

PROCEEDINGS

The 16th HKBU-CSD Postgraduate Research
Symposium

PG Day 2013



Department of Computer Science
Hong Kong Baptist University
June 18, 2013

The 16th HKBU-CSD Postgraduate Research Symposium (PG Day) Program

June 18 th 2013, Tuesday		
Time	Sessions	
09:00-09:20	On-site Registration (T909)	
09:20-09:30	Welcome: Prof. P. C Yuen (Head of Department of Computer Science, HKBU) (T909)	
09:30-11:00	Session I_A (Chair: Ms. Liang Fengfeng, T909) <ul style="list-style-type: none"> ➤ Zhao Kaiyong GPU-BLASTN: accelerating nucleotide alignment by graphic processors ➤ Lan Xiangyuan Multi-Cue Tracking via Joint Sparse Representation ➤ Chen Lei Spatial-keyword related works 	Session II_A (Chair: Ms. Chang He, T716) <ul style="list-style-type: none"> ➤ Bao Qing Incorporating Structural Diversity of Neighbors in a Diffusion Model for Social Networks ➤ Li Chen Literature Review of Process Mining ➤ Yang Xiaofei Inferring Disease Transmission Network
	Tea Break	
11:00-11:15	Session I_B (Chair: Mr. Chen Lei, T909) <ul style="list-style-type: none"> ➤ Chan Sheung wai Knowledge Extraction and Mining in Research Using Biomedical Network Model ➤ Fan Zhe Privacy Preserving Reachability Query Services ➤ Chen Qian Authenticating Location-based Services without Compromising Location Privacy 	Session II_B (Chair: Mr. Li Chen, T716) <ul style="list-style-type: none"> ➤ Peng Qinmu Salient Region Detection Using Local and Global Saliency ➤ Li Jiawei Domain Transfer Support Vector Ranking for Person Re-Identification without Target Camera Label Information ➤ Tang Chao Analysis of Healthcare Social Media Dissemination and Knowledge Extraction
	Noon Break	
12:45-14:30	Session III (Chair: Mr. Chen Lei, T714) <ul style="list-style-type: none"> ➤ Wong Yee man Coping with cyber-victimization: The role of cognition and negative affects ➤ Liu Fei A Theoretical Model and Empirical Investigation of Social Networking Site Users' Switching Intention ➤ Gu Fangqing On Solving Complex Optimization Problems with Objective Decomposition 	Session IV (Chair: Mr. Li Chen, T716) <ul style="list-style-type: none"> ➤ Mei Xinxin Can GPU DVFS Save Energy? ➤ Yu Lu Multiple Radios for Effective Rendezvous in Cognitive Radio Networks
	14:30-16:00	
16:30-17:00	Best Paper and Best Presentation Awards Announcement (T716)	
	Closing	

16th PG Day, Student Presentation List

30minute for both presentation and Q&A

			JL	WC	Clement	NGJ	PCY	BX	YMC	XJ	HH/ HL/ HZ	BC	CHU	CHL	YW	TAM
Jun. (Tuesday)																
09:30am-12:45pm T909																
Mr. Zhao Kaiyong	趙開勇	PhD								JX#	HZ◇	BC◇	CHU*			
Mr. Lan Xiangyuan	藍湘源	PhD					PCY*		YMC#				CHU◇			TAM◇
Mr. Chen Lei	陳磊	PhD								JX*	HH◇	BC#				
Mr. Chan Sheung wai	陳尚瑋	PhD	JL#		Clement*					JX◇						
Mr. Fan Zhe	樊哲	PhD								XJ#		BC*	CHU◇			
Mr. Chen Qian	陳乾	PhD								XJ*	HH#	BC#	CHU◇		YW◇	
9:30am-16:00pm T716																
Ms. Bao Qing	鮑青	PhD	JL#	WC*					YMC◇					CHL◇		
Mr. Li Chen	李辰	PhD	JL◇	WC*								BC#				
Mr. Yang Xiaofei	楊曉飛	MPhil	JL*	WC#	Clement◇				YMC◇							
Mr. Peng Qinmu	彭勤牧	PhD							YMC*			BC ◇	CHU◇	CHL#		
Mr. Li Jiawei	黎嘉偉	PhD					PCY*		YMC#					CHL◇		TAM◇
Mr. Tang Chao	湯超	MPhil		WC#		JNG◇		BX◇						CHL*		
Ms. Mei Xinxin	梅欣欣	PhD									HH◇		CHU*			TAM#
Ms. Yu Lu	余璐	PhD				JNG◇				JX◇	HL#		CHU#		YW*	
14:30pm-16:00pm T714																
Ms. Wong Yee man	王綺敏	MPhil			Clement◇			BX*				BC◇		CHL#		
Mr. Liu Fei	劉蜚	PhD			Clement◇			BX*				BC#		CHL◇		
Mr. Gu Fangqing	辜方清	PhD							YMC*			BC◇	CHU◇	CHL#		

Table of abstracts

The 16th HKBU-CSD Postgraduate Research Symposium (PG Day) Program.....	2
Section I_A:	5
GPU-BLASTN: accelerating nucleotide alignment by graphic processors	5
Zhao Kaiyong.....	5
Multi-Cue Tracking via Joint Sparse Representation	5
Lan Xiangyuan	5
Spatial-keyword related works.....	6
Chen Lei.....	6
Section I_B:	7
Knowledge Extraction and Mining in Research Using Biomedical Network Model	7
Chan Sheung wai	7
Privacy Preserving Reachability Query Services.....	7
Fan Zhe.....	7
Authenticating Location-based Services without Compromising Location Privacy	7
Chen Qian.....	7
Section II_A	9
Incorporating Structural Diversity of Neighbors in a Diffusion Model for Social Networks.....	9
Bao Qing.....	9
Literature Review of Process Mining.....	9
Li Chen.....	9
Inferring Disease Transmission Network.....	9
Yang Xiaofei	9
Section II_B.....	11
Salient Region Detection Using Local and Global Saliency Peng Qinmu	11
Domain Transfer Support Vector Ranking for Person Re-Identification without Target Camera Label.....	11
Li Jiawei	11
Analysis of Healthcare Social Media Dissemination and Knowledge Extraction.....	11
Tang Chao	11
Section III.....	13
Coping with cyber-victimization: The role of cognition and negative affects	13
Wong Yee man.....	13
A Theoretical Model and Empirical Investigation of Social Networking Site Users' Switching Intention.....	13
Liu Fei	13
On Solving Complex Optimization Problems with Objective Decomposition	13
Gu Fangqing.....	13
Section IV	15
Can GPU DVFS Save Energy?	15
Mei Xinxin.....	15
Multiple Radios for Effective Rendezvous in Cognitive Radio Networks.....	15
Yu Lu	15

Section I_A:

GPU-BLASTN: accelerating nucleotide alignment by graphic processors

Zhao Kaiyong

Abstracts: Motivation: Since 1990, Basic Local Alignment Search Tool (BLAST) has become one of the most popular and fundamental bioinformatics tools for sequence similarity searching. The two pioneering papers about BLAST has so far received over 87 000 citations. Considering the huge population of BLAST users, any improvement in the execution speed of BLAST would be of great importance to researchers, especially when the size of sequence databases becomes larger and larger. The existing GPU-BLAST is a promising software tool to accelerate protein sequence alignment by using graphics processing unit (GPU); and its speedups over single-thread NCBI-BLAST range between 3 and 4. Unfortunately, there is still no GPU-accelerated software tool for nucleotide sequence alignment.

Results: We developed a GPU-accelerated nucleotide alignment tool GPU-BLASTN based on the source code of widely used NCBI-BLAST. GPU-BLASTN can produce exactly the same results as NCBI-BLAST, and it also has the same user interface. In comparison to the sequential single-thread NCBI-BLAST, the speedups achieved by GPU-BLASTN on a single GTX580 graphic card (cost less than \$500) are mostly around 20 times. More impressively, in comparison to the parallel eight-thread NCBI-BLAST running on a dual-CPU server, GPU-BLASTN can achieve ten times speedup by using two GTX580 graphic cards.

Multi-Cue Tracking via Joint Sparse Representation

Lan Xiangyuan

Abstracts: In this paper, we address the problem of visual tracking with multiple features. Motivated by the recent success of joint sparse representation in image classification and recognition, we formulate this problem as a joint sparse representation model to combine the strength of multiple features for visual tracking. Compared with other tracking algorithm with multiple features, joint sparse representations can efficiently construct the appearance model with different features without handing them independently. Furthermore, by employing the joint sparse-inducing norm, our model exploits the interdependency of different representation of a target, and enforce them to be individually sparse and approximately share the feature representations of same representative set of templates, which enhance the representation ability by reducing the negative effect of redundant templates. The generative ability and discriminability of different features are well utilized for tracking in our model. Accelerated Proximal Gradient Method is also presented to optimize our model. Experiment results show improved tracking performance.

Spatial-keyword related works

Chen Lei

Abstracts: Web users and content are increasingly being geo-positioned. This development gives prominence to spatial-keyword queries, which involve both the locations and the textual descriptions of content.

This work first gives a brief survey of the spatial-keyword related works, which may include three main categories: spatial-keyword queries, reverse spatial-keyword queries, and answering why-not questions on spatial-keyword queries.

Finally, we materialize the problem of answering why-not questions on spatial-keyword queries and propose several algorithms to solve it

Section I_B:

Knowledge Extraction and Mining in Research Using Biomedical Network Model

Chan Sheung wai

Abstracts: With the advance of medical research, the publishing of biomedical literatures are increasing dramatically. Knowledge extraction from large amount of literatures becomes very time consuming. In this paper, data mining and natural language processing are applied in novel model, biomedical rule network mode. In this mode, the information of interaction between herb and diseases and the chemical constituents of herb are extracted. Also with the overlap chemical constituents of herbs, the promising effective constituents can be hypothesized.

Privacy Preserving Reachability Query Services

Fan Zhe

Abstracts: Privacy Preserving Reachability Query Services There have been a wide range of recent applications of graph data. Due to the massive volume of graph data and resources required to process numerous queries at large scale, it is becoming economically appealing to outsource graph data to a third-party service provider (SP) to provide query services. However, SP may not always be trusted. Hence, data owners and query clients may not prefer to expose their data graph and queries. This paper studies privacy preserving query services of a fundamental query for graph and semi-structured data namely the reachability query where both clients' queries and the structural information of the owner's data are protected. Due to the privacy requirements, many existing works on the reachability query do not seem to be applicable. We propose privacy preserving 2-hop labeling(pp-2-hop)where the queries are computed in an encrypted domain and the input and output sizes of any queries are indistinguishable, respectively. We analyze the security offered by pp-2-hop. We conduct an experimental study on pp-2-hop with both synthetic and real-world datasets.

Authenticating Location-based Services without Compromising Location Privacy

Chen Qian

Abstracts: The popularity of mobile social networking services (mSNSs) is propelling more and more businesses, especially those in retailing and marketing, into mobile and location-based forms. To address the trust issue, the service providers are expected to deliver their location-based services in an authenticatable manner, so that the correctness of the service results can be verified by the client. However, existing works on query authentication cannot preserve the privacy of the data being queried, which are sensitive user locations when it comes to location-based services and mSNSs. In this paper, we address this challenging problem by proposing a comprehensive solution that preserves unconditional location privacy when

authenticating range queries. Three authentication schemes for R-tree and grid-file index, together with two optimization techniques, are developed. Cost models, security analysis, and experimental results consistently show the effectiveness, reliability and robustness of the proposed schemes under various system settings and query workloads. .

Section II_A

Incorporating Structural Diversity of Neighbors in a Diffusion Model for Social Networks

Bao Qing

Abstracts: Diffusion (or called propagation) is the phenomenon that an action or information spread from node to node by edges through network. It is known to be an important process governing the behaviours observed in networked environments like social networks, contact networks for disease transmission and etc. We are all social entities engaged in social networks with interaction with friends, colleagues and so on. In the context of viral marketing, selecting the key users who can best spread the influence to the network based on the underlying mechanism for diffusion helps companies to gain most profit with limited cost is known as influence maximization. In the context of disease control, the operations are opposite, that is to find key patients to eliminate disease on them to prevent them from spreading the disease to as many people as possible, known as influence minimization. Thus, it is a very interesting work to model the mechanism of diffusion network.

In this paper, we argue that most of the related work on diffusion models does not consider the structural diversity of node's neighborhood which however has been known to be an important factor to be considered in social ties related studies. We propose a diffusion model with neighbors' structural diversity incorporated.

Literature Review of Process Mining

Li Chen

Abstracts: Process mining draws a lot of attentions recently for it provides a new method to analyze the processes from event logs. In this paper, we first present the motivation of process mining. Then we show some background of process mining and give a literature review of process model discovering algorithms. Lastly, we put forward a research plan and go to conclusion.

Inferring Disease Transmission Network

Yang Xiaofei

Abstracts: When investigating epidemic dynamics in real world, it is more realistic and normal to obtain the surveillance data rather than the epidemic models or the underlying disease diffusion patterns. Thus, based on available empirical surveillance data, inversely inferring the possible hidden epidemic transmission network is becoming a significant important research issue. Unveiling such a network would be helpful to (i) investigate the hidden impact factors for epidemic dynamics, (ii) reveal underlying reasons for epidemic outbreaks, and (iii) practically improve the controllability of infectious disease. Existing studies mainly focus on information diffusion network inference, which treat each individual as independent node. Specifically in this paper, to explore an epidemic transmission network, we propose an efficient and accurate inference algorithm that taking into account both the temporal information and the discrepancy of nodes' attributes, where the discrepancy includes both internal and external influences, such as the environmental impact

factors for specific disease, geographical distances between locations, import cases from other countries and etc.

Section II_B

Salient Region Detection Using Local and Global Saliency

Peng Qinmu

Abstracts: In this paper, we present a novel local-global salient region detection method. We first obtain the smoothed image via gradient minimization, resulting in more homogeneous background. Then, we partition the smoothed image into a set of regions and compute the region saliency by measuring the dissimilarity and spatial distance. Furthermore, we adopt the global color distribution, including the color coherence, to yield global saliency region. Finally, we combine the local-and-global salient regions and the composition information to obtain the overall salient regions. Experimental results show the efficacy of the proposed method in comparison with the existing methods.

Domain Transfer Support Vector Ranking for Person Re-Identification without Target Camera Label

Li Jiawei

Abstracts: Domain Transfer Support Vector Ranking for Person Re-Identification without Target Camera Label Information. This paper addresses a new person re-identification problem without the label information of persons under non-overlapping target cameras. Given the matched (positive) and unmatched (negative) image pairs from source domain cameras, as well as unmatched (negative) image pairs which can be easily generated from target domain cameras, we propose a Domain Transfer Ranked Support Vector Machines (DTRSVM) method for target domain cameras. To overcome the problems introduced due to the absence of matched (positive) image pairs in target domain, we relax the rank relation constraint to a necessary condition only relying on the positive mean in target domain. By estimating the target positive mean using source and target domain data, a new discriminative model with high confidence in target positive mean and low confidence in target negative image pairs is developed. Since the necessary condition may not truly preserve the rank relation, multi-task support vector ranking is proposed to incorporate the training data from the source domain with label information. Experimental results show that the proposed DTRSVM outperforms existing methods without using label information in target cameras. And the top 30 rank accuracy can be improved by the proposed method up to 9.40% on popular and public available person re-identification datasets.

Analysis of Healthcare Social Media Dissemination and Knowledge Extraction

Tang Chao

Abstracts: In this study, we investigate the challenges in disseminating and extraction of knowledge in healthcare social media.

The development of social media for healthcare faces significant challenges and dissemination of patient-oriented healthcare information are impeded by healthcare organizations and professionals. Patients who turn to Internet for complaints and feedbacks on healthcare services often face deletion of their postings and sometimes even harassments and retaliations. The complex issues of balancing the rights among different stakeholders like patients, hospitals,

healthcare professionals, pharmacy companies, and insurance companies are not completely resolved. We suggest a decoupled approach to develop healthcare review system by isolating subjective components and objective components and hence enhancing flexibility in reducing institutions and personal conflicts.

For knowledge extraction, we propose a method for finding discriminant features and discriminant topics in online reviews. The discriminant features are obtained by linear discriminant analysis (LDA) method, and the discriminant topics are then discovered by non-negative matrix factorization (NMF) applied on data matrix with multiplication of class discriminant feature weights. Experiments show such method outperforms standard unsupervised NMF, and is able to discover discriminant features and topics in terms of patients' satisfaction rating and gender of patients. Product features like treatment purpose and ingredients of herb medicine could be found with proposed method as well. We further analyze the trend and cycle in time series of healthcare reviews, in terms of number of reviews and satisfaction rating score.

Section III

Coping with cyber-victimization: The role of cognition and negative affects

Wong Yee man

Abstracts: Rising incidents of and tragedies from cyber-bullying have alerted researchers, educators, government officials, and parents to the severe consequence of this new form of bullying. This study aims to understand the consequences of cyber-bullying behaviors from victims' perspective. Drawing from the Stimulus–Organism–Response framework and General Strain Theory, this study examines how the victims' coping responses are influenced by the breadth of audience of a cyber-bullying incident and the victims' relational closeness with perpetrators. This study identifies five major coping responses, confrontation, retaliation, technical defense, support-seeking and inaction. Results show that both cognition and negative affects influence victims' coping responses. The study not only demonstrates the conceptual framework of cyber-victimization coping, but also provides valuable insights to practitioners to focus on the development of individuals' cognition about cyber-bullying and offer clear guidelines as to how the individuals engaged in cyber-bullying perpetration.

A Theoretical Model and Empirical Investigation of Social Networking Site Users' Switching Intention

Liu Fei

Abstracts: Customers' post-adoption switching behavior among competing service providers, particularly among different online Social Network Sites (SNSs), is gaining increasing attention from both information system researchers and practitioners, as the size of user base is both a selling point and a source of revenue of SNS service providers. In this study, we draw on the uses and gratification theory to identify factors motivating an individual to switch SNSs. In addition, informed by social identify theory, we explore how individuals' social identities on their current SNSs moderate the effects of the motivating factors on their intention to switch to a more attractive SNS. The results of our survey study reveal that the perceived relative values of a competing SNS (when compared with the user's current SNS) positively influence his/her intention to switch to the competing SNS. However, the positive impact of perceived relative values on switching intention is mitigated by the user's perception of his/her salient social identity on his/her current SNS. This study has significant implications for both academics and practitioners.

On Solving Complex Optimization Problems with Objective Decomposition

Gu Fangqing

Abstracts: This paper addresses the complex optimization problem, of which the objective

function consists of two parts: One part is differentiable and unimodal, meanwhile the other part is non-differentiable and multi-modal. Accordingly, we decompose the original objective function into several relatively simple sub-objective ones, which therefore formulate as a multiobjective optimization problem~(MOP). To solve this MOP, we propose a simulated water-stream algorithm (SWA) inspired by the natural phenomenon of water streams. The water streams with a hybrid process of downstream and penetration towards the basin are analogous to the process of finding the minimum solution in an optimization problem. The SWA featuring a combination of deterministic search and heuristic search generally converges much faster than the existing counterparts with a considerable accuracy enhancement. Experimental results show the efficacy of the proposed algorithm.

Section IV

Can GPU DVFS Save Energy?

Mei Xinxin

Abstracts: Nowadays, the many-core architecture, GPU appears more often in super computers. The participation of GPUs brings considerable speedup to a variety of scientific applications, anyhow, at the same time, brings larger energy consumption. First applied to CPU systems, DVFS is a commonly knob for saving energy. This paper studies the influence of GPU DVFS on GPU power consumption. We measure the energy cost of nine typical CUDA kernels with different GPU core voltage/ frequency pairs. We find that up to 28.57% energy can be saved when GPU runs at a core voltage than default setting. Based on this observation, we train a statistical model to estimate GPU energy consumption considers GPU DVFS. Our model can achieve average ration error. All experiments are conducted on NVIDIA GeForce GTX 560 Ti.

Multiple Radios for Effective Rendezvous in Cognitive Radio Networks

Yu Lu

Abstracts: Rendezvous is a fundamental operation in cognitive radio networks (CRNs) for establishing a communication link on a commonly-available channel between cognitive users. The existing works on rendezvous implicitly assume that each cognitive user is equipped with one radio (i.e., one wireless transceiver). As the cost of wireless transceivers is dropping, this feature can be exploited to significantly improve the rendezvous performance at low cost. In this study, we investigate the rendezvous problem in CRNs where cognitive users are equipped with multiple radios and different users may have different number of radios. We first study how the existing rendezvous algorithms can be generalized to use multiple radios for faster rendezvous. We then propose a new rendezvous algorithm, called role-based parallel sequence (RPS), which specifically exploits multiple radios for more efficient rendezvous. Our basic idea is to let the cognitive users stay in a specific channel in one dedicated radio and hop on the available channels with parallel sequences in the remaining general radios. We prove that our algorithm provides guaranteed rendezvous and derive the maximum time-to-rendezvous (TTR) and upper-bounds on the expected TTR. Extensive experiments are conducted to evaluate the proposed solutions.