthesis of a neural network for quadratic minimization, Neural Networks 4, 1991, 599–613.
[61] Tanabe, K., A geometric method in nonlinear programming, JOTA 30 (2), 1980, 181–210.
[62] Tank, D. W. and Hopfield, J. J., Simple neural optimization networks: An A/D convert, signal decision circuit, and a linear programming circuit, IEEE Trans. Circuits Syst. 33, 1986, 533–541.
[63] Teo, K. L., Wong, K. H. and Yan, W. Y., Gradient-flow approach for computing a nonlinear-quadratic optimal-output feedback gain matrix, JOTA 85, 1995, 75–96.
[64] Vincent, T. L., Goh, B. S. and Teo, K. L., Trajectory-following algorithms for min-max optimization problems, JOTA 75, 1992, 501–519.
[65] Wilde, N. G., A note on a differential equation approach to nonlinear programming, Management Science 15 (11), 1969, 739–739.
[66] Williems, J. L., Stability Theory of Dynamical Systems, Nelson, 1970.
[67] Wu, X., Xia, Y., Li, J. and Chen, W. K., A high performance neural network for solving linear and quadratic programming problems, IEEE Trans. Neural Networks, 7, 1996, 643–651.
[68] Xia, Y., A new neural network for solving linear programming problems and its applications, IEEE Trans. Neural Networks 7, 1996, 525–529.
[69] Xia, Y., A new neural network for solving linear and quadratic programming problems, IEEE Trans. Neural Networks 7, 1996, 1544–1547.
[70] Xia, Y. and Wang, J., A general methodology for designing globally convergent optimization neural networks, IEEE Trans. Neural Networks 9, 1998, 1331–1343.
[71] Yamashita, H., A differential equation approach to nonlinear programming, Math. Prog. 18, 1980, 155-168.
[72] Zabcayk, J ., Mathematical Control Theory: An Introduction, Birkhauser, Boston, 1992.
[73] Zak, S. H., Upatising, V. and Hui, S., Solving linear programming problems with neural networks: a comparative study, IEEE Trans. Neural Networks 6, 1995, 94-104.

Transaction Support on Flash Devices

Sai Tung ON

Abstract

Thanks to its superiority such as fast data access, low power consumption, high shock resistance, small dimensions and light weight, NAN Flash has become more and more popular in mobile computing devices and embeded systems. The out-of-place update nature of NAND Flash brings us an opportunity to propose new transaction recovery technique for flash-based database systems. In this paper, we propose a novel commit protocol called flag commit *which achieves better performance and lower space overhead over conventional protocols by eliminating the need of writing commit records. Moreover,* flag commit *incurs little overhead to the garbage collection algorithm and preserves efficient wear-leveling performance. We also show that the basic design of flag commit can be extended to support the no-force buffer management policy and a more fine-grained concurrent control mechanism.*

1 Introduction

NAND Flash has become the preferred storage alternative for mobile computing devices and embedded systems due to its superiority such as fast data access, low power consumption, high shock resistance, small dimensions and light weight. With the capacity rapidly increasing and the cost per bit continuously decreasing, NAND-based flash device is emerging to be an ideal replacement for traditional magnetic disks in personal computer and enterprise server domain [15].

Despite of the advantages mentioned above, compared with traditional magnetic disks, NAND Flash has a number of disadvantages. Firstly, write operations in NAND Flash can only clear bits (change their value from 1 to 0). The only way to set bits (change their value from 0 to 1) is to erase an entire region memory (called *erase units*) — a block consisting of a number of pages [14]. This is called *erase-before-write* constraint. Secondly, in the NAND Flash, the read and write operations take place on a page basis (typically, 528 bytes at a time for small-block NAND flash devices). Each page can be modified only a small number of times without performing an erase operation [6, 4, 5, 14]. Thirdly, the page write cost is much more expensive than the read cost, while the erase requirement makes write cost even higher. Fourthly, with no mechanical latency, NAND Flash has uniform random/sequential read access cost. Finally, each erase unit can bear a limited number of erase cycles (typically, 10,000$\sim$100,000 times), after which it will be worn out and becomes unreliable. Table 1 summarizes the features of a Samsung flash memory chip [5].

Table 1. Features of a Typical NAND Flash [5]

Parameter	Value	Unit
page size	512+16	Byte
block size	16K+512	Byte
page read time	15	μs
page write time	200	μs
block erase time	2	ms
partial program cycles in the same page	2	cycle

To avoid performance degradation caused by the aforementioned erase-before-write limitation, NAND-based flash devices adopt the out-place-update strategy to trade expensive in-place updates (along with the erase they incur) for cheaper writes onto free flash pages. The basic mechanism is to allocate a new free page, write the updated data onto the new page, and then, mark the original page as obsolete (they will later be reclaimed by garbage collection) [20, 14]. As a result, out-of-date versions and the latest copy of data might co-exist over flash memory simultaneously. Such a feature gives us a new opportunity to redesign database transaction recovery algorithms for NAND-based flash devices.

In the transaction-oriented database management system (DBMS), transaction recovery support is to enforce the atomicity and durability of each transaction [17]. That is, for each transaction, either all or none of write operations are performed, and once committed, the data written by the transaction should be made durable, even if there are system failures. Among various existing disk-specific recovery

schemes, write ahead log (WAL) [18, 26] and shadow paging [16, 27] are predominant. The central concept of WAL is that updates to data pages can be written only after they have been logged, i.e., when log records have been forced to stable storage. There are two drawbacks when directly applying WAL upon NAND-based flash devices. First, in consideration of the expensive write cost on flash, the overhead of transaction rollback & recovery is high (due to the requirement of explicit undo/redo operations). Second, as the write granularity of NAND Flash is a page, frequently writting log records is not only time-consuming but also space-consuming. Although group commit [13] can alleviate such log bottleneck, it does not improve the response time of individual transactions. Shadow paging is a recovery technique which aims to eliminate the need of writing logs. It bases on the out-place-update pattern and maintains two tables of page addresses (i.e., the current page table and the shadow page table) during the execution of any transaction. To commit/abort a transaction is done by switching between these two tables. Shadow paging has serveral drawbacks when running upon the traditional magnetic disk. First, the commitment of a single transaction under shadow paging involves outputting the updated data, the current page table and its disk address, while WAL-based schemes only need to output log records. Second, shadow paging causes database pages to change locations (therefore, no longer contiguous), turning most sequential read accesses into random ones (which are extremely expensive on magnetic disks). Third, garbage collection is required to reclaim space occupied by obsolete data, thereby incurring additional overhead and complexity. Fourth, it is hard to support fine-grained concurrency control mechanism (e.g., in record level). As previously described, out-place-update and uniform random/sequential read access cost are among the key features of NAND Flash. Therefore, only the first and the last ones are the additional problems brought by applying shadow paging upon NAND-based flash devices. Furthermore, as the current page table and the shadow page table are always persistent in NAND Flash (in the form of inverse mapping), the first problem can be simplified as how to choose the correct table for each transaction (i.e., to judge whether a transaction is committed).

For flash-specific recovery schemes, Lee and Moon [22] proposed a log-based scheme called in-page logging for flash-based DBMSs, while Prabhakaran et al. [29] developed a shadow paging-based scheme called cyclic commit for flash-based file systems. Prabhakaran's work mainly focused on reducing the overhead of committing transactions (see the first problem of shadow paging). Unlike the standard commit protocol which writes a commit record as a flag for each committed transaction, cyclic commit stores a link to the next updated page and creates a cycle among the pages updated by the same transaction. Hence, as whether a transaction is committed or not can be judged by the existence of the cycle, it successfully eliminates the need for a separate commit record for each transaction. However, in order to preserve these cycles, the garbage collection algorithm becomes very complicated and might potentially impair the wear-leveling performance. Moreover, when it is extended to support DBMSs, the cyclic commit protocol is required to hold the last page in the buffer before a transaction commits, which unavoidably leads to a transaction response time.

In this paper, based on shadow paging, we proposed a *flag commit* protocol for NAND-based flash devices to support transaction recovery in database systems. Compared to the cyclic commit, flag commit has a simpler garbage collection algorithm and achieves better response time and wear-leveling performance. In our flag commit protocol, the state of a transaction is indicated by the flag metadata stored along with the pages this transaction updates. To summarize, our contributions in this paper are as follows:

- We propose two designs of the flag commit protocol: commit-based flag commit and abort-based flag commit. The former suits for cases when the abort ratio is high, while the latter performs better when the abort ratio is low.
- We conduct a cost analysis on the proposed protocol and show it a better protocol over the existing protocols in terms of both performance and space overhead.
- To further enhance the transaction response time, we combine flag commit protocol with a logging scheme to support the no-force buffer management policy.
- We also extend the flag commit protocol to support a more fine-grained concurrency control mechanism.

The rest of the paper is organized as follows. Section 2 introduces the background of our research. We gives the system architecture in Section 3. Section 4 presents two designs of the flag commit protocol and their cost analysis. Section 5 develops the extensions of the flag commit protocol. In section 6, we review the related work on flash data management and transaction processing. Finally, we disuss the future work and conclude the paper in Section 7.

2 Background

In this section, we briefly discribe the key characteristics of NAND flash memory and the flash translation layer which emulates it as traditional magnetic disk.

2.1 NAND Flash Memory Characteristics

Like magnetic disk drives, NAND[1] flash memory is non-volatile and retains its contents even when the power is turned off. A NAND flash memory chip is organized in many blocks and each block is composed of a fixed number of pages. The typical block size and page size are (16K+512)bytes and (512+16)bytes, respectively. Each page consists of a data area and a spare area. For every page with 512 bytes data area, there is a corresponding 16 bytes spare area for storing metadata such as the error correction code (ECC) and logical block address (LBA). A block is the smallest unit for erase operations, while all read and write operations are at the page granularity.

There are two types of NAND flash memory: single-level cell (SLC) NAND and multi-level (MLC) NAND. While each flash memory cell stores 1 bit in SLC chips, MLC chips can store 2 or more bits in each cell. SLC NAND chips allow for partial page programming (i.e., each page in SLC can be partially programmed several times without performing an erase operation), while MLC NAND chips do not. In general, MLC NAND chips cost less and allow for higher storage density, and therefore they tend to be used in low-cost consumer applications, including media players, MP3 devices, media cards, and USB flash drives. Meanwhile, SLC Flash devices provide faster write performance and greater reliability, and hence are desirable for professional products and solid state drives (SSDs) [2, 8]. As performance and durability is essential for DBMS transaction applications, in this paper, we focus on SLC NAND flash memory. Hereafter, we use the term flash memory to refer to SLC NAND flash memory.

Flash memory exhibits a number of characteristics which distinguishes itself from magnetic disk drives: (1) asymmetric read/write cost, (2) the erase-before-write constraint, (3) the endurance issue — each block can be erased for only a limited number of cycles before it is worn out, and (4) no mechanical latency.

2.2 Flash Translation Layer

Flash devices usually access the embedded flash memory chips through a software layer called FTL (Flash Translation Layer) [20, 3]. The FTL provides a disk-like interface, which includes the capability to read and write a page directly without caring about the special characteristics of flash memory. The core data structures of FTL are two address mappings between blocks, represented by their logical block addresses (LBAs) and physical pages. A *direct* mapping from LBAs to physical pages is stored in RAM to speed up reads, and an *inverse* mapping is stored on flash, to re-build the direct mapping during boot time [24]. There are several mapping strategies such as page mapping [20], block mapping [7] and hybrid mapping [28]. By maintaining these mappings, the FTL can support out-of-place updates: when a logical page is updated, the FTL writes the data to a new physical page and updates the mappings. As a result, the erase-before-write constraint is overcome and expensive in-place updates are avoided.

As each out-of-place update leaves an obsolete flash page, gradually such pages will accumulate and use up the free space of flash memory. The FTL maintains a list of free blocks and has a *garbage collection* routine to reclaim those obsolete pages. Specifically, the FTL selects a block, moves valid pages to some free block, erases it and puts it to the free block list. To length the lifetime of flash memory, *wear-leveling* technique which uniformly distributes writes/erases across the entire storage space is adopted.

As discussed in Section 1, the out-of-place update and the fast random access natures of flash memory make it a perfect storage media to apply the shadow paging-based techniques when processing DBMS transaction recovery. In what follows, we will propose a novel transaction supporting system which eliminates the drawbacks of traditional shadow paging and achieve a competitive performance on various aspects such as transaction response time, transaction throughput and space overhead.

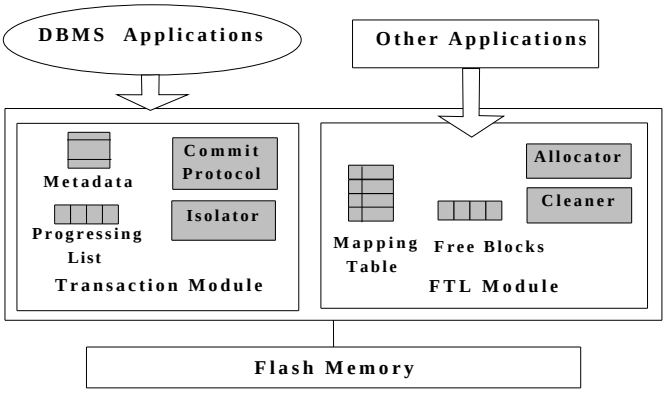


Figure 1. System Architecture

3 System Architecture

In this section, we describe the architecture of our system model. As shown in Figure 1, besides the well-known FTL module, we propose to add a transaction module in the firmware of flash-based devices. For non-transactional applications, they continue to access the embedded flash purely via the FTL. For DBMS transactional applications,

[1]There are two major architectures in flash memory design: NAND Flash and NOR Flash. NOR Flash has a faster random access speed but a lower storage capacity so it is preferred for code storage. NAND Flash has a denser architecture and is specially designed for data storage.

they rely on both modules to ensure the atomicity, isolation, and recovery of the transactions.

The transaction module consists of two major components: the commit protocol and the isolator. The former is the core of the whole system. It includes a commit policy which is executed when a transaction is committed, an abort policy which is excuted when a transaction is aborted and a recovery policy which is executed when system reboots. The last policy is actually a mechanism which rebuilds the correct mapping information by scanning all of the spare area of pages on the flash memory. The latter is to provide isolation among multiple transactions — when a translation is in progress, the isolator ensures that no conflicting reads/writes (i.e., writing/reading the same page updated by a progressing transaction) are issued. The FTL module includes an *allocator* and a *cleaner*. The allocator handles any translation of logical block addresses (LBAs) and their physical block addresses (PBA). The cleaner is to do garbage collection to reclaim pages of invalid data (see Section 2.2 for details). Note that the introduction of the transaction module will not harm the performance of the FTL module (e.g., the garbage collection policies which the cleaner bases on can still work well as usual in the proposed system model).

In addition to the in-memory data structures such as the direct mapping table and the free blocks (please refer to Section 2.2), the system should also maintain a list of progressing transactions. Once a transaction is committed, in the mapping table, the corresponding entries of those pages in this transaction are updated. On the other hand, when a transaction is aborted, no update on the mapping table is required.

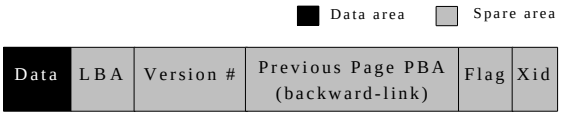


Figure 2. Page Format

4 Commit Protocols

In this section, we present the commit protocols our system adopts. The purpose of the commit protocol is to enforce the atomicity of transactions. Specifically, the protocol specifies the steps involved when committing/aborting a transaction, as well as a recovery procedure. The design principle of the proposed commit protocols is to eliminate the need for a separate commit record for each committed transaction (thereby reducing the space and performance overhead), while still preserving a simple and efficient garbage collection mechanism.

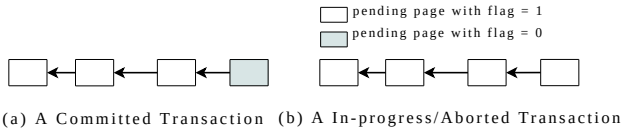


Figure 3. CFC Examples

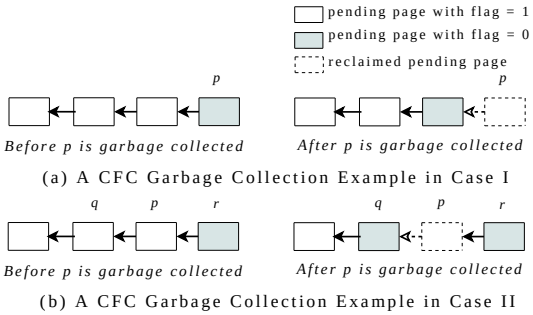


Figure 4. CFC Garbage Collection Examples

4.1 Commit-based Flag Commit(CFC)

In our proposed system, a transaction writes new data for each page to the flash devices as a *pending page*, which contains a data portion and a metadata portion. Figure 2 shows the format of the pending page, where LBA and version # represent the identity of the page, *xid* stores the transaction ID which the pending page belongs to, *backward-link* contains the physical address of the previous page of the same transaction, and *flag* is a bit value indicating whether the transaction is committed or not. As such, by storing a link in the metadata portion, the pending pages of the same transaction are chained together with backward pointers. In Commit-base Flag Commit(CFC), we say a transaction is committed if at least one of the pending pages on its linked list have the flag bit value equal to 0. Otherwise, this transaction is in processing or aborted. Any pending page belonging to a committed/aborted transaction is committed/uncommitted.

CFC Commit Policy. When a DBMS transaction T is in progress, its associated pending pages $\{p_1, \cdots, p_n\}$ are written to flash storage dynamically. For the first pending page p_1, it is written with a backward-link setting to null, while for succedent pages p_i, they are written with a link pointing to their previous page (i.e., $p_i.previous = p_{i-1}$). The xid is set as the $T.id$ and the version number is set based on how many time its logical page has been updated/committed. Initially, the value of the flag bit in each pending page is set to 1. Once the transaction is committed, the flag bit in the last pending page is clear (i.e., set to 0) to indicate that it has been committed. Meanwhile, the direct mapping table is updated for all pages in the transaction.

We call the flag bit in each committed transaction as *commit flag*. Note that we can clear bits on the written page without erasing is due to the nature of NAND SLC flash memory — it allows several cycles of (typically, 2∼8) consecutive partial page programming operation within the same page without an intervening erase operation. Figure 3(a) shows an example of a committed transaction.

CFC Abort Policy. The policy to abort a transaction is very simple in the CFC protocol. When a transaction is aborted, no extra actions are required. This implies the abort overhead of the CFC protocol is extremely low. Figure 3(b) shows an example of an aborted/in-progress transaction.

CFC Garbage Collection. When free space on flash memory becomes low, the cleaner is triggered to reclaim space occupied by obsolete data. With CFC protocol, any committed page can be garbage collected as long as a newer version of the same logical page is committed. For uncommitted pages, they can be garbage collected at any time. In order to keep the commit flag for committed transactions, before erasing a committed page p, the cleaner should have extra actions based on different cases of p.

I. $p.flag = 0$: In the first case, as p holds the commit flag for its associated transaction, before erasing p, the cleaner need to move this flag to p's previous page (by clearing the flag bit on the page located by $p.previous$). Such a mechanism ensure that the transaction consistently possesses a commit flag before and after the erasure of p. Figure 4(a) illustrates the garbage collection in case I.

II. $p.flag = 1$: In the second case, once p is garbage collected, the linked list of its associated transaction would break in two. Over time the original transaction may be represented by a number of linked lists of pending pages. Whenever such p is going to be reclaimed, the cleaner manages to clear the flag bit of p's previous page to guarantee that each such linked list owns a commit flag. The benefit is two-folded. First, each small linked list can be treated as a separate transaction, and therefore the recovery procedure can be simplified. Second, it can guarantee there always exist commit flags for each committed transaction. Consider the garbage collection example of this case shown in Figure 4(b), if the cleaner does not clear the flag bit on q before erasing p, then later if r is garbage collected as well, there is not any commit flag for this committed transaction and therefore would lead to misclassification of the transaction. Enforcing a commit flag for each linking piece can sucessfully avoid the above case.

CFC Recovery Policy. When system reboots, a recovery procedure should be conducted to recover the last committed version for each page. The target is actually to rebuild the correct direct mapping table by scanning the spare area of physical pages. During recovery, guided by the recovery policy, CFC classifies the pending pages as committed or uncommitted and identifies the lastest committed version for each page. The detail of the recovery algorithm is given below:

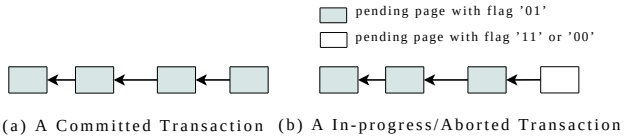


Figure 5. AFC Examples

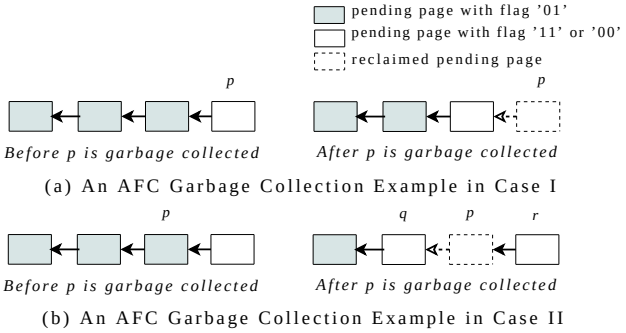


Figure 6. AFC Garbage Collection Example

We use S to refer to the set of pending pages and R for the set of pending pages that are referenced by the backward-link field of any pending pages in S. Both S and R can be obtained by scanning the spare area of physical pages. Obviously, $P \ominus R$ (denoted by U) is the set of pending pages on the tail of the transaction linked lists ($\ominus$ means set difference). From the proposed commit policy and garbage collection mechanism, it is easy to know that whether a transactions is committed can be judged simply by the flag of its corresponding pending page in U. Hence, we can divide U into two sets: U^+ for committed ones (whose flag is 0) and U^- for uncommitted ones (whose flag is 1). Start from each committed page in U^+, by following the backward-links to collect other committed pages, we can get the whole set of committed pages denoted by S^+. Mathematically, it can be calculated by:

$$S^+ = U^+ \cup \{p \in S | \exists u \in U^+, p \text{ is reachable from } u\},$$

where "p is reachable from u" means there is a path consisting of backward-links from u to p.

After the set of committed pages is obtained, the remaining recovery is quite straightforward. For each logical page, the recovery algorithm identifies the comitted page with the largest version number and updates the direct mapping table accordingly. As each flash page is visited at most twice during the entire recovery procedure, the time complexity of the CFC recovery is $O(n)$, where n is the number of flash pages.

4.2 Abort-based Flag Commit(AFC)

The Abort-based Flag Commit(AFC) adopts the same page format (see Figure 2) as the CFC protocol for those pending pages. The only difference is that, in AFC, the number of bits on the *flag* component is not one but two. While CFC relies on the commit flag to label committed transactions, AFC takes another approach: we say a transaction is committed if and only if the flag of all of its pending pages are '01'.

AFC Commit Policy. Same as the CFC protocol, pending pages of the same transaction in the AFC are also chained together. Initially, for each transaction, the flag of all pending pages except the tail are set to '01'. The flag of the pending page on the tail is set to '11' to indicate the transaction is still in progress. We call this flag with values '11' or '00' as *abort flag*, based on which we can simply judge the status of a transaction during the recovery procedure. Once the transaction is committed, the abort flag is clear (i.e., set to '01') to reflect the commitment. Meanwhile, the direct mapping table is updated accordingly. Figure 5(a) shows an example of a committed transaction.

AFC Abort Policy. Same as the CFC protocol, under the AFC protocol, when a transaction is aborted, no extra actions are required. Figure 5(b) shows an example of an aborted/in-progress transaction.

AFC Garbage Collection. With the AFC protocol, any committed page can be reclaimed freely as long as a newer version of the same logical page is committed, and no additional action is needed. For uncommitted pages, in order to keep their abort flag, additional actions are required. In detail, before an uncommitted page p is erased, similar as the CFC protocol, the actual actions are based on different cases of p.

I. $p.flag =$ '11' or '00' : In the first case, as p holds the abort flag for its associated transaction, before erasing p, the cleaner need to move this flag to p's previous page (by changing the flag bits from '01' to '00'). By enforcing such a mechanism, the aborted transaction is guaranteed to hold an abort flag before and after the erasure of p. Figure 6(a) illustrates the garbage collection in case I.

II. $p.flag =$ '01' : In the second case, same as discussed in CFC, once p is garbage collected, the corresponding linked list of its associated transaction would break in two. Based on the same rationales given in the CFC garbage collection (see Section 4.1), whenever such p is going to be reclaimed, the cleaner should set the flag bits of p's previous page to '00'. The purpose is to guarantee that each linked list owns an abort flag. Figure 6(b) illustrates the garbage collection in case II.

AFC Recovery Policy. The AFC protocol shares almost the same recovery policy with the CFC protocol. The only difference is how they tell apart committed&uncommitted pages in U. While CFC identifies the ones with the commit flag (i.e., '0') as committed pages, AFC determines the ones without the abort flag (i.e., '11' or '00') as committed pages. Due to space limitations, we omit it here. Please refer to Section 4.1 for details.

4.3 Cost Analysis: CFC vs. AFC

In this subsection, we will analyze the I/O cost of the CFC and the AFC protocols. As there are no overhead for aborting transactions in the proposed system, we focus on the cost of committing transaction. Before going further, for easy presentation, we give some assumptions and define a few notations. We assume the FTL module adopts the page-based mapping strategy, each time only one transaction is executing and the system adopts the force approach as the buffer management policy (i.e., all the pending pages written by a committing transaction have to be forced out to flash pages). Let δ denote the transaction abort ratio, l denote the average number of pending pages in a transaction, and n denote the average number of blocks to be reclaimed by the cleaner when allocating a fresh flash page. We use k to represent the average number of pending pages in a block, and use C_r, C_w, C_{bit} and C_e to represent the costs of one page reading, one page writing, clearing one bit and reclaiming one block, respectively.

With those assumptions and definitions, we can roughly estimate the average I/O cost when committing a transaction with the CFC protocol as:

$$C = l * C_w + C_{bit} + n * l * k * (1 - \delta) * C_{bit} + n * l * C_e,$$

where the first two items are the costs of outputing pending pages and setting the commit flag, respectively, the third item is the cost of extra actions when garbage collecting committed pages, and the last item is the cost of reclaiming blocks.

Similarly, we estimate the average I/O cost for the AFC protocol as:

$$C' = l * C_w + C_{bit} + n * l * k * \delta * C_{bit} + n * l * C_e.$$

Based on the above two equations, we can easily derive that AFC outperforms CFC when δ is below 0.5.

Compared with the traditional log-based approaches [18, 26, 22], CFC and AFC protocols eliminate the need of writing logs for each committed record and thus have advantages on both performance and space overhead. For existing shadow paging-based approaches such like Cyclic Commit [29], as they require not only a complex garbage collection policy (which is against conventional wisdom to uniformly distribute erases among all pages) but also a special memory management policy to buffer a page for each committing transaction, the proposed protocols are expected to outperform them in terms of wear-leveling performance, garbage collection overhead and transaction response time.

5 Extensions of Basic Commit Protocols

In this section, we discuss how the basic commit protocols can be augmented to support a more sophisticated buffer management policy and a more fine-grained concurrent control mechanism.

5.1 Supporting the No-Force Policy

To achieve better response time, most database systems adopt a *no-force* policy (i.e., once a transaction is committed, only the redo logs are forced to the stable storage) [18].In order to support this no-force buffer management policy, we combine the proposed protocols with the redo logging scheme.

We propose to record all the update activities in the database in the form of *log record*, whose format is represented by $\langle xid, rid, v_1, v_2 \rangle$, where xid is the unique identifier of the transaction that performed the write operation, rid is the unique identifier of the data item written, v_1 and v_2 are the values of the data item before and after the write. Initially, these log records are buffered in the memory. Later, when a pending page is swapped out from the memory and forced back to flash storage, its corresponding log records can be removed. This is because the update information can be reflected automatically by the pending pages of the new & old versions.

Commit Policy. When a transaction is committed, if some of its pending pages are still buffered in memory, we write a commit log record in the log buffer and then forces all the transaction's log records to flash storage (with the commit record being the last such record), so that the system is able to redo this transaction during the recovery. Otherwise, we follow the CFC/AFC commit policy to do the commitment.

Abort Policy. When a transaction aborts, the only thing need to do is to remove all of its log records from the log buffer.

System Checkpointing. To speed up the recovery procedure when system reboots, the system periodically performs checkpointing, which involves the following steps:

1. Stop to accept new transactions and force all pending pages in the memory buffer to flash storage (consequentially, their corresponding log records are removed from the log buffer).

2. Write a special log record $\langle checkpoint \rangle$ to indicate the end of the checkpointing and resume to accept new transactions.

After the checkpointing, the log records appear in flash storage before the $\langle checkpoint \rangle$ can be ignored, and can be garbage collected whenever desired. This is because all modifications recorded by these log records must have been written to the database, and consequently, during recovery, there is no need to perform redo operations for these log records.

Garbage Collection. The garbage collection procedure is exactly the same as that presented in the CFC/AFC protocol (see Section 4.1).

Recovery Policy. When system reboots, in addition to re-build the direct mapping table by following the CFC/AFC recovery policy, the system needs to perform redo operations according to the log records which appear after the most recent checkpointing.

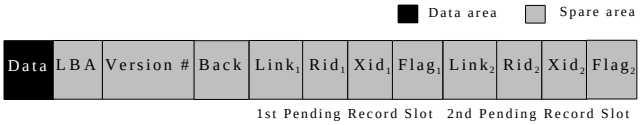


Figure 7. Extended Page Format

5.2 Supporting Fine-grained Concurrent Control

Currently, our system can only support the concurrent control on a granularity of page. In this subsection, we will extend the basic commit protocols to enable the concurrent control in the record level. In what follows, we present the extension based on the CFC protocol. The scheme to extend the AFC protocol is similar and is thus omitted from this paper.

Under the new concurrent control mechanism, multiple in-progress transaction are allowed to update the same page concurrently. Each of them write new data to the flash page as a *pending record*, which is represented by $\langle data, xid, rid, link, flag \rangle$, where data and rid are the value and identifier of the record respectively, xid is the transaction identifier, link stores LBA of its previous pending record of the same transaction, and *flag* stores a bit value indicating the status of the transaction. Specifically, the data portion of the pending records are placed on the data area of its residing page, while remaining portions are placed on the spare area. It is necessary to know that the maximum number of pending records allowed in each page depends on the size of the spare area of flash pages (typically, 16 bytes for small-block flash and 64 bytes for large-block flash). Same as discussed in CFC protocol, whether a transaction is committed depends on whether it owns a commit flag. Over time, the lastest committed records of a logical page might scatter over multiple physical pages of different version numbers. The current page content should be computed by merging those versions together. To facilitate the merging process, we additionally add a backward link *back* to the page metadata. That is, before writing a new version p_k of a logical page p, in addition to the identity of the page

and the metadata of those pending records, a pointer to the last version of p is also stored (i.e., $p_k.back = p_{k-1}$). Figure 7 shows the extended page format which contains slots for up to two pending records.

Commit and Abort Policies. The policy to commit/abort a transaction is similar to that introduced in the basic commit protocols. That is, once a transaction is committed, we label the transaction with a commit flag by clearing the flag bit on its last pending record. When a transaction is aborted, no extra actions are needed. Because every single commitment of a transaction requires clearing a bit on the flash page, the maximum number of transactions allowed to update the same page concurrently cannot exceed the maximum cycle of consecutive partial page programming operations within the same flash page.

Garbage Collection. For a logical page p, there are a set of physical pages $\{p_1, \cdots, p_n\}$ with the same LBA but different version number. Under the extended protocol, any page p_k can be reclaimed if and only if there exists a page $p_i (i > k)$ which satisfies the condition that any committed pending record in p_k is already committed in p_i. Before erasing such p_k, same as the basic protocols, the cleaner needs to conduct extra actions to ensure that every committed transaction on p_k consistently possesses a commit flag before and after the erasure. We omit the details of these actions as they are already described in Section 4.1.

Recovery Policy. As the latest committed records of a logical page might scatter over multiple flash pages of the same LBA, when system reboots, a recovery procedure only need to identify the last version of page having at least one committed pending record. The detailed process is similar to the recovery of the basic commit protocols. Please refer to Section 4.1 for details.

6 Related Work

Data management on flash-based media has received much attention from research community in recent years. To enable a quick deployment of flash-memory technology, early work attempted to hide the unique characteristics of flash memory. They focused on simulating traditional magnetic disks by flash memory chips. Kawaguchi et al. [20] proposed a software module called flash translation layer (FTL) to transparently access flash memory, so that conventional disk-based algorithms and access methods can work as usual. To overcome the erase-before-write constraint, an out-of-place update scheme was adopted and various garbage collection mechanisms [11, 20, 19] were proposed to reclaim invalidated space. To lengthen the lifetime of flash memory, wear-leveling algorithms that attempted to evenly distribute writes/erases across all pages were developed in [10, 12, 25]. Besides these fundamental achievements, recent work shifted to exploit the characteristics of flash memory to enhance the performance of file systems and DBMSs. In view of the slow write speed on flash memory, the log structure was adopted to reduce the number of write operations. Along this direction, some flash-aware log-based file systems like YAFFS [1] and JFFS [30] were proposed. For DBMSs running on flash-based media, Lee and Moon [22] presented a novel design of data logging called in-page logging (IPL) to further improve the logging performance. Kim and Ahn [21] proposed to use the in-device write buffer to improve the random write performance of flash storage. Lee et al. [23] conducted a case study to investigate how the performance of conventional database applications is affected by the new flash-based disk.

Research efforts have also been put into optimizing transaction-related algorithms. To enable efficient initialization and crash recovery for flash-based file systems, Wu et al. [31] proposed a log management scheme which commits log records onto a special check region to avoid scanning entire flash storage during recovery. To make efficient use of the out-of-place update nature of flash-based media, Prabhakaran et al. [29] developed a novel commit protocol called cyclic commit for flash-based file systems. By using additional metadata on physical pages to creat a cycle for each committed transaction, the cyclic commit protocol successfully eliminates the need of writing commit records and hence achieves comparatively high performance and low space overhead. However, it introduces a complicated garbage collection mechanism and potentially impairs the wear-leveling performance. There are also a few existing works for transaction support in the flash-based DBMSs. Byun [9] proposed a new scheme called flash two phase locking (F2PL) scheme for efficient transaction processing in a flash memory database environment. F2PL achieves high transaction performance by allowing previous version reads and efficiently handling slow write/erase operations in lock management processes. Lee and Moon [22] extended their basic design of IPL to realize a lean recovery mechanism for transactions.

7 Conclusions and Future Work

In this paper, we explored the opportunities to improve the performance of the conventional transaction recovery techniques on flash-based database systems. Based on the nature of flash memory (e.g., out-of-place update pattern), we proposed a novel commit protocol, which ensures the atomicity of transactions by using the flag metadata on physical pages, thereby eliminating the overhead incurred by writing commit records. We design two variants of the flag commit, CFC and AFC, each of which favors for different transaction abort ratios. We also presented extensions to support a better buffer management policy and a more fine-

grained concurrent control mechanism. In future work, we will implement the proposed commit protocol and its extensions. In addition, we will also evaluate the performance of our design in comparison with existing flash-based transaction recovery techniques such as IPL and cyclic commit.

References

[1] Aleph One Ltd., Yaffs: A NAND-Flash File system. http://www.aleph1.co.ukyaffs.

[2] Cactus Technology, CTAN010: SLC vs. MLC NAND. http://www.cactus-tech.com/download/CTAN010-SLCvsMLC-20080603.pdf.

[3] Intel Corporation, Understanding the flash translation layer (FTL) specification. Intel Technical Report. http://developer.intel.com.

[4] Numonyx Corp., NAND01GW3A2B-KGD Datasheet. http://www.numonyx.com/Documents/Datasheets.

[5] Samsung Corporation, K9F1208X0C Datasheet. http://www.samsung.com/Products/Semiconductor/.

[6] Samsung Corporation, K9XXG08UXA Datasheet. http://www.samsung.com/Products/Semiconductor/.

[7] SSFDC Forum, SmartMedia Specification. http://www.ssfdc.or.jp.

[8] Super Talent Technology, SLC vs. MLC: An Analysis of Flash Memory. http://www.supertalent.com/datasheets/SLCvsMLCWhitepaper.pdf.

[9] S. Byun. Transaction management for flash media databases in portable computing environments. *J. Intell. Inf. Syst.*, 30(2):137–151, 2008.

[10] L.-P. Chang and T.-W. Kuo. An efficient management scheme for large-scale flash-memory storage systems. In *SAC '04: Proceedings of the 2004 ACM symposium on Applied computing*, pages 862–868, New York, NY, USA, 2004. ACM.

[11] L.-P. Chang, T.-W. Kuo, and S.-W. Lo. Real-time garbage collection for flash-memory storage systems of real-time embedded systems. *Trans. on Embedded Computing Sys.*, 3(4):837–863, 2004.

[12] Y.-H. Chang, J.-W. Hsieh, and T.-W. Kuo. Endurance enhancement of flash-memory storage systems: an efficient static wear leveling design. In *DAC '07: Proceedings of the 44th annual conference on Design automation*, pages 212–217, New York, NY, USA, 2007. ACM.

[13] D. J. DeWitt, R. H. Katz, F. Olken, L. D. Shapiro, M. R. Stonebraker, and D. Wood. Implementation techniques for main memory database systems. *SIGMOD Rec.*, 14(2):1–8, 1984.

[14] E. Gal and S. Toledo. Algorithms and data structures for flash memories. *ACM Comput. Surv.*, 37(2):138–163, 2005.

[15] J. Gray and B. Fitzgerald. Flash disk opportunity for server applications. *Queue*, 6(4):18–23, 2008.

[16] J. Gray, P. McJones, M. Blasgen, B. Lindsay, R. Lorie, T. Price, F. Putzolu, and I. Traiger. The recovery manager of the system r database manager. *ACM Comput. Surv.*, 13(2):223–242, 1981.

[17] J. Gray and A. Reuter. *Transaction Processing: Concepts and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1992.

[18] T. Haerder and A. Reuter. Principles of transaction-oriented database recovery. *ACM Comput. Surv.*, 15(4):287–317, 1983.

[19] H. joon Kim and S. goo Lee. A new flash memory management for flash storage system. In *COMPSAC '99: 23rd International Computer Software and Applications Conference*, page 284, Washington, DC, USA, 1999. IEEE Computer Society.

[20] A. Kawaguchi, S. Nishioka, and H. Motoda. A flash-memory based file system. In *TCON'95: Proceedings of the USENIX 1995 Technical Conference Proceedings on USENIX 1995 Technical Conference Proceedings*, pages 13–13, Berkeley, CA, USA, 1995. USENIX Association.

[21] H. Kim and S. Ahn. Bplru: a buffer management scheme for improving random writes in flash storage. In *FAST'08: Proceedings of the 6th USENIX Conference on File and Storage Technologies*, pages 1–14, Berkeley, CA, USA, 2008. USENIX Association.

[22] S.-W. Lee and B. Moon. Design of flash-based dbms: an in-page logging approach. In *SIGMOD '07: Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pages 55–66, New York, NY, USA, 2007. ACM.

[23] S.-W. Lee, B. Moon, C. Park, J.-M. Kim, and S.-W. Kim. A case for flash memory ssd in enterprise database applications. In *SIGMOD '08: Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 1075–1086, New York, NY, USA, 2008. ACM.

[24] B. r. J. Luc Bouganim and P. Bonnet. uflip: Understanding flash io patterns. In *CIDR*, Asilomar, CA, USA, 2009.

[25] P. C. H. L. Mei-Ling Chiang and R. chuan Chang. Manage flash memory in personal communicate devices. In *Proceedings of the 1997 International Symposium on Consumer Electronics (ISCE'97)*, pages 177–182, 1997.

[26] C. Mohan, D. Haderle, B. Lindsay, H. Pirahesh, and P. Schwarz. Aries: a transaction recovery method supporting fine-granularity locking and partial rollbacks using write-ahead logging. *ACM Trans. Database Syst.*, 17(1):94–162, 1992.

[27] A. Silberschatz, H. F. Korth, and S. Sudarshan. *Database Systems Concepts*. McGraw-Hill Higher Education, 2001.

[28] B. soo Kim and G. young Lee. Method of driving remapping in flash memory and flash memory architecture. United States Patent No. 6,381,176, 2002.

[29] T. L. R. Vijayan Prabhakaran and L. Zhou. Transactional flash. In *OSDI '08: Proceedings of the 8th USENIX Symposium on Operating Systems Design and Implementation*, Berkeley, CA, USA, 2008. USENIX Association.

[30] D. Woodhouse. Jffs: The journalling flash file system. In *Proceedings of the of the Ottawa Linux Symposium*, pages 177–182, 2001.

[31] C.-H. Wu, T.-W. Kuo, and L.-P. Chang. Efficient initialization and crash recovery for log-based file systems over flash memory. In *SAC '06: Proceedings of the 2006 ACM symposium on Applied computing*, pages 896–900, New York, NY, USA, 2006. ACM.

Exploiting Fast Random Reads for Flash-based Joins

Yu LI

Abstract

Flash disks have been an emerging secondary storage media. There have been portable devices, multimedia players and laptop computers that are configured with no magnetic disks but flash disks. In this paper, we studied the core of query processing in RDBMSs – join processing – on flash disks. In particular, we present the evaluation results of a new join method, called DigestJoin, which exploits fast random reads of flash disks to RDBMSs. Experiments with datasets of TPC-H bechmark show that DigestJoin will generally boost relational joins under various system configurations.

1 Introduction

Flash disks have been widely used in portable devices such as PDAs, smartphones and multimedia players, as well as in some laptop and desktop computers in the form of Solid State Drive (SSD). As a candidate for mass storage media, recent research, *e.g.*, [15, 21, 22, 27], has investigated the possibilities for flash-based RDBMSs. It is remarked that flash disks have unique I/O characteristics. For instance, flash-based storage does not involve any mechanical components and hence there is a negligible seek time and rotational delay in reading or writing a page on a flash disk. The overhead on each I/O operation, which is caused by the encapsulated logic for such purposes as wear leveling and internal caching [5, 7], can be more than 20 times smaller than a mechanical seek in magnetic disks. Recall that query processing algorithms on magnetic disks often spend an effort to avoid random I/O operations but exploit sequential I/O operations whenever possible. The I/O characteristics of flash disks indicate that such an effort is no longer crucial in flash-based RDBMSs. However the state-of-the-art RDBMSs are mainly designed to operate on magnetic disks and therefore assume the I/O characteristics of magnetic disks. Among the operators of SQL (supported by all RDBMSs), joins are particularly I/O intensive, and computationally expensive [20]. In particular, when process joins in the absence of indexes, *i.e.*, non-index joins, the main objective of optimizing basic non-index join methods, *e.g.* nested-loop, sort-merge and hash joins, is to reduce the number of I/O opertions, especially the expensive seek operations. For example, sort-merge join algorithm first sequentially scans and sorts the tables (using an external sort algorithm if necessary) by the join attributes, and then sequentailly scans the sorted results in the merge phase of the algorithm, as illustrated in Fig. 1. In general, the state-of-the-art optimizating techniques for nested-loop, sort-merge and hash joins have done their best to minimize both the number of seek operations and total I/O operations on magnetic disks.

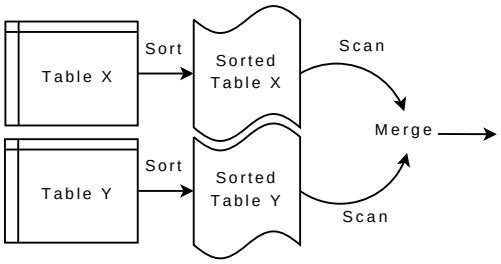


Figure 1: Sort-Merge Join Algorithm

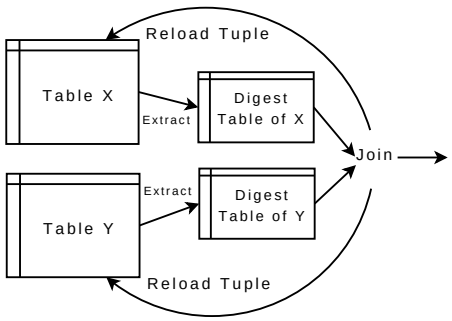


Figure 2: Overview of DigestJoin

Nevertheless, in the context of flash-based RDBMSs, minimizing the number of seek operations does not offer much advantage. Given flash disks, now it is desirable to study how the cheap seek operations, which in practice leads to cheap random read operations, can enhance the performance of joins. In last paper, we propose *DigestJoin*, a

join algorithm that exploits fast random reads, among others, on flash disks. *DigestJoin* consists of two phases. In the first phase, *DigestJoin* projects the tuple id (*tid*)[1] and only the attributes that are relevant to the join operator from the tables that participate the join. The projected tables are called the *"digest" tables*. The main intuition here is that flash disks are often installed inside mobile devices with limited main memory. It is often beneficial to reduce the I/O operations required by the join operator through a scan on the tables to obtain digest tables. A traditional join algorithm is then applied on the smaller digest tables to generate the *digest join results*. The digest join results are pairs of *tid*s together with the join attributes, to minimize the size of intermediate join results. It is worth remarking that the digest join results are similar to join indices [26, 25, 17] except that they are computed on-the-fly in the first phase of of *DigestJoin* because we do not assume the presence of indices on the join attributes.

2 DigestJoin

2.1 Overview

Fig. 2 gives an overview on *DigestJoin*. As described, it is divided into two phases: *digest join phase* and *page fetching phase*.

Digest Join Phase. This phase extracts the "digest tables" from original tables and joins them. Consider two tables, $X = \{tid_x, Attr_{x_1}, Attr_{x_2}, \ldots, Attr_{x_m}\}$ and $Y = \{tid_y, Attr_{y_1}, Attr_{y_2}, \ldots, Attr_{y_n}\}$, join under the drective $X \bowtie_{Attr_{x_1}=Attr_{y_1}} Y$. The *digest tables* are projections containing only the join attributes and the tuple ids. In our example, they are $X' = \{Attr_{x_1}, tid_x\}$ and $Y' = \{Attr_{y_1}, tid_y\}$. Such a projection obviously reduces much I/O in performing the actual join. Then, we apply a traditional join algorithm (*e.g.*, nested-loop, sort-merge, or hash join) to the digest tables to generate the *digest join results*, in the form of $\{Attr_{x_1}, tid_x, tid_y\}$. The digest join results are often small, and yet may be written to the flash disk sequentially if it is larger than the memory size. However the digest join results $\{Attr_{x_1}, tid_x, tid_y\}$ only tell us which tuples satisfy the join. In next phase, in order to construct the final join results, we have to fetch the corresponding tuples from the original tables.

Page Fetching Phase. Fetching tuples of the digest join results from a flash disk is necessary to *DigestJoin*. But its efficiently is critical. Suppose that we have the following digest join results: $(x_1, tid_{x_1}, tid_{y_1})$, $(x_2, tid_{x_2}, tid_{y_2})$, $(x_3, tid_{x_3}, tid_{y_3})$, and $(x_4, tid_{x_4}, tid_{y_4})$, where tuples of tid_{x_1} and tid_{x_3} are stored on page A, tuples of tid_{x_2} and tid_{x_4} are on page B, tuples of tid_{y_1} and tid_{y_3} are on page C, tuples of tid_{y_2} and tid_{y_4} are on page D. If we have sufficient memory space, we may fetch all those four pages A B C D and keep them in the memory to construct the final result tuples. However, when memory space is limited and therefore we need to carefully schedule the page fetching to minimize the read cost. Suppose that in the worst case the memory space can hold two pages only. If we construct the final result in the order of x_1, x_2, x_3, and x_4, we need to fetch pages A and C for x_1, B and D for x_2, then A and C again for x_3, and finally B and D again for x_4. In this scenario, each page is fetched twice. Alternatively, we may reschedule the order by swapping x_2 and x_3 when we construct the final join result. Then, each page can be fetched once. Therefore, the page fetching schedule can significantly affect the I/O cost. Hence, it is a key to the success of *DigestJoin*.

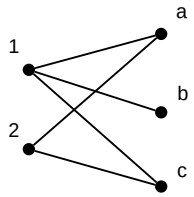


Figure 3: Example of Join Graph

We can use a graph model called *join graph* to model and analyze the page fetching phase. A join graph is used to represent the relationship between the disk pages specified by the join results. It is defined as an undirected bipartite graph $G = (V_1 \cup V_2, E)$, where V_1 and V_2 denote the set of pages from the two original tables, respectively, and $E \subseteq V_1 \times V_2$ denotes the set of page-pairs specified by the join result. Specifically, for each edge $(v_a, v_b) \in E$, there exists a tuple on page v_a which joins with a tuple on page v_b. The join graph can be used to dynamically represent the remaining pages to be fetched and joined: An edge (v_a, v_b) is removed from E if pages v_a and v_b have been fetched into the main memory. A vertex v is removed from G once its out-degree is zero. An example of a join graph is shown in Fig. 3, where vertices 1 and 2 represent the pages from one table, while vertices a, b and c represent the pages from the other table. An edge $(1, a)$ means there is/are tuple(s) on page 1 that can be joined with tuple(s) on page a. Similarly, edges $(1, b)$, $(1, c)$, $(2, a)$ and $(2, c)$ represent the "join relationship" between other pairs of pages.

A page fetching sequence is equivalent to a sequence for removing all edges of the join graph. And an optimal page fetching sequence is a sequence for removing all edges in the join graph with the minimum number of visting vertices, which eaquals to the page fetchings. When the mem-

[1]Throughout this paper, we assume that the tuple id is implemented as a page id and slot number.

ory space is limited, page swaps are needed and the problem becomes intractable, and unfortunately, it is proven to be a *NP-hard* problem. Furthermore, to the best of our knowledge, there has not been practical approximation algorithm to address it [4]. So we turn to propose heuristic strategies.

2.2 Page Fetching Strategies

2.2.1 Naive Fetching Strategy

The first strategy – called *naive fetching strategy* – is to fetch the pages of the tuples as soon as they are produced in the digest join phase. The intuition is that seek operations do not incur much overhead on a flash disk. Hence, though with predicable reduandent page fetching, the naive fetching strategy may still perform acceptable well. In practice, we can assign less input buffer pages in the merge phase of sort merge join and the probe phase of hash join, leaving rest available memory space to cache disk pages due to page fetching.

2.2.2 Page-based Fetching Strategy

When the digest join results are clustered with respect to the page address and the full tuples of the join results are clustered with respect to the join attributes, the naive fetching strategy would result in an optimal page fetching sequence which ensures each page is fetched at most once. This motivates us to propose the next *page-based fetching strategy*. Specifically, we build two kinds of temporary tables to assist page fetching. The first one is called *fetching instruction table*, which archives digest join results. After this table is filled with all digest join results, we sort its digest join results based on their page addresses. Thus, fetching tuples based on such sorted digest results avoids duplicated page fetching requests. However, tuples fetched according to their page addresses are generally not clustered on the join attribute. Hence, we have another temporary table called *join candidate table* to store the tuples fetched according to the fetching instruction table. Sort-merge join or hash join algorithm can be then applied on this table for producing final join results.

2.2.3 Graph-based Fetching Strategy

An alternative method is to archive the digest join results with the join graph. If the memory can hold all digest join results in the form of a join graph, we may find good heuristics to travel all edges to do the page fetching and joining near optimally. The *graph-based fetching strategy* follows the intuition. It adopts the technique behind two heuristics [4, 23] in literature, which originally focus on page fetching for index-based joins. The basic idea is to select a minimal subgraph of a vertex which contains all its adjacent edges and requires the fewest non-resident pages to be fetched. Here, a non-resident page is a page that is not currently stored in the main memory. Because iterating all subgraphs of the join graph makes the computational cost prohibitively high, we approximates it by only selecting the vertex with the fewest non-resident neighbors, together with its neighbors in the join graph, called *segment* [4].

However, there are some challenges in applying this heuristic in our scenario. First is the limited memory space is not likely to hold the entire join graph. Second, preallocating sufficient memory for caching pages for the join graph is not feasible in general because the heuristics we want to adapt only provide a impractical rough bound on the maximum required size of the page. Therefore, we propose to dynamically manage the memory to effectively hold both join graph and page cache, with of the aiming of achieving acceptable performance. We divided the memory space into two parts: one for storing the join graph (*i.e.*, join graph storage) and the other for caching fetched pages (*i.e.*, page cache). Intiailly, the join graph storage is configured to one, and all left memory will be used as page cache. When a digest join result comes, if there is space, we directly add it into the join graph. Otherwise, we try to adjust the space for bigger join graph. In particular, we calculate the *required storage size*(RSS) and *required cache size* (RCS). The RSS of a join graph is equal to the number of pages that are required to hold this join graph. The RCS of a join graph is the minimum cache size for fetching and joining any segment of this join graph. And then we perform the following.

- If $RCS <$ the current page cache size, we enlarge the memory space for join graph storage (correspondingly, shrink the space of the page cache by removing some cached pages) to insert that result.
- Otherwise, we try to select segments of the join graph to load and join using the page cache, until that we empty at least one page for new digest join result. After that, we insert the digest join result, and check RSS to see whether we need to enlarge the space for the page cache (correspondingly, shrink the memory space of the join graph).

3 Performance Evaluation

3.1 Experiment Setup

We implemented *DigestJoin* for both sort merge join and hash join algorithms with all three page fetching strategies in an experimental database system built on top of 16GB Mtron MSD-SATA3025 SSD. The experimental database system is designed to enable an easy evaluation on the

performance of different join algorithms. It is composed of three components: *raw storage manager*, *page-oriented buffer manager*, and *query executor*. The raw storage manager maintains a bunch of pre-allocated continuous space on the SSD and provides a page-oriented read/write interface. The other two components rely on this interface to perform I/O access to the SSD. The page-oriented buffer manager maintains a fixed number of memory pages, each of which have the same size as the ones in the raw storage manager. Any query processing in the query executor should interact with the memory pages in the buffer manager. The query executor provides facilities to implement and executes specific join algorithms. The system catalog maintains some statistics information, such as system parameters (*e.g.*, page size, memory size, etc.), table summaries (*e.g.*, # tuples, # pages, etc.), as well as join statistics (*e.g.*, join selectivity).

Tables are organized on page-oriented space of the SSD. The data tuples are stored on the pages following a row-based storage scheme. Since we study with non-index-based joins, we do not build any index for each table. To save space, we store data tuples in variable-size format. Finally, the data tuples are imported into the storage in a random order. This is to simulate a general-case join where the join attribute values could be in any arbitrary order in the original tables.

The test dataset is taken from TPC-H benchmark. In particular, we use two `CUSTOMER` tables of TPC-H and perform a natural join on them through the key `C_CUSTKEY`. The selectivity is 100%. In our implementation, we install a filter function before we are about to get a join result. The filter function will flip a coin to decide whether to drop the result or not. By controlling the probability of flipping, we can simulate different selectivities. In particular, we can also control the selectivity on a page basis to simulate skewed join distributions.

Tow categories of algorithms are implemented. On category consists of sort merge related algorithms, which are sort merge join (SM), *Digest sort merge join* with naive page fetching strategy (DigestSM(Naive)), *Digest sort merge join* with page-based fetching strategy (DigestSM(Page)), and *Digest sort merge join* with graph-based fetching strategy (DigestSM(Graph)). The other one consists of hash related algorithms, which are hash join (H), *Digest hash join* with naive page fetching strategy (DigestH(Naive)), *Digest hash join* with page-based fetching strategy (DigestH(Page)), and *Digest hash join* with graph-based fetching strategy (DigestSM(Graph)).

We run all the experiments on a desktop PC equipped with a Core 2 Quad Q6600 CPU and 4GB main memory. The default system parameters used in the evaluation are listed in Table 3.

Table 1: Default Parameter Settings

Parameter	Setting
Page size	4 KBytes
Ratio of digest entry size to tuple size (p)	9%
Average tuples per page (t)	25
Join attribute	`C_CUSTKEY`
Table size	512 MB $\times$ 2
Join selectivity	0.05
Memory size	32MB

3.2 Impact of Join Selectivity

We investigate how the join selectivity would affect their performance in this section. Fig. 4, Fig. 5 , Fig. 6 and Fig. 7 show the performance comparison with join selectivity changing from 0 to 0.1 and from 0.1 to 0.5, respectively. The results do not count the I/O cost to output final join results, since usually they are not output to the secondary storage and this cost is the same for all join algorithms. As such, the results of traditional join algorithms, *i.e.*, SM and H, remain the same over different selectivity settings. They are used as the baseline in the performance analysis.

As can be seen from Fig. 4 and Fig. 6, DigestSM(Naive) and DigestH(Naive) outperform traditional join algorithm only when the selectivity is very low (*i.e.*, < 0.05), because when the selectivity is higher than 0.05, the naive fetching brings too many duplicate page accesses, which wastes the IO saving in previous digest join stage. DigestSM(Page), DigestH(Page), DigestSM(Graph) and DigestH(Graph) shows similar performance trends to DigestSM(Navie) and DigestH(Naive), but their performance degrades much slower when the selectivity increases. DigestSM(Graph) outperforms SM until the selectivity exceeds 0.3(Fig. 5), and DigestH(Page) outperforms H until the selectivity exceeds 0.2(Fig. 7). After that, all digest join algorithms lose their superiorities as predicted.

Notice that DigestH(Page) generally outperforms DigestH(Naive) and DigestH(Graph). The reason can be drawn as follows: our implementation trends to decide the hash function by considering whether the result bucket size is smaller than usable memory, as following the GraceHash algorithm. However, applying this technique to digest hash algorithms makes most of the memory be occupied by the hash bucket. Only few of memory pages will be used by page cache and join graph. It affects DigestH(Graph) because with insufficient memory, the quality of segment selection degenrates and the false hit on page cache increases. On the other hand, it does not affect DigestH(Page) which uses temporary tables to optimizing the fetching process, because the temporary table could be sequentially access by careful implementation. So DigestH(Page) becomes the

best of digest hash join algorithms.

3.3 Impact of Page Size

Fig. 8, Fig. 9, Fig. 10 and Fig. 11 show the performance comparison with page size changing from 4KBytes to 32KBytes. By default we use 4KBytes page size, which is the smallest page size supported by our SSD, and there is totally 32MB memory (8192 pages) dedicated for join. When changing the page size, we maintain the same total amount of memory (*i.e.*, 32MB). Hence, there are 4096/2048/1024 memory pages that could be used when the page size is 8KB/16KB/32KB, respectively. The impacts of page size on join algorithms are different. Traditional algorithms(*i.e.*, SM, H), and digest algorithms with page fetching strategy(*i.e.*, DigestSM(Page), DigestH(Page)) become better when page size increases. The reason is that a big page size implies fewer pages of tables, and the external sort can be done in fewer runs. DigestSM(Graph) and DigestH(Graph) also benefit from a big page size, because fewer pages of tables result in a smaller join graph, and therefore we can make the in-memory join graph more informative, which leads to better page fetching schedules. On the other hand, DigestSM(Naive) and DigestH(Naive) perform worse when page size increases, and it is even worse than traditional algorithms when page size is bigger than 8KB with low selectivity(see Fig. 8 and Fig. 10). This is because no matter how few the pages are, the number of fetching requests remains the same as the selectivity does not changed. As such, when the cost of fetching a page become higher due to a bigger page size, naive fetching strategy becomes worse.

3.4 Impact of Memory Size

Fig. 12 and Fig. 13 show the performance comparison with memory size ranging from 8MB to 128MB (*i.e.*, 1.6% to 25% of the table size). As expected, the I/O cost decreases for all algorithms when more memory space is available. It can also be observed that when the memory size increases from 8MB to 128MB, the DigestSM(Page) outperforms and DigestSM(Graph)(Fig. 12). This is because a large memory is more helpful for the page-based strategy than for graph-based strategy. Page fetching strategy can utilize large memory to load and join the candidate join tables with fewer runs or even in memory, thereby achieving a more efficient memory usage. But the graph fetching strategy, which even may have the whole join graph in memory, still suffer from the penatlty of redundant loading caused by heuristically fetching pages.

3.5 Impact of Digest Entry Size

According to the TPC-H schema, the `C_CUSTKEY` is defined as a 4Bytes integer, the `C_ACCTBAL` is defined as a 8Bytes float number, and the `C_PHONE` is defined as a 15Bytes fix-length string. In our system, the digest entry consists of a 8Bytes *page-id*, a 8Bytes *rid* and the join attribute. So the size of digest entry with `C_CUSTKEY` is 9% of the size of database tuple in schema. The ratios of digest entry size to database tuple size will be 11% and 15%, when join attributes are `C_ACCTBAL` and `C_PHONE`, respectively. Fig. 14 and 15 show how the IO cost changes as we change join attribute from small one to big one, *i.e.*, from `C_CUSTKEY` to `C_ACCTBAL` and then to `C_PHONE`. As we can observe, when the digest entry size increases because of different join attributes, the IO cost also increases. This can be understand as follows: the foundation of digest join algorithms is that they can reduce table size to save IO in the first digest join phase. But now the size of digest entry is growing, so the saving in digest join phase is less. On the save time, the second fetching phase comsumes same IOs because the selectivity does not change, so therefore overall performance gain drops.

3.6 Impact of Join Result Distribution

In previous experiments, all algorithms are evaluated under the setting where join results are uniformly distributed over pages. In this section, we evaluate their performances when the join results follow non-uniform distributions. Specifically, we skewed the join results on disk pages based on a Zipf distribution. The Zipf distribution is controlled by a skewness parameter of θ. When θ is 0, the distribution is uniform. The larger is the value θ, the more skewed is the distribution. We apply the Zipf distribution to one table only, which emulates the case that one transaction table join with a part of the fact table, *e.g.*, some day's orders join with the whole customer table.

Fig. 16 and Fig. 17 show the results under the Zipf distribution. Due to the skewed distribution of join results, the selectivity cannot be very high. We set it to be 0.1. We make two observations showing that when θ becomes larger, digest join algorithms have a better performance improvement over traditional algorithms. In detail, one observation is that DigestSM(Naive) and DigestH(Naive) outperform SM and H in most cases. This is because fewer pages are required to be fetched when join results have high skewed distributions. The other observation is that DigestSM(Graph) and DigestH(Graph) save more I/O cost than DigestSM(Page) and DigestH(Page) when θ is larger. This is because in that case the join graph is generally small and hence an effective page fetching schedule is more likely to be achieved.

3.7 Magnetic Disk v.s. Flash Disk

Fig. 18 and Fig. 19 show the comparison of running traditional and digest join algorithms on magnetic disk (HD) and flash disk (SSD). First impression is that SSD's fast IO definitely boost all join algorithms. But the overall conclusion is that digest join algorithms are not suitable for HD. There are two reasons. First is the IO cost of running digest join algorithms on HD are significantly bigger than running them on SSD. As we described before, this is predicable as the random read in the page fetching phase is expensive on HD. Second, the IO cost of digest join algorithms are even bigger than their corresponding traditional versions, *i.e.*, SM and H. This is not only because the single random read is expensive on HD, but also because digest join algorithms introduce more randome read in the page fetching phase. In particular, the more random readings the algorithm need, the bigger IO cost on HD will there be. For example, we can observe that both IO cost of DigestSM(Naive) and DigestSM(Graph) on HD are bigger than IO cost of DigestSM(Page) on HDin Fig. 18, even though all of them is bigger than the IO cost of traditional sort merge join on HD. Similar phenomea can be observed between IO cost of DigestH(Naive), DigestH(Graph), DigestH(Page) and H in Fig 19.

4 Related Work

Relational database management on flash-based storage media has attracted increasing research attention in recent years. Because of the unique I/O characteristics of flash disks, early work focused on assembling flash chips to simulate traditional hard disk [13, 6, 14] and guaranteeing long life span of data [5, 7, 14]. Based on such research, recent work exploits the characteristics of flash disks to enhance the performance of RDBMSs. In view of the asymmetric read/write speed and the erase-before-write limitation, Wu et al. [27] proposed log-based indexing scheme for flash memory. Observing that the log-based indexing scheme is not suitable for read-intensive workload on some flash devices, Nath and Kansal [22] developed an adaptive indexing method that adapts to the workload and storage device. Lee and Moon [15] presented a novel storage design called in-page logging (IPL) for RDBMS. Lee et al. [16] investigated how the performance of standard RDBMS algorithm is affected when the conventional magnetic hard disk is replaced by the flash disk. Shah et al. [24] presented a fast scanning and joining method based on adapting PAX storage model [1] on flash disks, which falls in the same category of our work. The main difference is that our work utilizes fast random reads to optimize traditional join algorithms, while PAX presented in [1] is an alternative storage scheme of relations on flash disks.

Join has been one of the important query operators in RDBMSs. Extensive research efforts have been spent to optimization for join processing. Mihra and Eich [20] surveyed a number of join algorithms and their implementations. In this paper we focus on exploring the possibility of further improving on-indexed-based join algorithms by utilizing the random read of flash disks. In particular, our proposed algorithm is inspired by join indices and page fetching. The idea of join indices [26, 25, 17] is to precompute join results and record pairs of tuple ids that satisfy the join to speed up future join requests. We find that this idea is useful even when we generate the join indices on demand for joins on flash disks, *i.e.*, the first phase of *DigestJoin*. Next we need to tackle determining an optimal page fetching schedule in the second phase of *DigestJoin*. This is actually found in index-based join algorithms. Specifically, index-based join algorithms first compose a list of tuple pairs that participate in the join by using indexes, and then tuples themselves have to be fetched to construct the final results [3, 8]. Merrett et al. [19] proved the decision problem of optimally scheduling the page fetching to be *NP-complete*. A number of heuristics have been developed, *e.g.*, [11], [23] and [4]. We adopt the ideas behind those heuristics to design page-based and graph-based strategy in the second phase of *DigestJoin* while focusing on making them memory-conscious.

5 Conclusion

In this paper, we evaluates *DigestJoin* which exploits fast random reads of flash disks. *DigestJoin* is a generic join method as its implementation invokes traditional join algorithms. By implementing two categories of *DigestJoin* methods, *i.e.*, Digest Sort Merge Join algorithms and Digest Hash algorithms, with the three page fetching strategies in an experimental database system on top of a real flash disk, we show that the *DigestJoin* method generally improves the traditional join algorithm under various system parameter settings. We also show that *DigestJoin* is designed for flash disk by a comparsion of running *DigestJoin* algorithms on magnetic disk and flash disks.

References

[1] A. Ailamaki, D. J. DeWitt, M. D. Hill, and M. Skounakis. Weaving relations for cache performance. In *VLDB '01*, pages 169–180, 2001.

[2] F. Bancilhon, P. Richard, and M. Scholl. Verso: The relational database machine. *Advanced Database Machzne Architecture*, pages 1–18, 1983.

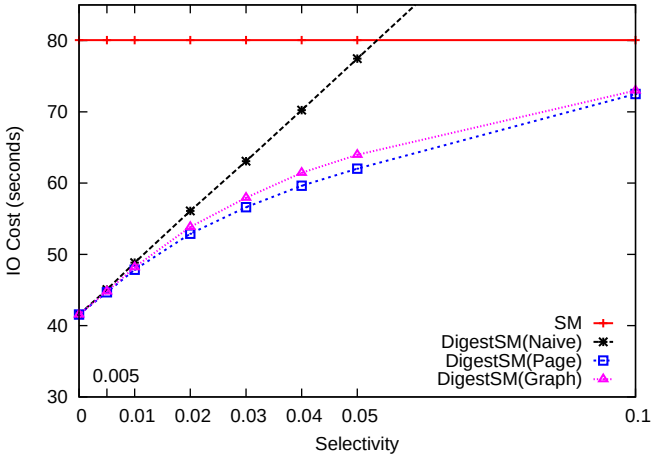


Figure 4: Digest Sort Merge Join vs. Traditional Sort Merge Join under Low Selectivities

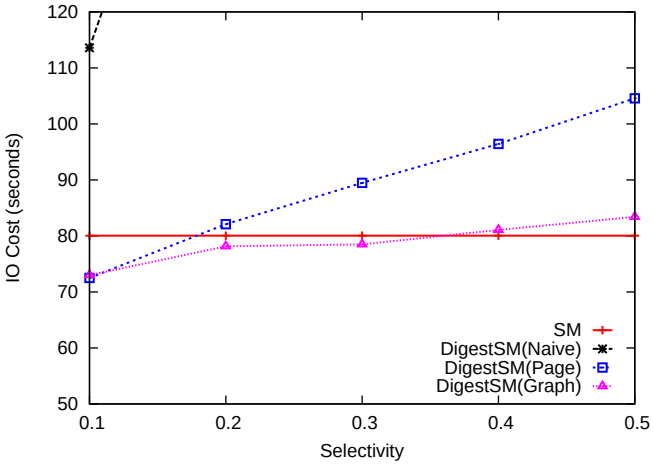


Figure 5: Digest Sort Merge Join vs. Traditional Sort Merge Join under High Selectivities

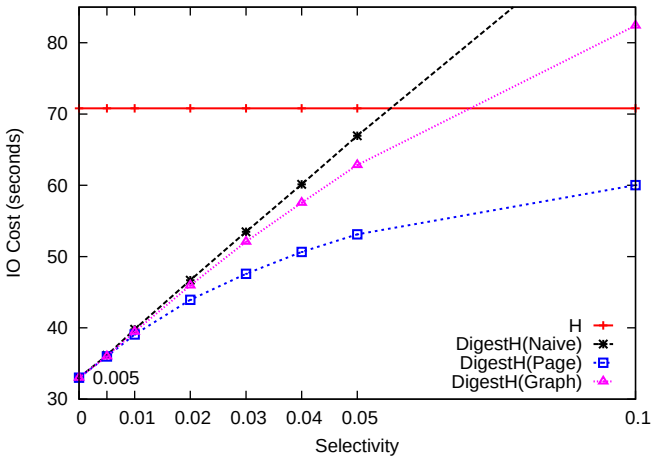


Figure 6: Digest Hash Join vs. Traditional Hash Join under Low Selectivities

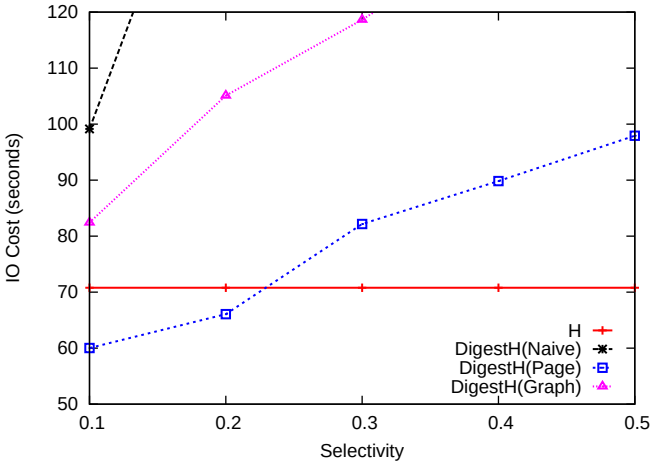


Figure 7: Digest Hash Join vs. Traditional Hash Join under High Selectivities

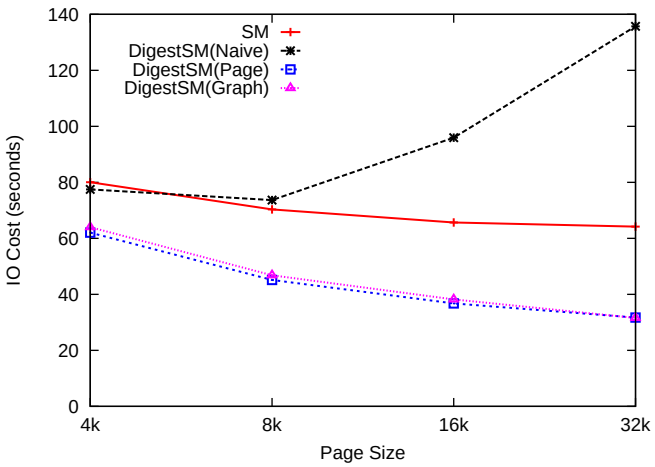


Figure 8: Digest Sort Merge Join vs. Traditional Sort Merge Join under Different Page Sizes (Selectivity=0.05)

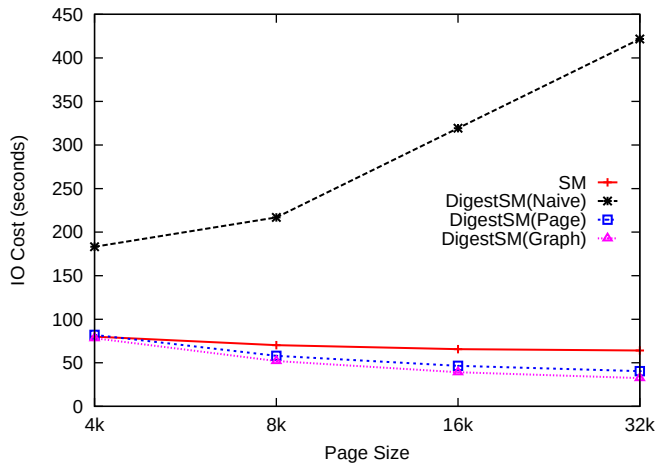


Figure 9: Digest Sort Merge Join vs. Traditional Sort Merge Join under Different Page Sizes(Selectivity=0.2)

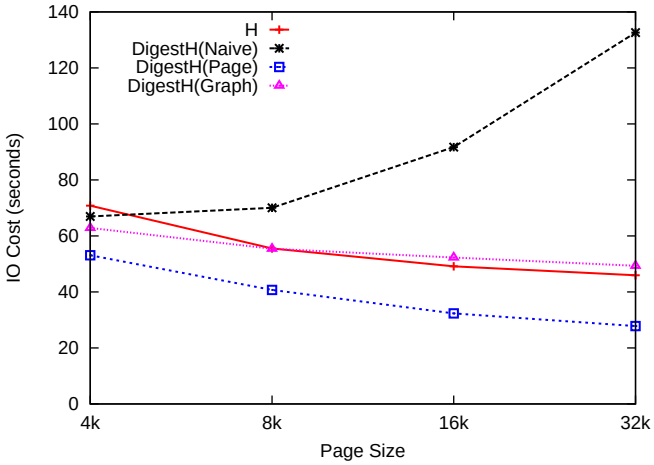


Figure 10: Digest Hash Join vs. Traditional Hash Join under Different Page Sizes(Selectivity=0.05)

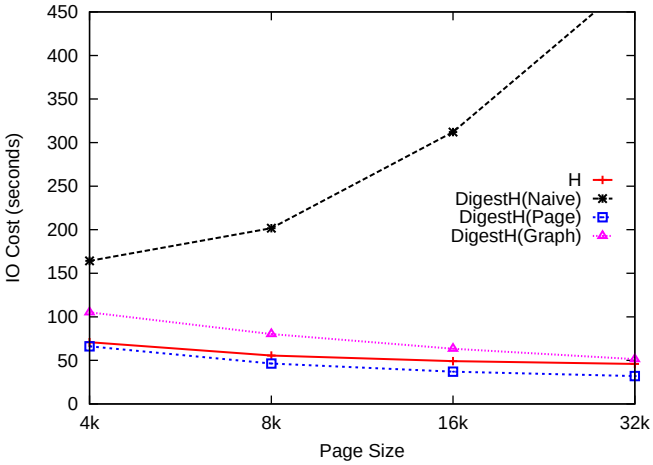


Figure 11: Digest Hash Join vs. Traditional Hash Join under Different Page Sizes(Selectivity=0.2)

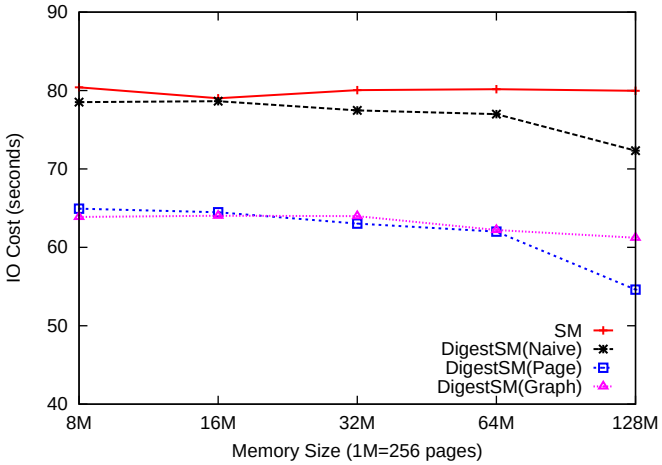


Figure 12: Digest Sort Merge Join vs. Traditional Sort Merge Join under Different Memory Sizes

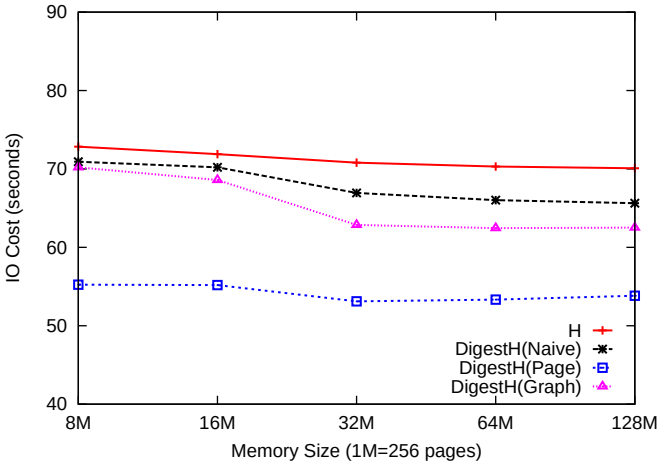


Figure 13: Digest Hash Join vs. Traditional Hash Join under Different Memory Sizes

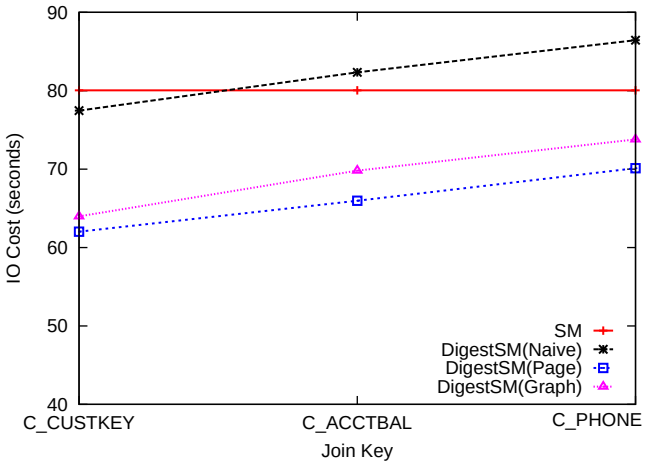


Figure 14: Digest Sort Merge Join vs. Traditional Sort Merge Join under Different Join Keys

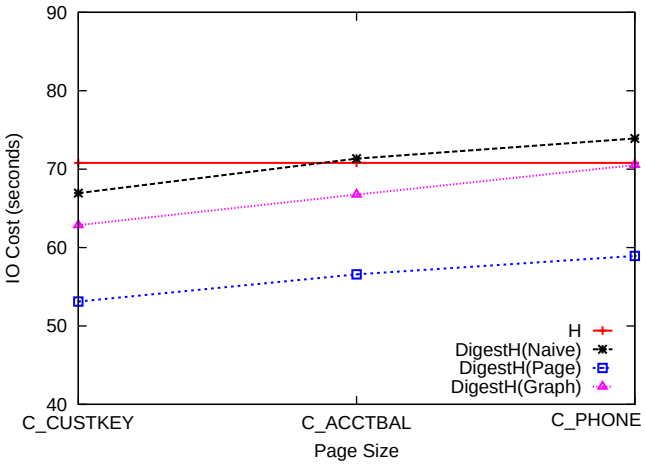


Figure 15: Digest Hash Join vs. Traditional Hash Join under Different Join Keys

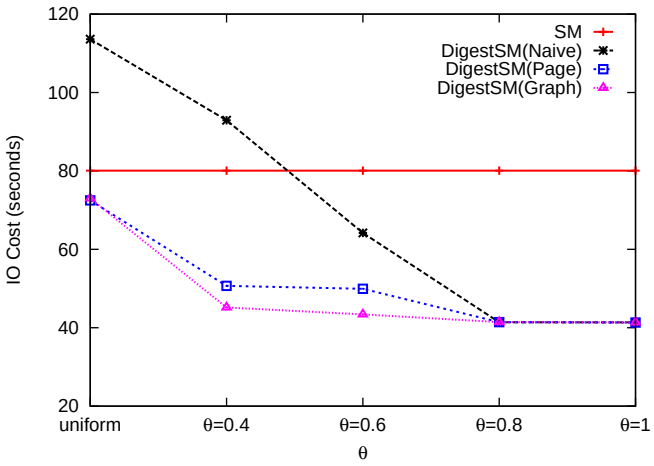


Figure 16: Digest Sort Merge Join vs. Traditional Sort-Merge Join under Zipf Distribution (Selectivity=0.1)

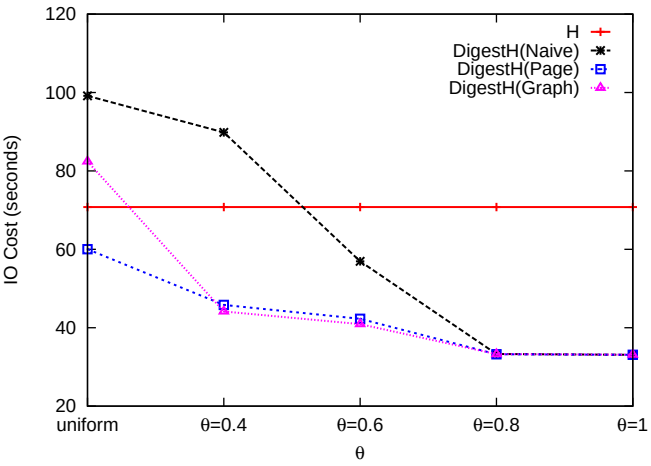


Figure 17: Digest Hash Join vs. Traditional Hash Join under Zipf Distribution (Selectivity=0.1)

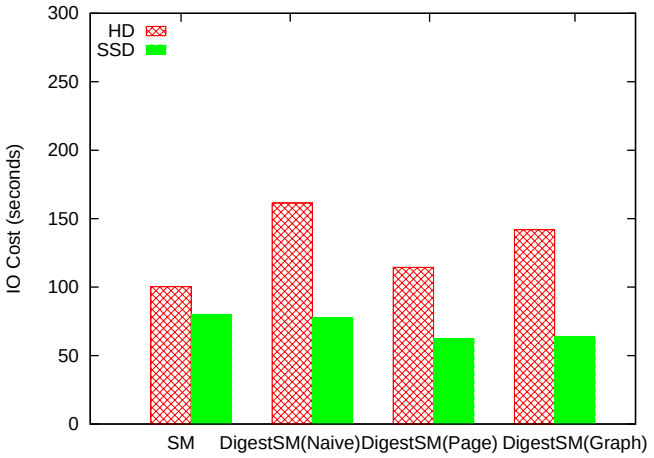


Figure 18: HD vs. SSD with Sort Merge Join Algorithms

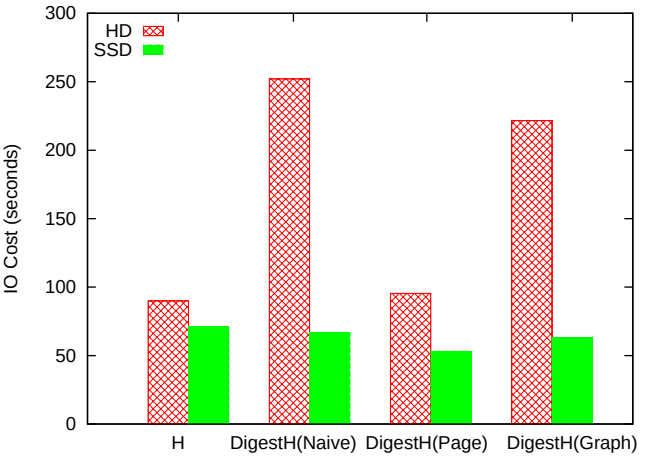


Figure 19: HD vs. SSD with Hash Join Algorithms

[3] M. W. Blasgen and K. Eswaran. Storage and access in relational databases. *IBM System Journal*, 16(4):363–377, 1977.

[4] C. Y. Chan and B. C. Ooi. Efficient scheduling of page access in index-based join processing. *IEEE Trans. on Knowl. and Data Eng.*, 9(6):1005–1011, 1997.

[5] L. Chang. On efficient wear leveling for large-scale flash-memory storage systems. In *SAC'07*, pages 1126–1130, 2007.

[6] L. Chang, T. Kuo, and S. Lo. Real-time garbage collection for flash-memory storage systems of real-time embedded systems. *Trans. on Embedded Computing Sys.*, 3(4):837–863, 2004.

[7] Y. Chang, J. Hsieh, and T. Kuo. Endurance enhancement of flash-memory storage systems: an efficient static wear leveling design. In *DAC'07*, pages 212–217, 2007.

[8] J. M. Cheng, D. J. Haderle, R. Hedges, B. R. Iyer, T. Messinger, C. Mohan, and Y. Wang. An efficient hybrid join algorithm: A db2 prototype. In *ICDE*, pages 171–180, 1991.

[9] D. J. DeWitt and R. H. Gerber. Multiprocessor hash-based join algorithms. In *VLDB'85*, pages 151–164, 1985.

[10] D. J. DeWitt, R. H. Katz, F. Olken, L. D. Shapiro, M. R. Stonebraker, and D. Wood. Implementation techniques for main memory database systems. *SIGMOD Record*, 14(2):1–8, 1984.

[11] F. Fotouhi and S. Pramanik. Optimal secondary storage access sequence for performing relational join. *IEEE Trans. on Knowl. and Data Eng.*, 1(3):318–328, 1989.

[12] G. Graefe. The five-minute rule twenty years later, and how flash memory changes the rules. In *DaMoN '07*, pages 1–9, New York, NY, USA, 2007. ACM.

[13] A. Kawaguchi, S. Nishioka, and H. Motoda. A flash-memory based file system. In *USENIX Winter*, pages 155–164, 1995.

[14] H. Kim and S. Lee. A new flash memory management for flash storage system. In *COMPSAC'99*, page 284, 1999.

[15] S. Lee and B. Moon. Design of flash-based dbms: an in-page logging approach. In *SIGMOD '07*, pages 55–66, 2007.

[16] S.-W. Lee, B. Moon, C. Park, J.-M. Kim, and S.-W. Kim. A case for flash memory ssd in enterprise database applications. In *SIGMOD*, pages 1075–1086, 2008.

[17] Z. Li and K. A. Ross. Fast joins using join indices. *The VLDB Journal*, 8:1–24, 1999.

[18] M.-L. Lo and C. V. Ravishankar. Towards eliminating random i/o in hash joins. In *ICDE '96*, pages 422–429, 1996.

[19] T. H. Merrett, Y. Kambayashi, and H. Yasuura. Scheduling of page-fetches in join operations. In *VLDB'81*, pages 488–498, 1981.

[20] P. Mishra and M. H. Eich. Join processing in relational databases. *ACM Comput. Surv.*, 24(1):63–113, 1992.

[21] D. Myers. On the use of nand flash memory in hgh-performance relational databases. Master's thesis, MIT CSAIL, Cambridge, MA, December 2007.

[22] S. Nath and A. Kansal. Flashdb: Dynamic self-tuning database for nand flash. Technical Report MSR-TR-2006-168, Microsoft Research, 2006.

[23] E. R. Omiecinski. Heuristics for join processing using nonclustered indexes. *IEEE Trans. Softw. Eng.*, 15(1):18–25, 1989.

[24] M. A. Shah, S. Harizopoulos, J. L. Wiener, and G. Graefe. Fast scans and joins using flash drives. In *DaMoN08*, pages 17–24, 2008.

[25] P. Valduriez. Optimization of complex database queries using join indices. *Database Engineering*, 9(16):10–16, Dec. 1986.

[26] P. Valduriez. Join indices. *ACM Trans. Database Syst.*, 12(2):218–246, 1987.

[27] C. Wu, T. Kuo, and L. P. Chang. An efficient b-tree layer implementation for flash-memory storage systems. *Trans. on Embedded Computing Sys.*, 6(3):19, 2007.

Lightweight Emulation to Study BitTorrent-like Systems

Xiaowei CHEN

Abstract

Peer-to-Peer (P2P) networks can reduce the distribution cost of large files for the original provider of the data significantly. Therefore, the BitTorrent protocol is widely used in the Internet today. The current simulation methods, flow-level or packet-level, used to test and study peer-to-peer systems (namely modeling, simulation, or execution on real testbeds) often show limits regarding scalability, realism and accuracy. This paper describes and evaluates Virtual BT, our framework to study BitTorrent by combining emulation (use of the real studied application within a configured synthetic environment). Virtual BT is efficient and scalable and has good virtualization characteristics (many virtual nodes can be executed on the same physical node by using process-level virtualization and nodes do not have disk I/O management). Experiments with the BitTorrent file-sharing system demonstrate the usefulness of this platform. Finally, we identify new areas of improvements for efficient BitTorrent protocol.

1. Introduction

Peer-to-peer systems have become more and more popular over the last few years, and this popularity often required changes that made them more and more complex. BitTorrent is the very popular peer-to-peer file distribution system. It provides very good performance by ensuring that downloaders cooperate by sharing parts they have already downloaded through a complex reciprocation system. Due to this ever increasing complexity, the development and the study of peer-to-peer systems have become more difficult: we need ways to ensure that a peer-to-peer application will work properly on thousands of nodes, or ways to understand applications running on thousands of nodes.

Distributed applications are traditionally studied using mathematical modeling, simulation, and execution on a real system. Simulation consists in using a model of the application's code in a synthetic environment. This method is widely used, and gives valuable results easily. However, it is often difficult to simulate efficiently a large number of nodes using a complex model: a trade-off between the realism of the model and the number of nodes always has to be made.

On the other side, one can run the real application to study on a real-world experimentation platform like PlanetLab [1]. But the environment is then difficult to control and modify (since it depends heavily on the real system itself), and results are often difficult to reproduce (since the environmental conditions may vary a lot between experiments). This kind of real-world experiments is needed when developing a peer-to-peer system, but it doesn't satisfy all needs.

BitTorrent has already been largely evaluated through analysis of large scale utilization [2, 3], analytical modeling [4] or simulation [5].

However, those works have never been compared to large scale studies on real world systems, or to studies using emulation. BitTorrent is an engineering work, not a research prototype, and several parts of its code are very complex. The large number of constants used as parameters of all the important algorithms makes it very hard to model accurately.

We present literature review in section 2. Then, we give motivation in section 3, and we give a detailed description of our emulator in section 4, and our experiment result in section 5. We conclude the paper with future work in section 6.

2. Literature Review

BitTorrent is an extremely popular P2P file distribution protocol that is not formally standardized. The original BitTorrent client, called the mainline client, was open-source. Other clients interoperated with the original client and each other by conforming to aspects of the "protocol" gleaned from its source code. A de facto standard consisting of dominant protocol features eventually emerged [6]. However, many clients make non-standard extensions to the protocol such as Ono [7], peer ISP-cache support, peer exchange protocol, and multitorrents. Interestingly, the mainline client is no longer open source and is now owned by a company called BitTorrent. Nonetheless, other existing clients have no incentive to adopt any new, official protocol changes.

We will review BitTorrent protocol and related literature from the following aspects: undertanding and evaluation, modification and improving, application and extension.

2.1 Understanding and Evaluation

The popularity of BitTorrent comes from its efficiency ensured by its peer and piece selection strategies. The peer selection strategy aims at enforcing the cooperation between peers while the piece selection strategy tends to maximize the variety of pieces available among those peers. The great success of BitTorrent has attracted the curiosity of the research community and several papers have appeared on this subject. Based on the official introduction about BitTorrent [6], [8], we now have a better idea on the strengths and weaknesses of the protocol [9], [10], [11], [12]. We also have a clear idea on the peers' behavior (i.e., arrival and departure processes), and on the quality of service they experience [13], [14], [15], [16]. Pawel et al. [17] first systematically investigated the optimal piece size in BitTorrent, explained why small pieces are the optimal choice for small-sized content, and why further dividing content pieces into subpieces is unnecessary for such content.

Peer oriented experiment result [18] and Hamra et al. [19] expanded on those results through simulation with some experimental confirmation. They also examined the diameter of the overlay created, and the robustness of the overlay in the presence of churn and attacks. Urvoy et al. [20] used a simulated BitTorrent overlay to look at the distance of peers from the initial seed and the matrix of peer connections. Their results were based on a homogeneous collection of peers, and were limited to the initial stage of a swarm. In addition, as compared to a chain, a full mesh overlay makes BitTorrent more robust to peers' departures and overlay partitions.

F. Benbadis [21] deeply analyzed bandwidth relationship among the server, leecher and seed, proposed a conservation law. Focused on greedy strategies in BitTorrent networks, D. Carra [22] took BitTyrant as case study its impact. C. Dale and J. Liu [23] focused on expreimental evaluation, suggest that the initial stage is not predictive of the overall performance; find no clear evidence of persistent clustering in any of the networks, precluding the presence of a small-world that is potentially efficient for peer-to-peer downloading. They first attempt to introduce clustering into BitTorrent. This approach is theoretically proven and makes minimal changes to the tracker only.

Until now, almost all the important BitTorrent default parameters and core algorithms are deeply studied. Many researchers still put effort to them to push the BitTorrent performance to limit by tuning these paramenters and algorithms.

2.2 Modification and Improving

The enhancement of BitTorrent maily focused on incentives and security.

2.2.1 Incentive Mechanisms

Cohen found strict tit-for-tat to come at too high a cost, and weakened the protocol's incentives to achieve better performance.

BitTyrant [29] is a modification of the Azureus BitTorrent client that exploits the "last place is good enough". BitTyrant is an empirical study of its effectiveness. BitThief [11] studies the feasibility of downloading in BitTorrent without uploading. A BitThief client attempts to enter as many peers' optimistic unchoke slots as possible. This strategy results in a tragedy of the commons.

Others have considered game-theoretic models of BitTorrent. Coupon replication [30] has been used as a way to model trading in BitTorrent and show that altruism does not play a critical role in file swarming systems' performance, nor that rarest-first block scheduling is of critical importance. Proportional share [31] has been studied in many contexts. Zhang and Wu show that proportional share in a BitTorrent-like system quickly achieves market equilibrium.

Much work has gone into encouraging cooperation among selfish BitTorrent participants. Tit-for-tat is a common incentive mechanism that peers provide blocks to those who have provided them blocks in the past. BitTorrent was originally described as using tit-for-tat [8]; Jun and Ahamad [32] propose removing optimistic unchoking from BitTorrent in favor of a k-TFT scheme, in which peers continue uploading to others until the deficit (blocks given minus blocks received) exceeds some niceness number k. Garbacki et al. [33] consider an amortized tit-for-tat scheme that effectively allows contributions made while downloading one file to apply in a tit-for-tat-like manner to other files in the future. Other solutions to this problem generally involve monetary mechanisms.

Dandelion [34] is a file distribution protocol that uses currency and key exchanges through a centralized server to provide incentive for sharing across different downloads.

Dave [35] provides newly joined peers an initial set of pieces of the file to trade. Their mechanism encourages peers to trade immediately, and ensures that new peers upload blocks at the same time as downloading. Further, it is resilient to Sybil attacks.

2.2.2 Security Problems

BitTorrent seurity maily includes churn and poisoning, ISP blocking, and attack to seeder or leecher.

Daniel [36] divides churn study into two groups. One is passive monitoring, Sen et al. [37] use passive measurement at several routers to monitor flows in FastTrack, Gnutella, and Direct-Connect. The other is active probing, which is using crawling to characterize P2P networks and present the behavior of session length across peers, [2], [3] and [14] are the representatives. Each of these studies show that session-lengths are not Poisson, and some of the studies further conclude that session lengths are heavy-tailed (or Pareto).

Recently, there have been a number of studies on pollution and poisoning attacks on second-generation P2P file sharing systems (such as Kazaa and eDonkey). Similarly, the "index poisoning" attack, wherein an attacker advertises an enormous number of bogus sources for a targeted content, was highly pervasive in the FastTrack and Overnet DHT (eDonkey) networks. For copyrith protection, X. Lou [38] apply content poisoning model into BitTorrent, eDonkey and eMule. They discover that BitTorrent is most resistant to content poisoning. Index poisoning could be a viable alternative to cope with copyright violation on BitTorrent.

Many ISPs are known to rate-limit the bandwidth consumed by BitTorrent traffic by deploying traffic shapers in their networks. However, it has been discovered recently that some ISPs do not just rate-limit BitTorrent flows but block them outright by injecting forged RST packets into the flows. When the end nodes of a BitTorrent transfer receive the RST packets, they immediately terminate the transfer. Marcel [39] developed the first tool (BTTest) to offer highly specific, reliable blocking detection to a large number of end users. It is widely used in public deployment.

BitTorrent swarms are susceptible to a number of different attack types. For leechers, Prithula [40] observed two attacks that are frequently deployed today, which we refer to as the fake-block attack and the uncooperative-peer attack. They present the results of both passive and active measurements. They developed a crawler that contacts all the peers in any given swarm, determines whether the swarm is under attack, and identifies the attack peers in the swarm. Using passive measurements, they performed a detailed analysis of a recent album that is under attack. While for seeders, Prithula [41] consider two natural seed attacks: the bandwidth attack and the connection attack.We take a three-prong approach to analyze these attacks. They created their own private torrents within PlanetLab; carefully analyzed the connection management and seeding algorithms in open-source BitTorrent seeds; constructed a simple fluid model which provides additional insights into the empirical results. They also studied how torrents can be discovered in their early stages. The observations and conclusions can help P2P developers to design highly-resilient P2P systems.

2.3 Application and Extension

2.3.1 Cross-ISP Traffic

The load peer-to-peer traffic creates on ISPs has been discussed for a few years. There are mainly two categories of work in peer-to-peer content replication: evaluation and architectural works.

The evaluation works focus on understanding the impact of peer-to-peer traffic on ISPs and on how locality can help out of the context of a specific implementation of locality.

Bindal et al. [24] present the impact of a deterministic locality policy on ISPs' peering links load and on end-users experience. They show that, in the scenario they consider, their biased neighbor selection significantly reduces inter-ISP traffic with a minor impact on the end-users download time, as long as the seed is four times faster than a leecher.

Karagiannis et al. [25] first introduced the notion of locality in the context of peer-to-peer content replication. First, they monitored the BitTorrent traffic flowing through the access link of an edge network. They show that 70 − 90% of all the contents downloaded on the local network was downloaded from external peers. Their conclusion is that peer-assisted locality distribution is an efficient solution for both the ISPs and the end-users.

The architectural works propose new architectures to implement locality policies. Those works build on the results provided by the evaluation works.

P4P [26] is a project whose aim is to provide a light-weight infrastructure to allow cooperation between peer-to-peer applications and ISPs. The P4P architecture is based on two components: the ISP tracker (itracker) that is deployed by each ISP and the application tracker (app-tracker) that is deployed by the content provider. Results derived from P4P suggest that locality reduces peer-to-peer traffic by up to 50 − 70% with increased performances.

Aggarwal et al. [27] present an architecture that is similar by some aspects to P4P. The authors define the notion of oracle that is supplied by ISPs in order to propose a list of neighbors to peers.

Another approach that requires no dedicated infrastructure is Ono. Ono clusters users based on the assumption that clients redirected to a same CDN server are close.

In all those approaches, only a fraction of the traffic is kept local in order preserve the robustness of the torrent.

Stevens et al. [28] provide new insights to content providers and ISPs. The overhead decreases linearly with locality. The capacity of the initial seed is critical to have a low peer download completion time and overhead with a high locality. High locality values enable a low overhead and slowdown in a large variety of scenarios. In

case of churn, a fraction of peers do not complete with high locality values, they identified the issue and proposed a solution that consist in a modification of the BitTorrent algorithm that manages the reconnection to the tracker.

2.3.2 Streaming Media Distribution

Prior work on peer-to-peer (or peer-assisted) streaming can be classified into either live streaming or on-demand streaming. These systems typically use either a tree-based or a data-driven approach. Tree-based approaches are typically based on application-level multicast architectures, in which the data is propagated through one or more relatively static spanning trees. Such application-level solutions have mainly been used for live streaming. Related treebased approaches using cache-and-relay have also been proposed for on-demand streaming. In cache-andrelay systems, each peer receives content from one or more parents and stores it in a local cache, from which it can later be forwarded to clients that are at an earlier playback point of the file. The tree-based approaches work best when peer connections are relatively stable.

Parvez et. al [42] provide insight into transient and steady-state system behavior, and help explain the sluggishness of the system with strict In-Order streaming. They also provide quantitative results on the startup delays and retrieval times for streaming media delivery. The results provide insights into the optimal design of peer-to-peer networks for on-demand media streaming.

2.3.3 Replication in Bandwidth-Symmetric Networks

Most papers on BitTorrent focus on heterogeneous end-users sharing files on the Internet. A few researchers focus in homogeneous local environments with symmetric bandwidth properties. A notable exception is the work presented in [43] and [44], where BitTorrent-based data distribution on LAN-based desktop grids is studied. The authors show by experiments that BitTorrent clearly outperforms FTP for the dissemination of large files over a LAN, and present an enhancement of the protocol that improves the performance for small file distribution as well. However, replication mechanisms such as the one we present in this paper are not considered.

Replication and caching have been widely researched in a variety of contexts. M. Meulpolder [45] aim to improve the performance of bandwidth-symmtric networks with a novel mechanism for replication using so-called replicators, which replicate a subset of the files in the system. The results show that Replicated BitTorrent significantly improves download times in local bandwidth-symmetric BitTorrent networks.

2.3.4 Wireless ad hoc networks

Several works tried to adapt BitTorrent to wireless ad hoc networks (e.g. [46] and [47]). They only focus on the tuning of the peer discovery phase without addressing the efficiency of the content sharing itself. Michiardi et al. study in [48] the performance of a cooperative mechanism to distribute content from one source to a potentially large number of destinations. They propose to deploy BitTorrent with a minor change allowing neighbor discovery and traffic locality. This is done by selecting only near neighbors as effective neighbors. The result is a decrease in the total download time and energy consumption. While M. K. Sbai [49] go beyond by focusing not only on the download time but also on the sharing among peers which we will show to further improve the download time as well.

In summary, there is prosperous development about BitTorrent performance enchancement, BitTorrent-like applications or extensions. Many simulators are developed to study all the possibilities which BitTorrent can be applied in, such as NS2, GPS [50], UTAPS [51], P2PLab [52], Top-BT [53], etc. But simulator has its inherent drawback that can not present approximately to real world as stated in section 1. Thereby, we need to design an emulator to solve this problem.

3. Motivation

As stated in the introduction section, though experiment in real environment can get realistic result, it is difficult and expensive to set up, limited in size and complexity. It will be interfered with production networks and restricted to existing technologies. Moreover, its reproducibility is not good.

As for simulation, it builds a synthetic enironment for running representations of code. It can be fully controled over target platform. Simulation provides a feasible method for investigation of complex network topologies and conditions. It is not limited by speed of simulation hardware, has low cost and good flexibility. But it is hard to model network traffic, might fail to mimic subtlties of real code.

Thereby, between simulation and real-world experimentation these two approaches, we need to find a compromise way, which can not only simulate efficiently a large number of nodes using a complex model, but can easy to control and modify, and the generated results are easy to reproduce.

This paper explores an intermediate solution using BitTorrent-like emulation (use of the real studied application within a configured synthetic environment) and virtualization (allows to share a resource between several instances of an application), and shows that such

an approach can provide interesting results when used to study BitTorrent-like systems.

Emulation and Simulation need to be defined and distinguished here.

Emulation consists in providing a modified environment to the studied application, to match the conditions of the experiment. It is combination of simulation and implementation. It provides semi-synthetic enviroment for running cod, that is, on one hand it offers real network implementation and supplementary means for introducing synthetic delays and faults. On the other hand, it provides a virtual network to networked devices and applications.

In emulation, applications can run on unmodified real devices or systems, can be deployed in a configurable Internet-like environment. Generally, emulation is an enviroment which is configurable, controlled and reproducible and can generate real network traffic.

For example, when studying peer-to-peer systems, network emulation is important, for instance, real network traffic can pass through emulator. While emulation of different types of hard disks is probably not necessary.

Simulation is the use of a model to represent over time essential characteristics of a system under study, attempts to predict aspects of the behaviour of some system by creating an approximate (mathematical) model of it. This can be done by physical modelling, by writing a special-purpose computer program or using a more general simulation package, probably still aimed at a particular kind of simulation (e.g. structural engineering, fluid flow). Normally, simulator is software runs in a single computer to simulate another system.

4. Virtual BT Design

4.1 Overview

Virtual BT (Virtual BitTorren) is our emulator for studying BitTorrent-like systems. It targets high efficiency (large number of virtual nodes can be studied on a low number of physical nodes without disk I/O operation), and scalability (experiments can be done with thousands of nodes).

Virtual BT virtualizes at the process level, not at the system level, to provide better scalability. It runs on Linux, which makes the emulator high efficient. A decentralized approach is used to emulate network topologies, allowing better scalability.

First, we will verify that Linux (2.6.22) is a suitable platform for Virtual BT by checking that its scheduler is better than Windows platform. The system architecture of Virtual BT is presented in section 4.4.

4.2 Linux Scheduler vs Windows Scheduler

While most virtualization systems virtualize on the operating system level, it is not mandatory here since the goal is to study peer-to-peer systems. It was therefore decided to virtualize the process' network identity by binding each process to its own IP address.

The Linux operating system was chosen for Virtual BT which is written by C language because of the availability and efficiency. But it was still necessary to test whether Linux was a suitable platform to run a very large number of processes without compromising our experiment's results. Every part of the scheduler is guaranteed to execute within a certain constant amount of time regardless of how many tasks are on the system. This allows the Linux kernel to efficiently handle massive numbers of tasks without increasing overhead costs as the number of tasks grows.

According to Johnaton Weare [54], we can get comparison results about scheduler between Linux and Windows, shown in the following tables.

Table 1 Timeslice – Uniprocessor

Scheduler	Linux	Windows
Timeslice Range	10ms-200ms	10-120ms (Client) 120ms (Server)
Timeslice Default	100ms	20-60ms (Client) 120ms (Server)

Table 2 Timeslice – Multiprocessor

Scheduler	Linux	Windows
Timeslice Range	10ms-200ms	15-180ms (Client) 180ms (Server)
Timeslice Default	100ms	30-90ms (Client) 180ms (Server)

Table 3 Performance

Scheduler	Linux	Windows
Scheduling Latency (average)	0.009ms	2ms
Scheduling Latency (worse)	0.3ms	16ms

From the table 1-3, we can see that for uniprocessor, multiprocessor, and system latency, Linux scheduler is much better than Windows sheduler. Further more, Linux is open source, there are many research projects are based on it, which can offer valuable references to our emulation.

4.3 Virtualization

Virtualization is made at the level of the process network identity: instances on the same physical system share all resources (filesystem, memory etc.) as normal processes do. However, each process has its own IP address on the network. The main IP address of each physical system is kept for administration purposes. The IP addresses of the virtual nodes are configured as interface aliases as shown in figure 1. On each physical node, IP addresses for virtual nodes are configured as interface aliases. Actually, most UNIX systems, including Linux and FreeBSD, allow each network interface to be assigned several IP addresses through an aliasing system.

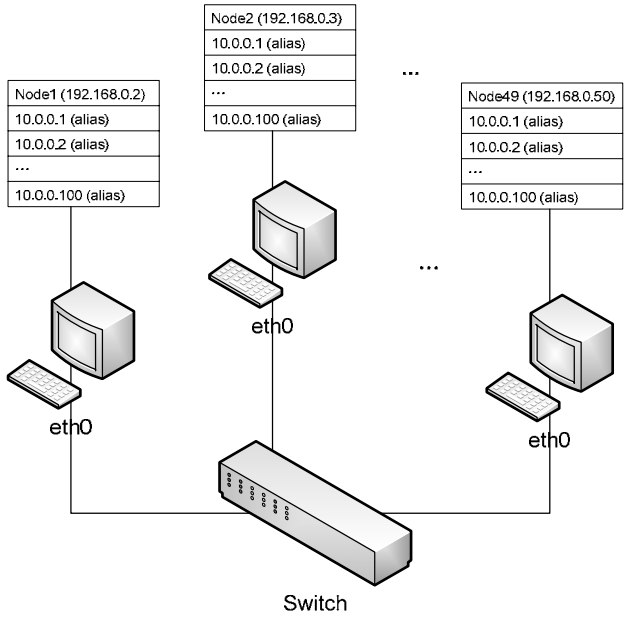


Fig. 1. IP Alias in Physical Node

Evaluation showed that interface aliases produced no overhead compared to the normal assignment of an IP address to an interface. To avoid namespace conflicts, the adresses of the virtual nodes were chosen in different subnet. Figure A shows an example configuration using the 192.168.0.2/24 network for administration and the 10.0.0.0/8 network for virtual nodes.

4.4 System Architecture

According to the BitTorrent Protocol Specification [6] and mainline BitTorrent application, the system architecture of Virtual BT is designed as figure 2 shown.

There are five main modules in the system: parse metainfo module, connecting to Tracker module, error handling module, running log module and exchanging data with peers module.

Note the system does not include disk buffer management. All the data exchange among memory in order to increase the data exchange efficieny and system performance.

Here are each module's functions:

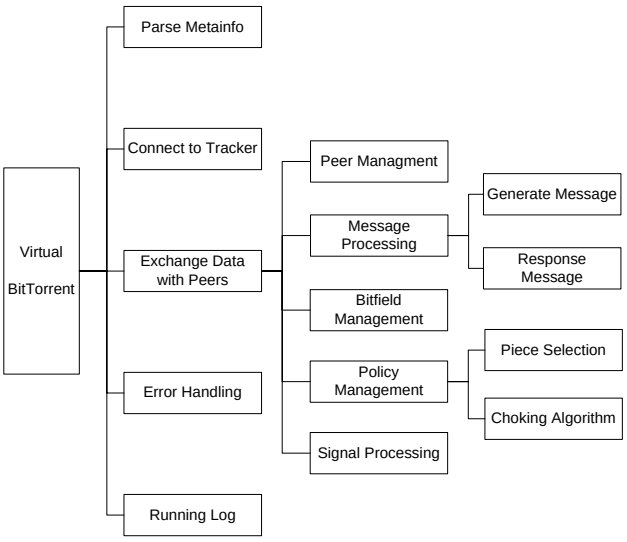


Fig. 2. Virtual BT System Architure

1. Parse Metainfo Module: Parse metainfo file, get tracker server's address, downloading file's name, piece length, and each piece's hash value.

2. Connecting to Tracker Module: According to HTTP protocol, get the request from peer's address, connect to tracker, parse tracker's response, and then get each peer's IP address and port number.

3. Error Handling Module: Define all the possible error types in the system, and process errors.

4. Running Log Module: Record running log, and save to files in order to view or analyze.

5. Exchanging data with Peers Module: According to the peer's IP address and port number, connect to peer, download data from peers, and upload data to peers. One of the main functions is responsible for sending and receiving messages, exchange messages in all peers. Core algorithms are choking algorithm and piece selection algorithm..

(1) Peer Management: Create peer struct type, manages peer linklist, adds and deletes peer node.

Peer struct type is the most important and complex data structure. We define seven statuses in header file. These statuses tranformation figure is shown in figure 3.

Halfshaed status means peer already sent handshaking message but still did not receive handshaking message from the other peer, or the peer already received the other peer's handshaking but did not send itself handshaking message. When two sides enter the Data status, they can exchange data, and four member variables (am_choking,

am_interested, peer_choking and peer_interested, as shown in figure 7 and 8) are available to use. These vaiables are the core members of BitTorrent choking algorithm.

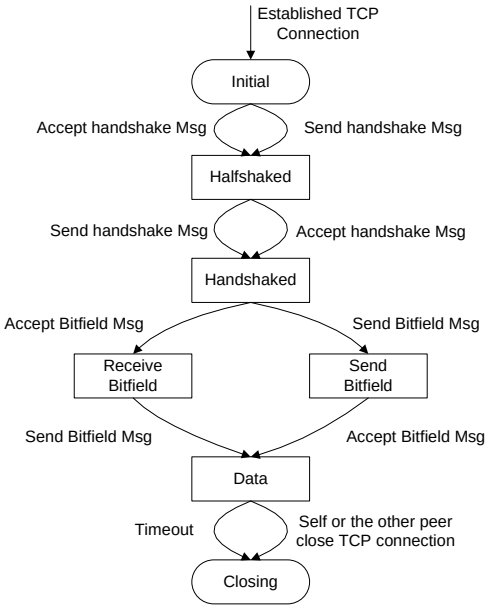


Fig. 3. Establish Connection Flowchart

(2) Message Processing: It is for peer to peer communication processed by sending and receiving message. This module will generate and send message, receive and process message according to the current status. Figure 7 and 8 are the transition diagrams for BitTorrent choking algothm in message processing module.

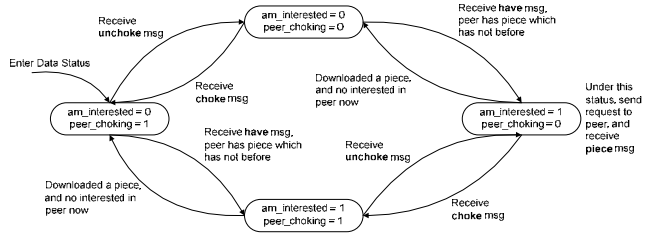


Fig. 4 Status Transition Diagram of a Connection from the Downloading Perspective

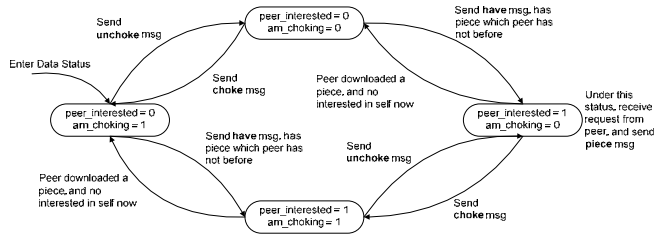


Fig. 5 Status Transition Diagram of a Connection from the Uploading Perspective

(3) Bitfield Management: Indicate which pieces are downloaded, and which pieces are not downloaed.

(4) Policy Management: Implement choking algorithm and piece selection algorithm.

(5) Signal Processing: Some processings before the program be terminated. E.g. Free memory, close file and close socket, etc.

In summary, this emulator is designed in C language, run on Linux operating system, and it is based on BitTorrent mainline 4.0.2. It enhances log module and takes out disk management module.

5. Vitual BT Network Emulation

5.1 Emulation Configuration

Current network topology emulators like a realistic emulation of the core network. But most peer-to-peer applications run on nodes on the edge of the Internet: while the traffic in the core of the Internet can influence the peer-to-peer system behaviour (congestion between providers can increase latency, for example), the main bottleneck for end nodes is often the link between the user and its Internet service provider (ISP). Therefore, it is possible to model the Internet by reproducing what the end node really sees, excluding what is less important from the end node point of view. Some features will take effect when we adopt ModelNet [55] in the future work.

Table 3 gives the emulation configuration which we successfully emulated using Virtual BT.

Table 3 Emulation Configuration

Hardware / Software	Configuration
Computer (Intel Core 2 1.86G/1G Memory)	4
Switch (D-Link DGS-1224T)	1
Operation System	Fedora 6
Seeding Software	Azureus 2.5
Content1 (Kung.Fu.Panda.rmvb)	333 MBytes
Content2 (Guns.N'Roses.rar)	52.7 MBytes
Content 1 Distribution	50 IP, 1 computer
Content 2 Distribution	200 IP, 2 computers

There are four computers in our experiments: one is for tracker, one is for seed, and the other two are for leechers. All computers installed with Fedora 6 are connected by a Gbps bandwidh switch. The whole emulation topology is shown in figure 6.

Network emulation is achieved in a decentralized way: each physical node is in charge of the network emulation

for its virtual nodes. Both incoming and outgoing packets are delayed by program control.

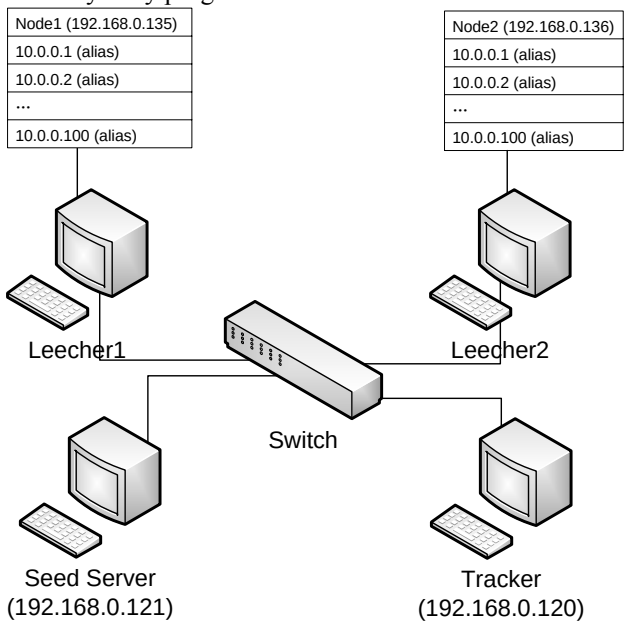


Fig. 6. Virtual BT Emulation Topoly

We implemented two experiments. Experiment one is for moive (Kung.Fu.Panda.rmvb) distribution. Leecher 1 computer generates 50 processes to get the content from seed server. Experiment two is for compressed file (Guns.N'Roses.rar) distribution. Leecher 1 and 2 generate 200 processes to get the content from seed server. 100 processes are generated by each leecher.

5.2 Virtual BT Performance

In this section, we implemented our emulation experiments, showing that Virtual BT is a suitable experimentation platform to study BitTorrent system.

5.2.1 Experiment 1

Experiment 1 is easy and fast to implement. Figure 7 shows the percentage of data transferred on each peer with time.

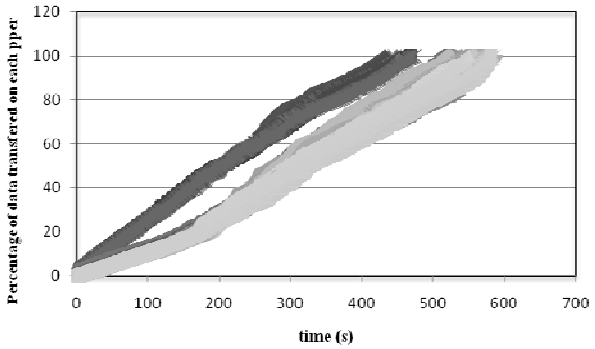


Fig. 7. Data Transferred Percentage in Experiment 1

From figure 7, we can see that peers distribute the movie quickly. We control peer only can receive data from seed with 500KBps rate, after a while, peer can exchange data with each other to complete download very fast. We will see more obvious in experiment 2.

5.2.2 Experiment 2

Experiment 2 adopts the same configuration. There are 200 processes join the file distribution. Figure 8 shows swarm in the early stage. Many file pieces are downloaded from seed server.

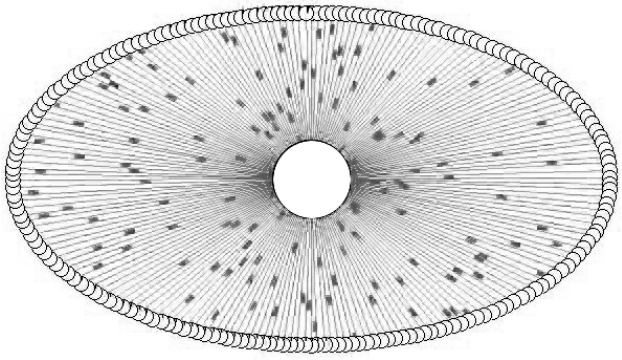


Fig. 8. Swarm in Early Stage in Experiment 2

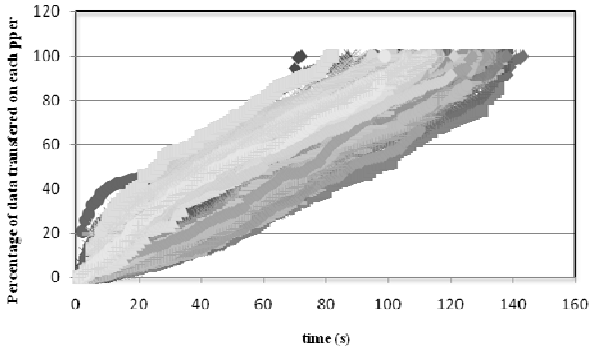


Fig. 9. Data Transferred Percentage in Experiment 2

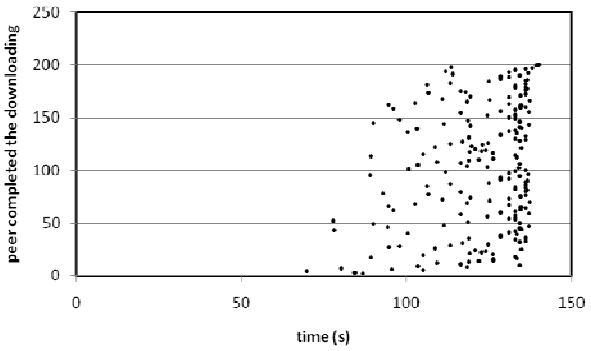


Fig. 10. Peer Completed the Downloading

In figure 9, due to more peers join the file distribution, a lot of peers speed up downloading fast. Data exchanging is more frequently than experiment 1, and it also help to decrease the distribution stress of seed server.

Figure 10 shows the peers completed downloading diagram. As the time goes by, more and more peers complted downloading and become to be seeds. The latter the time goes by, the more the seeds appear. This result is similar with the reference [19], which verified our emulation experiment is valid.

6. Conclusion and Future Work

6.1 Future Work

6.1.1 Emulator Performance Improvement

Virtual BT's emulation model is designed to control:

1. Bandwidth, latency and packet loss rate on the network link between the end node and its ISP. Now we only use *select()* function to control the packet transmission rate. We plan to apply ModelNet [55] to improve performance of our emulator.

2. Latency between group of nodes, allowing us to study problems involving locality.

Our decentralized network model doesn't consider the problem of congestion in the core links (between internet service providers). Its role in the performance of peer-to-peer systems needs to be determined through experiments on PlanetLab [5] or DSL-Lab [8]. Then, we will be able to modify our network model to include it.

We know that DummyNet plus IPFW based on FreeBSD can handle bandwidth control, too. DummyNet is a kind of modelling of networks as simple delay lines, which requires little hardware support. ModelNet is a kind of real-time network simulation and detailed modelling of virtual networks. We will compare two methods with Virtual BT in future to see which is more accurate.

We follow up the BitTorrent mainline development, test and study the important parameters, such as peer set size, torrent set size, interaction time period, etc., to optimize the performance to BitTorrent.

6.1.2 BitTorrent Performance Improvement

As the literature review stated, there are many aspects of applications improve their efficiency by using BitTorrent. But there exist phenomenon we can not neglect, that is, eMule or eDonkey is another killer application for P2P file sharing, it has several advantages that BitTorrent cannot achieve up to present. Here as the follows:

For searching, it is easier to search for files on the eD2k than BitTorrent. The only way to search for files on the BitTorrent network is through public tracker search engines or private trackers. There is no way for example to search multiple private trackers automatically, not to mention that you would have to be a registered member in all of them first.

For variety, perhaps the greatest advantage of the eD2k network is the variety of files. BitTorrent is most suited for new files, which are guaranteed to be released and downloaded very quickly. If you are looking for a good variety of files on the BitTorrent network, your best solution is to register on multiple private BitTorrent trackers that offer more specialized than generalized media.

For availability, the eD2k network by design is a sharing P2P network, not a trading one like BitTorrent. Files on the eD2k network stay alive almost forever, while files on the BitTorrent network usually die in a matter of few months or even weeks.

For sharing, as the Internet casual users, it is much easier to share files on the eD2k network. You just share the files on eMule and connect to a server, and then everyone connected to the eD2k network (not necessarily the same eD2k server) can find it. You also guarantee that the files will be shared and spread in a more efficient manner on the eD2k network, prolonging their life. On the other hand, with BitTorrent you need to run a tracker and publish the torrent, and even then users cannot find your file unless you post it on a website.

Despite these disadvantages that BitTorrent can not solve now, it also have some metris that eMule hardly achieve, either. For example, the speed, ease of use, etc. So how to combine the advantages of eMule with BitTorrent, make BitTorrent has more widely application space, increase "stickness" of user, these are open questiones for us. From the eMule's history record mechanism, mechanism design and social network, we can be inspired to boost BitTorrent protocol's performance in future. We need to find a better incentive

mechanism. DAMD (Distributed Algorithmic Mechanism Design) is a good way to try.

Finally, we need to extend Virtual BT to other peer-to-peer systems, to verify that it allows to answer wide range of questions on a wide range of systems. With the increase of the resources available on a single computer, emulation has been the target of a lot of research in the last years.

6.2 Conclusion

With the increase of the resources available on a single computer, emulation is used to build useful experimentation platforms, and are a promising tool to study peer-to-peer systems: they allow to use the real application on a large number of nodes in a configurable environment allowing reproduction of experiments.

This work shows our emulation platform only virtualizes what is needed to make the different virtual nodes look like real separate nodes from the outside: its network identity. This lightweight virtualization allows to maximize the system utilization ratio.

By detailed literature review, we see the extensive applications and improving for BitTorrent protocol. Our emulation platform, Virtual BT, enabled us to perform some experiments on the BitTorrent peer-to-peer system. During those experiments, Virtual BT proved to be scalable, efficient and useful. Through further development, it can be used to do more researches about BitTorrent performance improvement.

References

[1] B. Chun, D. Culler, T. Roscoe, A. Bavier, L. Peterson, M. Wawrzoniak, and M. Bowman. Planet-Lab: An Overlay Testbed for Broad-Coverage Services. *ACM SIGCOMM Computer Communication Review*, 33(3):00–00, July 2003.

[2] M. Izal, G. Urvoy-Keller, E. W. Biersack, P. A. Felber, A. Al Hamra, and L. Garces-Erice. Dissecting BitTorrent: five months in a torrent's lifetime. In *PAM'2004, 5th annual Passive & Active MeasurementWorkshop, April 19-20, 2004, Antibes Juan-les-Pins, France / Also Published in Lecture Notes in Computer Science (LNCS), Volume 3015, Barakat, Chadi; Pratt, Ian (Eds.) 2004, XI, 300p - ISBN: 3-540-21492-5*, Apr 2004.

[3] J. A. Pouwelse, P. Garbacki, D. H. Epema, and H. J. Sips. The Bittorrent P2P file-sharing system: Measurements and analysis. In *4th International Workshop on Peer-to-Peer Systems (IPTPS)*, Feb 2005.

[4] D. Qiu, R. Srikant. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 367–378, New York, NY, USA, 2004. ACM Press.

[5] A. R. Bharambe, C. Herley. Analyzing and improving BitTorrent performance. Technical Report MSRTR-2005-03, Microsoft Research, 2005.

[6] B. Cohen. The BitTorrent Protocol Specification, Feb 2008. http://www.bittorrent.org/beps/bep_0003.html

[7] Choffnes, D., F. Bustamante. Taming the torrent: a practical approach to reducing cross-ISP traffic in peer-to-peer systems. In *Proc. of ACM SIGCOMM*, 2008.

[8] B. Cohen. Incentive Build Robustness in BitTorrent, in *Conf. Electronic Commerce*, 2003.

[9] A. Legout, G. Urvoy-Keller, and P. Michiardi, Rarest first and choke algorithms are enough, in *Proc. of IMC*, Rio de Janeiro, Brazil, October 2006.

[10] A. Legout, N. Liogkas, E. Kohler, and L. Zhang, Clustering and sharing incentives in bittorrent systems, in *Proc. of SIGMETRICS*, San Diego, CA, USA, June 2007.

[11] T. Locher, P. Moor, S. Schmid, and R.Wattenhofer, Free riding in bittorrent is cheap, in *Proc. of HotNets-V*, Irvine, CA, USA, November 2006.

[12] Y. Tian, D. Wu, and K.-W. Ng, Modeling, analysis and improvement for bittorrent-like file sharing networks, in *Proc. of INFOCOM*, Barcelona, Spain, April 2006.

[13] M. Izal, G. Urvoy-Keller, E. W. Biersack, P. A. Felber, A. A. Hamra, and L. Garc´es-Erice, Dissecting bittorrent: Five months in a torrent's lifetime, in *Proc. of PAM*, Juan-les-Pins, France, April 2004.

[14] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang, Measurement, analysis, and modeling of bittorrent-like systems, in *Proc. Of IMC*, New Orleans, LA, USA, October 2005.

[15] D. Qiu and R. Srikant, Modeling and performance analysis of bittorrent-like peer-to-peer networks, in *Proc. of SIGCOMM*, Portland, Oregon, USA, August 2004.

[16] J.A. Pouwelse, P. Garbacki, D.H.J. Epema, H.J. Sips, A Measurement Study of the BitTorrent Peer-to-Peer File-Sharing System, Elsevier Science, 2004

[17] P. Marciniak, N. Liogkas and A. Legout, et al. Small is not always beautiful. In *IPTPS*, 2008.

[18] A. Legout, Understanding BitTorrent: An Experimental Perspective, INRIA, 2005.

[19] A. Al-Hamra, A. Legout, and C. Barakat, Understanding the properties of the bittorrent overlay, INRIA, Tech. Rep., 2007. [Online]. Available: http://arxiv.org/pdf/0707.1820

[20] G. Urvoy-Keller and P. Michiardi, Impact of inner parameters and overlay structure on the performance of bittorrent, in *Proc. of Global Internet Symposium*, Barcelona, Spain, April 2006.

[21] F. Benbadis, N. Hegde and F. Mathieu, Playing with the bandwidth: Conservation Law. In *P2P*, 2008.

[22] D. Carra, G. Neglia, P. Michiardi, On the impact of greedy strategies in BitTorrent networks: the case of BitTyrant. INRIA, in *P2P*, 2008.

[23] C. Dale, J. Liu, J. Peters and B. Li. Evolution and enhancement of bittorrent network topologies, in *Internet Workshop on QoS*, 2008.

[24] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates, and A. Zhang. Improving traffic locality in bittorrent via biased neighbor selection. In *Proc. Of ICDCS'06*, Lisboa, Portugal, July 2006.

[25] T. Karagiannis, P. Rodriguez, and K. Papagiannaki. Should internet service providers fear peer-assisted content

distribution? In *Proc. of IMC'05*, Berkeley, CA, USA, October 2005.
[26] H. Xie, Y. R. Yang, A. Krishnamurthy, Y. Liu, and A. Silberschatz. P4P: Provider portal for applications. In *Proc. of ACM SIGCOMM*, Seattle, WA, USA, August 2008
[27] V. Aggarwal, A. Feldmann, and C. Scheideler. Can ISPs and p2p users cooperate for improved performance? *Proc. of CCR*, July 2007.
[28] Stevens Le Blond, Arnaud Legout and Walid Dabbous, Pushing BitTorrent Locality to the Limit, INRIA, December, 2008.
[29] M. Piatek, T. Isdal, T. Anderson, A. Krishnamurthy, and A. Venkataramani. Do incentives build robustness in BitTorrent? In *NSDI*, 2007.
[30] L. Massoulié and M. Vojnovi´c. Coupon Replication Systems. In *ACM SIGMETRICS*, 2005.
[31] F. Wu and L. Zhang. Proportional response dynamics leads to market equilibrium. In *ACM STOC*, 2007.
[32] S. Jun and M. Ahamad. Incentives in BitTorrent induce free riding. In *P2PEcon*, 2005.
[33] P. Garbacki, D. H. Epema, and M. van Steen. An amortized tit-for-tat protocol for exchanging bandwidth instead of content in P2P networks. In *SASO*, 2007.
[34] M. Sirivianos, J. H. Park, X. Yang, and S. Jarecki. Dandelion: Cooperative content distribution with robust incentives. In *USENIX*, 2007.
[35] D. Levin, K. LaCurts, N.Spring and B. Bhattacharjee. BitTorrent is an Auction: Analyzing and Improving BitTorren's Incentives. In *SIGCOMM*, 2008.
[36] D. Stutzbach, R. Rejaie, Understanding churn in Peer-to-Peer Networks, In *IMC*, 2006
[37] S. Sen and J. Wang. Analyzing Peer-To-Peer traffic across large networks. *IEEE/ACM Transactions on Networking*, 12(2), 2004.
[38] X. Lou and K. Hwang. Adaptive content poisoning to prevent Illegal file distribution in P2P networks. *IEEE Trans. Computer*, April 2008.
[39] M. Dischinger, A. Mislove, A. Haeberlen, et al. Detecting bittorrent blocking. In *IMC*, 2008.
[40] P. Dhungel, D. Wu, B. Schonhorst and Keith W. Ross. A Measurement Study of Attacks on BitTorrent Leechers, IPTPS, 2008.
[41] P. Dhungel, X. Hei, D. Wu, and Keith W. Ross. The seed attack: can bittorrent be nipped in the bud? Technical Report. 2008.
[42] N. Parvez, C. Williamson, A. Mahanti, et al. Analysis of bittorrent-like protocols for on-demand stored media streaming. In *SIGMETRICS*, 2008.
[43] B. Wei, G. Fedak, and F. Cappello. Collaborative data distribution with bittorrent for computational desktop grids. In *Proc. of the The 4th International Symposium on Parallel and Distributed Computing (ISPDC'05)*, pages 250–257, Washington, DC, USA, 2005. IEEE Computer Society.
[44] B. Wei, G. Fedak, and F. Cappello. Scheduling independent tasks sharing large data distributed with bittorrent. In *Grid Computing Workshop*, pages 219–226. IEEE Computer Society, 2005.
[45] M. Meulpolder, D.H.J. Epema, H.J.Sips. Replication in bandwidth-symmtric bittorrent networks. In *IPDPS*, 2008.
[46] A. Nandan, S. Das, G. Pau, M. Gerla, Cooperative downloading in vehicular ad hoc networks, In *WONS*, Washington, USA, 2005.
[47] S. Rajagopalan, C-C. Shen, A cross-Layer Decentralized BitTorrent for Mobile Ad hoc Networks, In *MOBIQUITOUS*, San Jose, USA, 2006.
[48] Michiardi P., Urvoy-Keller G., Performance analysis of cooperative content distribution for wireless ad hoc networks, In *WONS 2007*, Obergurgl.
[49] M. K. Sbai, C. Barakat, J. Choi. Adapting BitTorrent to wireless ad hoc networks. *Ad hoc Now, Sophia Antipolis: France*, 2008
[50] W. Yang, N. Abu-Ghazaleh, GPS: A General Peer-to-Peer Simulator and its Use for Modeling BitTorrent, In *MASCOTS*, 2005
[51] W. Li, S. Chen, T. Yu. UTAPS: An Underlying Topology-Aware Peer Selection Algorithm in BitTorrent, *22nd Conf. Advanced Information Networking Application*, 2008
[52] P2Plab.http://perso.ens-lyon.fr/lucas.nussbaum/p2plab.php, available in December. 2008.
[53] Top-BT. http://www.cse.ohio-state.edu/~sren/topbt/, available in December. 2008.
[54] Kernel Comparison: Linux (2.6.22) vs. Windows (Vista). http://widefox.pbwiki.com/Kernel%20Comparison%20Linux%20vs%20Windows. Available in Jan. 2009.
[55] Diwaker Gupta, Kashi Vishwanath, and Amin Vahdat. Die Cast: Testing Distributed Systems with an Accurate Scale Model. In *Proceedings of the 5th ACM/USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, April 2008

Speeding Up Homomorphic Hashing Using GPUs

Kaiyong ZHAO

Abstract

Network coding has recently been widely investigated in the area of peer-to-peer networks in order to improve the system throughput and/or robustness. Such systems are prone to pollution attacks, however, because a single polluted data packet from any malicious peer will be encoded with other genuine data packets and the pollution will be propagated to the whole network at an exponential speed. Homomorphic hash functions are currently the only known approach to defending the pollution attacks, but unfortunately they are computationally expensive for contemporary CPUs. This paper proposes to exploit the computing power of Graphic Processing Units (GPUs) for homomorphic hashing. Specifically, we demonstrate how to use NVIDIA GPUs and the Computer Unified Device Architecture (CUDA) programming model to achieve more than 60 times of speedup over the CPU counterpart. We also developed a multi-precision modular arithmetic library on CUDA platform. The library is not only key to our specific application, but also very useful for a large number of cryptographic applications.

Keywords

Network coding, homomorphic hash function, GPU computing, CUDA

1. Introduction

In recent years, peer-to-peer (P2P) content distribution applications such as BitTorrent and ppLive, have become the most popular Internet applications due to their scalability and robustness. Network coding has been proposed as an effective mechanism to improve the performance of such P2P applications [14]. However, P2P applications with network coding suffer from the notorious pollution attacks: a malicious node can send out bogus packets which will be merged into other genuine packets and propagated into the whole network at an exponential speed. To resolve this problem, homomorphic hash functions have to be applied such that the hash of any encoded packet can be effectively derived from the hashes of the original packets, which enables the detection of bogus packets before a peer encodes it with other packets [12]. Unfortunately homomorphic hash functions rely on multiple-precision modular operations and are computationally expensive [10] [12].

Recent advances in Graphics Processing Units (GPUs) opens a new era of GPU computing [20]. For example, commodity GPUs like NVIDIA's GTX 280 has 240 processing cores and can achieve 933 GFLOPS of computational horsepower. More importantly, the NVIDIA CUDA programming model makes it easier for developers to develop non-graphic applications using GPU [1] [4]. In CUDA, the GPU becomes a dedicated coprocessor to the host CPU, which works in the principle of Single-Program Multiple Data (SPMD) where multiple threads based on the same code can run simultaneously.

In this paper, we propose to use GPU for homomorphic hashing. The homomorphic hash function needs to multiply a large number of exponentiations, so the critical part is to optimize the exponentiation operation for high-precision large integers. To this end, we focus our work on the design, implementation, and optimization of exponentiation operations on GPU. In order to fully utilize the computing power of GPU, it is highly desirable to create thousands of threads. A common configuration of homomorphic hash function requires calculating hundreds of exponentiations, hence we propose to perform multiple homomorphic hashing simultaneously.

The contribution of this work is threefold:

- First, we designed and implemented a fast exponentiation algorithm using Montgomery reduction and also precomputation, for the CUDA architecture.
- Second, we achieved 8979 Kbps of throughput using a contemporary graphic card, which is more than 60 times of the CPU counterpart with similar working frequencies.
- Third, we developed a multiple-precision modular arithmetic library for CUDA, which could be used in lots of security applications such as RSA, and ElGamal schemes.

The rest of the paper is organized as follows. Section 2 provides background information on network coding, homomorphic hash functions, GPU architecture, and CUDA programming model. Section 3 presents the design of multiple-precision modular arithmetic on GPU. Section 4 presents the parallel implementation of homomorphic hash function. Experimental results are presented in Section 5, and we conclude the paper in Section 6.

2. Background and Related Work

In this section, we provide the required background knowledge of network coding, homomorphic hash functions, GPU architecture and CUDA programming model.

2.1 Network Coding

In traditional communications networks, the intermediate nodes simply perform data forwarding. Recently, a large number of works focus on applying network coding to improve network performance. The seminal work of network coding has been studied in [7], which showed that a multicast session can achieve the data rate of multicast upper bound if network nodes are allowed to perform coding. Later it has been further shown that linear network coding is sufficient to achieve the multicast capacity [9]. Linear network coding regards the messages as vectors of elements in a finite field, and the encoding function is a simple linear combination over the finite field. The framework of random network coding was proposed in [8], which shifts network coding research from theory to practical applications.

The Avalanche system from Microsoft Research exploits the random linear network coding in P2P content distribution [11]. LAVA and R2 P2P live streaming systems also uses random linear network coding [15] [16]. Network coding has also been applied to distributed storage systems, wireless networks, and sensor networks [14]. The following shows a general framework of random linear network coding in P2P applications.

The data to be distributed is divided into n blocks $(b_1, b_2, \ldots, b_n)$, where each block b_i is further divided into m codewords $b_{i,k}$, $k \in \{1,\ldots,m\}$. An encoded block e_j is a linear combination of the n original blocks and it is also divided into m codewords $e_{j,k}$, $k \in \{1,\ldots,m\}$. The linear relationship between e_j and the original n blocks is described by its global coefficient vector $(c_{j,1}, c_{j,2}, \ldots, c_{j,n})$, i.e., $e_{j,k} = \sum_{i=1}^{n} c_{j,i} \cdot b_{i,k}$, $k \in \{1,\ldots,m\}$. In a P2P application, a peer receives encoded data blocks from upstream peers, and also creates and disseminates encoded data blocks to its downstream peers. Notice that the global coefficient vector should be sent along with the encoded data block. A peer can recover/decode the original n blocks as soon as it has received n linearly independent coded blocks $(e_1, e_2, \ldots, e_n)$, by solving the set of linear equations $e_{j,k} = \sum_{i=1}^{n} c_{j,i} \cdot b_{i,k}$, $k \in \{1,\ldots,m\}$, $j \in \{1,\ldots,n\}$.

A final remark is that, the above operations could be implemented in finite fields such as prime fields GF(p) or extension fields GF(pr) where p is a prime number.

2.2 Homomorphic Hash Functions

P2P networks are prone to the pollution attacks in which bogus data blocks are disseminated into the network by malicious peers. When network coding is not deployed, each peer will receive original data blocks directly from other peers, and hence it is possible to use hash functions such as SHA1 to verify the correctness of a data block simply by comparing the hash of each received data block to the corresponding hash provided by the source. A hash function maps a large bit stream to a shorter one with a fixed length. Given a hash value, it is computationally difficult to find another bit stream which can result in the same hash value.

For P2P networks with network coding, the effect of pollution attack becomes more serious and more difficult to handle [12] [13] [19]. First of all, each bogus block could be mixed with valid blocks and propagated throughout the network; secondly, the standard hash functions cannot be applied here because a peer receives random encoded packets which cannot be predetermined by the source. Homomorphic hash functions are currently the only solution to this security issue, which enable a peer to detect the bogus data block once it has been received. Homomorphic hash functions have the property that the hash value of a linear combination of the input blocks can be constructed by the hash values of those input blocks. One such homomorphic hash function, $h(\cdot)$, has been proposed in [10], which requires to decide a set of hash parameters $G = (p, q, \mathrm{g})$ first. The parameters p and q are large prime numbers of order λ_p and λ_q chosen such that $q \mid p-1$. The parameter g is a vector of m numbers, each of which can be written as $x^{(p-1)/q}$ mod p where $x \in Z_q$ and $x \neq 1$. The method of creating the parameter set can be found at [10]. The homomorphic hash of a block b_i is then calculated as

$$h(b_i) = \prod_{k=1}^{m} g_k^{b_{k,i}} \bmod p \quad . \qquad (1)$$

Following the notations in Section 2.1, the hash values of the original blocks $(b_1, b_2, \ldots, b_n)$ are $h(b_1), h(b_2), \ldots, h(b_n)$ respectively. Given an encoded block e_j with coefficient vector $(c_{j,1}, c_{j,2}, \ldots, c_{j,n})$, the homomorphic hash function $h(\cdot)$ can be shown to satisfy the condition that $h(e_j) = \prod_{i=1}^{n} h^{c_{j,i}}(b_i)$. This property can be used to verify the authenticity of an encoded block. Typical values of the parameters are summarized in Table I. The verification process is illustrated in Figure 1.

Although the homomorphic hash function can theoretically solve the pollution attack problem, it is unfortunately computationally expensive for today's desktop CPUs. A 3 GHz Pentium 4 CPU can only achieve around 300 Kbps of throughput for verifying a single 16 KB data block, using the parameters in Table I.

To avoid downloading all the hash values of original data blocks, the authors in [13] designed a trapdoor homomorphic hash function which is even more computationally complicated than Eq. (1). Some compromised solutions have been proposed to address the computational difficulty. In [12], a cooperative scheme is proposed to prevent the propagation of bogus packets by probabilistically verifying packets and informing other nodes when detected a malicious node. This scheme cannot totally remove bogus packets from the network, however.

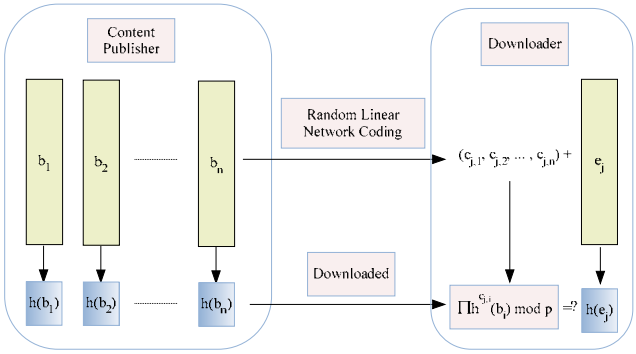


Figure 1: Data verification using homomorphic hash function in network coding enabled P2P applications

Table 1. Homomorphic hashing function parameters

Name	Description	Typical Value
λ_p	Discrete log security parameter	1024 bit
λ_q	Discrete log security parameter	257 bit
p	Random prime, $\lvert p \rvert = \lambda_p$	
q	Random prime, $\lvert q \rvert = \lambda_q$, $q \mid p-1$	
m	Number of codewords per data block	512
n	Number of data blocks	128
g	1 x m vector of order q	

2.3 GPU Computing and CUDA

GPUs are dedicated hardware for manipulating computer graphics. Due to the huge computing demand for real-time and high-definition 3D graphics, the GPU has evolved into a highly parallel, multithreaded, manycore processor. The advances of computing power in GPUs have driven the development of general-purpose computing on GPUs (GPGPU). The first generation of GPGPU requires that any non-graphics application must be mapped through graphics application programming interfaces (APIs).

Recently one of the major GPU vendors, NVIDIA, announced their new general-purpose parallel programming model, namely Compute Unified Device Architecture (CUDA) [1] [4], which extends the C programming language for general-purpose application development. Meanwhile, another GPU vendor AMD also introduced Close To Metal (CTM) programming model which provides an assembly language for application development [2]. Intel also exposed Larrabee, a new many-core GPU architecture specifically designed for the market of GPU computing this year [23].

Since the release of CUDA, it has been used for speeding up a large number of applications [17] [18] [20] [21] [22].

The NVIDIA GeForce 8800 has 16 Streaming Multiprocessors (SMs), and each SM has 8 Scalar Processors (SPs), resulting a total of 128 processor cores. The SMs have a Single-Instruction Multiple-Data (SIMD) architecture: At any given clock cycle, each SP of the SM executes the same instruction, but operates on different data. Each SP can support 32-bit single-precision floating-point arithmetic as well as 32-bit integer arithmetic.

Each SM has four different types of on-chip memory, namely registers, shared memory, constant cache, and texture cache. For GeForce 8800, each SM has 8192 32-bit registers, and 16 Kbytes of shared memory which are almost as fast as registers. Constant cache and texture cache are both read-only memories shared by all SPs. Off-chip memories such as local memory and global memory have relatively long access latency, usually 400 to 600 clock cycles [4]. The properties of the different types of memories have been summarized in [4] [17]. In general, the scarce shared memory should be carefully utilized to amortize the global memory latency cost. Shared memory is divided into equally-sized banks, which can be simultaneously accessed. If two memory requests fall into the same bank, it is referred to as bank conflict, and the access has to be serialized.

In CUDA model, the GPU is regarded as a coprocessor capable of executing a great number of threads in parallel. A single source program includes host codes running on CPU and also kernel codes running on GPU. Compute-intensive and data-parallel kernel codes run on GPU in the manner of Single-Process Multiple-Data (SPMD). The threads are organized into blocks, and each block of threads are executed concurrently on one SM. Threads in a thread block can share data through the shared memory and can perform barrier synchronization. Each SM can run at most eight thread blocks concurrently, due to the hard limit of eight processing cores per SM. As a thread block terminate, new blocks will be launched on the vacated SM. Another important concept in CUDA is warp, which is formed by 32 parallel threads and is the scheduling unit of each SM. When a warp stalls, the SM can schedule another warp to execute. A warp executes one instruction at a time, so full efficiency can only be achieved when all 32 threads in the warp have the same execution path. Hence, if the number of threads in a block is not a multiple of warp size, the remaining instruction cycles will be wasted.

3. Multiple-Precision Modular Arithmetic for CUDA

In this section, we present a set of library functions of multiple-precision modular arithmetic implemented on GPUs. These library functions are the cornerstones of the network coding system and homomorphic hash functions. It is of critical importance to implement these library functions efficiently. In modular arithmetic, all operations are performed in a group Z_m, i.e., the set of integers $\{0,1,2,\cdots,m-1\}$. In the following, the modulus m is represented in radix b as $(m_n m_{n-1}\cdots m_1 m_0)_b$ where $m_n \neq 0$. Each symbol $m_i, 0 \le i \le n$, is referred to as a radix b digit. Non-negative integers x and y, $x < m, y < m$, are represented in radix b as $(x_n x_{n-1}\cdots x_1 x_0)_b$ and $(y_n y_{n-1}\cdots y_1 y_0)_b$ respectively.

We have implemented the following multiple-precision library functions for CUDA:

- Multiple-precision comparison
- Multiple-precision subtraction
- Multiple-precision modular addition
- Multiple-precision modular subtraction
- Multiple-precision multiplication
- Multiple-precision division
- Multiple-precision multiplicative inversion

Due to the space limitation, we do not present the implementation details in this paper.

3.1 Montgomery Reduction

The classical modular multiplication is suitable for normal operations. However, when performing modular exponentiations, Montgomery multiplication shows much better performance advantage [5]. The following gives the Montgomery reduction and Montgomery multiplication algorithms.

Let m be a positive integer, and let R and A be integers such that $R > m$, gcd(m, R) = 1, and $0 \le A < m \cdot R$. The Montgomery reduction of A modulo m with respect to R is defined as $A \cdot R^{-1}$ mod m. In our applications, R is chosen as b^n to simply the calculation.

Algorithm 1 Multiple-precision Montgomery Reduction

INPUT: integer m with n radix b digits and gcd(m, b) = 1, $R = b^n$, $m' = -m^{-1} \bmod b$, and integer A with $2n$ radix b digits and $A < m \cdot R$.

OUTPUT: $T = A \cdot R^{-1} \bmod m$.

1: $T \leftarrow A$;
2: **for** (i **from** 0 **to** $n-1$)
3: $u_i \leftarrow T_i \cdot m' \bmod b$;
4: $T \leftarrow T + u_i \cdot m \cdot b^i$;
5: **end for**
6: $T \leftarrow T / b^n$;
7: **if** ($T \ge m$) **then** $T \leftarrow T - m$;
8: **return** T;

Algorithm 2 Multiple-precision Montgomery Multiplication

INPUT: non-negative integer m, x, y with n radix b digits, $x < m, y < m$, and gcd(m, b) = 1, $R = b^n$, $m' = -m^{-1} \bmod b$.

OUTPUT: $T = x \cdot y \cdot R^{-1} \bmod m$.

1: $T \leftarrow 0$;
2: **for** (i **from** 0 **to** $n-1$)
3: $u_i \leftarrow (T_0 + x_i \cdot y_0) \cdot m' \bmod b$;
4: $T \leftarrow (T + x_i \cdot y + u_i \cdot m) / b$;
5: **end for**
6: **if** ($T \ge m$) **then** $T \leftarrow T - m$;
7: **return** T;

3.2 Modular Exponentiation

Algorithm 3 Multiple-precision Montgomery Exponentiation

INPUT: integer m with n radix b digits and gcd(m, b) = 1, $R = b^n$, positive integer x with n radix b digits and $x < m$, and positive integer e = $(e_t \cdots e_0)_2$.

OUTPUT: $x^e \bmod m$.

1: $\tilde{x} \leftarrow Mont(x, R^2 \bmod m)$;
2: $A \leftarrow R \bmod m$;
3: **for** (i **from** n **down to** 0)
4: $A \leftarrow Mont(A, A)$;
5: **if** $e_i == 1$ **then** $A \leftarrow Mont(A, \tilde{x})$;
6: **end for**
7: $A \leftarrow Mont(A, 1)$;
8: **return** A;

4. Parallel Homomophic Hashing on GPUS

Fast exponentiation is critical to lots of cryptographic applications and has been extensively studied in the history. It is also the most important component of the homomorphic hash function. Some methods of fast exponentiation have been summarized in [6]. In this section, we present several parallel algorithms for homomorphic hashing based on different modular exponentiation methods.

4.1 Naïve Parallel Homomorphic Hashing

The homomorphic hash function has two steps: (1) perform m modular exponentiations; (2) perform m-1 modular multiplications. It is straightforward to implement the first step in parallel by distributing the m modular exponentiations to the GPU processing cores. Assume the GPU contains N cores, and each core takes time T_{exp} to calculate a single modular exponentiation, then step (1) will take time $(\lfloor m/N \rfloor + 1)T_{exp}$ to finish. It is also obvious to see that step (2) takes time $(\lfloor \log_2 m \rfloor + 1)T_{mul}$, where T_{exp} denotes the time of a single modular multiplication operation. Given the configuration of Table 1, T_{exp} is about two orders greater than T_{mul}. So our focus is to optimize step (1). Step(2) can be parallelized using standard reduction method. Due to the space limitation, we will not give the detailed algorithm for Step (2).

4.2 Parallel Homomorphic Hashing with Precomputation

When applying homomorphic hash function in network coding enabled P2P applications, the same homomorphic hash function, i.e., with the same set of parameters, will be used for a large data set such as a whole file or a video streaming session. Under this special circumstance, it is possible to speedup the modular exponentiations by precomputation [8]. To calculate g^e, we first represent the exponent e using radix $b = 2^k$: $e = \sum_{i=0}^{n-1} a_i b^i$, where $0 \le a_i < b$ and $a_{n-1} \neq 0$. It is easy to see that $n = \lceil (\lfloor \log_2 e \rfloor + 1)/k \rceil$. The fast modular exponentiation algorithm requires the precomputation of $g^{2^{ki}}$ mod m for $1 \le i \le n-1$. Then we can use the following algorithm to calculate g^e mod m.

Algorithm 4 Exponentiation with Precomputation

INPUT: integers *m*, *g*, $e = \sum_{i=0}^{m-1} a_i b^i$, R, and $Rg^{2^{ki}}$ mod *m* for $1 \le i \le n-1$

OUTPUT: g^e mod *m*.

1: $A \leftarrow R,\ B \leftarrow R$;
2: **for** (*j* **from** $b-1$ **down to** 1)
3: **for** *i* **from** 0 **to** $m-1$
4: **if** $a_i == j$ **then** $B \leftarrow Mont(B, Rg^{2^{ki}})$ mod *m*;
5: **end for**
6: $A \leftarrow Mont(A, B)$;
7: **end for**
8: $A \leftarrow Mont(A, 1)$;
9: **return** *A*;

The above algorithm takes $m + b - 3$ multiplications. For e with 257-bit, the optimal value of b is 16, which takes only 78 multiplications in the worst case, as compared with 512 multiplications required by the binary method. In theory, we can expect a speedup of 6.5 by using this algorithm.

5. Implementation and experimental Results

The CPU version of the homomorphic hash function is implemented in C language using the GNU MP arithmetic library, version 4.2.3 [3]. We have also implemented the different implementations of homomorphic hash function using CUDA. We tested these implementations on Inno3D GTX260 graphic card which contains an NVIDIA GeForce GTX260 GPU. The GTX260 GPU uses the GT200 architecture with 192 processing cores working at 1.24 GHz.

5.1 Homomorphic Hashing on CPU

Figure 1 shows the results of our CPU version of homomorphic hashing, running on a 1.6 GHz Quad-core Intel CPU. Since most of the computing task is the exponentiation, the throughput is almost independent of the value of m. The hashing throughput is around 130 Kbps, which is in accordance with the results reported in [10]. They used a 3.0 GHz CPU to achieve around 300 Kbps of throughput. Since the working frequency of our GPU is closer to the 1.6 GHz of our CPU, it is reasonable to use the hashing throughput of 1.6 GHz CPU for comparison purpose.

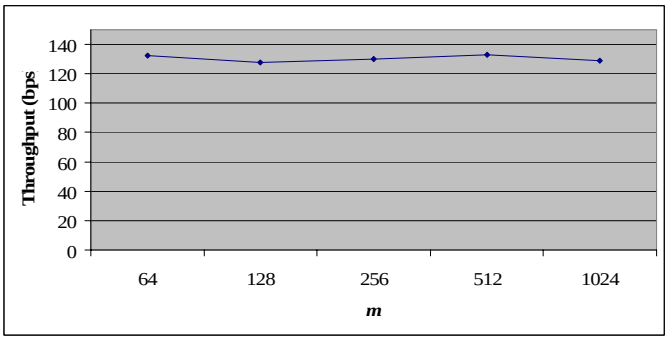


Figure 2. Throughput of homomorphic hashing on CPU

5.2 Naïve Parallel Homomorphic Hashing

The naïve parallel homomorphic hashing uses Algorithm 3 to calculate exponentiations. The CUDA architecture requires a large number of threads to hide the memory latency and to fully utilize the computing power. The number of threads per thread block (denoted by TB), and also the number of thread blocks (denoted by K), are the two main factors that affect the hashing throughput. We plot the throughput for different

configurations in Figure 3. It is easy to observe that more threads per block can generally achieve better throughput. When the number of threads per block is fixed, the throughput can be improved by creating more thread blocks. Since our GPU has 24 SMs, the number of thread blocks should be a multiple of 24. We found that using at least 6144 threads can achieve the best performance. This optimal configuration can be achieved by several combinations: TB=256 and K=24, or TB=128 and K=48, or TB=64 and K=96. To summarize, the best throughput can be achieved if the following conditions are satisfied: (1) the number of threads per block is a multiple of 32 (i.e., the warp size); (2) the number of thread blocks is a multiple of 24; (3) the total number of threads should be at least 6144. In order to create so many threads, it is necessary to perform the calculation of multiple homomorphic hashes together. For example, if m = 512, we should perform the homomorphic hashing for 12 different data blocks simultaneously.

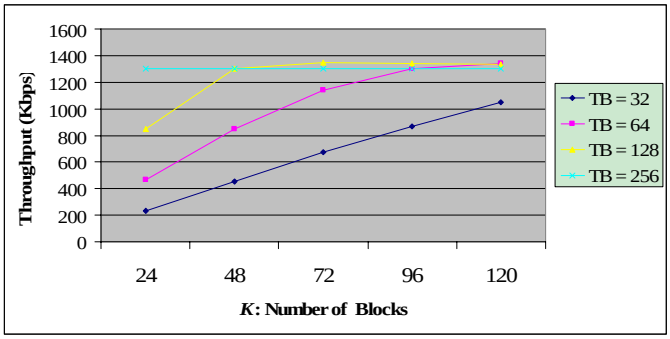


Figure 3. Throughput of homomorphic hashing on GPU usng Algorithm 3, with different number of threads per block (***TB***) and different number of thread blocks (***K***)

5.3 Parallel Homomorphic Hashing with Precomputation

The performance of the parallel homomorphic hashing with precomputation is shown in Figure 4. The effect of TB and K on the hashing throughput is very similar to the previous case. If we compare the results with those in Figure 3, we can observe a speedup of 6.8 for the highest throughput. This is very close to the theoretical speedup of 6.5 derived in Section 4.2. The highest hashing throughput is 8979 Kbps, which is achieved by using 256 threads per thread block and a total of 120 thread blocks. This is more than 60 times of the throughput achieved by our 1.6 GHz CPU!

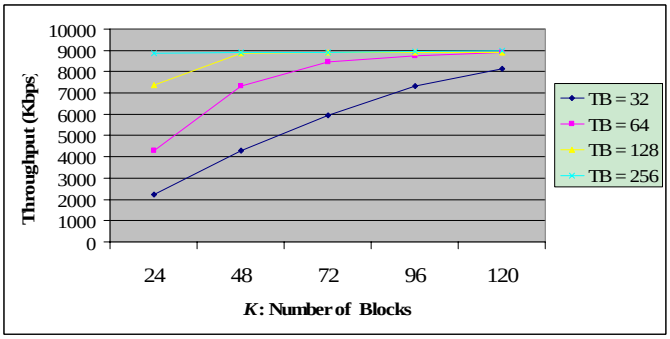


Figure 4. Throughput of homomorphic hashing on GPU using Algorithm 4, with different number of threads per block (*TB*) and different number of thread blocks (***K***)

6. Conclusions

Homomorphic hashing is an important tool for data authentication in p2p applications with erasure coding or network coding. Unfortunately it is computationally expensive to perform homomorphic hashing on today's CPU. In this paper, we describe the design and implementation of parallel algorithms for homomorphic hashing using GPU and CUDA. By using a contemporary graphic card, our parallel algorithm can achieve 8979 Kbps of hashing throughput, which is more than 60 times of the CPU counterpart.

7. References

[1] NVIDIA CUDA. http://developer.nvidia.com/object/cuda.html

[2] AMD CTM Guide: Technical Reference Manual. 2006. http://ati.amd.com/companyinfo/researcher/documents/ATI_CTM_Guide.pdf

[3] GNU MP Arithmetic Library. http://gmplib.org/

[4] NVIDIA CUDA Compute Unified Device Architecture: Programming Guide, Version 2.0beta2, Jun. 2008.

[5] Montgomery, P., 1985. Multiplication without trial division, Math. Computation, vol. 44, 1985, 519-521.

[6] Menezes, A., van Oorshot, P., and Vanstone S., 1996. Handbook of applied cryptography. CRC Press, 1996.

[7] Ahlswede, R., Cai, N., Li S. R., and Yeung, R. W. 2000. Network information flow. IEEE Transactions on Information Theory, 46(4), July 2000, 1204-1216.

[8] Ho, T., Koetter. R., Médard, M., Karger, D.R. and Effros, M. 2003. The benefits of coding over routing in a randomized setting. In Proceedings of IEEE ISIT, 2003.

[9] Li, S.-Y.R., Yueng, R.W., and Cai, N. 2003. Linear network coding. IEEE Transactions on Information Theory, vol. 49, 2003. 371-381.

[10] Krohn, M., FreedMan, M., and Mazieres, D. 2004. On-the-fly verification of rateless erasure codes for efficient content distribution. In Proceedings of IEEE Symposium on Security and Privacy, Berkeley, CA, 2004.

[11] Gkantsidis, C. and Rodriguez, P. 2005. Network coding for large scale content distribution. In Proceedings of IEEE INFOCOM 2005.

[12] Gkantsidis, C. and Rodriguez, P. 2006. Cooperative security for network coding file distribution. In Proceedings of IEEE INFOCOM'06, 2006.

[13] Li, Q., Chiu, D.-M., and Lui, J. C.S. 2006. On the practical and security issues of batch content distribution via network coding. In Proceedings of IEEE ICNP'06, 2006, 158-167.

[14] Chou, P. A. and Wu, Y. 2007. Network coding for the Internet and wireless networks. Technical Report. MSR-TR-2007-70, Microsoft Research.

[15] Wang, M. and Li, B. 2007. Lava: a reality check of network coding in peer-to-peer live streaming. In Proceedings of IEEE INFOCOM'07, 2007.

[16] Wang, M. and Li, B. 2007. R^2: random push with random network coding in live peer-to-peer streaming. In IEEE Journal on Selected Areas in Communications, Dec. 2007, 1655-1666.

[17] Ryoo, S., Rodrigues, C. I., Baghsorkhi, S. S., Stone, S. S., Kirk, D. B., and Hwu, W. 2008. Optimization principles and application performance evaluation of a multithreaded GPU using CUDA. In Proceedings of ACM PPoPP'08, Feb. 2008.

[18] Falcao, G., Sousa, L., and Silva, V. 2008. Massiv parallel LDPC decoding in GPU. In Proceedings of ACM PPoPP'08, Feb. 2008.

[19] Yu, Z., Wei, Y., Ramkumar, B., and Guan, Y. 2008. An efficient signature-based scheme for securing network coding against pollution attacks. In Proceedings of IEEE INFOCOM'08, Apr. 2008.

[20] Owens, J. D., Houston, M., Luebke, D., Green, S., Stone, J. E., and Phillips, J. C. 2008. GPU computing. IEEE Proceedings, May 2008, 879-899.

[21] Al-Kiswany, S., Gharaibeh, A., Santos-Neto, E., Yuan, G., and Ripeanu, M. 2008. StoreGPU: exploiting graphics processing units to accelerate distributed storage systems. In Proceedings of IEEE Symposium on High Performance Distributed Computing (HPDC), Jun. 2008.

[22] Silberstein, M., Geiger, D., Schuster, A., Patney, A., Owens, J. D. 2008. Efficient computation of sum-products on GPUs through software-managed cache. In Proceedings of the 22nd ACM International Conference on Supercomputing, Jun. 2008.

[23] Seiler, L., et. al., 2008. Larrabee: a many-core x86 architecture for visual computing. ACM Transactions on Graphics, 27(3), Aug. 2008.

Kernel Learning for Local Learning based Clustering

Hong ZENG

Abstract

For most kernel-based clustering algorithms, their performance will heavily hinge on the choice of kernel. In this paper, we propose a novel kernel learning algorithm particularly for the Local Learning based Clustering [15]. With multiple kernels available, we associate a non-negative weight with each Hilbert space for the corresponding kernel, and then extend our previous work on feature selection [18] to select the suitable Hilbert spaces for LLC. We show that it naturally renders a linear combination of kernels, accordingly, the kernel weights are estimated iteratively with the local learning based clustering. The experimental results demonstrate the effectiveness of the proposed algorithm in clustering benchmark document datasets and the unsupervised face detection task.

1. Introduction

Since the past few decades, the kernel methods have been widely applied to various learning problems, where the data is implicitly mapped into a nonlinear high dimensional space by kernel function [12]. Unfortunately, it is known that the performance will heavily hinge on the choice of kernel, and the most suitable kernel for a particular task is often unknown in advance. Thereby, learning an appropriate kernel, is critical to obtain a robust or even improved performance for the employed kernel-based inference method.

In this paper, we are particularly interested in the problem of kernel learning for clustering. In the literature, the kernel learning has been extensively studied for the supervised learning contexts. However, this issue remains less explored in unsupervised problems, due to the absence of ground truth class labels that could guide the learning for "ideal" kernels. Until very recently, several algorithms have been proposed to address this issue for clustering. Some approaches [17, 1] directly learn the kernel parameters of some specific kernels. Though improvement is often achieved, extension of the learning method to other kernel functions is often nontrivial. A more effective framework, termed as the multiple kernel learning [7], learns a linear combination of base kernels with different weights, which will be estimated iteratively with the inference process [14, 8]. This strategy may bring potential advantages over those which try to obtain a single best kernel, through exploiting the complementary information among different kernels. In [14], the algorithm tries to find a maximum margin hyperplane to cluster data (restricted to binary-class case), accompanied with learning a mixture of Laplacian matrices. In [8], clustering is phrased as a non-negative matrix factorization problem of a fused kernel matrices. Nevertheless, both approaches in [14, 8] are global learning based, their performance may get degraded in a difficult case where the similarities among samples are less discriminable from a global view.

To overcome such possible drawback, we propose a novel multiple kernel learning method in the Local Learning based Clustering [15] setting. The LLC algorithm aims at optimizing the local purity requirement of clustering assignment, thus it may be expected to produce a more reliable intermediate clustering result in the mentioned difficult case. We associate a non-negative weight with each Hilbert space (or called the feature space interchangeably) for the corresponding kernel, and then extend our previous work on feature selection [18] to select the suitable Hilbert spaces for LLC. It will be shown later that such strategy naturally leads to learn a linear combination of the available kernels at hand. Accordingly, an alternating algorithm is developed in which the combination coefficients of kernels are estimated iteratively with the local learning based clustering.

The remainder of the paper is organized as follows: After an overview of local learning based clustering algorithm is given in Section 2, we present the proposed method in Section 3. Section 4 illustrates the connection of the proposed method with a kind of related approaches. In Section 5, the experiments on several benchmark datasets are presented. The paper is concluded in Section 6.

2. Overview of the Local Learning based Clustering Algorithm

The indicator matrix that will be used later is introduced first. Given n data points $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^n (\mathbf{x}_i \in \mathbb{R}^d)$, the dataset will be partitioned into C clusters. The clustering

result could be represented by a *cluster assignment indicator matrix* $\mathbf{P} = [p_{ic}] \in \{0,1\}^{n\times C}$, such that $p_{ic} = 1$ if $\mathbf{x}_i$ belongs to the cth cluster, and $p_{ic} = 0$ otherwise. The *scaled cluster assignment indicator matrix* used in this paper is defined by: $\mathbf{Y} = \mathbf{P}(\mathbf{P}^T\mathbf{P})^{-\frac{1}{2}} = [\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^C]$, where $\mathbf{y}^c = [y_{1c}, \ldots, y_{nc}]^T \in \mathbb{R}^n (1 \le c \le C)$, is the c-th column of $\mathbf{Y} \in \mathbb{R}^{n\times C}$. $y_{ic} = p_{ic}/\sqrt{n_c}$ can be regarded as the confidence that $\mathbf{x}_i$ is assigned to the cth cluster, where n_c is the size of the cth cluster. It is easy to verify that $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$, where $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix.

The starting point of the LLC [15] is that the cluster assignments in the neighborhood of each point should be as pure as possible. Assume an arbitrary $\mathbf{Y}$ exists at first, for each $\mathbf{x}_i$, a regression model is built with the training data $\{(\mathbf{x}_j, y_{jc})\}_{\mathbf{x}_j\in\mathcal{N}_i} (1 \le c \le C, 1 \le i, j \le n)$, where $\mathcal{N}_i$ denotes the set of neighboring[1] points of $\mathbf{x}_i$ (not including $\mathbf{x}_i$ itself). The output of the local model is of the following form: $f_i^c(\mathbf{x}) = \mathbf{x}^T\boldsymbol{\theta}_i^c, \forall \mathbf{x} \in \mathbb{R}^d$, where $\boldsymbol{\theta}_i^c \in \mathbb{R}^d$ is the local regression coefficients vector. Here, the bias term is ignored for simplicity, assuming that one of the features is always 1. In [15], $\boldsymbol{\theta}_i^c$ is solved by:

$$\min_{\boldsymbol{\theta}_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n}\Big[\sum_{\mathbf{x}_j\in\mathcal{N}_i} \beta(y_{jc} - \mathbf{x}_j^T\boldsymbol{\theta}_i^c)^2 + \|\boldsymbol{\theta}_i^c\|^2\Big], \quad (1)$$

where β is a trade-off parameter. Denote the solution to the linear ridge regression problem (1) as $\boldsymbol{\theta}_i^{c*}$, the predicted cluster assignment for the test data $\mathbf{x}_i$ can then be calculated by:

$$\widehat{y}_{ic} = f_i^c(\mathbf{x}_i) = \mathbf{x}_i^T\boldsymbol{\theta}_i^{c*} = \boldsymbol{\alpha}_i^T\mathbf{y}_i^c, \quad (2)$$

where

$$\boldsymbol{\alpha}_i^T = \beta\mathbf{x}_i^T(\beta\mathbf{X}_i\mathbf{X}_i^T + \mathbf{I})^{-1}\mathbf{X}_i. \quad (3)$$

$\mathbf{X}_i = [\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_{n_i}}]$ with $\mathbf{x}_{i_k}$ being the k-th neighbor of $\mathbf{x}_i$, n_i is the size of $\mathcal{N}_i$, and $\mathbf{y}_i^c = [y_{i_1c}, y_{i_2c}, \ldots, y_{i_{n_i}c}]^T$.

After all the local predictors have been constructed, LLC aims to find an optimal cluster indicator matrix $\mathbf{Y}$ which could minimize the overall prediction errors:

$$\begin{aligned}
&\sum_{c=1}^{C}\sum_{i=1}^{n}(y_{ic} - \widehat{y}_{ic})^2 \\
&= \sum_{c=1}^{C}\|\mathbf{y}^c - \mathbf{A}\mathbf{y}^c\|^2 \\
&= trace[\mathbf{Y}^T(\mathbf{I}-\mathbf{A})^{\mathbf{T}}(\mathbf{I}-\mathbf{A})\mathbf{Y}] \\
&= trace(\mathbf{Y}^T\mathbf{T}\mathbf{Y}),
\end{aligned} \quad (4)$$

where $\mathbf{T} = (\mathbf{I}-\mathbf{A})^{\mathbf{T}}(\mathbf{I}-\mathbf{A})$, $\mathbf{A}$ is an $n \times n$ sparse matrix with its (i,j)-th entry a_{ij} being the corresponding element in $\boldsymbol{\alpha}_i$ by (3) if $\mathbf{x}_j \in \mathcal{N}_i$ and 0 otherwise.

[1]The k-mutual neighbors are adopted in order to well describe the local structure, i.e. $\mathbf{x}_j$ is considered as a neighbor of $\mathbf{x}_i$ only if $\mathbf{x}_i$ is also one of the k-nearest neighbors of $\mathbf{x}_j$.

As in the spectral clustering [10, 16], $\mathbf{Y}$ is relaxed into the continuous domain while keeping the property $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$ for (4). LLC then solves:

$$\min_{\mathbf{Y}\in\mathbb{R}^{n\times C}} \quad trace(\mathbf{Y}^T\mathbf{T}\mathbf{Y}) \ \ s.t.\ \mathbf{Y}^{\mathbf{T}}\mathbf{Y} = \mathbf{I} \quad (5)$$

A solution to $\mathbf{Y}$ is given by the first C eigenvectors of the matrix $\mathbf{T}$, corresponding to the first C smallest eigenvalues. The final partition result is obtained by discretizing $\mathbf{Y}$ via the method in [16] or by k-means as in [10].

3. Multiple Kernel Learning for Local Learning based Clustering

The LLC algorithm can be easily kernelized as in [15], by replacing the linear ridge regression with the kernel ridge regression. Hence selecting a suitable kernel function will be a crucial issue. In this section, we extend our previous work of feature selection for LLC [18] to learn a proper linear combination of several pre-computed kernel matrices.

In the kernel methods, the symmetric positive semi-definite kernel function $\mathcal{K} : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$, implicitly maps the raw input features into a high-dimensional (possibly infinite) *Reproducing Kernel Hilbert Space* (RKHS) $\mathcal{H}$, which is equipped with the inner product $< \cdot,\cdot >_{\mathcal{H}}$ via a nonlinear mapping $\phi : \mathcal{X} \to \mathcal{H}$, i.e., $\mathcal{K}(x,z) =< \phi(x), \phi(z) >_{\mathcal{H}}$. Assume there are altogether L different kernel functions $\{\mathcal{K}^{(l)}\}_{l=1}^L$ available for the clustering task at hand, with which L different feature spaces $\{\mathcal{H}^{(l)}\}_{l=1}^L$ are associated accordingly. We are unknown which feature space to use, an intuitive way is to use them all by concatenating all feature spaces into an augmented Hilbert space: $\widetilde{\mathcal{H}} = \bigoplus_{l=1}^L \mathcal{H}^{(l)}$, and associate each feature space a relevance weight τ_l ($\sum_{l=1}^L \tau_l = 1, \tau_l \ge 0, \forall l$), or equivalently the importance factor for kernel function $\mathcal{K}^{(l)}$. Later we will show that performing LLC in such feature space is equivalent to employing a combined kernel function: $\mathcal{K}^\tau(x,z) = \sum_{l=1}^L \tau_l\mathcal{K}^{(l)}(x,z)$ for LLC. A zero weight τ_l would correspond to *blend out* the feature space associated with the corresponding kernel similar as the feature selection in [18]. Our task is to learn the coefficients $\{\tau_l\}_{l=1}^L$ which can lead to a more accurate and robust performance. An alternating algorithm which iteratively performs clustering and estimates the kernel weight, is developed.

3.1. Update $\mathbf{Y}$ for a Given $\boldsymbol{\tau}$

First of all, given a $\boldsymbol{\tau}$, the nearest neighbors $\mathcal{N}_i$ for LLC algorithm will be re-found by the $\boldsymbol{\tau}$-weighted squared Euclidean distance in $\widetilde{\mathcal{H}}$, i.e.:

$$\begin{aligned}
d_\tau(\mathbf{x}_1,\mathbf{x}_2) &= \|\phi(\mathbf{x}_1) - \phi(\mathbf{x}_2)\|_\tau^2 \\
&= \mathcal{K}^\tau(\mathbf{x}_1,\mathbf{x}_1) + \mathcal{K}^\tau(\mathbf{x}_2,\mathbf{x}_2) - 2\mathcal{K}^\tau(\mathbf{x}_1,\mathbf{x}_2).
\end{aligned} \quad (6)$$

Then the local discriminant function in the $\widetilde{\mathcal{H}}$ can be written as follows:

$$f_i^c(\phi(\mathbf{x})) = \phi(\mathbf{x})^T \mathbf{w}_i^c + b_i^c, \tag{7}$$

where $\phi(\mathbf{x}) = [\phi_1(\mathbf{x})\ \phi_2(\mathbf{x})\ \cdots\ \phi_L(\mathbf{x})]^T \in \mathbb{R}^D$, $\phi_l(\mathbf{x}) \in \mathbb{R}^{D_l}$ is the sample mapped by the lth kernel function. D and D_l are the dimensionalities of $\widetilde{\mathcal{H}}$ and $\mathcal{H}^{(l)}$ respectively, $\sum_{l=1}^L D_l = D$. Taking the relevance of each feature space for clustering into account, the regression coefficient $\mathbf{w}_i^c \in \mathbb{R}^D$ and the bias $b_i^c \in \mathbb{R}$ now will be solved via the following weighted l_2 norm regularized least square problem:

$$\min_{\mathbf{w}_i^c, b_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n} \Big[\sum_{\mathbf{x}_j \in \mathcal{N}_i} \beta(y_{jc} - \phi(\mathbf{x}_j)^T \mathbf{w}_i^c - b_i^c)^2 + \mathbf{w}_i^{cT} \mathbf{\Lambda}_\tau^{-1} \mathbf{w}_i^c \Big], \tag{8}$$

where $\mathbf{\Lambda}_\tau$ is a diagonal matrix with the vector $\widetilde{\boldsymbol{\tau}} = (\underbrace{\tau_1, \ldots, \tau_1}_{D_1}, \ldots, \underbrace{\tau_L, \ldots, \tau_L}_{D_L})^T$ in the diagonal, and $\sum_{l=1}^L \tau_l = 1, \tau_l \geq 0 \forall l$. And similar to [18], the weighted l_2 norm (the second term in the square bracket of (8)) with $\boldsymbol{\tau}$ defined on the standard simplex, is able to provide adaptive regularization: a large penalty will be imposed on the elements of $\mathbf{w}_i^c$ corresponding to the feature spaces associated with irrelevant kernels. Thus an improved clustering result can be expected, since the vanishing elements in $\mathbf{w}_i^c$ will eliminate the feature spaces with irrelevant kernels from prediction (c.f. (7)).

After removing the bias term by plugging its optimal solution

$$b_i^c = \frac{1}{n_i} \mathbf{e}_i^T (\mathbf{y}_i^c - \phi(\mathbf{X}_i)^T \mathbf{w}_i^c), \tag{9}$$

into (8), we can reformulate the primal problem (8) as follows:

$$\min_{\mathbf{w}_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n} \Big[\beta \|\mathbf{\Pi}_i \mathbf{y}_i^c - (\phi(\mathbf{X}_i)\mathbf{\Pi}_i)^T \mathbf{w}_i^c\|^2 + \mathbf{w}_i^{cT} \mathbf{\Lambda}_\tau^{-1} \mathbf{w}_i^c \Big]. \tag{10}$$

Then we consider the dual formulation of the (10) in terms of $\mathbf{w}_i^c$. Denote

$$\boldsymbol{\zeta}_i^c = (\phi(\mathbf{X}_i)\mathbf{\Pi}_i)^T \mathbf{w}_i^c - \mathbf{\Pi}_i \mathbf{y}_i^c, \tag{11}$$

then the Lagrangian for problem (10) is

$$\mathcal{L}(\{\boldsymbol{\zeta}_i^c, \mathbf{w}_i^c, \boldsymbol{\gamma}_i^c\}) = \sum_{c=1}^{C}\sum_{i=1}^{n} \Big(\beta \|\boldsymbol{\zeta}_i^c\|^2 + \mathbf{w}_i^{cT} \mathbf{\Lambda}_\tau^{-1} \mathbf{w}_i^c \Big) \tag{12}$$

$$- \sum_{c=1}^{C}\sum_{i=1}^{n} \boldsymbol{\gamma}_i^{cT} \Big((\phi(\mathbf{X}_i)\mathbf{\Pi}_i)^T \mathbf{w}_i^c - \mathbf{\Pi}_i \mathbf{y}_i^c - \boldsymbol{\zeta}_i^c \Big),$$

where the $\boldsymbol{\gamma}_i^c$'s are the vectors of Lagrangian dual variables, $\boldsymbol{\gamma}_i^c \in \mathbb{R}^{n_i}$. Taking the derivatives of $\mathcal{L}$ with respect to the primal variables $\boldsymbol{\zeta}_i^c$ and $\mathbf{w}_i^c$, and setting them equal to zero, we obtain:

$$\boldsymbol{\zeta}_i^c = -\frac{\boldsymbol{\gamma}_i^c}{2\beta},\ \mathbf{w}_i^c = \frac{\mathbf{\Lambda}_\tau \phi(\mathbf{X}_i)\mathbf{\Pi}_i \boldsymbol{\gamma}_i^c}{2}, \tag{13}$$

and finally we arrive at the dual problem:

$$\begin{aligned} &\max_{\boldsymbol{\gamma}_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n} -\frac{1}{4\beta}\boldsymbol{\gamma}_i^{cT}\boldsymbol{\gamma}_i^c - \frac{1}{4}\boldsymbol{\gamma}_i^{cT}\mathbf{\Pi}_i\phi(\mathbf{X}_i)^T\mathbf{\Lambda}_\tau\phi(\mathbf{X}_i)\mathbf{\Pi}_i\boldsymbol{\gamma}_i^c \\ &+ \boldsymbol{\gamma}_i^{cT}\mathbf{\Pi}_i\mathbf{y}_i^c = \\ &\max_{\boldsymbol{\gamma}_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n} -\frac{1}{4\beta}\boldsymbol{\gamma}_i^{cT}\boldsymbol{\gamma}_i^c - \frac{1}{4}\boldsymbol{\gamma}_i^{cT}\mathbf{\Pi}_i\mathbf{K}_i^\tau\mathbf{\Pi}_i\boldsymbol{\gamma}_i^c + \boldsymbol{\gamma}_i^{cT}\mathbf{\Pi}_i\mathbf{y}_i^c. \end{aligned} \tag{14}$$

The last equality follows from:

$$\begin{aligned} \phi(\mathbf{X}_i)^T\mathbf{\Lambda}_\tau\phi(\mathbf{X}_i) &= \sum_{l=1}^{L} \tau_l \phi_l(\mathbf{X}_i)^T\phi_l(\mathbf{X}_i) \\ &= \sum_{l=1}^{L} \tau_l \mathbf{K}_i^{(l)} = \mathbf{K}_i^\tau. \end{aligned} \tag{15}$$

where $\mathbf{K}_i^{(l)}, \mathbf{K}_i^\tau \in \mathbb{R}^{n_i \times n_i}$ are the base and combined kernel matrices over $\mathbf{x}_j \in \mathcal{N}_i$ respectively, i.e., $\mathbf{K}_i^{(l)} = [\mathcal{K}^{(l)}(\mathbf{x}_u, \mathbf{x}_v)]$ and $\mathbf{K}_i^\tau = [\mathcal{K}^\tau(\mathbf{x}_u, \mathbf{x}_v)]$, for $\mathbf{x}_u, \mathbf{x}_v \in \mathcal{N}_i$. For fixed $\boldsymbol{\tau}$ constrained on the simplex, the convex combination of the positive semi-definite kernel matrices : $\mathbf{K}_i^\tau = \sum_{l=1}^L \tau_l \mathbf{K}_i^{(l)}$ is still a positive semi-definite kernel matrix. Therefore, the problem in (14) is an unconstrained concave quadratic program whose unique optimal solution can be obtained analytically:

$$\boldsymbol{\gamma}_i^{c*} = 2\beta(\mathbf{I}_i + \beta\mathbf{\Pi}_i\mathbf{K}_i^\tau\mathbf{\Pi}_i)^{-1}\mathbf{\Pi}_i\mathbf{y}_i^c. \tag{16}$$

Then altogether with (9), (13) and (16), the predicted indicator value at point $\mathbf{x}_i$ for the cth ($c = 1, \ldots, C$) cluster can be calculated by (7):

$$\widehat{y}_{ic} = f_i^c(\phi(\mathbf{x}_i)) = \phi(\mathbf{x}_i)^T\mathbf{w}_i^c + b_i^c = \boldsymbol{\alpha}_i^T\mathbf{y}_i^c, \tag{17}$$

with

$$\begin{aligned} \boldsymbol{\alpha}_i^T =& \beta(\mathbf{k}_i^\tau - \frac{1}{n_i}\mathbf{e}_i^T\mathbf{K}_i^\tau)\mathbf{\Pi}_i\Big[\mathbf{I}_i - \\ & (\beta^{-1}\mathbf{I} + \mathbf{\Pi}_i\mathbf{K}_i^\tau\mathbf{\Pi}_i)^{-1}\mathbf{\Pi}_i\mathbf{K}_i^\tau\mathbf{\Pi}_i\Big] + \frac{1}{n_i}\mathbf{e}_i^T, \end{aligned} \tag{18}$$

where $\mathbf{k}_i^\tau \in \mathbb{R}^{n_i}$ denotes the vector $[\mathcal{K}^\tau(\mathbf{x}_i, \mathbf{x}_j)]^T$ for $\mathbf{x}_j \in \mathcal{N}_i$.

To obtain $\mathbf{Y}$, we will first build the matrix $\mathbf{T}$ by (4) with $\boldsymbol{\alpha}_i$ defined in (18), using the combined kernel $\mathcal{K}^\tau(\mathbf{x}_i, \mathbf{x}_j) = \sum_{l=1}^L \tau_l \mathcal{K}^{(l)}(\mathbf{x}_i, \mathbf{x}_j)$. Then $\mathbf{Y}$ is given by the first C eigenvectors of $\mathbf{T}$ corresponding to the C smallest eigenvalues.

3.2. Update τ for a Given $\mathbf{Y}$

Subsequently, the L kernel combination coefficients $\{\tau_l\}_{l=1}^L$ will be recomputed based on the current estimation for $\mathbf{Y}$. We propose to estimate $\boldsymbol{\tau}$ using the *projected gradient descent* method as in [4, 11].

With fixed $\mathbf{Y}$ and neighborhood determined at each point, an optimal τ is expected to minimize:

$$\mathcal{P}(\boldsymbol{\tau}),\ s.t.\ \sum_{l=1}^{L}\tau_l = 1, \tau_l \geq 0, \forall l, \tag{19}$$

where

$$\mathcal{P}(\boldsymbol{\tau}) = \min_{\mathbf{w}_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n}\left[\beta\|\boldsymbol{\Pi}_i\mathbf{y}_{ic} - (\phi(\mathbf{X}_i)\boldsymbol{\Pi}_i)^T\mathbf{w}_i^c\|^2 + \mathbf{w}_i^{cT}\boldsymbol{\Lambda}_\tau^{-1}\mathbf{w}_i^c\right]. \tag{20}$$

In general, it could be solved by the projected gradient descent method through the recursive update equation $\boldsymbol{\tau}^{(new)} = \boldsymbol{\tau}^{(old)} - \eta\boldsymbol{\nabla}\mathcal{P}$, where the η is the step size, and $(-\boldsymbol{\nabla}\mathcal{P})$ is the descent direction. Nevertheless, since both the $\mathbf{Y}$ and $\mathcal{N}_i$ depend on $\boldsymbol{\tau}$ (c.f. Section 3.1), they need to be recomputed once the $\boldsymbol{\tau}$ applies one-step update. Therefore, we only take a single gradient descent step for $\boldsymbol{\tau}$, rather than repeated iterations. The $\boldsymbol{\tau}^{(new)}$ updated only once is still applicable to the problem, since with the $\mathbf{Y}$ and $\mathcal{N}_i$ fixed, and we make sure $\mathcal{P}(\boldsymbol{\tau}^{(new)}) \leq \mathcal{P}(\boldsymbol{\tau}^{(old)})$ holds, then the local regression model derived from the $\boldsymbol{\tau}^{(new)}$ is expected to be better than the one derived from $\boldsymbol{\tau}^{(old)}$. The feasibility of updating $\boldsymbol{\tau}$ in this manner has been verified by experimental results.

Then the key issue is to obtain the derivatives of $\mathcal{P}(\boldsymbol{\tau})$ in analytic forms. In order to do so, we resort to the dual of $\mathcal{P}(\boldsymbol{\tau})$ which has been investigated in subSection 3.1, and is rewritten below:

$$\mathcal{D}(\boldsymbol{\tau}) = \max_{\boldsymbol{\gamma}_i^c} \sum_{c=1}^{C}\sum_{i=1}^{n} -\frac{1}{4\beta}\boldsymbol{\gamma}_i^{cT}\boldsymbol{\gamma}_i^c - \frac{1}{4}\boldsymbol{\gamma}_i^{cT}\boldsymbol{\Pi}_i\mathbf{K}_i^\tau\boldsymbol{\Pi}_i\boldsymbol{\gamma}_i^c + \boldsymbol{\gamma}_i^{cT}\boldsymbol{\Pi}_i\mathbf{y}_i^c. \tag{21}$$

Note (10) is convex with respect to $\mathbf{w}_i^c$, by the principle of strong duality, we have $\mathcal{P}(\boldsymbol{\tau}) = \mathcal{D}(\boldsymbol{\tau})$. Furthermore, as $\{\boldsymbol{\gamma}_i^{c*}\}$ (c.f. (16)) maximizes $\mathcal{D}$, then according to [3], if $\{\boldsymbol{\gamma}_i^{c*}\}$'s are unique, $\mathcal{D}(\boldsymbol{\tau})$ is differentiable. Fortunately this unicity is guaranteed by the unconstrained concave quadratic program in (14). Moreover, as proved in Lemma 2 of [5], $\mathcal{D}(\boldsymbol{\tau})$ can be differentiated with respect to $\boldsymbol{\tau}$ as if $\{\boldsymbol{\gamma}_i^{c*}\}$ did not depend on $\boldsymbol{\tau}$. Finally, we have:

$$\begin{aligned}\frac{\partial\mathcal{P}}{\partial\tau_l} = \frac{\partial\mathcal{D}}{\partial\tau_l} &= -\frac{1}{4}\sum_{c=1}^{C}\sum_{i=1}^{n}\boldsymbol{\gamma}_i^{c*T}\boldsymbol{\Pi}_i\mathbf{K}_i^{(l)}\boldsymbol{\Pi}_i\boldsymbol{\gamma}_i^{c*} \\ &= -\frac{1}{4}\sum_{i=1}^{n} trace(\boldsymbol{\gamma}_i^{*T}\boldsymbol{\Pi}_i\mathbf{K}_i^{(l)}\boldsymbol{\Pi}_i\boldsymbol{\gamma}_i^{*}),\end{aligned} \tag{22}$$

where $\boldsymbol{\gamma}_i^* = [\boldsymbol{\gamma}_i^{1*}, \ldots, \boldsymbol{\gamma}_i^{C*}] \in \mathbb{R}^{n_i \times C}$.

Note the equality and non-negative constraints over the $\boldsymbol{\tau}$ have to be kept inviolated when updating $\boldsymbol{\tau}$ along the descent gradient direction. We used the same strategy as in [11] by first projecting the gradient to enforce the equality, and then ensuring that the descent direction does not lead to negative τ_l. Namely, each element of the reduced gradient $\boldsymbol{\nabla}\mathcal{P}$ is designed as follows:

$$(\boldsymbol{\nabla}\mathcal{P})_l = \begin{cases} \frac{\partial\mathcal{P}}{\partial\tau_l} - \frac{\partial\mathcal{P}}{\partial\tau_m}, & \text{if } l \neq m \text{ and } \tau_l > 0; \\ \sum_{\mu\neq m, \tau_\mu>0}\left(\frac{\partial\mathcal{P}}{\partial\tau_m} - \frac{\partial\mathcal{P}}{\partial\tau_\mu}\right), & \text{if } l = m; \\ 0, & \text{if } \tau_l = 0 \text{ and } \frac{\partial\mathcal{P}}{\partial\tau_l} - \frac{\partial\mathcal{P}}{\partial\tau_m} > 0, \end{cases} \tag{23}$$

where $m = \arg\max_l \tau_l$.

Then as we stated before, we only go one step along the descent direction: $\boldsymbol{\tau}^{(new)} = \boldsymbol{\tau}^{(old)} - \eta\boldsymbol{\nabla}\mathcal{P}$. We first try η with the maximal admissible step size η_{max} which sets τ_ν to zero, where

$$\nu = \arg\min_{\{l|(\boldsymbol{\nabla}\mathcal{P})_l>0\}} \frac{\tau_l^{(old)}}{(\boldsymbol{\nabla}\mathcal{P})_l}, \tag{24}$$

$$\eta_{max} = \frac{\tau_\nu}{(\boldsymbol{\nabla}\mathcal{P})_\nu}. \tag{25}$$

If $\mathcal{D}(\boldsymbol{\tau}^{(trial)}) \leq \mathcal{D}(\boldsymbol{\tau}^{(old)})$, where $\boldsymbol{\tau}^{(trial)} = \boldsymbol{\tau}^{(old)} - \eta_{max}\boldsymbol{\nabla}\mathcal{P}$, $\boldsymbol{\tau}$ gets updated; otherwise, a one-dimensional line search for $\eta \in [0, \eta_{max}]$ is applied. Algorithm 1 describes the steps to update $\boldsymbol{\tau}$.

3.3. The Complete Algorithm

The complete local learning based clustering algorithm with multiple kernel learning (denoted as LLC-mkl) is presented in Algorithm 2. The loop stops when the relative variation of the trace value in (5) between two consecutive iterations gets below a threshold (we set it at 10^{-4} in this paper), indicating the partitioning has almost stabilized. After the convergence, the $\mathbf{Y}$ is discretized to obtain the final clustering result with k-means as in [10].

4. Connection with the Multi-view Clustering

An important application of the multiple kernel learning is to fuse the information from heterogeneous sources as

Compute the projected gradient $\nabla\mathcal{P}$ by (23);
Compute the maximal admissible step size η_{max} by (25);
$\boldsymbol{\tau}^{(trial)} = \boldsymbol{\tau}^{(old)} - \eta_{max}\nabla\mathcal{P}$;
Compute $\mathcal{D}(\boldsymbol{\tau}^{(trial)})$ with $\{\boldsymbol{\gamma}_i^*\}$ calculated from $\mathbf{K}^{\tau^{(trial)}} = \sum_{l=1}^{L} \tau_l^{(trial)} \mathbf{K}^{(l)}$;
if $\mathcal{D}(\boldsymbol{\tau}^{(trial)}) \leq \mathcal{D}(\boldsymbol{\tau}^{(old)})$ **then**
  $\eta = \eta_{max}$;
else
  Perform line search for $\eta \in [0, \eta_{max}]$ along $\nabla\mathcal{P}$;
end
$\boldsymbol{\tau}^{(new)} = \boldsymbol{\tau}^{(old)} - \eta\nabla\mathcal{P}$;

Algorithm 1: Update kernel weight vector $\boldsymbol{\tau}$ with the current $\mathbf{Y}$ and $\mathcal{N}_i$.

input : L base kernel matrices $\mathbf{K}^{(l)}$'s, size of the neighborhood k, trade-off parameter β
output: $\mathbf{Y}, \boldsymbol{\tau}$

Initialize $\tau_l = \frac{1}{L}$, for $l = 1, \ldots, L$;
while *not converge* **do**
  Find k-mutual neighborhoods, using the metric defined in (6);
  Construct the matrix $\mathbf{T}$ by (4) with $\boldsymbol{\alpha}_i$ given in (18), and then solve the problem (5) to obtain $\mathbf{Y}$;
  Update $\boldsymbol{\tau}$ with the steps described in Algorithm 1;
end

Algorithm 2: Multiple kernel learning for local learning based clustering algorithm.

follows [7]: associating each source with a kernel function, then combining the set of prototype kernels generated from these sources to perform the inference. From this aspect, the multiview clustering is a related work, whose goal is to learn a *consensus* result from multiple representations [19, 9]. However, it implicitly treats all the sources equally, no matter the clustering result with each source is good or not. In contrast, our method is able to determine the weight for each source automatically according to its discriminating power, thus may be more robust in practice.

5. Experimental Results

Experiments on document clustering and unsupervised face detection were conducted with LLC-mkl. For comparison, the counterpart unsupervised multiple kernel learning algorithm based on NMF [8] (denoted as NMF-mkl) was conducted. We also compared with the self-tuning spectral clustering [17] (denoted as SelfTunSpec), which tries to build a single best kernel for clustering. Besides, the spectral clustering with multiple views [19] (denoted as SpecMV) was implemented as well, which generalizes the normalized cut from a single view to multiple views, and each view is represented by normalized adjacency matrix computed with some kernel function. The algorithm in [14] is not compared because the optimization software in [14] cannot deal with datasets with too many samples and will cause memory overflow on the datasets used in this paper. Since how to choose the optimal number of clusters is beyond the scope of this paper, we simply set the number of clusters equal to the number of classes in each dataset for all the algorithms. We evaluated the performance with the clustering accuracy (ACC) index [15] for all algorithms. The sensitivity of the proposed LLC-mkl algorithm with respect to k and β will be presented at the end of this section.

5.1. Document Datasets

The characteristics of the benchmark document datasets used in this experiment are summarized in Table 1.

Table 1. Characteristics of the document datasets

Dataset	#Sample (n)	#Class (C)
CSTR	476	4
WebACE	2340	20
tr11	414	9
tr31	927	7

- **CSTR**: This is the dataset of the abstracts of technical reports published in the Department of Computer Science at a university between 1991 and 2002. The dataset contains 476 abstracts, which are divided into four research topics.
- **WebACE**: This dataset is from WebACE project, and it contains 2340 documents consisting of news articles from Reuters news service with 20 different topics in October 1997.
- **tr11** and **tr31**: Both of the two datasets are from the CLUTO toolkit [6], they contain 414 and 927 articles categorized into 9 and 7 topics respectively.

To pre-process the CSTR and WebACE datasets, we remove the stop words using a standard stop list, all HTML tags are skipped and all header fields except subject and organization if the posted articles are ignored. Then the each document is represented by the term-frequency vector (Bag-of-Words). The datasets associated with the CLUTO toolkit have already been preprocessed. For all datasets, we use the top 1000 words by mutual information with class labels.

We applied the LLC-mkl with altogether 10 pre-computed base kernels, i.e., 7 RBF kernels

Table 2. Accuracies of various methods on the document datasets

Data Set	LLC-wkernel	LLC-bkernel	LLC-mkl	NMF-mkl	SelfTunSpec	Spec-mv
CSTR	0.3487	0.7374	**0.8508±0.0012**	0.6387	0.5210	0.3741
WebACE	0.2436	0.4885	**0.6316±0.0215**	0.4960	0.4880	0.3286
tr11	0.4251	**0.5966**	0.5609±0.0166	0.5145	0.4106	0.5242
tr31	0.5297	**0.6721**	0.6512±0.0007	0.5372	0.4412	0.5599

$\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\delta^2)$, with δ varying in $\{0.01, 0.05, 0.1, 1, 10, 50, 100\}$, 2 polynomial kernels $\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i^T\mathbf{x}_j)^d$ with degree $d = \{2, 4\}$, and a cosine kernel $\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T\mathbf{x}_j/(\|\mathbf{x}_i\| \cdot \|\mathbf{x}_j\|)$. All the kernels have been normalized through: $\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j)/\sqrt{\mathcal{K}(\mathbf{x}_i, \mathbf{x}_i)\mathcal{K}(\mathbf{x}_j, \mathbf{x}_j)}$. Besides, we also implemented the case where each time a single candidate kernel $\mathcal{K}^{(l)}(l = 1, \ldots, 10)$ was adopted in the LLC algorithm in which the local prediction is performed with kernel ridge regression, the best (denoted as LLC-bkernel) and the worst (denoted as LLC-wkernel) performance out of the 10 kernels are reported. NMF-mkl was applied on the same 10 base kernels. The adjacency matrix in SelfTunSpec [17] was built by its local scaling method [17] on the dataset. As for Spec-mv [19], since the combination weights are unknown *a priori* and it does not involve the re-estimation for them, without loss of generality, we applied the uniform weighting for the 10 kernels. For NMF-mkl, SelfTunSpec and Spec-mv, we only report the best accuracy among extensive trials of their free parameters. For LLC-mkl, the mean and standard deviation of ACC with $k = 30, \beta = 10$ over 10 runs are reported. The results are summarized in Table 2.

From Table 2, we could first observe that there is a big gap between the best and the worst performance of LLC with different choices of kernel. On the tr11 and tr31 datasets, the performance of LLC-mkl is close to that of the LLC with the best kernel, but obviously LLC-mkl is more sensible for practical application where we often do not know which kernel is the best *a priori*. On the CSTR and WebACE datasets, the LLC-mkl even outperforms the LLC with the best kernel. Namely, by combining multiple kernels and exploiting the complementary information contained in different kernels, the LLC-mkl indeed improves the robustness and accuracy of LLC. Compared to NMF-mkl which is derived globally, the LLC-mkl is consistently superior over it on these four datasets. A plausible reason is that the document datasets are very sparse, therefore the entries in the kernel matrix may resemble to each other from the global view or on a large scale. Thereby, finding the similar points locally may produce more reliable intermediate clustering result to guide the kernel learning. From Table 2, we could also observe that the LLC-mkl and NMF-mkl both outperform the selfTunSpec which tries to construct a single "best" kernel in this experiment. Moreover, the performance of the Spec-mv is generally inferior to that of the LLC-mkl, because it cannot update the combination weights, the algorithm with equal weights for all the adjacency matrices may tend to be affected by the improper kernel functions adopted.

5.2. Unsupervised Face Detection

This experiment was conducted on the MIT CBCL Face data set, which consists of altogether 31022 cropped face and non-face images. We randomly selected a subset of 1000 face images and 1000 non-face images, rescaled each image to 15 by 15 pixels and then processed with histogram equalization.

Based on the fragment idea of [13, 2], it is rational to assume that the different local regions in an image have different relevance in determining whether the image contains a face or not. Therefore, we divided each image into 9 non-overlapping patches of size 5 by 5. Each patch is considered as a different source which contains different spatial information. Note we only used non-overlapping patches for simplicity, but it is quite straightforward to apply the proposed method to use arbitrary, possibly overlapping patches. In addition to these intensity patches, we also computed the edge maps, i.e., the Sobel filter responses on each raw image for both the horizontal and vertical orientations. Then similar to the raw images, each edge maps image was divided into 9 patches. Therefore, there are 27 patches in total for each image: 9 patches from raw image, 9 patches from the horizontal edge maps and 9 patches from the vertical edge maps. In order to combine these fragments in a principled way, $\mathcal{K}^{(l)}(1 \leq l \leq 27)$ is defined as the cosine kernel restricted to the lth patch between each pair of images, and then the combination of $\mathcal{K}^{(l)}$'s which could lead to a more accurate unsupervised partition on these images was learned by LLC-mkl.

The obtained weight maps indicating the weights for different patches are shown in Figure 1(c). We could observe that, in general, the weights for edge maps patches dominate the weighting solution in this task, while the intensity patches seems to be less discriminative than the former. To confirm the rationality of this weighting result, we ran the LLC algorithm with the uniformly combined cosine ker-

Table 3. Confusion matrices and accuracies of various methods for the unsupervised face detection

	LLC-pix		LLC-tex		LLC-uniWei		LLC-mkl		NMF-mkl		SelfTunSpec		Spec-mv	
	face	non-face	face	non-face	face	non-face	face	non-face	face	non-face	face	non-face	face	non-face
face	963	37	999	1	1000	0	1000	0	962	38	935	65	952	48
non-face	31	969	51	949	50	950	1	999	558	442	16	984	4	996
Accuracy	0.9660		0.9740		0.9750		**0.9995**		0.7020		0.9595		0.9740	

nels computed from the 9 raw intensity patches (denoted as LLC-pix) and the 18 edge maps ones (denoted as LLC-tex), respectively. The results based on these two types of features are reported in Table 3. It can be easily found that such a weighting result is reasonable, because the edge maps features lead to higher accuracy than the intensity features (0.9740 vs. 0.9660). Besides, we also conducted the experiment where these 27 cosine kernels were uniformly combined then applying the LLC algorithm (denoted as LLC-uniWei). Though it outperforms the case where the intensity or edge maps features are used alone (i.e., LLC-pix and LLC-tex), its performance is still worse than that of the non-trivial weighting solution obtained by LLC-mkl, which has automatically assigned weight on each patch (see Figure 1(c)). A reasonable explanation is that the uniform weighting cannot make full use of the complementary information among these kernels. This is further confirmed by the experiment with Spec-mv, where 27 adjacency matrices were formed on each patch by the Gaussian kernel function with local scaling [17], and equal weights were associated with these matrices for general purpose. It is observed from Table 3 that the performance of such multiview spectral clustering falls behind that of LLC-mkl. Moreover, the LLC-mkl again yields a more accurate partition than NMF-mkl with the same 27 cosine kernels, as well as SpecTun-Spec performed on the data of 675 dimensions by simply stacking up the 27 patches into a "big" vector.

5.3. Parameter Sensitivity Study

The effects of these two parameters, i.e., k and β, on the performance of LLC-mkl are presented in Figure 2. From Figure 2(a) and 2(b), we could observe that for the document datasets used in this paper where there are no more than 300 samples per class on average, the proposed LLC-mkl algorithm with $k = 30 \sim 50$ and $\beta \in [0.01, 10]$ could produce considerably accurate results and the corresponding performance does not vary much. From Figure 2(c) and 2(d), for the face dataset used which has 1000 samples per class, the size of neighborhood k chosen from $60 \sim 100$ and the trade-off parameter β in the range $[1, 10]$ could result in considerably accurate and stable performance.

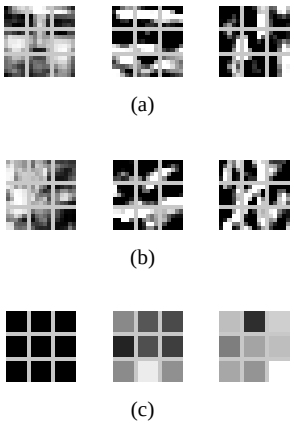


Figure 1. (a) a sample face image and its edge maps in horizontal and vertical orientations; (b) a sample non-face image and its edge maps in horizontal and vertical orientations; (c) weight maps obtained by LLC-mkl with $k = 60, \beta = 10$, patches with lighter intensities have larger weight values.

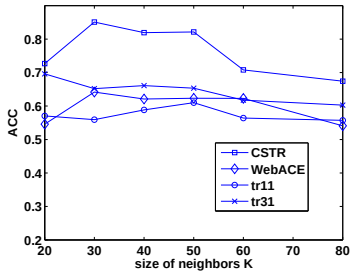
6. Conclusions

In this paper, a novel kernel learning approach has been proposed for the local learning based clustering, where a combination of kernels is jointly learned with the clustering. It is addressed under a regularization framework by taking the relevance of each kernel into account. Experimental results demonstrate that the proposed kernel learning method is able to improve the robustness and accuracy of the basic local learning clustering. Furthermore, it generally outperforms the state-of-the-art counterparts, especially when the similarities among samples are less discriminable.

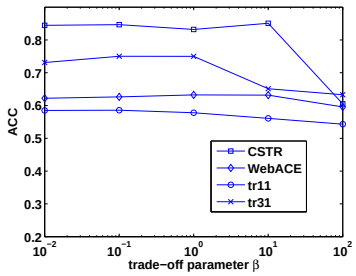
References

[1] F. R. Bach and M. I. Jordan. Learning spectral clustering, with application to speech separation. *The Journal of Machine Learning Research*, 7:1963–2001, 2006.

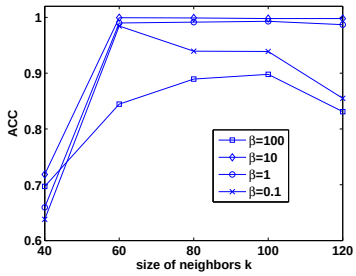
[2] G. BakIr, M. Wu, and J. Eichhorn. Maximum-margin feature combination for detection and categorization. *Technical Report*, 2005.



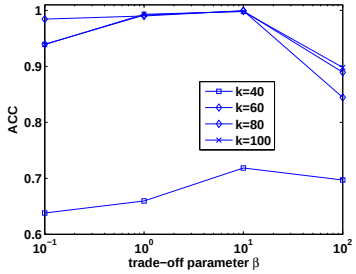
(a) document datasets



(b) document datasets



(c) MIT CBCL dataset



(d) MIT CBCL dataset

Figure 2. The parameter sensitivity studies of LLC-mkl algorithm. (a) varying the size of neighborhood with β being fixed at 10; (b) varying β with the size of neighborhood fixed at 30. (c) varying the size of neighborhood; (d) varying the trade-off parameter β. The values on each line represent the average ACC over 10 independent runs.

[3] J. Bonnans and A. Sharpiro. *Perturbation Analysis of Optimization Problems*. Springer, 2000.

[4] P. H. Calamai and J. J. Moré. Projected gradients methods for linearly constrained problems. *Mathematical Programming*, 39(1):93–116, 1987.

[5] O. Chapelle, V. Vanpnik, O. Bousquet, and S. Mukherjee. Choosing multiple parameters for surport vector machines. *Machine Learning*, 46(1):131–159, 2002.

[6] G. Karypis. *CLUTO-A Clustering Toolkit*. 2002.

[7] G. Lanckriet, N. Cristianini, P. Bartlett, M. Ghaoui, and M. Jordan. Learning the kernel matrix with semidefinite programming. *The Journal of Machine Learning Research*, 5:27–72, 2004.

[8] T. Lange and J. Buhmann. Fusion of similarity data in clustering. *Advances in Neural Information Processing Systems*, 18:723–730, 2006.

[9] B. Long, P. S. Yu, and M. Z. F. Zhang. General model for multiple view unsupervised learning. In *Proceedings of SIAM International Conference on Data Mining*, pages 822–833, 2008.

[10] A. Ng, M. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. *Advances in Neural Information Processing Systems*, 14:849–856, 2002.

[11] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet. More efficiency in multiple kernel learning. In *Proceedings of the International Conference on Machine Learning*, pages 775–782, 2007.

[12] B. Scholköpf and A. Smola. *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002.

[13] S. Ullman, M. Vidal-Naquet, and E. Sali. Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, 5(7):683–687, 2002.

[14] H. Valizadegan and R. Jin. Generalized maximum margin clustering and unsupervised kernel learning. *Advances in Neural Information Processing Systems*, pages 1417–1424, 2007.

[15] M. Wu and B. Schölkopf. A local learning approach for clustering. *Advances in Neural Information Processing Systems*, 19:1529–1536, 2007.

[16] S. Yu and J. Shi. Multiclass spectral clustering. In *Proceedings of IEEE International Conference on Computer Vision*, pages 313–319, 2003.

[17] L. Zelnik-Manor and P. Perona. Self-tuning spectral clustering. *Advances in Neural Information Processing Systems*, 17:1601–1608, 2005.

[18] H. Zeng and Y. Cheung. Feature selection for local learning based clustering. In *Proceedings of Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2009. To appear.

[19] D. Zhou and C. Burges. Spectral clustering and transductive learning with multiple views. In *Proceedings of the International Conference on Machine Learning*, pages 1159–1166, 2007.

Analysis of Recall Characteristics in Image Search Engines

Xiaoling WANG

Abstract

Precision and recall are important measures of retrieval effectiveness. While for finite databases such measures of performance can be observed in a straightforward manner, these measures are difficult to estimate for an infinite database as the relevant parameters are not directly observable. In this paper, we mainly focus on how to estimate recall on infinite databases. Estimating recall shows a challenge. We will use the hypergeometric distribution to help us to estimate the total relevant images in each image search engine (ISE).

Keywords: Image Search Engine, Recall, Precision, Hypergrometric Distribution, World Wide Web.

1. Introduction

Due to the increased importance of the Internet, the use of image search engines such as Google, Yahoo, Ask and Msn is becoming increasingly widespread in the world. Meanwhile, with the rapid developments in both digital cameras and digital recording devices, large collections of images have been made available on the web. However, for so many image search engines, a decision must be made by users which image search engine should be selected to search. In addition, major image search engines in use at present are limited in terms of the evaluation method (precision and recall); the need for a performance evaluation of current engines will be of great benefit to users. Therefore, retrieval effectiveness becomes one of the most important parameters to measure the performance of image retrieval systems.

In information retrieval, there are many evaluation measures, such as recall and precision, F measure, break-even point and 11-point averaged precision. More details about these evaluation criteria can be founded in 'Evaluation Techniques and Measures' in an appendix of TREC-8 [4]. However, in this paper, we focus on two primary evaluation measures: the first of the two primary evaluation measures is recall, which shows the ability of a retrieval system to present all relevant items, and the second is precision, which shows the ability of a retrieval system to present only relevant items. Measuring the recall presents a challenge because here we are dealing with an infinite image database, whose total number of images and total relevant images can be regarded as infinite.

2. Related Work

Several studies, such as [1, 2, 3], are done to evaluate ISE_k. Such as study [1], the recall has been computed like this:

$$\text{Recall of ISE}_k = \frac{|R_k|}{|R_1 \cup R_2 \cup \ldots \cup R_N|}$$

Where R_i is the set of relevant images relating to database i, with $|R_i|$ denoting the number of images in the set, and N is the number of ISE's under evaluation.

As we can see above, when calculate the total number of relevant images in the database, it combined all the returned results instead of using the summation of all search engines, which means that the overlap in results returned by all search engines will be excluded from the total number of relevant images across all the search engines.

An algorithm called sample-resample is presented in [2] that is extremely efficient; in environments containing resource descriptions already created by query- based sampling, the sample-resample method uses several additional queries to provide an estimate of database size. Therefore, if the database size has been known, then the distribution of relevant images can be estimated.

In the following, we will make use of a set of tagged relevant images, which will be tested on the search engine in question. And we will use the hypergeometric distribution [5] to estimate the database size and the total number of relevant images.

3. Analysis of Recall Characteristics Using the Tagged Relevant Image Method

3.1 Framework

The basic framework consists of a finite number of major image search engines $ISE_1,\ldots,$ ISE_k in the world, whose collection of images are viewed as infinite. While such infinite collection of images may be regarded as common and potentially available to all the search engines, each engines tend to have their own preferred subsets of the infinite collection $DB_1,\ldots, DB_k$, where we take

$|DB_i|>>1$ for $i=1, 2,\ldots, k$

in such a manner that it is assumed impossible or impractical to scan through the whole database. Different engines may activate different mechanisms, including any metadata extraction, content annotation, image object and relationship indexing for locating relevant images based on concept-based and content-based methods, and thus their search performance will exhibit differences in retrieval competence. This investigation will establish performance measures to enable the performance characteristics of different image search engines to be assessed.

3.2 Evaluation Methodology

For illustration purposes, we consider the following simple setting: each object is associated with a binary label l indicating whether the object to be relevant or not, where l=+ indicates relevant and l=- indicates non-relevant. In addition, the system produces a result s indicating whether the relevant images have been returned.

		Result(s)	
		+	-
Label (l)	+	TP	FN
	-	FP	TN

The experimental outcome may be conveniently summarized in a contingency table: where + and – stand for relevant and non-relevant, TP and TN stand for true positive and true negative respectively, while FP and FN for false positive and false negative respectively. While recall and precision are critical measures of retrieval effectiveness, they are, for finite collections, normally considered to be deterministic ratios. Therefore, one can compute the precision (p) and recall (r):

$$p = \frac{TP}{TP + FP} \qquad r = \frac{TP}{TP + FN}$$

But in the present context where parameters cannot be known precisely, we shall adopt a probabilistic interpretation of these and related measures, and apply stochastic analysis to study their performance characteristics. We can define precision as the conditional probability [6] that an image is relevant given that it is returned by the system; while recall is the conditional probability that a specific relevant image is returned:

$$p=P(l=+ \mid s=+)$$

$$r=P(s=+ \mid l=+)$$

We make use of a set of tagged relevant images, which may be obtained from a different search engine and tested on the search engine in question. Without loss of generality, we assume that this tagged set of relevant images is obtained from ISE_1 (AltaVista). Denoting this set by R_1, and letting $|R_1|=r_1$, then the probability that ISE_2 containing m $(<=r_1)$ of R_1 out of N returned images follows the Hypergeometric Distribution [5]:

$$\binom{r_1}{m}\binom{|DB_2|-r_1}{N-m} \div \binom{|DB_2|}{N}$$

According to the distributional property of the Hypergeometric Distribution, we learn that the expected value (mean) is $\frac{N * r_1}{|DB_2|}$, therefore, we shall determine N, m experimentally, and from the above distributional properties together with observed values, we can obtain an estimate of the infinite database size $|DB_2|=\frac{N * r_1}{m}$, where m indicates the number of tagged images be returned by the image search engine in question. Refer to the same query; m is the same, which means the expected value is constant. Because the result returned by image search engine is the same every time we do the image retrieval. Furthermore, if there is a total of x relevant images in the set of N returned images, the total number of relevant images q_2 for an image query Q in $|DB_2|$ may be estimated by xr_1/m, where x is again experimentally observed. Then the recall of each ISE_i, $i=2,\ldots, k$, can be estimated as

$$Recall=x/q_i, \text{ where } i=2,\ldots, k$$

Thus, the number of relevant images across all the database may be estimated by $q^*=max(q_i), i=1,2,\ldots, k$. From this, we can go back to iteratively re-compute the recall rate of ISE2. Likewise, we can determine the revised recall rate estimations for the remaining search engines $ISE_3,\ldots, ISE_k$.

$$Revised\ Recall=x/q^*$$

Global performance characterization for given search engines may then be developed through the execution and observation of returned results for a set of representative image search queries$Q_1,\ldots, Qs$.

3.3 Experiment and Results

The test query Q is “plane”

Firstly, we obtained 50 relevant images from the ISE_1 (AltaVista). Following that, we will test on some major image search engines($ISE_1,\ldots,$ ISE_k) as follows:

1. Google www.google.com
2. Yahoo www.yahoo.com
3. Msn www.msn.com
4. Ask www.ask.com
5. Lycos www.lycos.com
6. AltaVista www.altavista.com

Precision

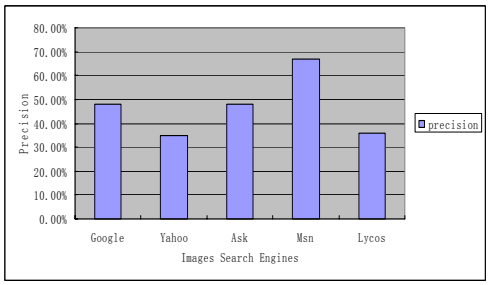


Figure 1: precision of five ISE

To compare the precision, Msn search engine have a top results 67%. Google and Ask have the same scores with 48%, while the engine has the lowest score is Yahoo.

Recall and Revised recall

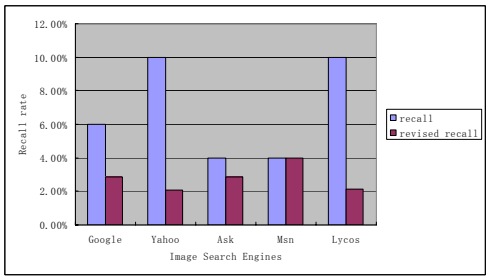


Figure 2: recall and revised recall of five ISE

For the recall, Yahoo and Lycos have the top result with 10%, while Google with 6% and MSN and Ask with the lowest score, which is 4%. However, after estimate the total number of relevant images across all the databases, we can obtain the revised recall as follows: all ISE are very low around 3% of relevant images returned across all the search engines.

4. Conclusion and Future Work

The objective of a good image search engine is to retrieve as many relevant items as possible meanwhile to reject as many irrelevant items as possible. Therefore, according to the experiment above, we can see that Msn provide the best result in precision and revised recall meanwhile the Lycos show the worst result in precision. Yahoo and Lycos have the best result in recall but provide the worst result in revised recall.

My future work is modeling cumulative page image relevance using stochastic analysis, and extrapolating the image relevance deterioration pattern. And then we will develop and formulate the optimal stopping rules and sequential stopping rules based on dynamic observable variable.

5. Reference

[1] K. Stevenson and C. Leung, "Comparative Evaluation of Web Image Search Engines for Multimedia Applications", *In Proceedings of IEEE International Conference on Multimedia and Expo*, July 2005

[2] Si, L. and Callan, J. 2003, "Relevant document distribution estimation method for resource selection", In *Proceedings of the 26th Annual international ACM SIGIR Conference on Research and Development in informaion Retrieval* Toronto, Canada, July 28 - August 01, 2003. SIGIR '03. ACM, New York, NY, 298-305

[3] Ishioka, T. 2003, "Evaluation of Criteria for Information Retrieval", In *Proceedings of the 2003 IEEE/WIC International Conference on Web Intelligence* (October 13 - 17, 2003). Web Intelligence IEEE Computer Society, Washington, DC, 425

[4] TREC-8, Evaluation Techniques and Measures, TREC-8 Results, page A-1, NIST Special Publication 500-246: The Eighth Text Retrieval Conference (TREC-8), Gaithersburg, Maryland, November 17-19, 1999. Available online http://trec.nist.gov/pobs/trec8/t8_proceedings.html

[5] Hypergeometric Distribution, http://en.wikipedia.org/wiki/Hypergeometric

_distribution
[6] Cyril Goutte and Eric Gaussier, “A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation”, *In Proceedings of Springer-Verlag Berlin Heidelberg, ECIR 2005, LNCS 3408, pp. 345-359, 2005*
[7] Raquel Kolitski Stasiu, Carlos A. Heuser, and Roberto da Silva, “Estimating Recall and Precision for Vague Queries in Databases”, *In Proceedings of Springer-Verlag Berlin Heidelberg, CAiSE 2005, LNCS 3520, pp. 187-200, 2005*
[8] Münevver Tuğçe Elagöz, Mehtap Mendeli, Remziye Zeden Manioğluları, Yıltan Bitirim. “An empirical Evaluation on Meta-Image Search Engines”, *In Proceedings of IEEE International Conference on Multimedia and Expo*,, June 2008

HMM-LDA Feature Extraction for Mining of Product Ownership of Online Forum Participants

Tianjie ZHAN

Abstract

There are great of amount of semantically feature in the online media which is useful for product accurate advertisement and product comments summary. This paper is to address the problem of mining of product ownership of online forum participants. As the problem is required to classify documents and make semantically analysis at the same time. We appeal the one of the most popular text mining model recently, Latent Dirichlet Allocation (LDA). However, there are many shortcomings for supervised setting LDA including not discriminating between syntactical words (functional words) and semantically words(content words). The latter one is great more important for supervised leaning. The syntactical words contribute little to the class discriminating and even confusing the classification. So we choose the HMM-LDA to improve this situation, which would generate syntactical class and semantically topics. What is more, we proposed an enhanced method for the topics modes to generate semantically topics of clearer, more interpretable and richer in semantically nature. Using such enhanced topics, the classification is improved lots. Based on the all the topic models, we extract proper feature for the document classification using a state of art classifier SVM. The experimental results show that enhanced HMM-LDA achieves a significant improvement in the classification accuracy than LDA and extracts richer semantically and more interpretable topics than LDA.

1. Introduction

Online forum, web blogs and other social media websites serve not only as important media for communications among individuals but also provide ample data mining opportunities. Product owners, potential product buyers and other interested parties often post messages to share information, seek for help and give suggestions and comments on products. There are often good business interests to capture such information to find potential buyers, product concerns or shortcomings, which will be helpful for advertising improvement, market analysis and product planning. Among such analysis, the primary one is the product ownership discovery of whether the person concerned has owned or not owned a particular product and the period of its presence. In this paper, we focused on this problem referred as "mining product ownership of online participants" (MPOOP). To solve the problem of discovery of product ownership, we employ text mining techniques and machine learning models.

In recent years, a popular model for text mining is posed as a generative probabilistic model, latent Dirichlet allocation model (LDA) [1]. There are various extensions such as (1) the multi-grain topic model [7] for customer review mining based on "local topics" and "background topics", (2) delta-LDA [2] for statistical program debugging. In supervised learning, supervised LDA[5] was proposed by introducing a label variable to model the dependence between label variables and topics, which was effective in web spam classification [6]. In review summary application, *] M. Hu and B. Liu* has the efficiency of topic model in [4].

The content feature of documents represented by the topics detected by LDA will be of use for classification of documents and provides great amount of semantically interpretation about the classification results. What is more, HMM-LDA[3] will provide features including functional words feature and content words feature, which will play different roles in the application of document classification and be of advance to improve the accuracy. To address the problem of ownership mining, a novel algorithm has to be introduced to address the supervised and semantically rich problem nature by extracting syntactical and semantically feature from the documents.

This paper is organized as follows. In Section 2, the mining of product ownership problem will be defined first, and then LDA, HMM-LDA will be introduced following the inference learning. In section 3, an enhanced learning for LDA and HMMLDA will be proposed to strengthen the power of feature extracted from the two models. Experimental results on a recent data set from a popular digital camera website will be given which will compare the performance with the state of art classifier $SVM^{multiclass}$ [9] in Section 4. We discuss the experiments in section 5 and the future work in Section 6.

2. Proposed methods

In this section we will define the problem addressed in this work. More over, the topics model will be introduced

with the inference learning, based on which the feature will be extracted for classification in next step.

2.1. The problem of mining of product ownership for online forum participants

The problem of "mining of product ownership of online forum participants"(MPOOP) includes two parts, (1) to detect product ownership from the messages posed on the online forum for each participated authors, (2) detect and analysis topics associated with different ownership.

2.1.1 Terms definition

(1) Author a_i is one unique valid poster in some online forum which is also referred to as participant.

(1) Message m_j^i is the j^{th} message posted by author a_i along the time, which contains some message of the target product g , e.g, the product name. In this paper, message is also referred as documents..

(3) Ownership: Given some sets of message m_j^i of author a_i , there are 3 types of unit ownership defined in this paper including, "positive", "negative" ownership and "unknown" ownership. Positive ownership is some kind of recognition that author a_i owns the product g only based on the text feature of the messages. "Negative" means "not owns". And the "unknown" ownership means that it is not clear about the ownership of author a_i from the messages.

2.1.2 Problem definition

Problem Definition: Given a set of authors $A = \{a_i\}_{i=1}^{N_A}$, messages $M = \{\{m_j^a\}\}_{a=1}^{N_A}$ associated wtih authors in A , of the messages, the problem is to determine the ownership class of each message, and detect the topics conditioned on different ownership.

Here Message m_j^a is made of work token $\{w_t\}$.

We divide the problem into three subtasks: (1) select the product related messages; (2) extract features from the messages including the semantically feature; (3) classify the messages into 3 groups with each associated with one class of ownership. In the first subtask, we used a simple scheme to select product related messages where only messages were selected that contain complete product name or its usual alias.

2.2. LDA

LDA is a popular generative model for document modeling and latent topics detection in recent years. The process corresponds to the graphical model shown in Figure 1.

In LDA, w denotes one work token in the document, z is the topic indicator variable, ϕ_i denotes parameters of the multinomial distribution of words conditioned topic i, which is drawn from the Dirichlet(β) prior independently. And θ is topic distribution of a document with a Dirichlet(α) prior.

Generative process could be generating as follows:

(1) For topic $i = 1, 2, .., T$

Draw $\phi_i \sim Dirichlet(\beta)$

(2) For each document d

Draw $\theta \sim Dirichlet(\alpha)$,

For each word token $t = 1, ..., N_d$

i. Draw $z_i \mid \theta \sim Discrete(\theta)$;

ii.Draw $w_i \mid z \sim Discretel(\phi_z)$.

Here T is the number of topics and N_d is the number of tokens in document d.

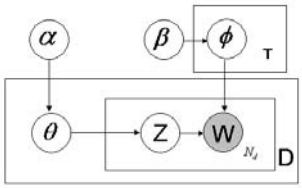


Figure 1: Graphical model of LDA

2.2.1 Inference

Algorithm such as variational inference, collapse Gibbs sampling [8] have been presented to estimate the parameters ϕ and θ. Efficient estimation of hyperparameter α has been given as a non-iterative method in [8].

The probability distribution of topic z conditioned on document d is given in (1) and token w conditioned on topic z in (2).

$$p(w = i \mid z = k) = \frac{\Gamma(W\beta)}{\prod_{j=1}^{W}\Gamma(\beta)} \frac{\Gamma(n_{k,i} + \beta)}{\Gamma(n_{k,*} + W\beta)} \quad (1)$$

$$p(z = k \mid d) = \frac{\Gamma(K\alpha)}{\prod_{k=1}^{K}\Gamma(\alpha)} \frac{\Gamma(n_k^d + \alpha)}{\Gamma(n_*^d + K\alpha)} \quad (2)$$

Here n_*^d is the sum of the number of words in document d. n_k^d is the sum of words assigned to the topic k in document d. And $n_{k,i}$ is the number of times of

word i assigned to topic k, $n_{k,*}$ is the number of times of all words assigned to topic k in all documents.

K is the number of topics and W is the size of the dictionary.

The conditional probability distribution of z_i on other information for Gibbs sampling approach is given as

$$p(z_i \mid z_{-i}, w) = \frac{n_k^d + \alpha}{n_*^d + K\alpha} \frac{n_{k,i} + \beta}{n_{k,*} + W\beta} \quad (3)$$

$-i$ means all other tokens except the current one i.

And the hyperparameters could be inferred from the topics indicator variable samples by

$$\hat{\phi}_{i,j} = \frac{n_{i,j} + \beta}{n_{i,*} + W\beta}, \ \hat{\theta}_{i,d} = \frac{n_i^d + \alpha}{n_*^d + K\alpha}. \quad (4,5)$$

Here $\hat{\phi}_{i,j}$ is the probability of word i conditioned on topic j and $\hat{\theta}_{i,d}$ is the probability of topics i conditioned on document d.

2.3. HMMLDA

In a document there are syntactical words which is also referred as functional words and semantically words referred as content words. However, in LDA, they are treated the same which will mess up the different important impact in different application. So Griffifths and Steyvers proposed the HMM-LDA models in [3], which will discriminate the syntactical class and semantically topics by using the characteristic of syntactic words in the word sequence, which will be modeled by HMM. The graphical model of HMM-LDA is in Figure 2.

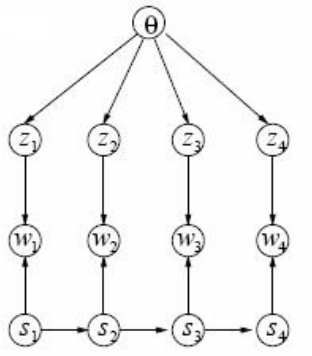


Figure 2: Graphical model of HMM-LDA.

2.3.1 Composite Process

In HMM-LDA, there are S syntactical classes $\{s_i\}$ with the first class s_0 denoting the sentence marker and the second one s_1 the semantically class which includes K semantically topics $\{k_i\}$ while $\{w_i\}$ indicates the word token, $\{c_i\}$ is the syntactical class indicator variable and $\{z_i\}$ is the semantically topic indicator variables. If token w_i is in the syntactical class $s_i (i > 1)$, it will follow ϕ_{s_i} while it will follow a word distribution ϕ_{k_i} if it in the semantically class s_1 and assigned to topic k_i.

It is shown the generative process as follows.

(1) For each syntactical class s_i

Draw $\phi_{s_i} \sim Dirichlet(\beta_s)$

(2) For each semantically topic k_i

Draw $\phi_{k_i} \sim Dirichlet(\beta_K)$

(3) For each document d

Draw $\theta_d \sim Dirichlet(\alpha)$,

For each word token $t = 1, ..., N_d$

Draw $z_i \mid \theta_d \sim Discrete(\theta_d)$

Draw $c_i \sim Discrete(\pi_{c_{i-1}})$

If $c_i = 1$,

Draw $w_i \sim \phi_{z_i}$

Else Draw $w_i \mid z \sim Discretel(\phi_{z_i})$.

Here π_i is the transfer probability from class i in HMM, and it is drawn from the Dirichlet prior $Dirichlet(\gamma)$. α, β_S, β_K and γ are the hypyerparameters.

2.3.2 Inference

Gibbs Sampling are used to approach the model and samples $\{c_i\}$, $\{z_i\}$ which will be used to infer the model parameters.

The conditional probability of syntactical c_i on other variables will given in (6) and semantically z_i on other variables is given in (7), both of which are used to get samples from Gibbs sampling.

$$P(z_i|\mathbf{z}_{-i}, \mathbf{c}, \mathbf{w}) \propto P(z_i|\mathbf{z}_{-i}) \quad P(w_i|\mathbf{z}, \mathbf{c}, \mathbf{w}_{-i})$$
$$\propto \begin{cases} n_{z_i}^{(d_i)} + \alpha & c_i \neq 1 \\ (n_{z_i}^{(d_i)} + \alpha) \quad \frac{n_{w_i}^{(z_i)} + \beta}{n^{(z_i)} + W\beta} & c_i = 1 \end{cases} \quad (5)$$

$$P(c_i|\mathbf{c}_{-i},\mathbf{z},\mathbf{w}) \propto P(w_i|\mathbf{c},\mathbf{z},\mathbf{w}_{-i}) \quad P(c_i|\mathbf{c}_{-i})$$

$$\propto \begin{cases} \frac{n_{w_i}^{(c_i)}+\delta}{n_{\cdot}^{(c_i)}+W\delta} & \frac{(n_{c_i}^{(c_i-1)}+\gamma)(n_{c_{i+1}}^{(c_i)}+I(c_{i-1}=c_i)\cdot I(c_i=c_{i+1})+\gamma)}{n_{\cdot}^{(c_i)}+I(c_{i-1}=c_i)+C\gamma} \quad c_i \neq 1 \\ \frac{n_{w_i}^{(z_i)}+\beta}{n_{\cdot}^{(z_i)}+W\beta} & \frac{(n_{c_i}^{(c_i-1)}+\gamma)(n_{c_{i+1}}^{(c_i)}+I(c_{i-1}=c_i)\cdot I(c_i=c_{i+1})+\gamma)}{n_{\cdot}^{(c_i)}+I(c_{i-1}=c_i)+C\gamma} \quad c_i = 1 \end{cases} \tag{6}$$

2.4. Enhanced topics models

In LDA, each word has the same power to guide the clustering process of word so that common words or stop words like 'the','a','an' will give a great impact to influence the topics generating and stay at a place of high probability in the word distribution of the topics. As the multinomial distribution treats all the word as the same, the word with dominant number of word counting will dominant the topic no matter what dose the word means and whether it is respective enough for the topics.

To enhance the word which is typical and respective to influence the topic forming process, we proposed a method by incorporating the TFIDF feature into the model.

The LDA is enhanced by redesigning the probability of word conditioned on topics as follows.

$$p(W \mid z,\beta) = \int p(\phi_z)\prod_{w\in W}(\phi_z)^{n_w * idf(n_w)} d\phi_z \qquad [7]$$

$idf(w) = \log(N_D/(df(w)+1))$ is the inverse document frequency (IDF) and $df(w)$ is the number of documents where word w occurs, N_D is the number of documents in corpus.

While $\phi_z \sim Dirichlet(\beta)$, [7] is recomputed and given in (8).

$$p(w=i \mid z=k) = \frac{\Gamma(W\beta)}{\prod_{j=1}^{W}\Gamma(\beta)}\frac{\Gamma(idfn_{k,i}+\beta)}{\Gamma(idfn_{k,*}+W\beta)} \tag{8}$$

Here $idfn_{k,i}$ is similar to LDA, but the counting of times of word i assigned to topic k with the weighting $idf(i)$ for each counting while $n_{k,i}$ is just counting without special weighting.

Other parts is identical to the LDA.

2.4.1. Approximated Inference. Similar to the standard LDA, exact parameters estimation is intractable. Gibbs sampling [9] is used to estimate the parameters including z, $idfn_{k.i}$.

From (8) and (2), the posterior distribution of topic distribution $p(z \mid w, z_{-i})$ can be inferred as in (9) where z_{-i} denotes all other topic indicator variables z except z_i.

$$p(z_i \mid z_{-i}, w) = \frac{n_k^d+\alpha}{n_*^d+K\alpha}\frac{\Gamma(idfn_{k,i}+\beta)}{\Gamma(idfn_{k,*}+W\beta)} \tag{9}$$

Here n_*^d is the sum of the number of words in document d. n_k^d is the sum of words assigned to the topic i in document d. And $idfn_{k,i}$ is the number of times of word i assigned to topic j in all documents with IDF weighting. $idfn_{k.*}$ is the number of times of all words assigned to topic k in all documents. $-i$ means all other tokens except current token i.

K is the number of topics and W is the size of the dictionary.

To approximate the second part of (9), we use the following method [10] simplify the calculation.

$$\frac{\Gamma(x+y)}{\Gamma(x)} \approx Fgmma(x,y) = \begin{cases} x, if(y<=1) \\ \prod_{i=1}^{[y]}(x+i-1) \end{cases} \tag{10}$$

Here $[\square]$ is the rounding operation.

So (9) could be simplified as

$$p(z_i \mid z_{-i}, w) = \frac{n_{-k,i}^{d_k}+\alpha}{n_{-k,*}^{d_k}+K\alpha}\prod_{l=1}^{[idf(w_i)]}\frac{idfn_{k,i}-l+1}{idfn_{k,*}+W\beta-l+1}. \tag{11}$$

And parameters θ, ϕ can be inferred as in (12),(13) from the samples generated by Gibbs sampling.

$$\hat{\phi}_{ij} = \prod_{l=1}^{[idf(w_i)]}\frac{idfn_{k,i}-l+1}{idfn_{k,*}+W\beta-l+1} \tag{12}$$

$$\hat{\theta}_{id} = \frac{n_i^d+\alpha}{n_*^d+K\alpha} \tag{13}$$

HMMLDA could be also included the IDF feature by similar operation into the semantically part.

3. Feature Extraction

From the models discussed in Section 2, we could inferred great rich semantically feature or even the syntactically feature for analysis. As for the problem of mining of product ownership, the task is to classify the documents. So feature used for classification will be focused and discussed in this section.

3.1. LDA

LDA model could gives the semantically feature by inferring the $\hat{\theta}$, which is the topic distribution of documents. As in different class, documents will focus on different topics and presents divergence distribution on topics. So it is expected that the semantically feature from LDA will be efficient for the ownership classification which will be discussed by experiments in Section 4.

3.2. HMM-LDA

The HMMLDA could gives the syntactical class information and also the semantically feature. It is argued that the syntactical class distribution of documents provides less discriminative power for documents classification as the distributions are almost the same for all documents which will be illustrated with experiment results in Section 4. So we extracted the semantically feature of HMMLDA and keeps the text feature as the word vector for the tokens which has not been assigned into the semantically class. The feature extraction approach is illustrated in Figure 3.

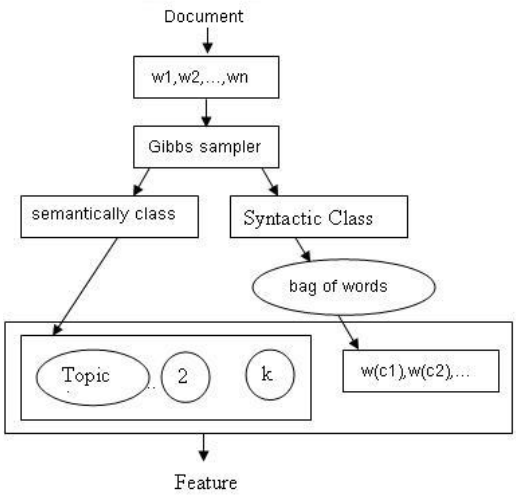


Figure 3 Flow char of feature extraction from HMM-LDA.

3.3. Enhanced topics model

In this paper, we develop the enhanced topic models to provide enhanced topic models. It will be illustrated that the enhanced LDA could generate topics more interpretable and improve the classification accuracy. As some very important word to discriminate the different class will be under estimated or even ignored by giving a very small probability to the word conditioned on topics.

Some important and representative word may have smaller number of word occurring relative to the common words such as “the”,”a”,”an” and etc. So the classification may be mistaken by the common words. And the discriminative word will be decreased its influence.

But enhanced topic model would increase the discriminate power of such important or representative word and decrease that of the common words. So it could increase the classification accuracy which will be supported by experiments in Section 4.

4. Experimental results

We prepare a dataset from an online forum in a popular digital camera website. There are 10 different camera models in this product set whose names $P = \{p_i\}_{i=1}^{10}$. In each message, we take word sequences of 10 words before and after each location of each camera model name. The data set totally contain 58342 messages of 7031 authors, 1100 of which are labeled. When training HMM-LDA, each “20 words sequence” are ended with an sentence marker..

We ran the Gibbs sampler on the dataset for LDA, (LDA + IDF) with the standard hyperparameters setting as $\alpha = 50/K$, $\beta = 0.01$, on the $K = 50$ topics. And for HMMLDA and (HMMLDA +IDF), we set $\gamma = 0.1$, $\alpha = 50/K$, $\beta_K = 0.01$, $\beta_S = 0.01$, on $S = 30$ syntactical class and $K = 50$ semantically topics. All samples are drawn from Gibbs sampler after 200 iterations.

4.1. Documents classification

After feature extracted from the generative models, we use the feature to train the SVM classifier and then classify the message in the test set. For comparison, we compare the features of (1) “Bag of words” vector, (2) LDA topic distribution $\hat{\theta}$, and also that of the (LDA + IDF), (3) the semantically topics distribution $\hat{\theta}$ and the syntactical “bag of word” word” vector including the words which are not assigned to the semantically class. Fo

We random draw some percent of labeled data from the 1100 labeled data, and test the classifier with the whole 1100 data. The classification results after 10 times random draw and testing is shown in Figure 4. The “SVM” means using just the “bag of word” feature to classify. From the classification result, it is shown that the (HMM-LDA + IDF) feature is nearly the same accurate to the classification using all text feature. At the same time, (HMM-LDA + IDF) can detect the semantically topics related to different ownership, which will be shown in Section 4.2.

We plot another picture to explain why (HMM-LDA + IDF) can improve the classification accuracy in Figure 5. The Figure 5(a) is syntactical class weighting of different ownership, which is the number of words assigned to different syntactical class except the “sentence marker” and “semantically class”. The Figure 5(b) is the statistical results of number of words assigned to different topics in documents labeled as different ownership. All the statistical results are normalized to sum to one. It could be seen that it is the statistical weighting of syntactical class

is so similar among the two ownership class that it could not be used to discriminate the different ownership class. That is why we do not use the syntactical class weighting as the classification feature like that of the semantically topics, but use the original "bag of word" as the classification feature for the words assigned to the syntactically class. So we keep the words assigned to syntactical class not changed, but represent the the words assigned to the semantically class as the topics distribution which is correspond to the $\hat{\theta}$ in LDA. In this experiments there are 159023 tokens out of 309215 tokens are assigned to the semantically class from HMM-LDA Gibbs sampler and 167908 tokens for (HMM-LDA + IDF) Gibbs sampler. So around half of the words are assigned to the semantically class and transformed to be the topics weighting feature by inferring $\hat{\theta}$.

The (HMM-LDA + IDF) could do better because it promotes the influence of the important and representative words of the topics which is more discriminate than the common words. The experimental results support our arguments, (1) LDA+ IDF beyond the LDA, HMM-LDA + IDF beyond the HMM-LDA (2) HMM-LDA beyond the LDA models.

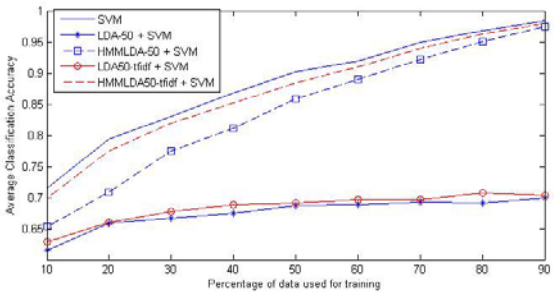
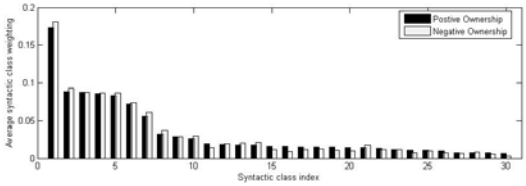
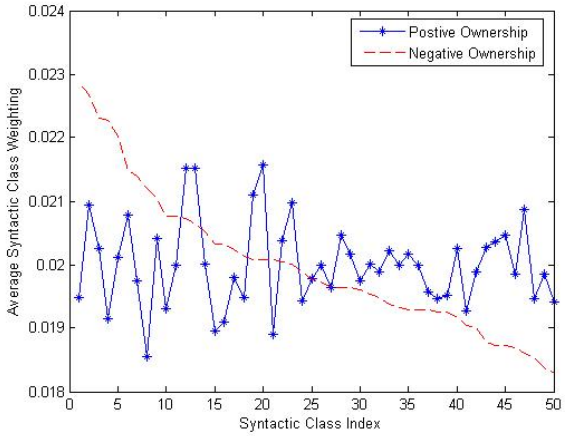


Figure 4 Classification results of the 5 models



(a) Comparison of the average syntactical class weighting of documents labeled as "positive ownership" and "negative ownership".



(b) Comparison of the average semantically topics weighting of documents labeled as "positive ownership" and "negative ownership"

Figure 5 Statistical analysis of syntactical class and semantically topics samples from HMM-LDA Gibbs sampler.

4.2. Class related topics

We select some representative topics related to "positive ownership" and "negative ownership" with the topics represented by the 10 most probable words conditioned on the topic. The results are shown in Table 1, which is topics related to different ownership selected by hands.

LDA	♦ 'the and to use only now also shots d rebel' ♦ 'the in on is image shot a for of and' ♦ 'with for flash raw no or and a shoot using'
LDA+ IDF	♦ 'only rebel use shots also now taken around with' ♦ 'flash raw shooting shoot images mode no user fast' ♦ 'in image shot quality quite photo bit is 800 of'
HMM LDA+ IDF	♦ 'picture different system lower 800 help seen enough kit choose' ♦ 'shot images shots test almost 200 myself black owner combo' ♦ 'exposure raw iso comparison many close jpeg looking likely anyway'

(a) Topics selected for "unknown" ownership class

LDA	♦ 'that I have the not to but think so it' ♦ 'be to will or would 40d a buy should upgrade' ♦ 'the for is a and price it great me difference' ♦ 'it t not that but don know s do work'

LDA+ IDF	♦ 'that 400d of people about is s who say this' ♦ 'be will d40 40d would should price lenses d40x buy' ♦ '30d 20d or upgrade would to over buy get either' ♦ 't don know he doesn they want that didn re'
HMM LDA+ IDF	♦ '40d probably feel stop ll buying second issues least version' ♦ 'say before small choice really cost being about phil low' ♦ 'better light keep extra consider hands show backup 500 coming' ♦ 'bought fine sold problems question anything real color gt 10'

(b) Topics selected for "negative" ownership class

LDA	♦ 'i my got ve just bought had for been have' ♦ '300d my from and still with a used gallery -' ♦ 'i m for am have a about sure looking happy'
LDA+ IDF	♦ 'own since back before again e ve two couple hope' ♦ '300d from am my gallery new 10d about to this' ♦ 'got my bought just first had after ve was when'
HMM LDA+ IDF	♦ 'work love thinking last yet took own purchased upgraded clean' ♦ 'much years faster two cameras iq works deal every now' ♦ 'mode shooting auto p decided upgrading xt experience amp buffer'

(c) Topics selected for "positive" ownership class

Table 1 Topics related to different ownership class selected by hands.

From Table 1 , The (HMM-LDA + IDF) captured the most interpretable words and discover topics of richer semantically nature. And then (LDA+IDF) gives a moderate performance, LDA worse.

For example, in Table 1(a) , it is easy to read the topics of (HMM-LDA + IDF) discussing the "picture processing by choosing different kits", "sharing photography by the camera owner to test its new camera", "exposure and data storing with different data format". But it is just known from LDA to find topics related to "photo shooting". In Table 1(b), (HMM-LDA + IDF) found that authors discussed about "the issue of new camera model releasing", "considering taking a new camera as backup one", "the experience of selling its last camera", all of which related to the "negative". In Table 1(c), (HMMLDA-IDF) found the topics of "owning a new camera just for upgrading", "owning two camera for works", "decide to buy a auto shooting camera".

5. Discussion

This paper has proposed a novel method to build enhanced topic model by incorporating the TFIDF information and a feature extraction method for the problem of mining of production ownership of online forum participants. It is shown that from the experiment results, (HMM-LDA) not only improve the classification accuracy which is almost approach to that of SVM classifiers with "Bag of word" feature, and also mining richer and more interpretable topics related to different ownership class.

The syntactical class and semantically topics holds rich information for text processing, but the LDA does not discriminate the important and representative words from
the common words, generating topics less meaningful and separable. HMMLDA are capable to samples the syntactical class and semantically topics, getting further to make finer grained semantically analysis. The IDF information enhanced the quality of the topics for both LDA and HMM-LDA. It is primarily to argue that the give words different importance will be efficient to generate clearer, and richer semantically topics.

6. Future work

We will try to find more sophisticated method to incorporate more statistical information to give words more accurate importance weighting. And we try to modify the HMM-LDA to capture more interesting and complex text structure. Inn this paper, it is a pity to selecte the class related topics by hands. We will try to find some measure to select the class related topics automatically.

References

[1] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet allocation. Journal of Machine Learning Research, 3:993–1022, 2003.

[2] Andrzejewski D., Mulhern, A., Liblit, B. Zhu J. Statistical Debugging using Latent Topic Models. ECML 2007.

[3]Thomas L. Griffiths, Mark Steyvers, David M. Blei, and Joshua B. Tenenbaum. Integrating Topics and Syntax. *Advances in Neural Information Processing Systems 17*, pp. 537--544, 2005

[4] M. Hu and B. Liu. Mining and summarizing customer reviews. In Proceedings of the 2004 SIGKDD international conference on Knowledge discovery and data mining, pages 168–177.

[5] David Blei and Jon McAuliffe. Supervised topic models. In Advances in Neural Information Processing Systems 21. 2008.
[6] I. B´ır´o, J. Szab´o, A. A. Bencz´ur. Latent Dirichlet Allocation in Web Spam Filtering. In Proc. 4th AIRWeb, 2008.
[7] I. Titov and R. McDonald. Modeling online reviews with multi-grain topic models. In Proceedings of the Annual World Wide Web Conference (WWW), 2008.
[8] Gregor.Heinrich. Parameter estimation for text analysis. URL http://www.arbylon.net/publications/text-est.pdf
[9] Thorsten Joachims, SVM[multiclass] : Support Vector Machine for multi-class classification, Version: 2.20. Software available at http://svmlight.joachims.org/svm_struct.html
[10] Vijay Krishnan, Short comings of latent models in supervised settings, SIGIR2005:625-626

Opinion Mining: A Survey of the State-of-the-Art and Possible Extensions

Kwan Wai LEUNG

Abstract

With the wide spread use and the new generation of the Web, people interact through the Internet in many forms. Blogs, forums and online review systems are popular platforms for people to express how they experience about a service or a product. People always express their opinions in the content of texts they leave on those sites or systems and those opinions are kind of important information on the Web. This paper presents an overview of opinion mining by surveying and analyzing the developed methods for accomplishing the main tasks in opinion mining as well as different limitations of the studied techniques. Followed the survey, some possible applications of the sentiment analysis results are discussed. The idea of examining sentiments on a popular restaurant review system, OpenRice, is also introduced in the paper. As investigated, PMI would be less effective for examining sentiments on OpenRice as it dose not have a consistent performance for multi-aspect reviews. Publicly available linguistic resources like WordNet are also not suitable for the system since the system is Chinese based. As such, possible new methods are proposed for handling the OpenRice reviews on sentence level.

Keywords
Opinion mining, sentiment analysis, sentiment classification, subjective word identification

1 Introduction

The World Wide Web is enlarging in a startling rate where the information grows in an exponential rate. With the evolution of the Web, Web 2.0 allows users to contribute websites content. With the interactivity provided by the unique features of Web 2.0, networks are built with blogs or online review systems in different domain. It is also a common practice for merchants and service providers to enable their customers to have online reviews on their products or services. Among the contents in such blogs or review systems, there are enormous quantities of customer opinions expressed. As there are lots of products and services a pouser can choose from, users always face the problem of decision making. Thus, the feedbacks on the Web can make them have an easier decision making. Since those customer feedbacks influence other customer's decisions, it is important for business to take into accounts the feedbacks on Web for their market planning and business development.

With the increasing number of people writing review, the number of reviews grows rapidly. Additionally, reviews are always very long and have many sentence not referring customers' opinions. These make potential customers to read them to make decisions. Merchants are also hard to retrieve opinion information about the products or services from heaps of such long reviews. Reading a few reviews would give a biased view on products or services, and thus gather information from multiple reviews would be much reliable than from one. Therefore, systems which can automatically process various reviews and give an appropriate summary showing the sentiments are welcomed.

To allow such a system to be developed, techniques for sentiment analysis should be studied, which sentiment refers to the orientation of opinion about a subject. Subjective words or phrases extraction is yet another essential step for opinion mining as not all the sentences and words in a review convey customer's opinion. For different goals of opinion mining, some other techniques should also be investigated, such as feature extraction for opinion mining on product of service features, sentiment classification for review orientation classifying.

Due to the importance of sentiment analysis, it is now a hot topic for research. Many methods for feature extraction, subjective word or phrase extraction, sentiment classification, sentiment analysis and opinion summarization have been studied. In this paper, I examine the investigated methods of the key tasks which are involved in opinion mining.

2 Opinion Mining Tasks

As mentioned in Section 1, there can be many tasks involved in opinion mining depending on the goals and how the summarized information is presented. However, the focus of opinion mining is the sentiments expressed in corpus. The major tasks for opinion mining can be roughly divided into identifying subjective words or phrases, and analyzing sentiment orientation.

3 Identifying Subjective Words or Phrases

There are basically two types of approaches for opinion words or phrases identification: (1) corpus-based approaches, and (2) dictionary-based approaches.

3.1 Corpus-Based

Corpus-based approach algorithms determine sentiments of words or phrases by discovering the co-occurrence patterns of words or phrases. Works in [1, 5] study this approach.

3.1.1 PMI-IR

PMI-IR uses Pointwise Mutual Information (PMI) and Information Retrival (IR) to measure similarity of pairs of words or phrases. Semantic association between two words or phrases can be measured by Pointwise Mutual Information (PMI) [6]. Semantic orientation can be inferred using the mutual information calculated. The Pointwise Mutual Information (PMI) of two words is defined as:

$$PMI(word_1, word_2) = log_2 \left[\frac{p(word_1 \ \& \ word_2)}{p(word_1)p(word_2)} \right] \quad (1)$$

where p($word_1$ & $word_2$) is the probability of $word_1$ and $word_2$ co-occur.

In [5], Turney presented a work on inferring semantic orientations using PMI-IR with unsupervised classification on reviews. The method introduced can handle phrases, adverbs and isolated adjectives. PMI-IR is employed to determine the semantic orientation. Phrases containing adjectives or adverbs are considered.

To identify subjective phrases, it firstly tags the words in a review by a POS tagger [8]. Adjective or adverb itself can state subjectivity, however, adjective or adverb alone may not have sufficient context to be used to estimate semantic orientation. For example, word 'unpredictable' in "unpredictable steering" is negative in an automotive review but positive in "unpredictable plot" in a movie review. Thus, pairs of words containing an adjective or an adverb and a consecutive context are extracted to be the subjective phrases.

PMI-IR is employed to estimate semantic orientation of each extracted phrase. It uses mutual information as a measure the strength of semantic association between the extracted phrase and a positive and negative reference word. The PMIs are computed via IR by counting the number of hits when queries are issued to a search engine. And the chosen positive and negative words are 'excellent' and 'poor' respectively. The Sentiment Orientation (SO) of a phrase, *phrase*, is:

$$\begin{aligned} SO(phrase) &= PMI(phrase, 'excellent') \\ &- PMI(phrase, 'poor') \end{aligned} \quad (2)$$

The sentiment of phrase is positive if SO(*phrase*) is positive, when SO(*phrase*) is negative, sentiment of phrase is then negative.

One of the limitations of the proposed method is that it is time consuming with IR which launches queries to search engine.

3.2 Dictionary-Based

Dictionary-based approach algorithms base on a set of seed opinion words, use the synonyms and antonyms in dictionary to indicate sentiment orientations. Both [2, 3] use WordNet as the dictionary and thesaurus.

3.2.1 WordNet-Based

In [2], Hu et al. identified opinion sentences and decided if each opinion expressed in each sentence is positive or negative with the assistance of the WordNet [7]. A set of adjective words is identified using natural language processing method. These words are the opinion words as they are usually be used in commenting a product, e.g. 'amazing', 'great', 'poor'. An opinion sentence is a sentence contains one or more product features and one or more opinion words. Semantic orientation of each opinion word is determined using a bootstrapping technique with WordNet [7]. The idea of adjectives share the same orientation as their synonyms and opposite orientations as their antonyms is used with a set of seed orientation-known adjectives to predict the orientations of all adjectives. The simple WordNet-based approach in [2] yields the context dependent opinion words problem.

> **Context dependent opinion words problem**
> There are many opinion words whose orientations are context dependent. For example, the word 'long' can indicate a positive or a negative opinion on a product feature depending on the product feature and the context. Take the digital camera as an example.
>
> "The battery lasts very long" - positive
> "The camera takes a long time to start up" - negative
>
> However, no effective mechanism is provided in this WordNet-based sentiment analysis method for dealing with such context dependent opinion words. This problem can be solved by holistic lexicon-based method proposed in [3] which will be discussed afterward.

Ding et al. proposed a holistic lexicon-based opinion mining technique which uses WordNet for building opinion lexicon in [3]. As opinions can be indicated by adjective, adverb, verb and noun, for instance 'good', 'fast', 'hate', 'rubbish', opinion lexicon constructed in [3] not only included adjectives and adverbs, but also verbs and nouns. A list of context dependent opinion words is also built and used for sentiment analysis. To create the lists, part-off-speech (POS) tagging [4] is performed. The different sets of opinion lexicon are obtained via a bootstrapping process using the WordNet [7]. Annotated opinion expressing idioms are also been identified and most of them express stronger opinions. Non-opinion phrases containing opinion words would have their opinion words be overwritten under the mechanism. Sentiments are estimated with the idea similar to [2] that a word has the same orientation with its synonym and opposite orientation as its antonym. Moreover, some linguistic rules are used together with the WordNet-based orientation identification to enhance the accuracy and handle the context dependent opinions.

3.3 Linguistic Rules

There are some relationships between semantic orientation and linguistic features. Semantic orientations of words can be indirectly indicated by conjunctions. Conjoined adjectives usually have the same orientation [10]. Work in [10] used a supervised learning algorithm to infer the semantic orientation of adjectives conjoined by conjunctions. It achieved a high accuracy, but highly relied on large corpus and a large amount of manually tagged training data is needed.

A holistic lexicon-based work [3] uses the global information involving the linguistic rules together with the dictionary-based method to deal with the context dependent opinion words and phrases which achieves a higher precision with the orientations of context dependent words. Below are the set of rules applied in [3].

1. **Intra-sentence conjunction rule**

   Only one opinion is expressed in a sentence unless there is a 'but' word that changes the orientation direction.

2. **Pseudo intra-sentence conjunction rule**

   The orientation of opinion remains the same even if no explicit conjunction 'and' is used in the sentence [10].

   If there is an opinion word that its semantic orientation is unclear, infer the orientation by other reviewers using the above two rules. e.g. "The battery life is very long", where semantic orientation of 'long' can be inferred to be positive if there is another reviewer said "The camera takes great pictures and have long battery life", which 'great' is positive and 'long' is then positive by rule 1. If there is a reviewer wrote "The battery life is long, it's great", 'long' is positive by rule 2.

   If confliction in orientation occurs for the same feature with the same opinion word, the majority view is followed.

3. **Inter-sentence conjunction rule**

   Neighbor sentence expresses the same opinion unless the word 'however' or 'but' is used to indicate the direction changes. If the above two rules fails to determine the opinion orientation, this rule serves the function.

Experiments showed that system with the context dependency handling in [3] improved the F-score compared to FBS [2] which did not deal with the context dependency.

4 Approaches for Opinion Analysis

Opinion analysis has been studied in two research directions. The directions are: (1) feature-based opinion mining, and (2) sentiment classification.

4.1 Feature-Based Opinion Mining

Feature-based opinion mining aims to indicate semantic orientations of opinions on each feature. It usually involves identification of features from input text in advance of deciding opinions orientations. Representative works on feature-based sentence level subjectivity classification includes [2, 3].

Hu et al. proposed a feature-based opinion mining method in [2]. The introduced method firstly identifies the features of a product customers have expressed opinions on, it then identifies, for each feature, the review sentences which containing opinions, the last step is to produce a sentiment summary with the discovered information. To mine the product features which have been commented on, both data mining and natural language processing techniques are used. Part-of-speech tags of each sentence of all reviews are produced by using NLProcesor linguistic parser [4]. A transaction file which contains only the identified nouns and noun phrases is created for hot feature (or frequent feature) generation. Association mining is then employed to find all frequent itemsets that are likely to be product features. Compactness pruning and redundancy pruning are used to remove unlikely features. After identifying the commented features, the method identifies the opinion sentences and deciding if each opinion expressed in each sentence is positive or negative using the WordNet.

Opinion words are used for extracting the infrequently commented features out. The heuristic identifies the nearest noun/noun phrase the opinion word modifies as the infrequent features. With the semantic orientations of opinion words identified, the task of deciding the orientation of each opinion sentence is done by counting the frequency of positive and negative opinion words. The opinion sentence orientation is determined by the dominant orientation of the opinion words in the sentence.

A feature-based review summary is then generated. For easier reading and decision making for customers, the summary is generated with all the features ranked according to their appearance frequency in reviews. For each feature, there are two categories, positive and negative, with the related opinion sentences regarding their orientations.

Experiment was done on a system, called Feature-Based Summarization (FBS), which was built based on the proposed techniques. The experiment results showed that the precision and recall of feature extraction are both significantly higher than the well-known and publicly available term extraction and indexing system, FASTR [11]. The system also has a good accuracy in sentence orientation prediction.

Although the technique introduced in [2] yields a relatively high precision, there is the multiple conflicting opinion words problem. The problem can be addressed by the opinion aggregating function deployed in [3].

> **Multiple conflicting opinion words problem**
> Some sentences are having multiple conflicting opinion words. The proposed technique simply sums up the opinion words' orientation regardless the distance of the opinion words and the feature. This lower the accuracy as far away opinion words may not modifies the current feature.

Ding et al. also worked on feature-based opinion analysis [3]. The proposed method firstly buils a set of opinion lexicon lists as mentioned before in Section 3.2.1. Opinion orientations of each feature are then aggregated regarding the distance of the opinion phrase and the product feature. Given a sentence *s* containing a set of features *F* and a set of identified opinion words *V*, an orientation score is computed for each feature *f*.

$$score(f) = \sum_{w_i : w_i \in S \wedge w_i \in V} \frac{w_i SO}{dis(w_i, f)} \qquad (3)$$

where w_i is an opinion word, *V* is the set of all opinion words and idioms. *dis*(w_i, *f*) is the distance between opinion word w_i and feature *f* in sentence *s*. w_i*SO* is semantic orientation of w_i.

With the above scoring formula, a lower weight is given to opinion words which are far away from the feature. As such, solved the multiple conflicting opinion words problem existed in [2].

Orientation of opinion on a feature in a sentence depends on the final score. The opinion is positive if the score is positive, the opinion is negative if the score is negative, and the opinion is neutral if the score is neither.

Several rules are employed for distinguishing semantic orientation:

1. Negation Rules: reverse the opinion's semantic orientation if there is a negation word or phrase in the sentence.

2. 'But' Clause Rules: if the opinion expressed in the 'but' clause is neutral in it's semantic orientation, the orientation follows the negation of the semantic orientation before the 'but' clause.

A holistic approach using different contextual information from the global with some linguistic rules is proposed for dealing the context dependent opinions as described in Section 3.3.

A system, called Opinion Observer, was built based on the proposed technique for experiment. From the experiment results, the fact that equation (3), which calculates the orientation scores regarding the distance of the opinion word and the product feature, and the context dependency handling are useful for improving the F-score is shown. With either one of the functions, Opinion Observer outperformed the FBS in [2]. Opinion Observer performed the best that improved the recall dramatically with almost no less in precision when having both the equation (3) and context dependency handling employed.

4.2 Sentiment Classification

Sentiment classification aims to classify each review document into positive, negative or neutral classes according to the semantic orientations. [5, 9, 13, 14] are both works on classification at document level.

4.2.1 Machine Learning Methods

Many machine learning methods are popular for topic-based classification. Those techniques can also be employed for sentiment classification where the classification is treated as a special case of topic-based categorization with just a few 'topics' which are the sentiment orientations.

[9] examines the effectiveness of applying machine learning techniques to the sentiment classification problem on document level. Techniques including Nave Bayes, Maximum Entropy and Support Vector Machines are compared. Among these three methods, Nave Bayes performed

the worst whilst SVMs performed the best in the experiment.

Moreover, [9] points out a problem that a "thwarted expectations' narrative is a common phenomenon in documents in whatever domains, where the review author makes a deliberate contrast to the earlier discussion. It is easy for human to detect the true sentiment of the review, but would be difficult for classifier. Regarding the problem, as the whole is not necessarily the sum of parts [5], it suggests that some sophisticated techniques should be used to determine the focus of each sentence and to identify the on-topic words as some thwarted expectations are expressing opinion about some other topic.

4.2.2 Natural Language Processing (NLP)-Based

WebFountain system [12] adopts bBNP (Beginning definite Base Noun Phrases) greedy approach for the extraction of product features. Definite base noun phrases which located at the start of sentences and followed by verb phrases are extracted. Reviews are dissembled and traversed with two linguistic resources, sentiment lexicons and lexicon pattern, in order to assign sentiments to the features. The polarity of terms are defined by the sentiment lexicon. A sentiment pattern database storing the sentiment assignment patterns of predicates is also used. As such, a simple Web interface for listing the sentiment bearing sentences of a particular product can be easily implemented.

4.2.3 PMI-IR

Turney [5] classified reviews into recommended and not recommended groups according to their sentiment using PMI. Phrases containing adjectives or adverbs are extracted and their sentiment orientations are calculated using PMI with "excellent" and "poor" as the positive and negative reference words. Unsupervised classification is then performed to classify each review based on the average semantic orientation of phrases in a review. Review that carries a positive average orientation is classified as recommended, otherwise not recommended.

The experimental results did not show a consistent performance with PMI-IR. The techniques would have high accuracies for some domains only, such as automobile or bank reviews. It achieves relatively low precisions with the movie and travel reviews. The reason for the poor performance within these domains is that there are different aspects to such domains. For example, good actor does not necessarily mean the movie is good.

In addition to time required for queries, the algorithm performs poorly for some domains as it does not consider features comprehensively.

5 Opinion Mining on OpenRice

OpenRice [17] is a well-known restaurant review system with a dense reviewers producing over 280,000,000 reviews which comment on more than 18,000 restaurants. With the vast amount of reviews the system carries, it is valuable to have sentiment analysis on the reviews to provide an easier environment for users to make decisions with the reviews.

5.1 Problem Formulation

Restaurant reviews are multi-aspect as a reviewer can comment on many aspects of a restaurant in a single review, such as food, service, and ambience. Therefore features should be firstly identified before analyzing the sentiment expressed. Sentiment analysis on OpenRice differs greatly to that of other review systems as the reviews it carries are mainly written in Chinese. Thus the opinion words cannot be extracted by using publicly available linguistic resource and must be learnt by machine automatically.

5.2 Features and Opinion Words Learning

The task of feature identification can be done by the help of Non-negative Matrix Factorization (NMF) or Latent Dirichlet allocation (LDA). For NMF, given a matrix X with reviews as rows and words as columns, the full composition of X is amounted to W and H such that:

$$X = WH + U \tag{4}$$

where W and H are two non negative matrices and U is a residual, which W and H minimize the function

$$F(W, H) = \|X - WH\|_F^2 \tag{5}$$

NMF or LDA groups corpus of similar topic together and extracts representative words of each group. As such, the extracted representative words can be the frequently commented features among a group of reviews. With the feature words, we can study the sentiments of each feature in a review by performing analysis on sentence level.

Opinion words can be learnt as well by using NMF or LDA with the similar manner of identifying the feature words. Groups resulted from the algorithms can be sentiment-oriented but not cuisine type dominating. Which their representative words are probably opinion words.

6 Experiment

Experiment has been done with a dataset of 4680 restaurant reviews from OpenRice review system. The dataset firstly underwent data preprocessing with Stanford Chinese Word Segmenter [16]. Term frequency inverse document

frequency (tf-idf) and NMF is then performed for clustering the reviews into 12 groups of different cuisine type. Table 1 shows the representative words of each group of reviews. After evaluation by human expert, the class labels are given as shown in Table 2.

The user tagged cuisine style information is used as the ground truth to evaluated the experiment result. The experiment yields a quite good result. For example, in group 3, 69.70% of reviews are tagged as Japanese cuisine. 76.69% reviews in group 12 are tagged as Guang Dong and Hong Kong style. Since the user tagged cuisine information is just a simple cuisine style of a restaurant, some of the groups, like group 4 and 10, cannot be evaluated by the user tagged information. However, the representing words do clearly indicate the cuisine type of the group.

6.1 Discussion

As shown in the tables, NMF can accurately cluster restaurant reviews into groups and the representative words of each group can be treated as the frequent commented objects within the group. Moreover, experiments shows that opinion words can also be learnt by clustering reviews into more groups. There exist some clusters with sentiment words dominated in the resulting groups.

It may also be a possible way to the sentiment of a review by multiplying a weight to the non-sentiment dominated groups to hide such groups. As such, reviews will then be assigned to a sentiment dominated group and the sentiment of the review can then be examined.

7 Possible applications based on sentiment analysis result

One of the potential applications is to build search engines for particular type of products or shops (e.g. digital camera, restaurants, movies, etc). By using the sentiment analysis result, the query result can be sorted according to the semantic orientation of each item. As such, those items associated with lots of positive reviews can be put to the top of the result to ease the decision making for customers.

On the other hand, it can be used for helping reviewers of academic paper to use appropriate tone in writing comments. Normally, reviewers are assumed to be objective and the tone used in writing comments should be neutral. But very often they will accidentally criticize the submitted papers by harsh wordings. With the help of the sentiment analysis result, it is possible to highlight those harsh wordings and reviewers can modify their comments accordingly.

Finally, sentiment analysis result can be used for building better recommender systems. Traditionally, recommender systems were heavily relied on collaborative filtering techniques [15]. However, in this traditional approach, only discrete ratings are used to represent the preference of users. In many cases, opinions and preference of users are just too complicated to be denoted by numbers. As such, the sentiment analysis results can be considered as a more accurate representation of user preference and thus a more sophisticated recommender system can be built.

Group	1	2	3	4
Words	紅豆	芝士	壽司	雪糕
	好多	麵包	刺身	士多啤梨
	唔係	文治	文魚	芒果
	仲有	芒果	日本	tiramisu
	叉燒	蕃茄	帶子	開心
	幾好	芝士味	海膽	奶味
	唔會	cheese	天婦羅	綠茶
	少少	牛油	吞拿魚	美味
	今日	火腿	烏冬	香蕉
	係度	薯條	新鮮	cream
	綠茶	cream	軟殼蟹	濃郁
	cream	漢堡	鰻魚	口味

Group	5	6	7	8
Words	意粉	豆腐	not	咖哩
	白汁	糖水	very	日式
	海鮮	餃子	good	起來
	肉醬	魚肉	quite	羊肉
	蕃茄	芝麻	too	埋單
	主菜	滑溜	so	吉列
	茄汁	豆漿	food	什麼
	餐廳	幾好	dinner	不少
	下午	仍然	cheese	午市
	午餐	杏仁	fresh	還有
	蟹肉	小食	risotto	日本
	侍應	個人	pasta	來吃

Group	9	10	11	12
Words	豬扒	湯底	早餐	腸粉
	吉列	麻辣	煎蛋	點心
	奶茶	米線	奶茶	燒賣
	香茅	拉麵	火腿	蝦餃
	咖喱	餃子	牛油	叉燒
	男友	配料	通粉	蘿蔔糕
	豬潤	牛腩	早上	下午
	飲品	雲吞	多士	春卷
	2008	河粉	健康	皮蛋
	文治	豬肉	茶餐廳	豉油
	多士	酸辣	炒蛋	酒樓
	下午	下午	茶記	豬肉

Table 1. Top 12 frequent keywords in each cluster

Group 1	Miscellaneous
Group 2	Sandwich and American Fast Food
Group 3	Japanese
Group 4	Western Dessert
Group 5	Western
Group 6	Street Snack
Group 7	English
Group 8	Miscellaneous
Group 9	Miscellaneous
Group 10	Noodles in Soup
Group 11	Local Cafe Breakfast
Group 12	Chinese Restaurant

Table 2. Latent Cuisine Type

8 Conclusion

Sentiment analysis has been studied by many researchers. There are techniques investigated to perform different tasks involved in automatic opinion mining systems with diverse objectives. In this paper, I focus on surveying the techniques developed for the major tasks of opinion mining, includes subjective words or phrases identification, and sentiment orientation estimation. Various limitations of the surveyed methods are reviewed. Further improvements are required to make sentiment analysis yields a better performance with a higher accuracy and more effective for a broader range of domains. Moreover, sentiment analysis is important for not only customer and merchants, but also organizations, like government agencies, education institutes, etc. Collective opinions on services, products, and policies can substantially improve our life. Besides the survey, possible applications of sentiment analysis results are introduced as well as the possible methods for mining opinions on OpenRice restaurant reviews.

References

[1] J. Wiebe, and R. Mihalcea. Word Sense and Subjectivity. Proceedings of the 21^{st} International Conference on Computational Linguistics and the 44^{th} annual meeting of the Association for Computational Linguistics. *ACL'06*, 2006.

[2] M. Hu and B. Lui. Mining and Summarizing Customer Reviews. Proceedings of the 10^{th} ACM SIGKDD international conference on Knowledge discovery and data mining. *KDD'04*, 2004.

[3] X. Ding, B. Lui and P. S. Yu. A Holistic Lexicon-Based Approach to Opinion Mining. Proceedings of the international conference on Web search and web data mining. *ACM-WSDM'08*, 2008

[4] NLProcessor- *Text Analysis Toolkit*. 2000. http://www.infogistics.com/textanalysis.html

[5] P. Turney. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Proceedings of the 40^{th} Annual Meeting of the Association for Computational Linguistics. *ACL'02*, 2002

[6] K. W. Churchand P. Hanks. Word Association Norms, Mutual Information and Lexicography. Proceedings of the 27^{th} annual meeting on Association for Computational Linguistics. *ACL'89*, 1989

[7] C. Fellbaum. WordNet: an Electronic Lexical Database. MIT Press, 1998

[8] E. Brill. Some Advances in Transformation-Based Part of Speech Tagging. Proceedings of the 12^{th} national conference on Artificial intelligence. *AAAI'94*, 1994

[9] B. Pang, L. Lee and S. Vaithyanathan. Thumbs up? Sentiment Classification using Machine Learning Techniques. Proceedings of the Conference on Empirical Methods in Natural Language Processing. *EMNLP'02*, 2002

[10] V. Hatzivassiloglou and K. R. McKeown. Predicting the semantic orientation of adjectives. Proceedings of the 8^{th} conference on European chapter of the Association for Computational Linguistics. *ACL'97*, 1997

[11] FASTR. http://www.limsi.fr/Individu/jacquemi/FASTR/

[12] J.Yi and W.Niblack. Sentiment Mining in WebFountain. Proceedings of the 21^{st} International Conference on Data Engineering. *ICDE'05*, 2005

[13] M. Hearst. Direction-based Text Interpretation as an Information Access Refinement. In. P.Jacobs, editor, *Text-Based Intelligent Systems*. Lawrence Erlbaum Associates, 1992.

[14] B. Pang and L. Lee, Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales. roceedings of the 43^{rd} Annual Meeting on Association for Computational Linguistics. *ACL'05*, 2005.

[15] G. Adomavicius and A. Tuzhilin, Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering(TKDE), pages 734 - 749, 2005

[16] Sanford Chinese Word Segmenter- *The Stanford Natural Language Processing Group*. 2008. http://nlp.stanford.edu/software/segmenter.shtml

[17] OpenRice. http://www.openrice.com.hk