THE IEEE
Intelligent Informatics
BULLETIN

IEEE Computer Society
Technical Committee
on Intelligent Informatics

November 2005 Vol. 6 No. 2 (ISSN 1727-5997)

Special Issue (Guest Editor: Shichao Zhang)

Feature Articles

Book Review

Announcements

On-line version: http://www.comp.hkbu.edu.hk/~iib (ISSN 1727-6004)

The Technical Committee on Intelligent Informatics (TCII) of the IEEE Computer Society deals with tools and systems using biologically and linguistically motivated computational paradigms such as artificial neural networks, fuzzy logic, evolutionary optimization, rough sets, data mining, Web intelligence, intelligent agent technology, parallel and distributed information processing, and virtual reality. If you are a member of the IEEE Computer Society, you may join the TCII without cost. Just fill out the form at http://computer.org/tcsignup/.

The IEEE Intelligent Informatics Bulletin

Aims and Scope

The IEEE Intelligent Informatics Bulletin is the official publication of the Technical Committee on Intelligent Informatics (TCII) of the IEEE Computer Society, which is published twice a year in both hardcopies and electronic copies. The contents of the Bulletin include (but may not be limited to):

1) Letters and Communications of the TCII Executive Committee
2) Feature Articles
3) R&D Profiles (R&D organizations, Interview profiles on individuals, and projects etc.)
4) Book Reviews
5) News, Reports, and Announcements (TCII sponsored or important/related activities)

Materials suitable for publication at the IEEE Intelligent Informatics Bulletin should be sent directly to the Associate Editors of respective sections.

Technical or survey articles are subject to peer reviews, and their scope may include the theories, methods, tools, techniques, systems, and experiences for/in developing and applying biologically and linguistically motivated computational paradigms, such as artificial neural networks, fuzzy logic, evolutionary optimization, rough sets, and self-organization in the research and application domains, such as data mining, Web intelligence, intelligent agent technology, parallel and distributed information processing, and virtual reality.

Publisher: *The IEEE Computer Society Technical Committee on Intelligent Informatics*
Address: *Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong (Attention: Prof. JimingLiu; Email: jiming@comp.hkbu.edu.hk)*
ISSN Number: *1727-5997(printed) 1727-6004(on-line)*
Abstracting and Indexing: *All the published articles will be submitted to the following on-line search engines and bibliographies databases for indexing—Google (www.google.com), The ResearchIndex (citeseer.nj.nec.com), The Collection of Computer Science Bibliographies (liinwww.ira.uka.de/bibliography/index.html), and DBLP Computer Science Bibliography (www.informatik.uni-trier.de/~ley/db/index.html).*

The Definition of Optimal Solution and an Extended Kuhn-Tucker Approach for Fuzzy Linear Bilevel Programming

Guangquan Zhang and Jie Lu

Abstract— **Bilevel decision techniques are mainly developed for solving decentralized management problems with decision makers in a hierarchical organization. Organizational bilevel decision-making, such as planning of land-use, transportation and water resource, all may involve uncertain factors. The parameters shown in a bileved programming model, either in the objective functions or constraints, are thus often imprecise, which is called fuzzy parameter bilevel programming (FPBLP) problem. Following our previous work [1, 2], this study first proposes a model of FPBLP. It then gives the definition of optimal solution for an FPBLP problem. Based on the definition and related theorems, this study develops a fuzzy number based Kuhn-Tucher approach to solve the proposed FPBLP problem. Finally, an example further illustrates the power of the fuzzy number based Kuhn-Tucher approach.**

Index Terms— **Linear bilevel programming, Kuhn-Tucker approach, Fuzzy set, Optimization.**

I. INTRODUCTION

THE execution of many decisions in businesses is sequential, from a higher level (leader) to a lower level (follower); each unit independently optimizes its own objective, but is affected by other unit's actions through externalities. This is called bilevel programming (BLP) problem (also called bilevel decision or bilevel optimization problems). BLP was first introduced by Von Stackelberg [3] in the context of unbalanced economic markets [4, 5]. In a BLP problem, each decision maker (leader or follower) tries to optimize his/her own objective function with partially or without considering the objective of the other level, but the decision of each level affects the objective optimization of the other level [6].

There have been nearly two dozen algorithms [5, 7-10] proposed for solving BLP problems since the field caught the attention of researchers in the mid-1970s [11-19]. Although BLP theory and technology have been applied with remarkable success in different domains [20-22], existing approaches mainly support the decision situation in which the objective functions and constraints are characterized with precise parameters. Therefore, the parameters are required to be fixed at some values in an experimental and/or subjective manner through the experts' understanding of the nature of the parameters in the problem-formulation process. It has been observed that, in most real-world situations, particularly in critical resource planning, such as planning of land-use, transportation and water resource, the possible values of these parameters are often only imprecisely or ambiguously known to these experts. It results in a difficulty to fix parameters in the objective functions or constraints of a bilevel programming model. With this observation, it would be certainly more appropriate to interpret the experts' understanding of parameters as fuzzy numerical data which can be represented by means of fuzzy sets theory [23]. A bilevel programming problem in which the parameters either in objective function or in constrains are described by fuzzy values is called a fuzzy parameter bilevel programming (FPBLP) problem.

The FPBLP problem was first explored by Sakawa et al. in 2000 [24]. Sakawa et al. formulates bilevel programming problems with fuzzy parameters from the perspective of experts' imprecision and proposes a fuzzy programming method for fuzzy bilevel programming problems. However, Sakawa's work is mainly based on the definition of solution for bilevel programming proposed by Bard [5, 15]. One deficiency of Bard's linear BLP theory is that it could not well solve a linear bilevel programming problem when the upper-level constraint functions are of arbitrary linear form. Our recent research work has extended Bard's theory of bilevel programming by proposing a new definition of optimal solution for linear bilevel programming which can overcome the arbitrary linear form problem indicated above [1]. We have then proposed an extended Kuhn-Tucher approach, based on our definition of optimal solution, for solving linear bilevel problems [2].

Following our previous research results shown in [1, 2], this study aims at solving a FPBLP problem by transferring it into a non-fuzzy bilevel programming problem. This paper first proposes a model of FPBLP problem, then gives a definition of the optimal solution for the FPBLP problem. Based on the definition and related theorems, this paper develops a fuzzy number based Kuhn-Tucher approach to solve the proposed FPBLP problem. As this paper only deals with linear bilevel problem, so bilevel programming means linear bilevel programming in this paper.

Following the introduction, Section 2 reviews related definitions, theorems and properties of fuzzy number, BLP solution and Kuhn-Tucher approach for solving an BLP problem. A definition of optimal solution and a fuzzy number based Kuhn-Tucher approach for solving FPBLP problems are presented in Section 3. A numeral example is shown in Section 4 for illustrating the proposed fuzzy number based Kuhn-Tucher approach. Conclusion and further study are discussed in Section 5.

Authors are with Faculty of Information Technology, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia (e-mail: {jielu, zhangg}@it.uts.edu.au).

II. PRELIMINARIES

A. Fuzzy Numbers

In this section, we present some basic concepts, definitions and theorems that are to be used in the subsequent sections. The work presented in this section can also be found from our recent paper in [25].

Let R be the set of all real numbers, R^n be n-dimensional Euclidean space, and $x = (x_1, x_2, \ldots, x_n)^T$, $y = (y_1, y_2, \ldots, y_n)^T \in R^n$ be any two vectors, where $x_i, y_i \in R$, $i = 1,2, \ldots, n$ and T denotes the transpose of the vector. Then we denote the inner product of x and y by $\langle x, y \rangle$. For any two vectors x, $y \in R^n$, we write $x \geqq y$ iff $x_i \geqq y_i, \forall\, i = 1,2,\cdots,n$; $x \geq y$ iff $x \geqq y$ and $x \neq y$; $x > y$ iff $x_i > y_i$, $\forall\, i = 1,2, \ldots, n$.

Definition 2.1 A fuzzy number $\tilde{a}$ is defined as a fuzzy set on R, whose membership function $\mu_{\tilde{a}}$ satisfies the following conditions:

1. $\mu_{\tilde{a}}$ is a mapping from R to the closed interval [0, 1];
2. it is normal, i.e., there exists $x \in R$ such that $\mu_{\tilde{a}}(x) = 1$;
3. for any $\lambda \in (0, 1]$, $a_\lambda = \{x;\ \mu_{\tilde{a}}(x) \geq \lambda\}$ is a closed interval, denoted by $[a_\lambda^L, a_\lambda^R]$.

Let $F(R)$ be the set of all fuzzy numbers. By the decomposition theorem of fuzzy set, we have

$$\tilde{a} = \bigcup_{\lambda \in [0,1]} \lambda [a_\lambda^L, a_\lambda^R], \tag{2.1}$$

for every $\tilde{a} \in F(R)$.

Let $F^*(R)$ be the set of all finite fuzzy numbers on R.

Theorem 2.1 *Let $\tilde{a}$ be a fuzzy set on R, then $\tilde{a} \in F(R)$ if and only if $\mu_{\tilde{a}}$ satisfies*

$$\mu_{\tilde{a}}(x) = \begin{cases} 1 & x \in [m, n] \\ L(x) & x < m \\ R(x) & x > n \end{cases},$$

where $L(x)$ is the right-continuous monotone increasing function, $0 \leqq L(x) < 1$ and $\lim_{x \to -\infty} L(x) = 0$, $R(x)$ is the left-continuous monotone decreasing function, $0 \leqq R(x) < 1$ and $\lim_{x \to \infty} R(x) = 0$.

Corollary 2.1 *For every $\tilde{a} \in F(R)$ and λ_1, $\lambda_2 \in [0, 1]$, if $\lambda_1 \leqq \lambda_2$, then $a_{\lambda_2} \subset a_{\lambda_1}$.*

Definition 2.2 For any $\tilde{a}, \tilde{b} \in F(R)$ and $0 \leqq \lambda \in R$, the sum of $\tilde{a}$ and $\tilde{b}$ and the scalar product of λ and $\tilde{a}$ are defined by the membership functions

$$\mu_{\tilde{a}+\tilde{b}}(t) = \sup_{t=u+v} \min\{\mu_{\tilde{a}}(u), \mu_{\tilde{b}}(v)\}, \tag{2.2}$$

$$\mu_{\tilde{a}-\tilde{b}}(t) = \sup_{t=u-v} \min\{\mu_{\tilde{a}}(u), \mu_{\tilde{b}}(v)\}, \tag{2.2}$$

$$\mu_{\lambda\tilde{a}}(t) = \sup_{t=\lambda u} \mu_{\tilde{a}}(u). \tag{2.3}$$

Theorem 2.2 *For any $\tilde{a}, \tilde{b} \in F(R)$ and $0 \leqq \alpha \in R$,*

$$\tilde{a} + \tilde{b} = \bigcup_{\lambda \in [0,1]} \lambda [a_\lambda^L + b_\lambda^L,\ a_\lambda^R + b_\lambda^R],$$

$$\tilde{a} - \tilde{b} = \tilde{a} + \left(-\tilde{b}\right) = \bigcup_{\lambda \in [0,1]} \lambda [a_\lambda^L - b_\lambda^R,\ a_\lambda^R - b_\lambda^L],$$

$$\alpha\tilde{a} = \bigcup_{\lambda \in [0,1]} \lambda [\alpha a_\lambda^L,\ \alpha a_\lambda^R].$$

Definition 2.3 Let $\tilde{a}_i \in F(R), i = 1,2,\cdots,n.$ We define $\tilde{a} = (\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_n)$

$$\mu_{\tilde{a}} : R^n \to [0,1]$$
$$x \mapsto \bigwedge_{i=1}^{n} \mu_{\tilde{a}_i}(x_i),$$

where $x = (x_1, x_2, \ldots, x_n)^T \in R^n$, and $\tilde{a}$ is called an n-dimensional fuzzy number on R^n. If $\tilde{a}_i \in F^*(R), i = 1,2,\cdots,n,$ $\tilde{a}$ is called an n-dimensional finite fuzzy number on R^n.

Let $F(R^n)$ and $F^*(R^n)$ be the set of all n-dimensional fuzzy numbers and the set of all n-dimensional finite fuzzy numbers on R^n respectively.

Proposition 2.1 *For every $\tilde{a} \in F(R^n)$, $\tilde{a}$ is normal.*

Proposition 2.2 *For every $\tilde{a} \in F(R^n)$, the λ-section of $\tilde{a}$ is an n-dimensional closed rectangular region for any $\lambda \in [0,1]$.*

Proposition 2.3 *For every $\tilde{a} \in F(R^n)$ and λ_1, $\lambda_2 \in [0,1]$, if $\lambda_1 \leqq \lambda_2$, then $a_{\lambda_2} \subset a_{\lambda_1}$.*

Definition 2.4 *For any n-dimensional fuzzy numbers $\tilde{a}, \tilde{b} \in F(R^n)$, we define*

1. $\tilde{a} \succeqq \tilde{b}$ iff $a_{i\lambda}^L \geqq b_{i\lambda}^L$ and $a_{i\lambda}^R \geqq b_{i\lambda}^R$, $i = 1,2,\cdots,n,\ \lambda \in (0,1]$;
2. $\tilde{a} \succeq \tilde{b}$ iff $a_{i\lambda}^L \geq b_{i\lambda}^L$ and $a_{i\lambda}^R \geq b_{i\lambda}^R$, $i = 1,2,\cdots,n,\ \lambda \in (0,1]$;
3. $\tilde{a} \succ \tilde{b}$ iff $a_{i\lambda}^L > b_{i\lambda}^L$ and $a_{i\lambda}^R > b_{i\lambda}^R$, $i = 1,2,\cdots,n,\ \lambda \in (0,1]$.

We call the binary relations $\succeqq, \succeq$ and $\succ$ a fuzzy max order, a strict fuzzy max order and a strong fuzzy max order, respectively.

B. The Extended Kuhn-Tucker Approach for Linear Bilevel Programming

Let write a linear programming (LP) as follows.

$$\min f(x) = cx$$
$$\text{subject to } Ax \leqq b$$
$$x \geqq 0,$$

where c is an n-dimensional row vector, b an m-dimensional column vector, A an $m \times n$ matrix with $m \leq n$, and $x \in R^n$.

Let $\lambda \in R^m$ and $\mu \in R^n$ be the dual variables associated with constraints $Ax \geqq b$ and $x \geq 0$, respectively. Bard [5] gave the following proposition.

Proposition 2.4 [5] *A necessary and sufficient condition that (x^*) solves above LP is that there exist (row) vectors λ^*, μ^* such that (x^*, λ^*, μ^*) solves:*

$$\lambda A - \mu = -c$$
$$Ax - b \geqq 0$$
$$\lambda(Ax - b) = 0$$
$$\mu x = 0$$
$$x \geqq 0, \lambda \geqq 0, \mu \geqq 0$$

For $x \in X \subset R^n$, $y \in Y \subset R^m$, $F: X \times Y \to R^1$, and $f: X \times Y \to R^1$, a linear BLP problem is given by Bard [4]:

$$\min_{x \in X} F(x,y) = c_1 x + d_1 y \quad (2.5a)$$

$$\text{subject to } A_1 x + B_1 y \leqq b_1 \quad (2.5b)$$

$$\min_{y \in Y} f(x,y) = c_2 x + d_2 y \quad (2.5c)$$

$$\text{subject to } A_2 x + B_2 y \leqq b_2 \quad (2.5d)$$

where c_1, $c_2 \in R^n$, d_1, $d_2 \in R^m$, $b_1 \in R^p$, $b_2 \in R^q$, $A_1 \in R^{p \times n}$, $B_1 \in R^{p \times m}$, $A_2 \in R^{q \times n}$, $B_2 \in R^{q \times m}$.

Definition 2.5 [1]

(a)Constraint region of the linear BLP problem:
$S = \{(x,y): x \in X, y \in Y, A_1 x + B_1 y \leqq b_1, A_2 x + B_2 y \leqq b_2\}$

(b) Feasible set for the follower for each fixed $x \in X$:
$S(X) = \{x \in X : \exists y \in Y, A_1 x + B_1 y \leqq b_1, A_2 x + B_2 y \leqq b_2\}$

(c)Projection of S onto the leader's decision space:
$S(X) = \{x \in X : \exists y \in Y, A_1 x + B_1 y \leqq b_1, A_2 x + B_2 y \leqq b_2\}$

(d) Follower's rational reaction set for $x \in S(X)$:
$P(x) = \{y \in Y : y \in \arg\min[f(x,\hat{y}) : \hat{y} \in S(x)]\}$
where
$\arg\min[f(x,\hat{y}) : \hat{y} \in S(x)] = \{y \in S(x) : f(x,y) \leqq f(x,\hat{y}), \hat{y} \in S(x)\}$

(e)Inducible region:
$IR = \{(x,y) : (x,y) \in S, y \in P(x)\}$

Definition 2.6 [1] (x^*, y^*) is said to be a complete optimal solution, if and only if there exists $(x^*, y^*) \in S$ such that $F(x^*, y^*) \leqq F(x,y)$ and $f(x^*, y^*) \leqq f(x,y)$ for all $(x,y) \in S$.

However, in general, such a complete optimal solution that simultaneously minimizes both the leader' and follower's objective functions does not always exist. Instead of a complete optimal solution, a new solution concept, called Pareto optimality, is introduced in linear BLP.

Definition 2.7 [1] (x^*, y^*) is said to be a Pareto optimal solution, if and only if there does not exist $(x,y) \in S$ such that $F(x,y) \leqq F(x^*, y^*)$, $f(x,y) \leqq f(x^*, y^*)$ and $F(x,y) \neq F(x^*, y^*)$ or $f(x,y) \neq f(x^*, y^*)$.

Definition 2.8 A topological space is compact if every open cover of the entire space has a finite subcover. For example, $[a,b]$ is compact in R (the Heine-Borel theorem) [26].

To ensure that (2.5) has a Pareto optimal solution, Bard gave the following assumption.

Assumption 2.1

(a) S is nonempty and compact.

(b) For decisions taken by the leader, the follower has some rooms to respond; i.e, $P(x) \neq \phi$.

(c) $P(x)$ is a point-to-point map.

To ensure that (2.5) is well posed we assume that S is nonempty and compact, and that $P(x)$ is a point-to-point map. The rational reaction set $P(x)$ defines the response while the inducible region IR represents the set over which the leader may optimize his objective. Thus in terms of the above notations, the linear BLP problem can be written as

$$\min\{F(x,y) : (x,y) \in IR\} \quad (2.6)$$

We also present the following theorem to characterize the condition under which there is a Pareto optimal solution for a linear BLP problem.

Theorem 2.3 [1] *If S is nonempty and compact, there exists a Pareto optimal solution for a linear BLP problem*

Theorem 2.4 [2] [Extended Kuhn-Tucher Theorem] *A necessary and sufficient condition that (x^*, y^*) solves the linear BLP problem (2.5) is that there exist (row) vectors u^*, v^* and w^* such that $(x^*, y^*, u^*, v^*, w^*)$ solves:*

$$\min F(x,y) = c_1 x + d_1 y \quad (2.7a)$$

$$\text{subject to } A_1 x + B_1 y \le b_1 \quad (2.7b)$$

$$A_2 x + B_2 y \le b_2 \quad (2.7c)$$

$$u B_1 + v B_2 - w = -d_2 \quad (2.7d)$$

$$u(b_1 - A_1 x - B_1 y) + v(b_2 - A_2 x - B_2 y) + wy = 0 \quad (2.7e)$$

$$x \ge 0, y \ge 0, u \ge 0, v \ge 0, w \ge 0 \quad (2.7f)$$

III. Fuzzy Parameter Linear Bilevel Programming Problem

Consider the following fuzzy parameter linear bilevel programming (FPBLP) problem:

For $x \in X \subset R^n$, $y \in Y \subset R^m$, $F: X \times Y \to F^*(R)$, and $f: X \times Y \to F^*(R)$,

$$\min_{x \in X} F(x,y) = \tilde{c}_1 x + \tilde{d}_1 y \quad (3.1a)$$

$$\text{subject to } \tilde{A}_1 x + \tilde{B}_1 y \precsim \tilde{b}_1 \quad (3.1b)$$

$$\min_{y \in Y} f(x,y) = \tilde{c}_2 x + \tilde{d}_2 y \quad (3.1c)$$

$$\text{subject to } \tilde{A}_2 x + \tilde{B}_2 y \precsim \tilde{b}_2 \quad (3.5d)$$

where $\tilde{c}_1, \tilde{c}_2 \in F^*(R^n)$, $\tilde{d}_1, \tilde{d}_2 \in F^*(R^m)$, $\tilde{b}_1 \in F^*(R^p)$, $\tilde{b}_2 \in F^*(R^q)$, $\tilde{A}_1 = (\tilde{a}_{ij})_{p \times n}$, $\tilde{a}_{ij} \in F^*(R)$, $\tilde{B}_1 = (\tilde{b}_{ij})_{p \times m}$, $\tilde{b}_{ij} \in F^*(R)$, $\tilde{A}_2 = (\tilde{e}_{ij})_{q \times n}$, $\tilde{e}_{ij} \in F^*(R)$, $\tilde{B}_2 = (\tilde{s}_{ij})_{q \times m}$, $\tilde{s}_{ij} \in F^*(R)$.

Associated with the FPBLP problem, we now consider the following linear multi-objective multi-follower bilevel programming (LMMBLP) problem:

For $x \in X \subset R^n$, $y \in Y \subset R^m$, $F: X \times Y \to F^*(R)$, and $f: X \times Y \to F^*(R)$,

$$\begin{aligned} &\min_{x \in X} (F(x,y))^L_\lambda = c_{1\lambda}^L x + d_{1\lambda}^L y, \quad \lambda \in [0,1] \\ &\min_{x \in X} (F(x,y))^R_\lambda = c_{1\lambda}^R x + d_{1\lambda}^R y, \quad \lambda \in [0,1] \end{aligned} \quad (3.2a)$$

$$\text{subject to } A_{1\lambda}^L x + B_{1\lambda}^L y \leqq b_{1\lambda}^L,\ A_{1\lambda}^R x + B_{1\lambda}^R y \leqq b_{1\lambda}^R, \lambda \in [0,1] \quad (3.2b)$$

$$\begin{aligned} &\min_{y \in Y} (f(x,y))^L_\lambda = c_{2\lambda}^L x + d_{2\lambda}^L y, \quad \lambda \in [0,1] \\ &\min_{y \in Y} (f(x,y))^R_\lambda = c_{2\lambda}^R x + d_{2\lambda}^R y, \quad \lambda \in [0,1] \end{aligned} \quad (3.2c)$$

$$\text{subject to } A_{2\lambda}^L x + B_{2\lambda}^L y \leqq b_{2\lambda}^L,\ A_{2\lambda}^R x + B_{2\lambda}^R y \leqq b_{2\lambda}^R, \lambda \in [0,1] \quad (3.5d)$$

where $c_{1\lambda}^L, c_{1\lambda}^R$, $c_{2\lambda}^L, c_{2\lambda}^R \in R^n$, $d_{1\lambda}^L, d_{1\lambda}^R$, $d_{2\lambda}^L, d_{2\lambda}^R \in R^m$, $b_{1\lambda}^L, b_{1\lambda}^R \in R^p$, $b_{2\lambda}^L, b_{2\lambda}^R \in R^q$, $A_{1\lambda}^L = (a_{ij\lambda}^L), A_{1\lambda}^R = (a_{ij\lambda}^R) \in R^{p \times n}$, $B_{1\lambda}^L = (b_{ij\lambda}^L), B_{1\lambda}^R = (b_{ij\lambda}^R) \in R^{p \times m}$, $R^{q \times n}$, $B_{2\lambda}^L = (s_{ij\lambda}^L), B_{2\lambda}^R = (s_{ij\lambda}^R) \in R^{q \times m}$.

Theorem 3.1 *Let (x^*, y^*) be the solution of the LMMBLP* *problem* (3.2). *Then it is also a solution of the FPBLP problem defined by* (3.1).

Proof. The proof is obvious from Definition 2.4.

Lemma 3.1 *If there is* (x^*, y^*) *such that* $cx + dy \geqq cx^* + dy^*$, $c_0^L x + d_0^L y \geqq c_0^L x^* + d_0^L y^*$ *and* $c_0^R x + d_0^R y \geqq c_0^R x^* + d_0^R y^*$, *for any* (x, y) *and isosceles triangle fuzzy numbers* $\tilde{c}$ *and* $\tilde{d}$, *then*

$$c_\lambda^L x + d_\lambda^L y \geqq c_\lambda^L x^* + d_\lambda^L y^*,$$
$$c_\lambda^R x + d_\lambda^R y \geqq c_\lambda^R x^* + d_\lambda^R y^*,$$

for any $\lambda \in (0,1)$, *where c and d are the centre of* $\tilde{c}$ *and* $\tilde{d}$ *respectively.*

Proof. As λ-section of isosceles triangle fuzzy numbers $\tilde{c}$ and $\tilde{d}$ are

$$c_\lambda^L = c_0^L(1-\lambda) + c\lambda \text{ and } c_\lambda^R = c_0^R(1-\lambda) + c\lambda$$
$$d_\lambda^L = d_0^L(1-\lambda) + d\lambda \text{ and } d_\lambda^R = d_0^R(1-\lambda) + d\lambda .$$

Therefore, we have

$$\begin{aligned} c_\lambda^L x + d_\lambda^L y &= c_0^L(1-\lambda)x + c\lambda x + d_0^L(1-\lambda)y + d\lambda y \\ &= (c_0^L x + d_0^L y)(1-\lambda) + (cx+dy)\lambda \\ &\geqq (c_0^L x^* + d_0^L y^*)(1-\lambda) + (cx^* + dy^*)\lambda \\ &= c_\lambda^L x^* + d_\lambda^L y^*, \end{aligned}$$

from $cx + dy \geqq cx^* + dy^*$ and $c_0^L x + d_0^L y \geqq c_0^L x^* + d_0^L y^*.$, we can prove $c_\lambda^R x + d_\lambda^R y \geqq c_\lambda^R x^* + d_\lambda^R y^*$ from similar reason.

Theorem 3.2 *For* $x \in X \subset R^n$, $y \in Y \subset R^m$, *If all the fuzzy coefficients* $\tilde{a}_{ij}, \tilde{b}_{ij}, \tilde{e}_{ij}, \tilde{s}_{ij}, \tilde{c}_i$ *and* $\tilde{d}_i$ *have triangle membership functions of the FPBLP problem* (3.1).

$$\mu_{\tilde{z}}(t) = \begin{cases} 0 & t < z_0^L \\ \dfrac{t - z_0^L}{z - z_0^L} & z_0^L \leqq t < z \\ \dfrac{-t + z_0^R}{z_0^R - z} & z \leqq t < z_0^R \\ 0 & z_0^R \leqq t \end{cases}, \qquad (3.3)$$

where $\tilde{z}$ *denotes* $\tilde{a}_{ij}, \tilde{b}_{ij}, \tilde{e}_{ij}, \tilde{s}_{ij}, \tilde{c}_i$ *and* $\tilde{d}_i$ *and z are the centre of* $\tilde{z}$ *respectively. Then, it is the solution of the problem (3.1) that* $(x^*, y^*) \in R^n \times R^m$ *satisfying*

$$\min_{x\in X} (F(x,y))_c = c_1 x + d_1 y,$$
$$\min_{x\in X} (F(x,y))_0^L = c_{10}^L x + d_{10}^L y, \qquad (3.4a)$$
$$\min_{x\in X} (F(x,y))_0^R = c_{10}^R x + d_{10}^R y,$$

subject to $A_1 x + B_1 y \leqq b_1,$

$$A_{10}^L x + B_{10}^L y \leqq b_{10}^L, \qquad (3.4b)$$
$$A_{10}^R x + B_{10}^R y \leqq b_{10}^R,$$
$$\min_{y\in Y}(f(x,y))_c = c_2 x + d_2 y,$$
$$\min_{y\in Y}(f(x,y))_0^L = c_{20}^L x + d_{20}^L y, \qquad (3.4c)$$
$$\min_{y\in Y}(f(x,y))_\lambda^R = c_{20}^R x + d_{20}^R y,$$

subject to $A_2 x + B_2 y \leqq b_2,$

$$A_{20}^L x + B_{20}^L y \leqq b_{20}^L, \qquad (3.4d)$$
$$A_{20}^R x + B_{20}^R y \leqq b_{20}^R.$$

Proof. From Lemma 3.1, if (x^*, y^*) satisfies (3.4a) and (3.4c), then it satisfies (3.2a) and (3.2c). Then we need only prove, if (x^*, y^*) satisfies (3.4b) and (3.4d), then it satisfies (3.2b) and (3.2d). In fact, for any $\lambda \in (0,1)$,

$$a_{ij\lambda}^L = a_{ij}\lambda + a_{ij0}^L(1-\lambda),$$
$$b_{ij\lambda}^L = b_{ij}\lambda + b_{ij0}^L(1-\lambda) \quad \text{and}$$
$$b_{1\lambda}^L = b_1\lambda + b_{10}^L(1-\lambda),$$

we have

$$\begin{aligned} A_{1\lambda}^L x^* + B_{1\lambda}^L y^* &= (a_{ij\lambda}^L)x^* + (b_{ij\lambda}^L)y^* \\ &= \left(a_{ij}\lambda + a_{ij0}^L(1-\lambda)\right)x^* + \left(b_{ij}\lambda + b_{ij0}^L(1-\lambda)\right)y^* \\ &= (a_{ij})x^*\lambda + (a_{ij0}^L)x^*(1-\lambda) + (b_{ij})y^*\lambda + (b_{ij0}^L)y^*(1-\lambda) \\ &= \left((a_{ij})x^* + (b_{ij})y^*\right)\lambda + \left((a_{ij0}^L)x^* + (b_{ij0}^L)y^*\right)(1-\lambda) \\ &= \left(A_1 x^* + B_1 y^*\right)\lambda + \left(A_{10}^L x^* + B_{10}^L y^*\right)(1-\lambda) \\ &\leqq b_1\lambda + b_{10}^L(1-\lambda) = b_{1\lambda}^L, \end{aligned}$$

from (3.4b). Similarly, we can prove

$$A_{1\lambda}^R x^* + B_{1\lambda}^R y^* \leqq b_{1\lambda}^R,$$
$$A_{2\lambda}^L x^* + B_{2\lambda}^L y^* \leqq b_{2\lambda}^L,$$
$$A_{2\lambda}^R x^* + B_{2\lambda}^R y^* \leqq b_{2\lambda}^R,$$

for any $\lambda \in (0,1)$ from (3.4b) and (3.4d). The proof is complete.

Theorem 3.3 [Extended Kuhn-Tucher Theorem] *A necessary and sufficient condition that* (x^*, y^*) *solves the FPBLP problem* (3.1) *with triangle fuzzy numbers is that there exist* (*row*) *vectors* u^*, v^* *and* w^* *such that* $(x^*, y^*, u^*, v^*, w^*)$ *solves:*

$$\min_{x\in X} (F(x,y)) = (c_1 x + d_1 y) + \left(c_{10}^L x + d_{10}^L y\right) + \left(c_{10}^R x + d_{10}^R y\right) \qquad (3.5a)$$

subject to $A_1 x + B_1 y \leqq b_1,$

$$A_{10}^L x + B_{10}^L y \leqq b_{10}^L, \qquad (3.5b)$$
$$A_{10}^R x + B_{10}^R y \leqq b_{10}^R,$$
$$A_2 x + B_2 y \leqq b_2,$$
$$A_{20}^L x + B_{20}^L y \leqq b_{20}^L, \qquad (3.5c)$$
$$A_{20}^L x + B_{20}^L y \leqq b_{20}^L,$$
$$u_1 B_1 + u_2 B_{10}^L + u_3 B_{10}^R + v_1 B_2 + v_2 B_{20}^L + v_3 B_{20}^R - w = -\left(d_2 + d_{20}^L + d_{20}^R\right) \qquad (3.5d)$$
$$\begin{aligned} &u_1\left(b_1 - A_1 x - B_1 y\right) + u_2\left(b_{10}^L - A_{10}^L x - B_{10}^L y\right) + \\ &u_3\left(b_{10}^R - A_{10}^R x - B_{10}^R y\right) + v_1(b_2 - -A_2 x - B_2 y) + \\ &v_2\left(b_{20}^L - A_{20}^L x - B_{20}^L y\right) + v_3\left(b_{20}^R - A_{20}^R x - B_{20}^R y\right) + wy = 0 \end{aligned} \qquad (3.5e)$$
$$x \ge 0, y \ge 0, u \ge 0, v \ge 0, w \ge 0 \qquad (3.5f)$$

Proof: (1) From Theorem 3.2, we know that we need only to solve the problem (3.4). In fact, to solve the problem (3.4), we can use the method of weighting [27] to this problem, such that it is the following problem:

$$\min_{x\in X} (F(x,y)) = (c_1 x + d_1 y) + \left(c_{10}^L x + d_{10}^L y\right) + \left(c_{10}^R x + d_{10}^R y\right) \qquad (3.6a)$$

subject to $A_1 x + B_1 y \leqq b_1,$

$$A_{10}^L x + B_{10}^L y \leqq b_{10}^L, \qquad (3.6b)$$
$$A_{10}^R x + B_{10}^R y \leqq b_{10}^R,$$
$$\min_{y\in Y}(f(x,y)) = c_2 x + d_2 y + c_{20}^L x + d_{20}^L y + c_{20}^R x + d_{20}^R y \qquad (3.6c)$$

subject to $A_2x + B_2y \leqq b_2$,

$$A_{20}^{L}x + B_{20}^{L}y \leqq b_{20}^{L}, \quad (3.6d)$$

$$A_{20}^{R}x + B_{20}^{R}y \leqq b_{20}^{R}.$$

Therefore, the linear BLP problem can be written as

$$\min\{F(x,y):(x,y)\in IR\} \quad (3.7)$$

Let us get an explicit expression of (3.7) and rewrite (3.7) as follows:

$$\min F(x,y)$$
subject to $(x,y)\in IR$.

We have

$$\min F(x,y)$$
subject to $(x,y)\in S$
$y\in P(x)$

by Definition 2.5(e). Then, we have

$$\min F(x,y)$$
subject to $(x,y)\in S$
$y\in \arg\min[f(x,\hat{y}):\hat{y}\in S(x)]$

by Definition 2.5(d). We rewrite it as:

$$\min F(x,y)$$
subject to $(x,y)\in S$
$\min f(x,y)$
subject to $y\in S(x)$.

We have

$$\min F(x,y)$$
subject to $(x,y)\in S$
$\min_{y\in Y} f(x,y)$
subject to $(x,y)\in S$,

by Definition 2.5(c). Consequently, we can have

$$\min_{x\in X}\left(F(x,y)\right) = \left(c_1x + d_1y\right) + \left(c_{10}^{L}x + d_{10}^{L}y\right) + \left(c_{10}^{R}x + d_{10}^{R}y\right) \quad (3.8a)$$

subject to

$$\begin{aligned} &A_1x + B_1y \leqq b_1,\\ &A_{10}^{L}x + B_{10}^{L}y \leqq b_{10}^{L},\\ &A_{10}^{R}x + B_{10}^{R}y \leqq b_{10}^{R},\\ &A_2x + B_2y \leqq b_2,\\ &A_{20}^{L}x + B_{20}^{L}y \leqq b_{20}^{L},\\ &A_{20}^{R}x + B_{20}^{R}y \leqq b_{20}^{R}. \end{aligned} \quad (3.8b)$$

$$\min_{y\in Y}\left(f(x,y)\right) = c_2x + d_2y + c_{20}^{L}x + d_{20}^{L}y + c_{20}^{R}x + d_{20}^{R}y \quad (3.8c)$$

subject to

$$\begin{aligned} &A_1x + B_1y \leqq b_1,\\ &A_{10}^{L}x + B_{10}^{L}y \leqq b_{10}^{L},\\ &A_{10}^{R}x + B_{10}^{R}y \leqq b_{10}^{R},\\ &A_2x + B_2y \leqq b_2,\\ &A_{20}^{L}x + B_{20}^{L}y \leqq b_{20}^{L},\\ &A_{20}^{R}x + B_{20}^{R}y \leqq b_{20}^{R}. \end{aligned} \quad (3.8d)$$

by Definition 2.5(a).

This simple transformation has shown that solving the fuzzy linear BLP (3.1) is equivalent to solving (3.8).

(2) Necessity is obvious from (3.8).

(3) Sufficiency. If (x^*,y^*) is the optimal solution of (3.6), we need to show that there exist (row) vectors u_1^*,u_2^*,u_3^*, v_1^*,v_2^*,v_3^* and w^* such that $(x^*,y^*,u_1^*,u_2^*,u_3^*,v_1^*,v_2^*,v_3^*,\ w^*)$ to solve (3.5). Going one step farther, we only need to prove that there exist (row) vectors u_1^*,u_2^*,u_3^*, v_1^*,v_2^*,v_3^* and w^* such that $(x^*,y^*,u_1^*,u_2^*,u_3^*,v_1^*,v_2^*,v_3^*,\ w^*)$ satisfy the follows

$$u_1B_1 + u_2B_{10}^{L} + u_3B_{10}^{R} + v_1B_2 + v_2B_{20}^{L} + v_3B_{20}^{R} - w = -(d_2 + d_{20}^{L} + d_{20}^{R}) \quad (3.9a)$$

$$u_1(b_1 - A_1x - B_1y) = 0 \quad (3.9b)$$

$$u_2(b_{10}^{L} - A_{10}^{L}x - B_{10}^{L}y) = 0 \quad (3.9c)$$

$$u_3(b_{10}^{R} - A_{10}^{R}x - B_{10}^{R}y) = 0 \quad (3.9d)$$

$$v_1(b_2 - A_2x - B_2y) = 0 \quad (3.9e)$$

$$v_2(b_{20}^{L} - A_{20}^{L}x - B_{20}^{L}y) = 0 \quad (3.9f)$$

$$v_3(b_{20}^{R} - A_{20}^{R}x - B_{20}^{R}y) = 0 \quad (3.9g)$$

$$wy = 0, \quad (3.9h)$$

where $u_1,u_2,u_3\in R^p$, $v_1,v_2,v_3\in R^q$, $w\in R^m$ and they are not negative variables.

Because (x^*,y^*) is the optimal solution of (3.6), we have

$$(x^*,y^*)\in IR,$$

by (3.7). Thus we have

$$y^*\in P(x^*),$$

by Definition 2.5(e). y^* is the optimal solution to the following problem

$$\min(f(x^*,y):y\in S(x^*)), \quad (3.10)$$

by Definition 2.5(d). Rewrite (10) as follows

$$\min f(x,y)$$
subject to $y\in S(x)$
$x = x^*$.

From Definition 3.2(b), we have

$$\min_{y\in Y}\left(f(x,y)\right) = c_2x + d_2y + c_{20}^{L}x + d_{20}^{L}y + c_{20}^{R}x + d_{20}^{R}y \quad (3.11a)$$

subject to

$$A_1x + B_1y \leqq b_1, \quad (3.11b)$$

$$A_{10}^{L}x + B_{10}^{L}y \leqq b_{10}^{L}, \quad (3.11c)$$

$$A_{10}^{R}x + B_{10}^{R}y \leqq b_{10}^{R}, \quad (3.11d)$$

$$A_2x + B_2y \leqq b_2, \quad (3.11e)$$

$$A_{20}^{L}x + B_{20}^{L}y \leqq b_{20}^{L}, \quad (3.11f)$$

$$A_{20}^{R}x + B_{20}^{R}y \leqq b_{20}^{R}. \quad (3.11g)$$

$$x = x^* \quad (3.11h)$$

$$y \geqq 0 \quad (3.11i)$$

To simplify (3.11), we can have

$$\min g(y) = (d_2 + d_{20}^{L} + d_{20}^{R})y \quad (3.12a)$$

subject to

$$-B_1y \geqq -(b_1 - A_1x^*), \quad (3.12b)$$

$$-B_{10}^{L}y \geqq -(b_{10}^{L} - A_{10}^{L}x^*), \quad (3.11c)$$

$$-B_{10}^{R}y \geqq -(b_{10}^{R} - A_{10}^{R}x^*), \quad (3.12d)$$

$$-B_2y \geqq -(b_2 - A_2x^*), \quad (3.12e)$$

$$-B_{20}^{L}y \geqq -(b_{20}^{L} - A_{20}^{L}x^*), \quad (3.12f)$$

$$-B_{20}^{R}y \geqq -(b_{20}^{R} - A_{20}^{R}x^*), \quad (3.12g)$$

$$y \geqq 0. \quad (3.12h)$$

Let we note

$$B=\begin{pmatrix} B_1 \\ B_{10}^L \\ B_{10}^R \\ B_2 \\ B_{20}^L \\ B_{20}^R \end{pmatrix},\quad A=\begin{pmatrix} A_1 \\ A_{10}^L \\ A_{10}^R \\ A_2 \\ A_{20}^L \\ A_{20}^R \end{pmatrix},\quad \text{and}\quad b=\begin{pmatrix} b_1 \\ b_{10}^L \\ b_{10}^R \\ b_2 \\ b_{20}^L \\ b_{20}^R \end{pmatrix}. \tag{3.13}$$

We rewrite (3.12) by using (3.13) and we get

$$\min g(y)=(d_2+d_{20}^L+d_{20}^R)y \tag{3.14a}$$

$$\text{subject to } -By \geqq -(b-Ax^*) \tag{3.14b}$$

$$y \geqq 0. \tag{3.14c}$$

Now we see that y^* is the optimal solution of (3.14) which is a LP problem. By Proposition 2, there exists vector λ^*, μ^*, such that (y^*, λ^*, μ^*) satisfy a system below

$$\lambda B-\mu=-(d_2+d_{20}^L+d_{20}^R) \tag{3.15a}$$

$$-By+(b-Ax^*)\geqq 0 \tag{3.15b}$$

$$\lambda(-By+(b-Ax^*))=0 \tag{3.15c}$$

$$\mu y=0, \tag{3.15d}$$

where $\lambda \in R^{3p+3q}$ and $\mu \in R^m$.

Let $u_1, u_2, u_3 \in R^p$, $v_1, v_2, v_3 \in R^q$ and $w \in R^m$ and define

$$\lambda=(u_1, u_2, u_3, v_1, v_2, v_3)$$

$$w=\mu.$$

Thus we have $(x^*, y^*, u_1^*, u_2^*, u_3^*, v_1^*, v_2^*, v_3^*, w^*)$ that satisfy (3.9). Our proof is completed.

Theorem 3.3 means that the most direct approach to solving (3.1) is to solve the equivalent mathematical program given in (3.5). One advantage that it offers is that it allows for a more robust model to be solved without introducing any new computational difficulties

IV. An Illustrative Example

Example 1 Consider the following FPBLP problem with $x \in R^1$, $y \in R^1$, and $X=\{x \geq 0\}$, $Y=\{y \geq 0\}$,

$$\min_{x\in X} F(x,y)=\tilde{1}x-\tilde{2}y \tag{4.1a}$$

$$\text{subject to } -\tilde{1}x+\tilde{3}y \preceq\!\!\!\!= \tilde{4} \tag{4.1b}$$

$$\min_{y\in Y} f_1(x,y)=\tilde{1}x+\tilde{1}y \tag{4.1c}$$

$$\text{subject to } \tilde{1}x-\tilde{1}y \preceq\!\!\!\!= \tilde{0} \tag{4.1d}$$

$$-\tilde{1}x-\tilde{1}y \preceq\!\!\!\!= \tilde{0} \tag{4.1e}$$

where

$$\mu_{\tilde{1}}(t)=\begin{cases} 0 & t<0 \\ t & 0\leqq t<1 \\ 2-t & 1\leqq t<2 \\ 0 & 2\leqq t \end{cases},$$

$$\mu_{\tilde{2}}(t)=\begin{cases} 0 & t<1 \\ t-1 & 1\leqq t<2 \\ 3-t & 2\leqq t<3 \\ 0 & 3\leqq t \end{cases},$$

$$\mu_{\tilde{3}}(t)=\begin{cases} 0 & t<2 \\ t-2 & 2\leqq t<3 \\ 4-t & 3\leqq t<4 \\ 0 & 4\leqq t \end{cases},$$

$$\mu_{\tilde{4}}(t)=\begin{cases} 0 & t<3 \\ t-3 & 3\leqq t<4 \\ 5-t & 4\leqq t<5 \\ 0 & 5\leqq t \end{cases},$$

$$\mu_{\tilde{0}}(t)=\begin{cases} 0 & t<-1 \\ t+1 & -1\leqq t<0 \\ 1-t & 0\leqq t<1 \\ 0 & 1\leqq t \end{cases}.$$

Step 1 The problem is transferred to the following LMMBLP problem by using Theorem 3.2

$$\min_{x\in X}(F(x,y))_c = 1x-2y$$
$$\min_{x\in X}(F(x,y))_0^L = 0x-3y$$
$$\min_{x\in X}(F(x,y))_0^R = 2x-1y$$
$$\text{subject to } -1x+3y\leq 4$$
$$-2x+2y\leq 3$$
$$0x+4y\leq 5$$
$$\min_{y\in Y}(f(x,y))_c = 1x+1y$$
$$\min_{y\in Y}(f(x,y))_0^L = 0x+0y$$
$$\min_{y\in Y}(f(x,y))_0^R = 2x+2y$$
$$\text{subject to } 1x-1y\leq 0$$
$$0x-2y\leq -1$$
$$2x-0y\leq 1$$
$$-1x-1y\leq 0$$
$$0x-0y\leq 0$$
$$-2x-2y\leq -1$$

Step 2. The problem is transferred to the following linear BLP problem by using method of weighting [27].

$$\min_{x\in X} F(x,y)=3x-6y$$
$$\text{subject to } -1x+3y\leq 4$$
$$-2x+2y\leq 3$$
$$0x+4y\leq 5$$
$$\min_{y\in Y} f(x,y)=3x+3y$$
$$\text{subject to } 1x-1y\leq 0$$
$$0x-2y\leq -1$$
$$2x-0y\leq 1$$
$$-1x-1y\leq 0$$
$$-2x-2y\leq -1.$$
$$0x-0y\leq 1$$

Step 3 Solve this linear BLP problem

$$\min_{x\in X} F(x,y)=3x-6y$$
$$\text{subject to } -1x+3y\leq 4$$
$$-2x+2y\leq 3$$
$$0x+4y\leq 5$$
$$1x-1y\leq 0$$
$$0x-2y\leq -1$$
$$2x-0y\leq 1$$
$$-1x-1y\leq 0$$

$$
\begin{aligned}
& -2x-2y \le -1 \,. \\
& 0x-0y \le 1 \\
& 3u_1+2u_2+4u_3-u_4-2u_5-0u_6-u_7-2u_8-0u_9-u_{10}=-3 \\
& u_1(4+1x-3y)+u_2(3+2x-2y)+u_3(5-4y)+ \\
& u_4(-x+y)+u_5(-1+2y)+u_6(1-2x)+ \\
& u_7(x+y)+u_8(-1+2x+2y)+u_9+u_{10}y=0 \\
& x\ge 0, y\ge 0, u_1\ge 0,\ldots,u_{10}\ge 0 \,.
\end{aligned}
$$

Step 4 The result is

$$\min_{x\in X}(F(x,y))_c = 1x-2y=-1$$

$$\min_{x\in X}(F(x,y))_0^L = 0x-3y=-1.5$$

$$\min_{x\in X}(F(x,y))_0^R = 2x-1y=-0.5$$

and

$$\min_{y\in Y}(f(x,y))_c = 0.5$$

$$\min_{y\in Y}(f(x,y))_0^L = 0$$

$$\min_{y\in Y}(f(x,y))_0^R = 1$$

$$x=0, y=0.5$$

Consequently, we have the solution of the problem (4.1)

$$\min_{x\in X} F(x,y)=\tilde{1}x-\tilde{2}y=\tilde{c}$$

$$\min_{y\in Y} f_1(x,y)=\tilde{1}x+\tilde{1}y=\tilde{d}$$

and

$$x=0, y=0.5\,,$$

where

$$\mu_{\tilde c}(t)=\begin{cases}0 & t<-1.5\\ \dfrac{t+1.5}{0.5} & -1.5\le t<-1\\ \dfrac{-0.5-t}{0.5} & -1\le t<-0.5\\ 0 & -0.5\le t\end{cases},\quad \mu_{\tilde d}(t)=\begin{cases}0 & t<0\\ \dfrac{t}{0.5} & 0\le t<0.5\\ \dfrac{1-t}{0.5} & 0.5\le t<1\\ 0 & 1\le t\end{cases}.$$

V. Conclusion

Many organizational decision problems can be formulated by bilevel programming models. Following our previous research [1, 2], this paper proposes the definition of optimal solution and related theorems for fuzzy parameter based linear bilevel programming. By using the proposed definition and theorems, this study develops a fuzzy number based Kuhn-Tucher approach to solve proposed FPBLP problem. A numeral example illustrates the power and details of the proposed approach. Further study includes the development of the model and related solving approaches for fuzzy parameter based multi-follower bilevel programming problems.

Acknowledgments

The work presented in this paper was supported by Australian Research Council (ARC) under discovery grants DP0557154 and DP0559213.

References

[1] C. Shi, G. Zhang, and J. Lu, "On the definition of linear bilevel programming solution", Applied Mathematics and Computation, vol. 160, 169-176, 2005.

[2] C. Shi, J. Lu, and G. Zhang, "An extended Kuhn-Tucker approach for linear bilevel programming", Applied Mathematics and Computation, vol. In press, 2004.

[3] H. Stackelberg, The Theory of the Market Economy. New York, Oxford: Oxford University Press, 1952.

[4] G. Anandalingam and T. Friesz, "Hierarchical optimization: An introduction", Annals of Operations Research, vol. 34, pp. 1-11, 1992.

[5] J. Bard, Practical Bilevel Optimization: Algorithms and Applications.: Kluwer Academic Publishers, 1998.

[6] Y. J. Lai, "Hierarchical optimization: A satisfactory solution", Fuzzy Sets and Systems, vol. 77, pp. 321-335, 1996.

[7] J. Bracken and J. McGill, "Mathematical programs with optimization problems in the constraints", Operations Research, vol. 21, pp. 37-44, 1973.

[8] E. Aiyoshi and K. Shimizu, "Hierarchical decentralized systems and its new solution by a barrier method", IEEE Transactions on Systems, Man, and Cybernetics, vol. 11, pp. 444-449, 1981.

[9] W. Bialas and M. Karwan, "Multilevel linear programming. Technical Report 78-1," State University of New York, Buffalo, Operations Research Program 1978.

[10] W. Candler and R. Townsley, "A linear twolevel programming problem", Computers and Operations Research, vol. 9, pp. 59-76, 1982.

[11] W. Bialas and M. Karwan, "Two level linear programming", Management Science, vol. 30, pp. 1004-1020, 1984.

[12] Y. Chen, M. Florian, and S. Wu, "A descent dual approach for linear bilevel programs. Technical Report CRT866," Centre de Recherche sur les Transports 1992.

[13] S. Dempe, "A simple algorithm for the linear bilevel programming problem." Optimization, vol. 18, pp. 373-385, 1987.

[14] G. Papavassilopoulos, "Algorithms for static Stackelberg games with linear costs and polyhedral constraints", presented at the 21st IEEE Conference on Decisions and Control, 1982.

[15] J. Bard and J. Falk, "An explicit solution to the programming problem", Computers and Operations Research, vol. 8, pp. 77-100, 1982.

[16] J. F. Amat and B. McCarl, "A representation and economic interpretation of a twolevel programming problem", Journal of the Operational Research Society, vol. 32, pp. 783-792, 1981.

[17] P. Hansen, B. Jaumard, and G. Savard, "New branchandbound rules for linear bilevel programming", SIAM Journal on Scientific and Statistical Computing, vol. 13, pp. 1194-1217, 1992.

[18] W. Bialas, M. Karwan, and J. Shaw, "A parametric complementary pivot approach for twolevel linear programming. Technical Report 802," State University of New York at Buffalo, Operations Research Program 1980.

[19] D. White and G. Anandalingam, "A penalty function approach for solving bilevel linear programs", Journal of Global Optimization, vol. 3, pp. 397-419, 1993.

[20] L. Leblanc and D. Boyce, "A bilevel programming algorithm for exact solution of the network design problem with useroptimal flows", Transportation Research, vol. 20, pp. 259-265, 1986.

[21] P. Marcotte, "Network optimization with continuous control parameters", Transportation Science, vol. 17, pp. 181-197, 1983.

[22] T. Miller, T. Friesz, and R. Tobin, "Heuristic algorithms for delivered price spatially competitive network facility location problems", Annals of Operations Research, vol. 34, pp. 177-202, 1992.

[23] L. A. Zadeh, "Fuzzy sets", Inform & Control, vol. 8, pp. 338-353, 1965.

[24] M. Sakawa, I. Nishizaki, and Y. Uemura, "Interactive fuzzy programming for multilevel linear programming problems with fuzzy parameters", Fuzzy Sets and Systems, vol. 109, pp. 3-19, 2000.

[25] G. Zhang, Y. Wu, M. Remia, and J. Lu, "Formulation of fuzzy linear programming problems as four-objective constrained problems", Applied Mathematics and Computation, vol. 139, pp. 383-399, 2003.

[26] University of Cambridge http://thesaurus.maths.org/dictionary/map/word/10037, 2001.

[27] M. Sakawa, "Interactive multiobjective linear programming with fuzzy parameters", in Fuzzy sets and interactive mulitobjective optimization. New York: Plenum Press, 1993.

XML Query Based on Ontology

Tao-Shen Li, Ting Han, Guo-Ning Chen

Abstract—**Document Type Definition (DTD) is a fundamental tool that enables users to constrain the structure of XML documents. This tool does not support semantic query. Therefore, it has recently become an active research topic in Web intelligence to endow XML with semantics for query quality. In this paper, we develop an extension of the DTD using for the provision of formal semantics.. It is implemented using an entity declaration in the DTD to describe the ontology and recurring this description to the form of Frame Logic. After validated by our extended DTD, a semantic valid XML document will be produced, so as to be queried using existing query languages (such as XQL). Our extended DTD is based on two main principles (1) maximizing the sharing of meta-data on the Web and (2) possibly using the DTD's provisions for reducing development expenses.**

Index Terms—**XML, DTD, Ontology, Frame logic, Semantic web**

I. Introduction

THE XML is widely known in the Internet community and has become the lingua franca for data dissemination, exchange and integration on the World Wide Web. Nearly every data management-related application now supports the import and export of XML, and standard XML schemas and DTD(Document Type Definition) are being developed promoted for all types of data sharing. XML implements the requirements of the universal expression and syntactic interoperability because anything for which a grammar can be defined can be encoded in XML and an XML parser can parse ang XML data. When it comes to semantic interoperability, however, XML has disadvantage as follows[1]:.

• XML just describes grammars and can't recognize a semantic unit from a particular domain, because it is designed only for markup in document structure but did not consider the common interpretation of the data contained in the document.

• XML is useful for data interchange between applications, but not for situation where new communication partners are frequently added. Because new information sources continually become available and new business partners join existing relationships on the Web, it is important to reduce the costs of adding communication partners as much as possible.

Recently, more attentions have been paid to endowing XML with the semantic property for higher quality of query. There are several approaches of defining ontology representations for semantic purposes, such as extending query languages, developing ontology languages or linking to an existent ontology. The work in [2] has shown how abstract models of agent communication and content languages are strongly related to the notion of domain ontology, and has presented a common framework for these types of model. The integration of XML query languages with IR-style similarity comparisons for ranked retrieval of XML data on the Web is presented in [3]. And the work of [4] enrich ontology using specialization processes based on some heuristics in order to offer to the expert of the domain a decision-making aid concerning its field of application. These projects ultimately adopt the method of linking to an existent ontology or extending query languages. The imaginable trouble for them may be the complicated mapping process, which influences the efficiency of the query greatly.

In [5], the ontology language OIL as an extension of RDFS is described. As a result, a full knowledge representation (KR) language can be expressed in RDFS and the extended language can be a maximal backward compatibility with RDFS. [6] defines an ontology language OWL to escape the limitation brought by RDF and RDFS, and this language is built upon RDF and RDFS. The work of [5] and [6] is to develop a kind of new ontology language and such language may be integrated and self-governed, but the workload seems vast and for users, they have to learn a new language over again. The SemanticMiner project[7] uses Frame Logic to define their ontology, and indicates that F-Logic covers most parts of OWL and allows specifying axioms freely. Additionally F-Logic uses the same syntactical constructs for both modeling and querying the ontology.

In this paper, we discuss a ontology model based on F-Logic, which this model is the template for our semantic description. We then briefly present an extension of the DTD using for the provision of formal semantics. It is implemented using an entity declaration in the DTD to describe the ontology and recurring this description to the form of Frame Logic. Our approach uses DTD to translate XML document from users, only and such translation doesn't include the change of document type and structure, it just changes the name of relevant elements or its attributes partially and the output is still a XML document. On the other hand, DTD also embodies the concept of domain but lack of semantic layer, which is a big limitation for its use in semantic web. So, our work is a consummation for DTD. By comparison with [5] and [6], our method we simplifies the development process and embodied the kernel of ontology successfully. SemanticMiner project also extended the query language for their ontology defined by F-Logic, while we just appeal our semantic purpose to an extended DTD which is similar to F-Logic in form, or the method of extending DTD just adopts the ideology and form of F-Logic.

The rest of this paper is structured as follows: Section 2 presents the basic concepts of the XML language and ontology. In section 3, the issues existed in the XML query will be discussed. Then ontology will be proposed as a solution. An implementation of our

This work was supported by the Natural Science Foundation of Guangxi,China under Grand Number 0229008.

Tao-Shen Li works at the Department of Computer Science and Technology, Guangxi University, Nanning, 530004 China (phone: 86-0771-3236627; fax: 86-0771-3236627; e-mail: tshli@ gxu.edu.cn).

Ting Han is a MS candidate in computer application technology from the Guangxi University, Nanning, 530004 China (e-mail: bobobye@126.com).

Guo-Ning Chen works at the Department of Computer Science and Technology, Guangxi University, Nanning, 530004 China (e-mail: chengu@cs.orst.edu).

approach will be shown in section 4, which uses entity declaration to extend the DTD and follows the form of Frame-Logic to describe ontology. A part of codes of our experiment and a query example will be also provided later. Finally, our conclusions and future work are presented.

II. BASIC CONCEPTS

A. XML Language

XML has been a new standard for information exchange on the World Wide Web. It allows users to define labels according to their interests and there are no syntax criterions when people write them, so it has a high flexibility. However, extensible characteristic of XML is its merit, but a biggest limitation of itself as well. The reason for this amphibious problem is that there is no way for server to understand labels defined by user himself. Traditional DTD is a basic tool that enables users to constrain the structure of XML documents. It can enforce constraints on which tags to use and how they should be nested within a document, but no help to the semantic aspect. Now, people are all making great efforts to find a validation mechanism for XML documents. One validation mechanism is XML Schema, which can offer more data types and also support naming space. The other is RDF (Resource Description Framework) which adopts URI to locate resources on the web accurately. These mechanisms are all trying to find an end-result for every resource, but they cannot establish semantic relations between elements or resources inside documents. Fortunately, philosophical ontology provides us a chance to solve this problem..

B. Ontology

In philosophy, ontology is the study of the kind of things that exist[8]. In the computer field, ontology is a sort of specific representation and description for conceptualization object using certain language, so it depends on the adopted language. According to the formalization degree of representation and description, ontology can be divided into absolute informalization, half formalization and rigid formalization. Ontology with higher formalization degree will be more favorable for computer to deal with automatically. From the definition of conceptualization, we can conclude that the terms, definition of terms and semantic web between terms in certain domain are information that should be included in the domain ontology.

Ontologies has played a key role in many fields[1,9], such as knowledge processing based web, share and Reusable Software. Ontologies are used in e-commerce to enable machine-based communication between buyers and sellers; vertically integration of markets; and description reuse between different marketplaces. Search engines also use ontologies to fine pages with words that are syntactically different but semantically similar.

If we use ontologies to define the shared conception hierarchy of certain domain, it will provide simple and comprehensible subjects which are used for communicating between the person and the application system. According to the limitation existed in the process of XML validation using traditional DTD, this paper presents an effective validation methodology based on ontology, which defines the terminology of a domain, provides a sound semantics, and formalizes relationships between the terms, i.e. it provides rich background knowledge, and implements the description via frame logic. Our method makes it possible that XML document could be semantic valid, therefore, improves the quality of query by a long way.

III. ISSUE

Traditional DTD, using simple syntax, provides effective validation mechanism for user. User can customize the DTD himself to limit XML document structure conveniently, and defines labels he likes consequently. If his cooperative fellows agree on this common DTD, then all documents can be kept consistently in the process of building, transferring, importing or translating the documents.A simple DTD document (Example1.dtd) is as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<!ENTITY % P "(#PCDATA)">
<!ELEMENT general (introduction, cooperation?)>
  <!ELEMENT introduction (corporation*)>
    <!ELEMENT corporation (name, support*)>
      <!ELEMENT name %P;>
      <!ELEMENT support %P;>
    <!ATTLIST corporation corporation_ID ID
      #REQUIRED>
  <!ELEMENT cooperation (item*)>
    <!ELEMENT item (technology, partner*)>
      <!ELEMENT technology %P;>
      <!ELEMENT partner %P;>
<!ATTLIST item item_ID ID #REQUIRED>
```

The Example1.dtd above can make validity validation for following example of XML document (example1.xml):

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE test SYSTEM "test.dtd">
<general>
  <introduction>
    <corporation corporation_ID ="_001">
      <name>BenTeng Computer Corporation</name>
      <support>software development</support>
    </corporation>
    <corporation corporation_ID ="_002">
      <name>AnKang Pharmaceutical group</name>
      <support>bio_pharmacy</support>
    </corporation>
  </introduction>
  <cooperation>
    <item item_ID ="_1001">
      <technology>software development</technology>
        <partner>BenTeng Computer Corporation</partner>
        <partner>ShengLi Computer Corporation</partner>
    </item>
  </cooperation>
</general>
```

From example above we can see that the basic function of traditional DTD validation is to prescribe logic structures of documents. It defines elements of page, attributes of element and relations between elements and attributes, such as the order or appearing times for the sub-elements of certain element. It also defines whether an element has value and what the data type of its value is, etc. So, the action of the DTD is to interpret all details that a valid XML document needs to be knew, just like a syntax analyzer.

Well then, is server able to meet the requirement of user when user asks for processing? First, let's come to a query request as follow:

//corporation/name[../support="software development"]

This query sentence is written according to the syntax of XQL. Its requirement is to find out the corporations which have technology of software development and return the list of corporation name to user. However, if we analyze the semantic of document further, we will easily find that *ShengLi computer* has cooperative relation with *BenTeng computer*. That is, *ShengLi computer* has the technology of software development too. Its name should be also returned to user. But, such result can only be obtained using following query sentence:

//corporation/name[../support="software development"]
$union$
//item/partner[../technology ="software development"]

Obviously, server cannot build query sentence above if we only rely on the traditional DTD. How can we solve such problem? Practical test shows that we can make such query easy to accomplish if we use an ideology of semantic validation based on ontology introduced in the following section.

IV. Implementation

A. Frame Logic

People use semantic network and frame logic to describe hierarchy relations and association relations between concepts in artificial intelligence (AI). Hierarchy relations use the way of the diagram example to describe different and relative views between concepts, while association relations are structured representations about individual class.

Considering the goal of trying to form a corresponding relation with the DTD hierarchy, we use frame logic to describe domain ontology. "F-Logic is a deductive, object oriented database language which combines the declarative semantics and expressiveness of deductive database languages with the rich data modeling capabilities supported by the object oriented data model."[8] In the following parts, we will present an ontology document example2.dtd. It is an ontology description based on frame logic. It describes a domain including people and corporation. The content of this document is composed of the hierarchy of the domain, relations between concepts and some axioms or rules. There are three sections in it. First section of the ontology describes the notional level of the domain, where "::" shows a inclusive relation that the concept on the right is the upper one to the concept on the left. Second section describes the unification of concepts. It introduces the concepts into the attributes definition, defines attributes of the concepts in the "[ ]", and explains the data type definition of the attributes with "=>>". The last section describes some accepted rules or relations, where "<->" shows these relations.

```
(example2.dtd)
domain[ ].
   people :: domain.
      employee :: people.
         technician :: employee.
            programmer :: technician.
               bachelor :: programmer.
      student :: people.
         bachelor :: student.
   corporation :: domain.
      State corporation :: corporation.
      foreign capital corporation :: corporation.
         IT company :: foreign capital corporation
      joint-stock corporation :: corporation.

people[name=>>STRING;email=>>STRING;company
   =>>corporation; address =>>STRING].
employee[employee ID=>>bachelor].
technician[supervise=>>programmer].
programmer[cooperate  with =>> programmer].
bachelor[administer =>>technician].
student[student ID=>>NUM].
corporation[incorporator=>>people;name=>>STRING;
   property=>>STRING;summary=>>STRING].
foreign capital corporation[linkman=>>people; corporation
   ID=>>NUM; number of employee =>>NUM; calling=>>IT
   company].
IT company[property=>>corporation].

FORALL Jerry, Tom
   Jerry: programmer[cooperate with ->> Tom] <->
      Tom: programmer[cooperate with ->> Jerry].
FORALL Jerry, certain corporation
   Certain corporation:corporation[incorporator->>Jerry]<->
      Jerry: people[company ->> certain corporation].
FORALL Jerry, certain foreign capital corporation
   certain foreign capital corporation: foreign capital
corporation[linkman ->> Jerry] <->
      Jerry: people[ company ->> certain corporation].
FORALL Jerry, Tom
   Jerry: bachelor[administer ->> Tom] <->
      Tom: technician[supervise ->> Jerry].
FORALL corporation A, corporation B
   corporation B: foreign capital corporation[calling ->>
      corporation A] <->
corporation A: IT company[property ->> corporation B].
```

B. Semantic Validation Using the DTD

Example2.dtd exhibits the advantage of ontology when describing conceptual hierarchy. According to this advantage, two concepts which have no direct relation could be associated via their common upper concept. Thus, the semantics of document can be understood on the higher level.

The substance of the ontology is to show a kind of inherited relation between concepts. If a concept owns a upper concept, then it will inherit the attributes of its upper concept naturally. We can also conclude that if there are two concepts having some common attributes, it is possible that they have the same upper concept. Thus, concepts are constrained to a semantic intersection chain which the concept has itself. Regardless the methods to express the concept are different, all concepts should be found as long as they have same semantic. However, whether need to define a new language that carry on formalization description to the ontology? Obviously, this task is very complicated and enormous. Considering that traditional DTD has had a powerful function for structure validation, if we use the DTD to carry on formalization description to the ontology and introduce ontology to the DTD, it will be the most valid and efficient way. That is the basic idea of this paper. Figure 1 shows the framework description of the tasks to need to be complete for implementing this idea.

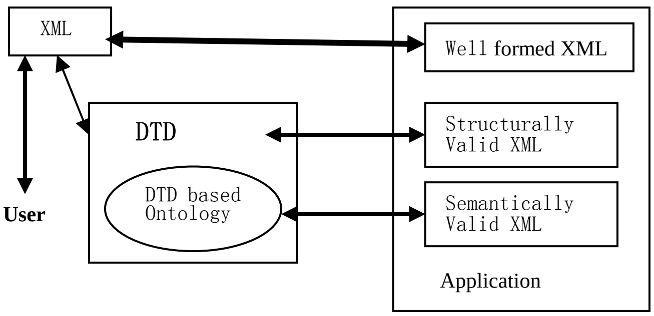


Figure 1 framework description of the task for implementing DTD semantic validation

Before introducing ontology to the DTD, first problem we need to solve is what should be used to define ontology in the DTD. It perhaps is the method that people first thought of to define every ontology as an element. However, concepts often exist in the definition of attributes in the actual circumstance, or we can say more constitutionally that some concepts contain some attributes defined by another concept. So, it is not enough that just defining every ontology as an element simply. Therefore, we consider adopting an extended method that using element to define ontology and introducing the defined concept to the attribute definition at the same time. The result for doing so is that we can find the semantic root of every concept. This ideology seems like the definition of class in OOP (Object-Oriented Programming). Conclusion of our analysis is that inherited relations between concepts are produced at the same time they are defined, in other words, such relations are inherent.

The next problem to solve is how to implement inherited relations between concepts in the DTD. Because what we need to hold is the semantic contents of the document, and these contents are defined only in the values of element or attribute, so the problem will be translated into the definition of values. In the DTD, there are ten kinds of inner data types and an entity declaration method for values definition. Among them entity declaration makes values definition to become more flexible. It can be used to define reused data blocks or to cite non-XML data for simplifying the DTD and enhancing the readability. However，it is to be noticed that entity is a placeholder representing content and has inherent substitution ability, and the substituted content could be a meaningful phrase or concept. The key point of this paper is to develop and emphasize this powerful ability contained in entity declaration of the DTD, and apply this ability to the attribute definitions that contain some concepts. Consequently we can implement associations between concepts through ontology. From this point of view, entity transfers the common information between concepts just as an excellent carrier of ontology element, and also produces more meaningful ontology elements. In fact, this process is a kind of evolution for concepts. Certainly, in order to implement the inherited relation between concepts, the concept transferred by entity and the concept to be defined must be in the same semantic chain.

The special point of our semantic validation is that what entity declares is an ontology element which has been defined. With that we can define other elements that are semantic interrelated with it and produce new ontology. For example, in <!ELEMENT branch (#PCDATA | %tree;)* >, new ontology element branch is defined by ontology element tree which has been defined and inherits its attributes at the same time. If we continue to associate all the concepts like that, semantic chains between concepts will be established. It is obvious that the description process of whole document is the conformation process of ontology actually.

C. Example of Semantic Validation Using the DTD

According to the design idea above, we performed some researches and experiments. In the following parts we shall list a part of codes and illustration. In these DTD documents, the three parts of ontology description given in example2.dtd can also be expressed in the document.

```
<!-- entities for realizing the is-a hierarchy -->
<!ENTITY % people "people | employee | student | technician |
   programmer | bachelor " >
<!ENTITY % programmer "programmer | bachelor " >
<!ENTITY % corporation " corporation | state corporation |
   foreign capital corporation | joint-stock corporation | IT
   company " >
<!ENTITY % foreign capital corporation " foreign capital
   corporation | IT company " >

<!-- element declarations for ontology concepts -->
<!ELEMENT people (#PCDATA | name | email | company |
   address)*>
<!ELEMENT programmer (#PCDATA | name | email |
   company | address | employee ID | administer | cooperate
   with)*>
<!ELEMENT corporation (#PCDATA | incorporator | name |
   property | summary)*>
<!ELEMENT foreign capital corporation (#PCDATA |
   incorporator | name | property | summary | linkman |
   corporation ID | number of employee | calling)*>
<!ELEMENT IT company (#PCDATA | incorporator | name |
   property | summary | linkman | corporation ID | number of
   employee | calling | property)*>

<!-- ATTLIST declatation for ontology attributes -->
<!ATTLIST people
   name     CDATA #IMPLIED
   email    CDATA #IMPLIED
   company  CDATA #IMPLIED
   address   CDATA #IMPLIED>

<!-- element declaration for ontology attributes -->
<!ELEMENT incorporator (#PCDATA | %people;)* >
<!ELEMENT linkman (#PCDATA | %people;)* >
<!ELEMENT address (#PCDATA) >
<!ELEMENT company (#PCDATA | %corporation;)* >
<!ELEMENT cooperate with (#PCDATA|%programmer,)*>
```

The first part of this DTD document is the entity declaration which translates inherited relation between concepts in ontology into substitution relation between concepts. The second part defines the value and attribute of upper ontology elements which are corresponding to the entities declared in entity declarations and belong the outmost layer of the frame logic. The third part implements inherited relation of concepts actually and uses entity declaration of defined ontology elements to define new ontology, which consequently makes new ontology inherit values and attributes defined in the second part and the new ones become the nether concept of the defined ontology.

Though the DTD above is a subset of this domain description, we can still find that nether concept inherit its upper concept very well through declaration of parameter entity, so that it contains either the commonness of the domain which it is affiliated to or its own individuality. This method presents a valid way to help extracting common information between concepts or searching for concepts through some information.

D. Query Based Ontology

The functionality of ontology in our approach can be generalized as defining a common vocabulary and improving the quality of query answers. Using the DTD based ontology discussed above, we can validate XML documents much further from semantics. Our experimental system implemented the process of translating users' XML documents into the semantic valid XML documents according to our extended the DTD. After translating, elements having semantic association in XML will be found through a common element which may be a unattached element or be showed in element attributes.

Let's come back to the example1.dtd presented in section two. We know that the meanings of partner in element item indicate certain corporation, and the technology of the cooperation is just the attribute of partners. So, we need to define substitution of corporation for partner, then naturally, the technology of partner becomes the attribute of corporation named support. Following entity definition implements this target:

```
<!ENTITY % corporation "corporation | partner">
<!ENTITY % support "support | technology">
```

Certainly, we should define element to support the attribute of corporation, just like:

```
<!ATTLIST corporation
    corporation_ID ID #REQUIRED
    name CDATA #IMPLIED
    support CDATA #IMPLIED>
```

In succession, we can define technology and partner using defined entity -- %support and %corporation:

```
<!ELEMENT technology (#PCDATA | %support;)*>
<!ELEMENT partner (#PCDATA | %corporation;)*>
```

According to amended example1.dtd, example1.xml was translated into the new one following after our processing:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE test SYSTEM "test.dtd">
<general>
  <introduction>
     <corporation corporation_ID ="_001" name = "BenTeng
        Computer Corporation" support = "software
        development"/>
     <corporation corporation_ID ="_002" name = "AnKang
        Pharmaceutical group" support = "bio_pharmacy"/>
  </introduction>
  <cooperation>
     <item item_ID ="_1001">
        <corporation name = "BenTeng Computer Corporation"
             support = "software development"/>
        <corporation name = "ShengLi Computer Corporation"
             support = "software development"/>
     </item>
  </cooperation>
</general>
```

From the above-mentioned translation result, we can see that effect is remarkable The element partner has been replaced by the element corporation and has the attribute support. Now, we can implement the simple query for XML document as we wish and the query answers will never miss some information correlated with the request. Our query is just request for the name of corporations who have software development technology. For well output we limit the format:

```
<result>
{
 for $c in doc("H:/xml/testdata/test.xml")//corporation
 where $c/@support = "software development"
 return
  <corporation>
   { $c/@name }
  </corporation>
}
</result>
```

Just as we expect before, we get the corporation name not only "*BenTeng Computer Corporation*" but also "*ShengLi Computer Corporation*"

```
<result>
   <corporation name="BenTeng Computer
       Corporation" />
   <corporation name="ShengLi Computer
       Corporation" />
</result>
```

The experiment shows that the method we presented in this paper can improve veracity of information retrieval remarkably. This trait is a big help for users who need precise query. In additions, ontology using the DTD description has well maintainability. The introduction of new ontology is accomplished with definition of its elements and has no influence to the defined ontology elements, namely new ontology is just linked to the semantic chain simply and the link point is the upper concept associated with it directly. After this, we can define offspring belonging to this new ontology itself and extend the semantic chain.

Certainly, the unification of ontology and XML is not just limited to the DTD, but the simpleness of the DTD and its strong suit for describing structured data provide ontology with a simplest and easiest carrier. With the perfectness of relevant criterion of XML, we believe that the unification of ontology and XML schema will be deepened continually, so will be the semantic comprehension for XML document based content.

E. Intelligent XML Query System

As we showed above, our method has several important advantages for XML files querying. First, concepts with the same meaning can be translated into a formal expression which exists in Semantic Validation the DTD where a common vocabulary of system has been defined. Second, concepts with no relationships can be also associated in terms of rules defined in Semantic Validation the DTD. On the other hand, XML files after translation have uniform labels and coherent contexts, which means that users can get same format answers that are more understandable and follow the habits of users. Based on these features, putting our method to application will be helpful. We present here an implementation frame for XML query system based on ontology. It contains a Semantic Validation the DTD as semantic model of system.

The infrastructure of System takes charge to accept XML files' register, put XML files into format validation, structure validation from the DTD included and finally semantic validation from Semantic Validation the DTD of system. Then, XML files rejected by semantic validation will be translated into formal expressions and stored into database with all the eligible XML files. We can set a validation module and a translation module to implement these tasks through interacting with the Semantic Validation the DTD, and let them give their attentions to two points, one is whether the concepts contained in the labels are formal, and the other is whether the nested structure of concepts is coincident with the successive relationships between them.

Based on the data stored in the database, we can establish our applications. We start this from User Interface which receives users' query using a series of forms (XForms is a recommended standard now). After parsing, we translate these forms into queries written in XQuery language, then they can be performed by XQuery Engine directly. From database, finally, XQuery Engine gets all the information user want and displays the answers on the User Interface.

Here, we present a common framework for XML query application. Actually, because our method can be designed into an unattached module, many existing information systems can be integrated into their infrastructure to provide semantic support expediently. The only problem is that the model of the Semantic Validation DTD needs to be accepted widely. Now the easiest and soonest way we can take is to write it according to existing ontology (or domain ontology). We think that our method will be more compactly for intellectualized upgrade of information system.

V. CONCLUSION

The research on XML Query is currently increasing because of the wide dissemination of business data over the Web. We argue that the main problem related to the Query of XML is that the query result may neglect the content having semantic relevancy to the user request. That is, when there are several names for the same information, the trouble omes. Because of this, very heterogeneous ontology models arise for XML Query when we compare different models.

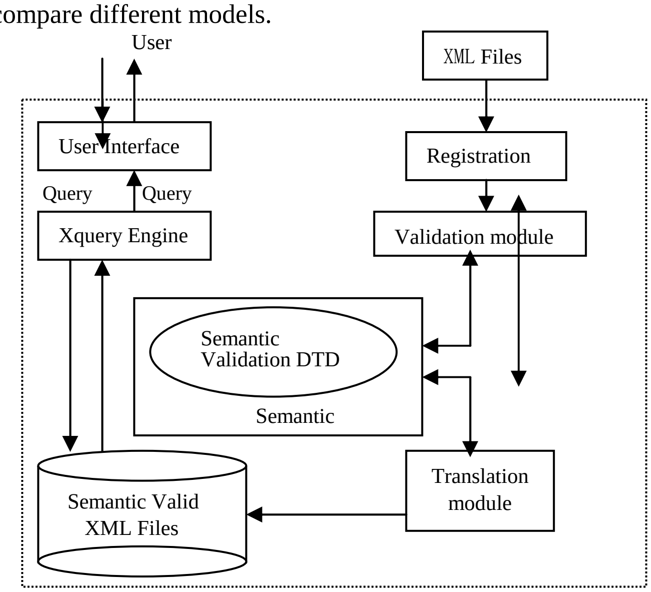


Figure 2 XML Query System Based Ontology

Our contribution to this problem is a method using ontology to enrich the DTD. Particularly, recurring to entity declarations, we deal with the problem of how to represent ontology in the DTD.

Another strong point of our approach is that we use the fabric of f-logic to establish ontology, therefore make this abstract concept comprehensible.

Currently, we are extending our method with an analysis of XML instances in order to improve the definition of query language. After finishing the extension, we will continue to optimize the quality of query.

Other current works are related to the quest for a better carrier of ontology, as well as the research for DAML+OIL and OWL in vogue.

REFERENCES

[1] S. Becker, S. Melnik, F. Van Harmelen, et al. The Semantic Web: the Roles of XML and RDF. IEEE Internet Computing, September-October 2000, 63—75
[2] Stephen Cranefield, Martin Purvis, Mariusz Nowostawski. Is it an Ontology or an Abstract Syntax? Modelling Objects, Knowledge and Agent Messages. The Information Science Discussion Paper Series, Number 2000/08, April 2000, ISSN 1172-6024
[3] Anja Theobald, Gerhard Weikum. Adding Relevance to XML[J]. Modern Information Retrieval. WebDB (Informal Proceedings) 2000: 35-40
[4] E. Desmontils, C. Jacquin, L. Simon. Ontology enrichment and indexing process. Research Report, 2003, No 03.05
[5] Jeen Broekstra,Michel Klein,Stefan Decker, et al. Enabling knowledge representation on the Web by extending RDF Schema. Computer Networks 39 (2002) 609﹣634
[6] Grigoris Antoniou1, Frank van Harmelen2. Web Ontology Language: OWL. Handbook on Ontologies in Information Systems. Springer-Verlag 2003
[7] E. Moench, M. Ullrich, Hans-Peter Schnurr, et al. SemanticMiner–Ontology-Based Knowledge Retrieval. j-jucs, volume 9 (2003) issue 7, 682-696
[8] K. Hori. An ontology of strategic knowledge: key concepts and applications. Knowledge-Based Systems, volume13 (2000), 369-374
[9] B. Chandrasekaran, J. R. Josephson, V. Richard Benjamins. What Are Ontologies, and Why Do We Need Them?. IEEE Intelligent Systems, January-February 1999, 20-26
[10] Liao Minghong. Ontology and Information Retrieval. Computer Engineering, 2000,26(2): 56－58
[11] Wan Jie, Teng Zhiyang. Application of Ontology in Content-based Information Retrieval. Computer Engineering, 2003,29(4):122－123,152

A New Similarity Measure for Vague Sets

Jingli Lu, Xiaowei Yan, Dingrong Yuan, Zhangyan Xu

***Abstract*--Similarity measure is one of important, effective and widely-used methods in data processing and analysis. As vague set theory has become a promising representation of fuzzy concepts, in this paper we present a similarity measure approach for better understanding the relationship between two vague sets in applications. Compared to existing similarity measures, our approach is far more reasonable, practical yet useful in measuring the similarity between vague sets.**

***Index Terms*-- Fuzzy sets, Vague sets, Similarity measure**

I. INTRODUCTION

In the classical set theory introduced by Cantor, a German mathematician, values of elements in a set are only one of 0 and 1. That is, for any element, there are only two possibilities: in or not in the set. Therefore, the theory cannot handle the data with ambiguity and uncertainty.

Zadeh proposed fuzzy theory in 1965 [1]. The most important feature of a fuzzy set is that fuzzy set *A* is a class of objects that satisfy a certain (or several) property. Each object *x* has a membership degree of *A*, denoted as $\mu_A(x)$. This membership function has the following characteristics: The single degree contains the evidences for both supporting and opposing *x*. It cannot only represent one of the two evidences, but it cannot represent both at the same time too.

In order to deal with this problem, Gau and Buehrer proposed the concept of vague set in 1993 [2], by replacing the value of an element in a set with a sub-interval of [0, 1]. Namely, a true-membership function $t_v(x)$ and a false-membership function $f_v(x)$ are used to describe the boundaries of membership degree. These two boundaries form a sub-interval $[t_v(x), 1 - f_v(x)]$ of [0, 1]. The vague set theory improves description of the objective real world, becoming a promising tool to deal with inexact, uncertain or vague knowledge. Many researchers have applies this theory to many situations, such as fuzzy control, decision-making, knowledge discovery and fault diagnosis. And the tool has presented more challenging than that with fuzzy sets theory in applications.

In intelligent activities, it is often needed to compare and couple between two fuzzy concepts. That is, we need to check whether two knowledge patterns are identical or approximately same, to find out functional dependence relations between concepts in a data mining system. Many measure methods have been proposed to measure the similarity between two vague sets (values). Each of them is given from different side, having its own counterexamples. Such as Shyi-Ming Chen proposed a similarity measure M_C in [3], whereas from the M_C model we can gain the similarity of vague values [0.5, 0.5] and [0, 1] is 1, obviously their similarity should not be 1. In reference [5] Dug Hun Hong put forward another similarity measure M_H, according to the formulae of M_H, we can obtain the similarity of vague values [0.3, 0.7] and [0.4, 0.6] is $M_H([0.3,0.7],[0.4,0.6]) = 0.9$, in the same model, we can get $M_H([0.3,0.6],[0.4,0.7]) = 0.9$. In a voting model, the vague value [0.3, 0.7] can be interpreted as: "the vote for a resolution is 3 in favor, 3 against and 4 abstentions"; [0.4, 0.6] can be interpreted as: "the vote for a resolution is 4 in favor, 4 against and 2 abstention". [0.3, 0.6] and [0.4, 0.7] can have similar interpretation. Intuitively, [0.3, 0.7] and [0.4, 0.6] may be more similar than [0.3, 0.6] and [0.4, 0.7]. Therefore sometimes the results of M_H model are not accordant with our intuition. After analyzing most existing vague sets and vague values similarity measures, we find out that almost each measure has its defect. In section 2.2, we will illustrate them with more examples.

Then we have to make a choice according to the applications. A more reasonable approach is proposed to measure similarity in this paper, after analyzing existing methods.

The remaining of this paper is organized as follows. In Section 2, several methods of similarity measure for vague set are discussed. An improved similarity measure method and its properties are given in Section 3. Section 4 concludes this paper.

II. PRELIMINARIES

A. Vague set

In this section, we review some basic definitions of vague values and vague sets from [2], [3], [4].

Definition 1 Vague Sets [2]: Let *X* be a space of points (objects), with a generic element of *X* denoted by *x*. A vague set *V* in *X* is characterized by a truth-membership function t_v and a false-membership function f_v. t_v is a lower bound on the grade of membership of *x* derived from the evidence for *x*, and f_v is a lower bound on the negation of *x* derived from the evidence against *x*, t_v and f_v both associate a real number in the interval [0,1] with each point in *X*, where $t_v + f_v \le 1$. That is

$$t_v : X \to [0,1] \;\; ; \;\; f_v : X \to [0,1]$$

This approach bounds the grade of membership of *x* to a sub-interval $[t_v(x), 1 - f_v(x)]$ of [0,1]

When *X* is continuous, a vague set *V* can be written as

$$V = \int_X [t_V(x), 1 - f_V(x)]/x, \; x \in X .$$

When *X* is discrete, a vague set *V* can be written as

$$V = \sum_{i=1}^{n} [t_V(x_i), 1 - f_V(x_i)]/x_i, \; x_i \in X .$$

Department of Computer Science, Guangxi Normal University, Gulin, 541004, P.C.China.
lujingli502@sohu.com; yanxw@mailbox.gxnu.edu.cn.

$x = [t_x, 1 - f_x]$ and $y = [t_y, 1 - f_y]$. If $t_x = t_y$ and $f_x = f_y$, then the vague values x and y are called equal (i.e., $[t_x, 1 - f_x] = [t_y, 1 - f_y]$).

Definition 3: Let A and B be vague sets of the universe of discourse U, $U = \{u_1, u_2, \cdots u_n\}$, where

$$A = [t_A(u_1), 1 - f_A(u_1)]/u_1 + [t_A(u_2), 1 - f_A(u_2)]/u_2 + \cdots + [t_A(u_n), 1 - f_A(u_n)]/u_n$$

$$B = [t_B(u_1), 1 - f_B(u_1)]/u_1 + [t_B(u_2), 1 - f_B(u_2)]/u_2 + \cdots + [t_B(u_n), 1 - f_B(u_n)]/u_n$$

If $\forall i$, $[t_A(u_i), 1 - f_A(u_i)] = [t_B(u_i), 1 - f_B(u_i)]$, then the vague sets A and B are called equal, where $1 \le i \le n$.

B. Research into similarity measure

Currently, there have been many similarity measurements for vague set (value). Suppose that $X = [t_x, 1 - f_x]$ and $Y = [t_y, 1 - f_y]$ are two vague values over the discourse universe U. Let $S(x) = t_x - f_x$, $S(y) = t_y - f_y$, the M_C, M_H, M_L and M_O models are defined respectively in [3], [5], [6] and [7] as follows,

$$M_C(x,y) = 1 - \frac{|S(x) - S(y)|}{2} = 1 - \frac{|(t_x - t_y) - (f_x - f_y)|}{2} \tag{1}$$

$$M_H(x,y) = 1 - \frac{|t_x - t_y| + |f_x - f_y|}{2} \tag{2}$$

$$M_L(x,y) = 1 - \frac{|S(x) - S(y)|}{4} - \frac{|t_x - t_y| + |f_x - f_y|}{4} = 1 - \frac{|(t_x - t_y) - (f_x - f_y)| + |t_x - t_y| + |f_x - f_y|}{4} \tag{3}$$

$$M_O(x,y) = 1 - \sqrt{\frac{(t_x - t_y)^2 + (f_x - f_y)^2}{2}} \tag{4}$$

Comparisons among the M_C, M_H, M_L, M_O models can also be found in [7].

From the definition of the M_C model, we know that

$$t_x - f_x = t_y - f_y \Rightarrow M_C \equiv 1,$$

i.e. the M_C model is too rough when $t_x - f_x = t_y - f_y$.

The M_H model pays equal attention both to the difference of two true-membership degrees and to the difference of two false-membership degrees, between two vague values. Pairs of vague values, which have both the same difference of true-membership degrees and the same difference of false-membership degrees, have the same similarity. But it does not distinguish the positive difference and negative difference between true- and false-membership degrees.

The M_L model inherits the advantages of the M_C and M_H models, paying equal attentions to the support of vague value, true-membership degree, and false-membership degree, respectively. But it uses absolute values, and hence increases the possibility of similarity coincidence. For example, it cannot distinguish between pair of ([0.4, 0.8], [0.5, 0.7]) and pair of ([0.4, 0.8], [0.5, 0.8]). According to our intuition, pair of ([0.4, 0.8], [0.5, 0.8]) is more similar than pair of ([0.4, 0.8], [0.5, 0.7]), but in the model of M_L, the two pairs of vague values have the same similarity.

The M_O model also reflects the equal concerns between the difference of true-membership degrees and the difference of false-membership degrees. But similar to the M_H model, the M_O model does not consider whether the differences are positive or negative.

The above methods of similarity measure can be used to solve the problem of how to determine the similarity between two vague values in a certain extent. But each of them focuses on different aspects. There are three factors which affect the similarity of vague values: true-membership function t_x, false-membership function f_x, and $1 - t_x - f_x$. The reason there are many counterexamples under the measures of the M_H, M_L and M_O models is that the weights of $|t_x - t_y|$, $|f_x - f_y|$ and $|(t_y + f_y) - (t_x + f_x)|$ in the above methods are constants. Its explicit characteristic is that it is not considered whether the difference is positive or negative. Based on this idea, a new weighted and variable similarity measure is proposed. It can considerably reduce the possibility of similarity coincidence.

III. Measuring the Similarity between Vague Sets

This section constructs a new approach for measuring the similarity between vague sets, analyzes the properties and illustrates the use by examples.

A. A New Similarity Measure

We first give an example. Assume that there are four candidates A, B, C, D, and ten voters. One voter supports A, one opposes A; two support B, one opposes B; seven support C, one opposes C; eight support D, and one opposes D. The voting results of A, B, C, and D can be viewed as four vague values, A[0.1, 0.9], B[0.2, 0.9], C[0.7, 0.9], and D[0.8, 0.9]. Now we compare the similarities between A and B, and between C and D. By formulae (1), (2), (3) and (4), we obtain the similarities as shown in Table 1.

Table 1.
An example of similarity calculation

	x	y	Mc	M_H	M_L	M_O	M'
A,B	[0.1,0.9]	[0.2, 0.9]	0.95	0.95	0.95	0.929	0.968
C,D	[0.7, 0.9]	[0.8, 0.9]	0.95	0.95	0.95	0.929	0.953

From Table 1, we can see that the similarities between A and B (and between C and D) are all the same by using the Mc, M_H, M_L and M_O models. If only a candidate can be selected and renunciation is considered, D is most possible to be selected. The possibility of selecting C is smaller than that of selecting D. Selecting A or B has rather low possibility. Intuitively, it should be easier to say that A and B are similar than C and D, because A and B are all impossible options, D might be selected, and C might not be selected. Among A, B, C, and D, we would be most concerned with the similarity between C and D. We need to enlarge the difference of similarities where we are concerned.

For another example, assume that there are other four candidates E, F, G, H, and ten voters. One voter supports E, nine oppose E; one supports F, eight oppose F; one supports G, two oppose *G*; one supports *H*, and one opposes *H*. The voting results of *E*, *F*, *G*, and *H* can also be viewed as four vague values, *E*[0.1, 0.1], *F*[0.1, 0.2], *G*[0.1, 0.8], and *H*[0.1, 0.9]. Now we compare the similarities between *E* and *F*, and between *G* and *H*. By formulae (1), (2), (3) and (4), we obtain the similarities as shown in Table 2.

TABLE 2.
AN EXAMPLE OF SIMILARITY CALCULATION

	x	*y*	*Mc*	M_H	M_L	M_O	*M'*
E,F	[0.1,0.1]	[0.1, 0.2]	0.95	0.95	0.95	0.929	0.948
G,H	[0.1, 0.8]	[0.1, 0.9]	0.95	0.95	0.95	0.929	0.931

We definitely know only the minority of voters support the candidates *E*, *F*, *G* or *H*, and the majority of voters oppose *E* or *F*, whereas we have little information about *G* or *H* because there are so many abstainers. Intuitively compared with *E* and *F*, *G* and *H* should have less similarity. But from the Table 2, we can see the similarities between *E* and *F* (and between *G* and *H*) are all the same by using the *Mc*, M_H, M_L and M_O models.

Then we need a new similarity measure which can magnify what we are concerned. That is, if $t_x - t_y$ is equal to $f_x - f_y$, the similarities can still be different. For example, the larger the support ($t_x + t_y$) is, the smaller the similarity should be. Analogically, the smaller the opposition ($f_x + f_y$) is, the smaller the similarity should be.

Based on the above discussion, we propose a weight-varied similarity measure *M*', i.e.

$$M' = 1 - \frac{(t_x + t_y)\,|\,t_x - t_y\,|}{(t_x + t_y) + (2 - f_x - f_y) + 2} - \frac{(2 - f_x - f_y)\,|\,f_x - f_y\,|}{(t_x + t_y) + (2 - f_x - f_y) + 2} - \frac{|\,(t_x - t_y) - (f_x - f_y)\,|}{(t_x + t_y) + (2 - f_x - f_y) + 2} \qquad (5)$$

Coefficient ($t_x + t_y$) of $|t_x - t_y|$ implies that when $|t_x - t_y|$ is the same, similarity should be smaller if ($t_x + t_y$) is larger. Coefficient ($2 - f_x - f_y$) of $|f_x - f_y|$ means that when $|f_x - f_y|$ is the same, similarity should be smaller if ($f_x + f_y$) is smaller.

Let $t_x + t_y = p$, $t_x - t_y = q$, $f_x + f_y = m$, $f_x - f_y = n$. Then, formula (5) can be reduced as

$$M' = 1 - \frac{|\,pq\,| + |\,(2-m)n\,| + |\,q - n\,|}{p + (2 - m) + 2} \qquad (6)$$

Where the *M'* model pays attention to both the difference of true-membership degrees and the difference of false-membership degrees between vague values. It also implies the attention to the support of vague value. Because of the introduction of $|(t_x - t_y) - (f_x - f_y)|$, The *M'* model can distinguish positive difference and negative difference. The strategy of varied-weight leads to the reduced possibility of similarity coincidence, and the weights meet the requirement that we are concerned with those similarities where supports are high and oppositions are low. For the sake of comparison, we enlarge the difference of similarities when the support is large and the opposition is small.

From the above definition, we obtain the following properties.

Property 1: $M'(x, y) \in [0, 1]$.

Proof: Since $t_x \in [0, 1]$, $t_y \in [0, 1]$, $f_x \in [0, 1]$, $f_y \in [0, 1]$, we have

$|t_x - t_y| \in [0, 1]$, $|f_x - f_y| \in [0, 1]$, $|(t_x - t_y) - (f_x - f_y)| \in [0, 2]$

$$M' \le 1 - \frac{(t_x + t_y)\cdot 0 + (2 - f_x - f_y)\cdot 0 + 0}{(t_x + t_y) + (2 - f_x - f_y) + 2} = 1$$

$$M' \ge 1 - \frac{(t_x + t_y) + (2 - f_x - f_y) + 2}{(t_x + t_y) + (2 - f_x - f_y) + 2} = 0 \qquad \blacksquare$$

Property 2: $M'(x, y) = M'(y, x)$.

It is obtained directly from the definition of the *M'* model.

Property 3: $M'(x, y) = 0 \Leftrightarrow x = [0, 0]$ and $y = [1, 1]$; or $x = [1, 1]$ and $y = [0, 0]$.

Proof: For $x = [0, 0]$ and $y = [1, 1]$ (or for $x = [1, 1]$ and $y = [0, 0]$), by the definition, we obviously have $M'(x, y) = 0$; and

If $M'(x, y) = 0$, we have $t_x - t_y = 1$ and $f_x - f_y = -1$;
or $t_x - t_y = -1$, $f_x - f_y = 1$

Hence, $x = [0, 0]$ and $y = [1, 1]$;
or $x = [1, 1]$, $y = [0, 0]$ $\blacksquare$

Property 4: $M'(x, y) = 0 \Leftrightarrow x = y$.

Proof: If $x = y$, from the definition, it is clear that

$$M'(x, y) = 1.$$

If $M'(x, y) = 1 \Rightarrow t_x - t_y = 0$, $f_x - f_y = 0$, that is, $x = y$. $\blacksquare$

Example 1: In table 3, seven groups of vague values (*x*, *y*) are given. Intuitively, the similarity of the first pair vague values should be larger than the second pair, namely M(x1, y1) > M(x2, y2). And experientially M(x4, y4) < M(x5, y5); M(x6, y6) < M(x7, y7). Consider the 7 groups of data pairs (*x*, *y*) in the second and third rows of Table 2. We compare our measure method with others. The results are shown in 4th — 8th rows of Table 3.

TABLE 3.
COMPARISONS OF VARIOUS SIMILARITY MEASURES

	x	*y*	M_C	M_H	M_L	M_O	*M'*
1	[0.3, 0.7]	[0.4, 0.6]	1	0.9	0.95	0.9	0.95
2	[0.3, 0.6]	[0.4, 0.7]	0.9	0.9	0.9	0.9	0.9
3	[0.3, 0.8]	[0.4, 0.7]	1	0.9	0.95	0.9	0.948
4	[1, 1]	[0, 1]	0.5	0.5	0.5	0.3	0.6
5	[0.5, 0.5]	[0, 1]	1	0.5	0.75	0.5	0.75
6	[0.4, 0.8]	[0.5, 0.7]	1	0.9	0.95	0.9	0.945
7	[0.4, 0.8]	[0.5, 0.8]	0.95	0.95	0.95	0.929	0.958

From the Table 3, we can see sometimes the similarities gained by formulae of M_C, M_H, M_L and M_O are counterintuitive.

For example, $M_C(x_1, y_1) = M_C([0.3,0.7],[0.4,0.6]) = 1$, apparently we know the similarity of [0.3,0.7] and [0.4,0.6] is absolutely not 1.

Another example, compare the similarity of the first group data pair ([0.3, 0.7], [0.4, 0.6]) with the second group data pair ([0.3, 0.6], [0.4, 0.7]) using the several different similarity measures. We get

$M_H(x_1, y_1) = M_H(x_2, y_2)$, $M_O(x_1, y_1) = M_O(x_2, y_2)$,

whereas intuitively the first data pair should be more similar than the second data pair, namely $M(x_1, y_1) > M(x_2, y_2)$, but only M_L and M' can be accordant with our intuition — $M(x_1, y_1) > M(x_2, y_2)$.

Then compare the similarity of the sixth group data pair ([0.4, 0.8], [0.5, 0.7]) with the seventh group data pair ([0.4, 0.8], [0.5, 0.8]). Intuitively, the similarity of the sixth group and the seventh group should satisfy $M(x_6, y_6) < M(x_7, y_7)$, but from the above result we can see only M_H, M_O, and M' can satisfy the limitation.

To sum up, none but M' can distinguish those groups vague values, to some extend, according with our intuition.

In table 3, we give more comparison of different similarity measures.

In the new similarity measure M', the three factors are considered equally which affect the similarity of vague values: true-membership function t_x, false-membership function f_x, and $1 - t_x - f_x$. The weights of $|t_x - t_y|$, $|f_x - f_y|$ and $|(t_y + f_y) - (t_x + f_x)|$ in new similarity measure M' are variable, and the variable weights reduce the possibility of similarity coincidence. Simultaneously, the new similarity measure M' enlarges the difference of similarities where we are concerned then it is easier for us to do some decision.

B. Similarity measure between Vague Sets

Assume that A and B are two vague sets over the discourse universe $U = \{u_1, u_2, ..., u_n\}$. $V_A(u_i) = [t_A(u_i), 1 - f_A(u_i)]$ is the membership value of u_i in vague set A, and $V_B(u_i) = [t_B(u_i), 1 - f_B(u_i)]$ is the membership value of u_i in vague set B. Let

$$A = \sum_{i=1}^{n} [t_A(u_i), 1 - f_A(u_i)] / u_i$$

$$B = \sum_{i=1}^{n} [t_B(u_i), 1 - f_B(u_i)] / u_i$$

Then, the similarity between vague sets A and B can be obtained by the following function T'.

$$T'(A, B) = \frac{1}{n} \sum_{i=1}^{n} M'(V_A(u_i), V_B(u_i))$$
$$= \frac{1}{n} \sum_{i=1}^{n} \left(1 - \frac{|p(u_i)q(u_i)| + |(2 - m(u_i))n(u_i)| + |q(u_i) - n(u_i)|}{p(u_i) + (2 - m(u_i)) + 2}\right) \quad (7)$$

where,

$$t_x(u_i) + t_y(u_i) = p(u_i)$$

$$t_x(u_i) - t_y(u_i) = q(u_i)$$

$$f_x(u_i) + f_y(u_i) = m(u_i)$$

$$f_x(u_i) - f_y(u_i) = n(u_i)$$

From the above definition, we have the following properties.

Property 5: $T'(A, B) \in [0, 1]$.

Property 6: $T'(A, B) = T'(B, A)$.

Property 7:

$$T'(A, B) = 0 \Leftrightarrow A = \sum_{i=1}^{n} [0,0] / u_i, B = \sum_{i=1}^{n} [1,1] / u_i$$

$$\text{or } A = \sum_{i=1}^{n} [1,1] / u_i, B = \sum_{i=1}^{n} [0,0] / u_i$$

Property 8: $T'(A, B) = 1 \Leftrightarrow A = B$.

Example 2: Let A and B be two vague sets over the discourse universe $U = \{u_1, u_2, u_3, u_4\}$, where

$A = [0.3, 0.7] / u_1 + [0.5, 0.5] / u_2 + [0.4, 0.8] / u_3 + [1.0, 1.0] / u_4$

$B = [0.4, 0.6] / u_1 + [0.0, 1.0] / u_2 + [0.5, 0.7] / u_3 + [0.0, 1.0] / u_4$

From formula (7), we have the following similarity between A and B.

$$T'(A, B) = \frac{1}{4} \sum_{i=1}^{4} M'(V_A(u_i), V_B(u_i))$$
$$= [(1 - 0.05) + (1 - 0.25) + (1 - 0.055) + (1 - 0.4)] / 4$$
$$= 0.811$$

C. Weighted Similarity Measure between Vague Sets

Suppose that A and B are two vague sets over the discourse universe $U = \{u_1, u_2, ..., u_n\}$, w_i is the weight of u_i, $w_i \in [0, 1]$, $1 \le i \le n$. Then, the weighted similarity between A and B can be obtained by calculating the following $W(A, B)$.

$$W'(A, B) = \sum_{i=1}^{n} w_i M'(V_A(u_i), V_B(u_i)) \Big/ \sum_{i=1}^{n} w_i$$
$$= \sum_{i=1}^{n} w_i \left(1 - \frac{|p(u_i)q(u_i)| + |(2 - m(u_i))n(u_i)| + |q(u_i) - n(u_i)|}{p(u_i) + (2 - m(u_i)) + 2}\right) \Big/ \sum_{i=1}^{n} w_i \quad (8)$$

Example 3: Let A and B be the same as that in Example 2, the weights of elements u_1, u_2, u_3, and u_4 in discourse universe U are 0.4, 0.2, 0.8, and 0.6, respectively. From (8), we have the weighted similarity between A and B

$$W'(A, B) = \frac{0.4(1 - 0.05) + 0.2(1 - 0.25) + 0.8(1 - 0.055) + 0.6(1 - 0.4)}{0.4 + 0.2 + 0.8 + 0.6}$$
$$= (0.38 + 0.15 + 0.756 + 0.36) / 2.0$$
$$= 0.823$$

IV. CONCLUSIONS

After analyzing the limitations in current similarity measures for vague sets, we have proposed a new method for measuring the similarity between vague sets in this paper. The basic idea is to deeply understand the support, the difference of true-membership and the difference of false-membership, to significantly distinguish the directions of difference (positive and negative), and properly use varied-weights in the differences of true- and false-membership, for two vague sets. The examples have illustrated that our approach is effective and practical, and presents much better discernibility than existing ones at measuring the similarity between vague sets.

REFERENCES

[1] Zadeh, L A. Fuzzy sets. *Information and Control*, 1965, 8: 338-356.
[2] Gau, W. L. and Buehrer, D. J. Vague sets. *IEEE Transactions on Systems, Man and Cybernetics* (Part B), 1993, 23(2): 610-614.
[3] Chen, S. M. Measures of similarity between vague sets. *Fuzzy Sets and Systems*, 1995, 74(2): 217-223.
[4] Chen, S. M. Similarity measures between vague sets and between elements. *IEEE Transactions on Systems, Man and Cybernetics* (Part B), 1997, 27(1): 153-158.
[5] Hong, D. H. and Kim, C. A note on similarity measures between vague sets and elements. *Information Sciences*, 1999, 115: 83-96.
[6] Li, F. and Xu, Z. Y. Similarity measures between vague sets. *Chinese Journal of Software*, 2001, 12(6): 922-927.
[7] Li, Y. H., Chi, Z. X., and Yan, D. Q. Vague similarity and vague entropy. *Computer Science* (Chinese journal), 2002, 29(12): 129-132.
[8] Li, F. and Xu, Z. Y. Vague sets. *Computer Science* (Chinese journal), 2000, 27(9): 12-14.

XEvent: An Event Notification System over Distributed Hash Table (DHT) Networks

Ruili Wang, Weixiong Rao and Chengqi Zhang

***Abstract*—In this paper, we propose a novel event notification system, named as XEvent. The system is built over a Distributed Hash Table (DHT)-based Peer-to-Peer (P2P) system by combining a content-based system with a specific event topic. XEvent can support basic topic-based message subscription by use of XML schema as event topic, and further filter the whole message contents using XPath as the filtering language. The unique features of XEvent include: (1) XEvent inherits the excellent features provided by DHT, including scalability, adaptation and self-maintenance; (2) XEvent provides the expressive topic-based and content-based event filtering; (3) XEvent can filter and deliver event messages to subscribers with high efficiency by building a delivery tree based on the subscriber's XPath filters, and (4) XEvent can provide fault-tolerance for node failure.**

According to the results of simulation and tests on our XEvent prototype built over a new Bamboo-DHT system, XEvent can achieve the scalability, expressiveness, high efficiency and reliability for the event notification service.

***Index Terms*— Distributed algorithms, Computer networks, Query languages, Message systems**

I. Introduction

An event notification service can be implemented by a centralized event broker or by a network of distributed event brokers. The first case, every client application can act as a publisher, a subscriber, or both, and connect to a single central broker server, which acts simply as a message filtering and routing engine. Obviously, this centralized manner, lack of scalability, results in the problem of single point of failure. The second case, multiple servers can act as event brokers and are cooperatively to form a distributed, coherent message routing, matching and filtering engine for clients. This distributed manner can be implemented on a wide-area network (WAN) and be scalable to a large number of clients.

Recently, a new event filtering model, called a content-based event filtering model [1, 2], allows subscriptions to evaluate the entire content of an event message, not just a traditional event topic. The content-based event system provides a more powerful and flexible event filtering than traditional topic-based event systems [1, 2]. However, what makes content-based event system challenging is that the system must be scalable to a large number of subscribers, publishers and event messages. Also, at the same time, the system must support a flexible and expressive event filtering mechanism, i.e. scalability and flexibility [1].

Distributed Hash Table (DHT) [3, 4, 5] provides unique features for a large scalable distributed system: scalability, adaptation and self-maintenance. Thus, the distributed applications built over DHT can inherit such features provided by DHT including completely decentralized, scalable, and self-organizing; and maintain guaranteed routing efficiency under the joining, departure and failure of nodes in a network. Although some event notification systems [9, 10] built over DHT systems [4, 5] have already been implemented, such systems are topic-based and cannot provide flexibility and expressiveness as they lack a content-based filtering mechanism.

The emergence of XML (eXtensible Markup Language) as a standard for inform4ation exchange on the Internet has led to an increased interest in content-based publish/subscribe system [6, 7, 8]. Using XML as the message format can allow structural information within message documents. Using XPath as an XML query language can provide the filtering on both the structure and the contents of published XML information. Although some XML-based filtering systems like [6, 7, 8] can provide a flexible message filtering and matching mechanism, such systems are all a centralized solution. They lack scalability and cannot be extended to a distributed environment.

In this paper, we propose a novel event notification system, named as XEvent. The system is built over a Peer-to-Peer (P2P) DHT networking by combining a content-based system with a specific event topic. XEvent can support basic topic-based message subscription by use of XML schema for event topic, and further filter the entire message contents using XPath as the filtering language. The unique features of XEvent include: (1) XEvent inherits the excellent features provided by DHT, including scalability, adaptation and self-maintenance; (2) XEvent provides the expressive topic-based and content-based event filtering; (3) XEvent can filter and deliver event messages to subscribers with high efficiency by building a delivery tree based on the subscriber's XPath filters, and (4) XEvent can provide fault-tolerance for node failure. According to the results of simulation and tests on our XEvent prototype built over a new Bamboo-DHT system [14] (developed by Berkeley), XEvent can achieve the scalability, expressiveness and high efficiency and reliability for the type-based event notification service.

R. Wang is with the Institute of Information Sciences and Technology, Massey University, New Zealand, Private Bag 11222, Palmerston North, New Zealand. (corresponding author's telephone: 0064-6-3569099 ext 2548; fax: 0064-6-3502259 e-mail: r.wang@massey.ac.nz).

W. Rao is with the Information System and Service Department, Shanghai General Motor, China 201102. (e-mail: rweixiong@yahoo.com.cn).

C. Zhang is with the Faculty of Information Technology, University of Technology, Sydney, Broadway NSW 2007, Australia. (e-mail: chengqi@it.uts.edu.au).

The remainder of the paper is organized as follows. Section II introduces the basic model of XEvent. Section III gives an overview of XEvent. In section IV we describe the basic operations in XEvent. Our XEvent prototype built on Bamboo[14] and experimental results are shown in section V. Related work is introduced in section VI, and section VII gives the conclusion.

II. Basic Model in XEvent

View: In XEvent, we decompose an XML message of tree structure into multiple sequential elements from the root to all leaves. We call the sequential elements with the element content as Data View V_d, which includes an element path V_{dp} and element content V_{dc}. V_{dp} is the sequential elements from a root to a leaf, and all elements is separated by the parent-child connector '/' to

(a) an XML message	(b) XML view
<SigmodRecord> <issue> <volume>11</volume> <articles> <article> <title>Process synchronization in Database Systems </title> <initPage>9</initPage> <endPage>29</endPage> <authors> <author position="00"> Philip A. Bernstein</author> <author position="01"> Marco A. Casanova</author> </authors> </article> </articles> </issue> </SigmodRecord>	**Element Path**
	/SigmodRecord/issue/volume
	/SigmodRecord/issue/articles/articl
	/SigmodRecord/issue/articles/articl e/initPage
	/SigmodRecord/issue/articles/articl e/ endPage
	/SigmodRecord/issue/articles/articl e/ authors/author
	Element Contents
	volume@text=11
	title@text=Process Synchronization in Database Systems
	initPage @text=9
	endPage @text=11
	author@position="00"@text=Phili p A. Bernstein author@position="01"@text=Mar co A.Casanova

Fig. 1. An XML message and its XML views

show the parent-child relationship between two elements. V_{dc} represents the element contents including the element texts and attributes. Fig. 1 shows an XML message and its V_d.

Similarly with V_d, we can define the schema view (V_s) based on an XML schema, which can be also modeled as a tree-like structure. An XML schema only specifies the structure information without content, while V_s only defines the sequential elements from the root to the leaves as the element paths (V_{sp}). Fig. 2 shows an XML schema (DTD) and view V_s.

XPath Model: we use XPath as the filtering language providing an expressive filtering against XML messages in publication. XPath can support both path structure and element content filtering. The path structure filtering in XPath supports a parent-child operator '/' or an ancestor-descendant operator '//'. Also XPath allows the use of a wildcard operator '*'. XPath also allows element content filtering against both element attributes and element texts. We call the sequential elements in an XPath filter F from the head element of F to the first element followed by a relative path separator in F as a key K of F. For example,

- If F1 = /SigmodRecord/issue/articles/article/title then K1 = F1 = /SigmodRecord/issue/articles/article/title
- If F2 = /SigmodRecord/issue/articles//title, then K2 = /SigmodRecord /issue /articles

We define two operations for the XPath filter F: "+"and "-". If $F_1 = E_1 \wedge \ldots \wedge E_{k-1} \wedge E_k \wedge \ldots \wedge E_n$ and $F_2 = E_k \wedge \ldots \wedge E_n$, then $F_3 = P_1 - F_2 = E_1 \wedge \ldots E_{k-1}$ and $F_3 + F_2 = F_1$. Also, if $F_1 = E_1 \wedge \ldots \wedge E_{k-1} \wedge E_k \wedge \ldots \wedge E_n$ and $F_3 = E_1 \wedge \ldots \wedge E_{k-1}$, then $F_4 = P_1 - F_3 = E_k \wedge \ldots \wedge E_n$ and $F_4 + F_3 = F_1$.

A covering relation between two XPath filters is defined that if F_1 is more specific than F_2, then F_2 covers F_1. The relation is also written as $F_2 \supseteq F_1$ or $F_1 \subseteq F_2$, *if and only if* both the path structure filtering P_2 and the element content filtering C_2 of F_2 covers the path structure filtering P_1 and the content filtering C_1 of F_1. Based on the XPath covering relationship, we can construct an XPath covering graph. In the graph, a node represents an XPath filter and an edge represents the covering relationship between XPath filters. If $F_1 \subsetneq F_2$, the node representing F_1 is the parent of the node representing F_2. As a result, the most specific filter is the root of the graph, and the least specific filter is the leaf node of the graph. Given the filters shown in Fig. 3(b), we know the following XPath covering relations, and the XPath covering relation graph is constructed as Fig. 3(a):

- $F_1 \subseteq F_3 \subseteq F_5 \subseteq F_2 \subseteq F_6 \subseteq F_7 \subseteq F_8$;
- $F_4 \subseteq F_6 \subseteq F_7 \subseteq F_8$.

III. Overview of the XEvent System

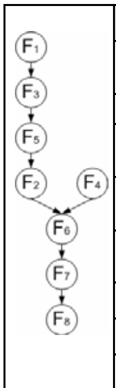
In XEvent, we define the XML topic (schema) as the event topic. The event messages in XML format are instances of the

(a) an XML DTD	(b) view
<!ELEMENT SigmodRecord (issue)* > <!ELEMENT issue (volume,number,articles) > <!ELEMENT volume (#PCDATA)> <!ELEMENT number (#PCDATA)> <!ELEMENT articles (article)* > <!ELEMENTarticle(title,initPage,endPage,aut hors)> <!ELEMENT title (#PCDATA)> <!ELEMENT initPage (#PCDATA)> <!ELEMENT endPage (#PCDATA)> <!ELEMENT authors (author)* > <!ELEMENT author (#PCDATA)> <!ATTLIST author position CDATA #IMPLIED>	/SigmodRecord/issue/vo me
	/SigmodRecord/issue/nu ber
	/SigmodRecord/issue/art les/ article/ title
	/SigmodRecord/issue/art les/article/ initPage
	/SigmodRecord/issue/art les/article/ endPage
	/SigmodRecord/issue/art les/article/ authors/author

Fig. 2. Event schema (DTD) and element path of schema views

event schema. XEvent supports subscriptions depending on both the event topic and event content filtering. The introduction of event schema into the event system can be useful for both publishers and subscribers. The schema can validate an XML message and enforce the validity of XML messages; Event subscribers can directly subscribe a certain topic of event messages based on an XML schema, and further subscribe the message using XPath as the filtering language.

We build XEvent (an event notification system) over DHT by arranging a set of cooperating peers in a distributed topology. Each peer in XEvent can act as three roles: publisher, subscriber or



Filter	XPath
F1	/SigmodRecord/issue/articles/article/authors/author [@position="00"]
F2	/SigmodRecord//author[@position = "00"]
F3	/SigmodRecord/issue/articles/article/*/author[@position = "00"]
F4	/SigmodRecord/issue/articles/article/authors/author [@position>"00"]
F5	/SigmodRecord/issue/articles/article//author [@position = "00"]
F6	/SigmodRecord//author[@position >= "00"]
F7	/SigmodRecord//author
F8	//author

(a) (b)

Fig. 3. XPath Relation Graph

broker. The basic operations in XEvent include: schema registration, event subscription and message publication. These operations are the three steps for an event message to be delivered from a publisher to subscribers. Firstly, the registration of an XML schema as event topic into XEvent is mandatory. Secondly, the subscribers can subscribe an interested event by use of the event schema as event topic and an XPath filter as content filter. Thirdly, when a message producer publishes an XML message after the validation of the XML schema, the event message is filtered over XEvent and forwarded by intermediate peers until the message is delivered to its interested subscribers.

In XEvent, we build two layers of overlay network over the IP network (Fig. 4), i.e. event sketch tree and delivery tree. The event sketch tree (described in section 4.1) is built based on the event schema (topic) view. The event delivery tree (described in section 4.2) is organized logically based on the XPath relation graph while physically within the event sketch tree. Using the delivery tree, the event is directly forwarded to its most specific and matched XPath filter node in the event delivery tree (described in section 4.3). If the message matches the most specific XPath filter, all descendant filters of the most specific filter are also matched with the message. Such a delivery tree can greatly reduce the overhead of message routing and matching. Concerning the reliability, an event sketch tree can self-heal its node failure and provide the reliability for the upper event delivery tree (described in section 4.4), and the event delivery tree can devote itself to message filtering and delivery without care of node maintenance.

IV. Basic Operations in XEvent

In this section, we give the description of XEvent operations: schema registration, event subscription, message publication and maintenance of event delivery tree.

A. Schema Registration

During schema registration, an event skeleton tree $\boldsymbol{T}_k$ is built. The event topic (schema) $\boldsymbol{S}$, which is a tree structure, can generate views $\boldsymbol{V}_s$ as defined in section 2. The destination host of Hash($\boldsymbol{S}$) in DHT namespace is termed as $\boldsymbol{H}_t$. As the definition of XPath key in section 2, we term $\boldsymbol{F}_{root}$ as the element paths $\boldsymbol{V}_{sp}$, and $\boldsymbol{K}_{root}$ as the key of $\boldsymbol{V}_{sp}$. As a result, $\boldsymbol{K}_{root} = \boldsymbol{F}_{root} = \boldsymbol{E}_1/\ldots/\boldsymbol{E}_n$ since all path separators in $\boldsymbol{F}_{root} = \boldsymbol{E}_1/\ldots/\boldsymbol{E}_n$ are '/'. For an event schema $\boldsymbol{S}$ in XEvent and the element paths $\boldsymbol{F}_{root}$ of $\boldsymbol{V}_s$, we introduce the destination of Hash($\boldsymbol{K}_{root}$) in DHT namespace as the root node of $\boldsymbol{T}_k$, termed as $\boldsymbol{H}_{root}$.

When building $\boldsymbol{T}_k$, $\boldsymbol{F}_{root}$ can be further decomposed into multiple sub-filters $\boldsymbol{F}_s$ and each of $\boldsymbol{F}_s$ is the sequential elements from the head element of $\boldsymbol{F}_{root}$ to other elements of **Froot**. For $\boldsymbol{F}_{root} = \boldsymbol{E}_1/\ldots/\boldsymbol{E}_n$, the sub-filters $\boldsymbol{F}_s$ can be $\boldsymbol{E}_1/\ldots/\boldsymbol{E}_{n-1}/\boldsymbol{E}_n$, $\boldsymbol{E}_1/\ldots/\boldsymbol{E}_{n-1}$, …, $\boldsymbol{E}_1/\boldsymbol{E}_2$, $\boldsymbol{E}_1$. For each sub-filter $\boldsymbol{F}_s$, the key $\boldsymbol{K}$ of $\boldsymbol{F}_s$ = $\boldsymbol{F}_s$ because all elements in sub-filter $\boldsymbol{F}_s$ are connected by '/'. Then the skeleton tree $\boldsymbol{T}_k$, rooted at $\boldsymbol{F}_{root}$, is built based on the filter covering the relation graph of all sub-filters $\boldsymbol{F}_s$. Each sub-filter $\boldsymbol{F}_s$ is physically located the destination host of Hash ($\boldsymbol{K}$) in the DHT namespace. The connection of two sub-filters represents the covering relationship of two sub-filters, the parent sub-filter is more specific than its children, and the root sub-filter is the most specific. Fig. 5 shows the event sketch tree and its sub-filters of the event schema (topic) in Fig. 2.

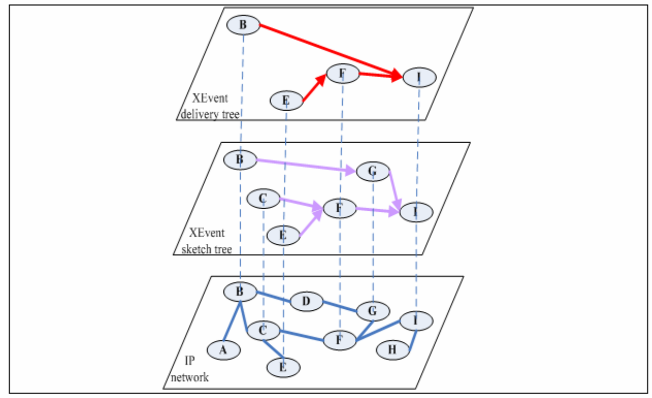


Fig. 4. XEvent Overlay Network

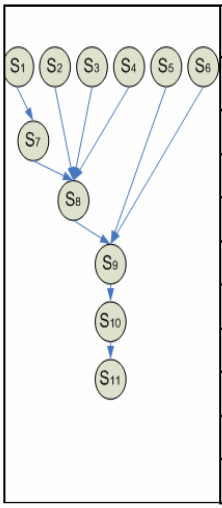
TABLE I
XPath Instantiation

XPath Filter	XPath instantiation Result
F1=//artitles//title	/SigmodRecord/issue/articles// title
F2=/*/issue//article/*/author	/SigmodRecord/issue//article/*/ author

B. Event Subscription

In XEvent, event subscribers can subscribe their interested message by an event schema (topic) and by an XPath filter. As the topic subscription is simple, we only discuss the XPath subscription in this paper. The XPath subscription can be implemented in two phases: XPath instantiation and event delivery tree building.

XPath Instantiation: we have defined the absolute path separator ('/'), and the relative path separator ('/*/' and '//'). Depending on whether an absolute path separator '/' at the head of an XPath filter or not, we category an XPath filter $\boldsymbol{F}$ into $\boldsymbol{F}_a$ or $\boldsymbol{F}_r$, where $\boldsymbol{F}_a$ has an absolute path separator '/' at the head of the XPath filter, and $\boldsymbol{F}_r$ has a relative path separator '//' or '/*/' at the head of the XPath filter. For example, *//author* or */*/issue/articles/*



S1	/SigmodRecord/issue/articles/article/authors/author
S2	/SigmodRecord/issue/articles/article/endPage
S3	/SigmodRecord/issue/articles/article/initPage
S4	/SigmodRecord/issue/articles/article/title
S5	/SigmodRecord/issue/number
S6	/SigmodRecord/issue/volume
S7	/SigmodRecord/issue/articles/article/authors
S8	/SigmodRecord/issue/articles/article
S9	/SigmodRecord/issue/articles
S10	/SigmodRecord/issue
S11	/SigmodRecord

Fig. 5. Event sketch tree and its sub-Filters

are $\boldsymbol{F}_r$. $\boldsymbol{F}_r$ can decompose into $\boldsymbol{F}_r$ except the relative path separator at the head, i.e. $\boldsymbol{F}_r = \{\wedge \boldsymbol{F}_a\}$ where $\wedge$ = '/*/' or '//'. The XPath instantiation actually replaces a relative path separator at the head of $\boldsymbol{F}_r$ into an absolute path separators, i.e. replace the '/*/' or '//' with '/'. For these relative path separators, which are located at the intermediate or the end part of an XPath filter, like */SigmodRecord//author* or */SigmodRecord/issue/articles/article/*/author*, no instantiation is needed.

XPath instantiation can be done as follows. Given an even schema $\boldsymbol{S}$ and XPath filter $\boldsymbol{F}_r = \{\wedge \boldsymbol{F}_a\}$ where $\wedge$ = '/*/' or '//', the instantiation result of $\boldsymbol{F}_r$ against $\boldsymbol{S}$, termed as $\boldsymbol{F}_i$, is to replaced $\boldsymbol{F}_r$ with $\boldsymbol{F}_i = \{\boldsymbol{F}_s \wedge \boldsymbol{F}_a\}$ where $\wedge$ = '/', if there exists one element path $\boldsymbol{V}_{ep}$ is satisfied with $\boldsymbol{F}_r$ and a sub-filter $\boldsymbol{F}_s$ of $\boldsymbol{V}_{ep}$ where $(\boldsymbol{V}_{ep} - \boldsymbol{S})^*$ is satisfied with $\boldsymbol{F}_a$. XPath instantiation can be implemented in the similar way as [6] by use of XPath as Finite State Machines (FSM) and $\boldsymbol{V}_{ep}$ as the message. Table 1 shows two XPath filter, and the result of XPath instantiation against the event schema in Fig. 2.

Building event delivery tree: When an XPath filter $\boldsymbol{F}$ is submitted to XEvent by a subscriber, the XPath filter $\boldsymbol{F}$ is first checked that $\boldsymbol{F}$ is $\boldsymbol{F}_a$ or $\boldsymbol{F}_r$. If $\boldsymbol{F}$ is $\boldsymbol{F}_a$, $\boldsymbol{F}$ is directly used to build an event delivery tree; otherwise $\boldsymbol{F}$ needs to instantiate locally. After instantiation, the XPath instantiated result $\boldsymbol{F}_i$ is used to build an event delivery tree.

The key of building an event delivery tree is to make use of the fact: for event schema $\boldsymbol{S}$, $\boldsymbol{F}_{root}$ of $\boldsymbol{S}$, and an XPath filter $\boldsymbol{F}$ ($\boldsymbol{F}_i$ or $\boldsymbol{F}_a$), there always exists a sub-filter $\boldsymbol{F}_s$ of $\boldsymbol{F}_{root}$, i.e. a sub-filter $\boldsymbol{F}_s$ is XPath key $\boldsymbol{K}$ of $\boldsymbol{F}$ ($\boldsymbol{K} = \boldsymbol{F}_s$) if $\boldsymbol{F}$ is satisfied with the XML event message that is instance of $\boldsymbol{S}$. For such an XPath filter $\boldsymbol{F}$, the destination of Hash($\boldsymbol{K}$) in DHT namespace, as a result, is the same as the destination of Hash($\boldsymbol{F}_s$) where is located at the event sketch tree. The event delivery tree $\boldsymbol{T}_d$ for an event topic (schema) $\boldsymbol{S}$ consists up the XPath filters for the event topic $\boldsymbol{S}$, which are organized based on the XPath filter relation graph. As a result, $\boldsymbol{T}_d$, logically organized by an XPath filter covering relation graph, is physically located within $\boldsymbol{T}_k$. For a new XPath filter $\boldsymbol{F}$, the subscription into XEvent can be done in both folders: physically inserted into $\boldsymbol{T}_k$ and logically inserted into $\boldsymbol{T}_d$. For the former, it can be easily implemented by locating the destination host of Hash($\boldsymbol{K}$) since $\boldsymbol{K} = \boldsymbol{F}_s$; for the latter, it can be implemented by finding $\boldsymbol{F}$'s parent (child) filter at $\boldsymbol{T}_d$, termed as $\boldsymbol{F}_p$ ($\boldsymbol{F}_c$), and inserting $\boldsymbol{F}$ between $\boldsymbol{F}_p$ and $\boldsymbol{F}_c$.

The process of finding $\boldsymbol{F}_p$ is described as follows. Given an XPath filter $\boldsymbol{F}$, the destination $\boldsymbol{H}_k$ of Hash($\boldsymbol{K}$) in DHT is located at $\boldsymbol{T}_k$. Based on the definition of XPath covering relation and the building of $\boldsymbol{T}_k$, $\boldsymbol{F}_p$ is physically located at $\boldsymbol{H}_k$, or at parent node of $\boldsymbol{H}_k$ in $\boldsymbol{T}_k$, or at ancestor node of of $\boldsymbol{H}_k$ in $\boldsymbol{T}_k$, and impossible at children nodes of $\boldsymbol{H}_k$ in $\boldsymbol{T}_k$. Then $\boldsymbol{F}_p$ can be found by first traversing all existing XPath filters at $\boldsymbol{H}_k$ to determine the least specific XPath filter as $\boldsymbol{F}_p$. If no XPath filters exist at $\boldsymbol{H}_k$, then the same procedure is done to find $\boldsymbol{F}_p$ at parent node of $\boldsymbol{H}_k$, …, ancestor node of $\boldsymbol{H}_k$ until the root of $\boldsymbol{T}_k$. If no XPath filters exist at the parent node of $\boldsymbol{H}_k$, ancestor node of $\boldsymbol{H}_k$ and the root of the event skeleton tree $\boldsymbol{T}_k$, $\boldsymbol{F}_p$ returns null, which means $\boldsymbol{F}$ being the most specific filter and becoming the new root of $\boldsymbol{T}_k$. Finding $\boldsymbol{F}_c$ can be done in the similar way as finding $\boldsymbol{F}_c$. After insert, the link between $\boldsymbol{F}$ and $\boldsymbol{F}_p$ is established. If both $\boldsymbol{F}$ and $\boldsymbol{F}_p$ are located at the same host, the link within the local host is actually a pointer from $\boldsymbol{F}_p$ to $\boldsymbol{F}$; otherwise, a physical link between different hosts is needed. Since multiple XPath filters share the same XPath filter key, these XPath filters with the same key are located at one node in $\boldsymbol{T}_k$. A physical link can contain multiple links between XPath filters. As a result a message (for example a Heartbeat message) from a parent node to a child node in $\boldsymbol{T}_k$ can aggregate the information of multiple XPath filter links to decrease the communication traffic.

In order to improve the performance of message delivery, we introduce a repair mechanism for the event delivery tree $\boldsymbol{T}_d$ (see section 4.4). Since there exists the case that the root XPath filter of $\boldsymbol{T}_d$ is not located at the root node of $\boldsymbol{T}_k$, in the repair mechanism a virtual XPath filter located at the root node of $\boldsymbol{T}_k$ is used to link the current root XPath filter of $\boldsymbol{T}_d$; Otherwise, the repair mechanism is not needed when the current root XPath filter of $\boldsymbol{T}_d$ is not located at the root node of $\boldsymbol{T}_k$.

Concerning the efficiency of finding $\boldsymbol{F}_p$, the IP hops of finding $\boldsymbol{F}_p$ ($\boldsymbol{F}_c$) is limited within the depth of $\boldsymbol{T}_d$ not the number of XPath filters, which can greatly decrease the communication cost. Thus, the cost of finding the least specific filter in a local host is negligible compared with the communication cost in DHT-based P2P system. Fig. 6(a) shows the XPath filter covering relation graph after instantiation of the XPath filters in Fig. 3(b), and Fig. 6(b) shows the event delivery tree $\boldsymbol{T}_d$ and its physical location in $\boldsymbol{T}_k$. Here note: (1) $\boldsymbol{F}_8$ (*//author*) after XPath instantiation becomes */SigmodRecord/issue/articles/article/authors/author*. Then the new XPath covering relation graph in Fig. 6(a) is twisted from the one shown in Fig. 3(a). (2) The filters in $\boldsymbol{T}_d$ are located within $\boldsymbol{T}_k$, but the physical link between filters in $\boldsymbol{T}_d$ does not follow the physical link in $\boldsymbol{T}_k$. $\boldsymbol{S}_1$ and $\boldsymbol{S}_8$ in $\boldsymbol{T}_d$ is directly linked not via $\boldsymbol{S}_7$.

C. *Message Publication*

In XEvent, only the valid XML message against the event topic (schema) can be published into XEvent. A valid XML message

* About definition of -, see the XPath model in section 2

can be decomposed into XML views ***V***d with element contents ***V***dc and element path ***V***dp. For ***V***d, the destination of Hash(***V***dp) in the DHT is a root node ***H***root of the event sketch tree ***T***k for the event schema as described in section 4.1. Based on the repair mechanism of the event delivery tree ***T***d, there is always a root XPath filter ***F***r at ***H***root. If ***F***r is virtual, the incoming ***V***d is delivered to its child XPath filter for further message filtering. Otherwise, ***V***d is matched with ***F***r. If ***V***d is satisfied with the ***F***r, ***V***d is also satisfied with all of its children. As a result, the XML message can be directly routed to the subscribers who subscribe ***F***r and all children of ***F***r. If the XML view ***V***d is unsatisfied with ***F***r, the traversing the event delivery tree from ***F***r to its children is needed until the XML view ***V***d is satisfied with an XPath filter ***F*** in ***T***d. Then, we can determine that the XPath filter ***F*** and its all children filters in ***T***d are interested in the XML message. Also, the XML message is delivered to the subscribers who subscribe ***F*** and its children filters. If an XPath filter ***F*** receives a previous duplicated ***V***d, the XML view ***V***d will not be forwarded and it will be discarded. In section 4.2, we introduce the repair mechanism of ***T***d. We argue that if a virtual XPath filter exists at the root of ***T***k, the matching can directly be forwarded to its child filter without traversing the child node to find the root filter of ***T***d.

About efficiency of message publication, since XPath filters with the same key are located at the same node of ***T***k, the forwarding cost within the local node is negligible compared with the forwarding cost between physical nodes. Also, the forwarding hops are limited within the depth of ***T***k not of ***T***d, which can greatly decrease the communication cost.

D. Maintenance of an Event Delivery Tree

The requirement of reliability needs XEvent to be robust against node failure. To repair the node failure in ***T***k, a node in ***T***k is required to periodically exchange Heartbeat messages with its parent/child nodes to detect failure. Each XPath filter in the node has a soft state with leases for its parent/child XPath filter.

We describe failures and their repair mechanisms as follows. Firstly, failure of ***H***t can be detected by the Heartbeat message from the root node ***H***root in the event sketch tree ***T***k, but the failure cannot cause any hamper of message publication or event subscription except for those XPath filters which need instantiation at ***H***t. The failed ***H***t can be healed by any publisher, subscriber or third-party to re-register the event topic (schema) while without re-building the entire event delivery tree ***T***d. Secondly, failure of ***H***root in ***T***k can be detected by the Heartbeat message from ***H***t, and can be healed by ***H***t to re-lookup the destination host of Hash(***F***root) at DHT as the new ***H***root. Thirdly, failure of any other node in ***T***k except ***H***root can be detected by its parent node. The parent node, destination of Hash(***E***1/…/***E***k-1/***E***k) in DHT, can re-lookup the destination of Hashing (***E***1/…/***E***k-1) to heal the failed child.

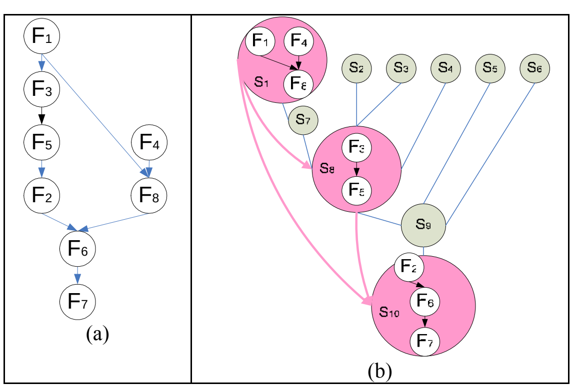


Fig. 6. Event Delivery Tree

Even before the node in ***T***k is recovered, XEvent provides a reliable message forwarding mechanism. The XML view ***V***d can be first routed to the children node of ***T***k, not the failure node where its children XPath filter is located. This begins the message matching against the XPath filters at the children nodes.

V. Experimental Evaluation

In order to evaluate XEvent performance, we implement the prototypes of basic SCRIBE [9] and XEvent using the discrete event simulator over Bamboo [14], a recent DHT-based P2P system. The simulation environment is a network topology with 5050 routers, which were generated by the Georgia Tech random graph generator using the transit-stub model [15]. The prototype codes including SCRIBE and XEvent run on 100,000 end nodes that were randomly assigned to routers in the core with uniform probability. Each end system was directly attached by a LAN link to its assigned router. During the evaluation, we designed the following test scenario: (1) *SigmodRecord.dtd.* [16] as an event topic (schema) is registered into XEvent; (2) 4096 XPath filters are generated by the IBM XPath query generator[17] to subscribe their interested messages; (3) XML-based event messages, which are validate with the XML event topic (schema), are published into XEvent continuously as a message stream. Finally, XEvent routes the published messages to all satisfied subscribers. From the time that the messages are published to the time that all satisfied subscribers get the messages, we evaluate XEvent and SCRIBE. For prototype implementation of SCRIBE, we filter the coming message content at the end host of SCRIBE, which is the destination of Hash (XPath key).

To facilitate our comparison, we use two matrices: Relative Delay Penalty (RDP) and physical link stress, which were originally proposed in [13] and were sequentially used in [10]. In addition to these two matrices, stability is also evaluated.

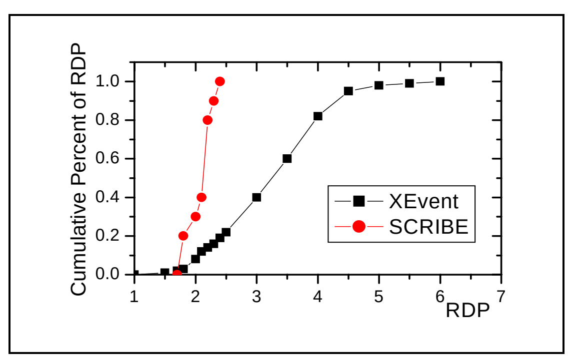


Fig. 7. The cumulative distribution of Relative Delay Penalty (RDP) of both XEvent and SCRIBE.

Relative Delay Penalty (RDP) is a measure of the increase in delay that applications incur while using overlay routing [13, 10]. Using RDP on both XEvent and SCRIBE [10], it is the ratio of DHT overlay unicast routing distances to IP unicast routing distances. Assuming symmetric routing, IP Multicast and naive unicast both have a RDP of 1. We use the ratio of the amount of total overlay DHT unicast routing hops to the amount of all naïve unicast hops as the RDP. Fig. 7 shows the cumulative distribution of RDP of both XEvent and SCRIBE. The X-axis represents the RDP value and y-axis represents the cumulative percent of RDP. XEvent has an average RDP value of 1.9 and SCRIBE has the value of 4.2. It can be explained that all communication in XEvent is within the event sketch tree, while the communication in SCRIBE is scattered through out the whole network.

Link Stress is a measure of how effective SCRIBE and XEvent prototypes are in a distributed network load across different physical links. It refers to the number of identical copies of a packet carried by a physical link. IP multicast has a stress of 1, and naïve unicast has a worst case stress equal to number of receivers. We calculate the ratio of the total number of messages that are sent over links to the total number of links as the average link stress. Fig. 8 shows the link stress comparison of XEvent and Scribe. The X-axis represents the Log scale of Link stress, i.e. the duplicate message copies across the physical link. The Y-axis represents the Log scale of link number. Obviously, XEvent has much less link stress than SCRIBE. It can be easily explained that the link between the nodes in XEvent event sketch tree can contain multiple logical links between XPath filters of XEvent event delivery tree. Also the logical link of XPath filters within the same local host does not require any physical communication.

Stability: To evaluate the stability of XEvent, we need to design an unstable scenario by re-designing step 3 in normal scenarios. We separately design 10%, 20% and 30% nodes crashed in the network model from normal message publication in the time interval of 1000 seconds, then we count the ratio of the nodes, which can receive the message, to all live nodes. In Fig. 9, X-axis represents time and Y-axis represents the ratio in percent. From the figure, the time interval for XEvent (with the self-healing mechanism) to recover 100% message receiving is shorter than for SCRIBE. Also, the failure makes the data lost of some nodes in SCRIBE since we simply implement the basic SCRIBE without the group maintenance mechanism.

VI. Related Works

Recent work on the scalable design of structured P2P overlay networks has introduced a new class of structured networks called Distributed Hash Tables (DHT). Well known representatives include [3, 4, 5]. All of these systems were built to allow efficient key lookups. Nodes in a P2P network send messages to each other based on a unique name, generated from a secure one-way hash of some unique string. Assigned a unique ID to each node by DHT, a message is delivered to the destination host in a fault-resilient fashion. The P2P routing and locating infrastructure can provide scalability, fault-tolerance, self-maintenance and adaptation for upper applications.

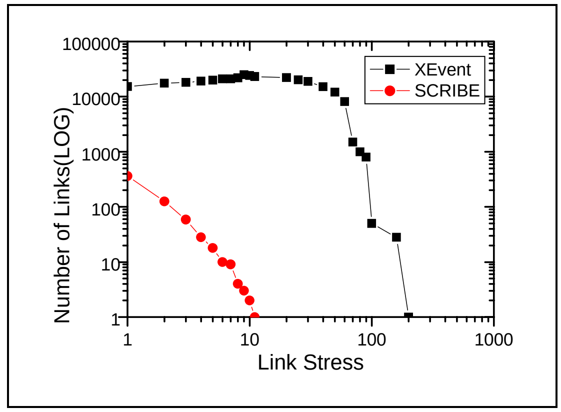


Fig. 8. The link stress comparison of XEvent and Scribe.

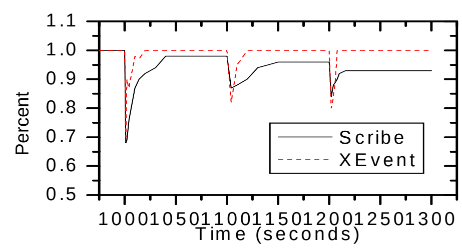


Fig. 9. XEvent stability

Several event notification systems have already developed over DHT. SCRIBE [9] is a topic-based event notification system that is built on top of "Pastry" [4]. "Pastry" is a DHT system developed at Rice University. Subscribers join topics of interest, where each topic is identified with a Pastry-level key. Bayeux [10] is built on top of Tapestry [5]. However, neither SCRIBE [9] nor Bayeux [10] supports content-based subscriptions and publications. Hermes [11] provides a type- and attribute-based routing schema over DHT, but the introduction the rendezvous node into Hermes can result in limited scalability and a single point of failure, also the content filtering is rather limited. An extension of traditional Bloom filters, called multi-level Bloom filters [12], can be used to route path queries in a P2P system, and build content-based overlay networks by linking together peers with similar content. However, the P2P system is totally different from the DHT-based structured P2P systems. Therefore, the system will not or cannot implement a flexible filtering mechanism like XEvent.

Some pioneer researchers in database field have incorporated XML into publish /subscribe system [6, 8]. An XML document filtering system, for Selective Dissemination of Information (SDI) has been proposed in [6], which allows users to define their interests using XPath query language. Using a modified Finite State Machine (FSM), a novel indexing mechanism and matching algorithms are developed. YFilter, combines all of the queries into a single Non-deterministic Finite Automaton (NFA) [8]. The approach exploits commonality among path expressions by merging common prefixes of the query paths so that they are processed at once. All of these approaches are based a single centralized filtering engine and are hard to scale to a large distributed network.

TABLE II
ABBREVIATIONS LIST

$\boldsymbol{V}_{\mathrm{d}}$	XML Data View
$\boldsymbol{V}_{\mathrm{dp}}$	Element path in $\boldsymbol{V}_{\mathrm{d}}$
$\boldsymbol{V}_{\mathrm{dc}}$	Element content in $\boldsymbol{V}_{\mathrm{d}}$
$\boldsymbol{V}_{\mathrm{s}}$	Schema View
$\boldsymbol{V}_{\mathrm{sp}}$	Element path in $\boldsymbol{V}_{\mathrm{s}}$
$\boldsymbol{F}$	XPath filter
$\boldsymbol{S}$	Event schema (topic)
$\boldsymbol{K}$	Key of $\boldsymbol{F}$
$\boldsymbol{F}_{\mathrm{root}}$	Element paths $\boldsymbol{V}_{\mathrm{sp}}$
$\boldsymbol{K}_{\mathrm{root}}$	Key of $\boldsymbol{F}_{\mathrm{root}}$
$\boldsymbol{H}_{\mathrm{root}}$	Destination Host of $\boldsymbol{F}_{\mathrm{root}}$ in DHT namespace
$\boldsymbol{T}_{\mathrm{k}}$	Event sketch tree
$\boldsymbol{T}_{\mathrm{d}}$	Event delivery tree
$\boldsymbol{F}_{\mathrm{s}}$	XPath Sub-Filter of $\boldsymbol{V}_{\mathrm{sp}}$
$\boldsymbol{F}_{\mathrm{a}}$	XPath Absolute Filter
$\boldsymbol{F}_{\mathrm{r}}$	XPath Relative Filter
$\boldsymbol{F}_{\mathrm{i}}$	XPath Instantiation result
$\boldsymbol{F}_{\mathrm{p}}$	XPath Parent filter in $\boldsymbol{T}$d
$\boldsymbol{F}_{\mathrm{c}}$	XPath child filter in $\boldsymbol{T}$d
$\boldsymbol{H}_{\mathrm{t}}$	Destination Host of Hash($\boldsymbol{S}$) in DHT namespace

VII. CONCLUSION

We have presented XEvent, a novel hybrid event notification system built over Distributed Hash Table (DHT)-based Peer-to-Peer (P2P) system. XEvent can support both event topic and event content subscription, which can provide expressiveness for subscribers by the use of XML as event message format and use of XPath as subscription filter.

XEvent make use of DHT to build two layers of an overlay network: event sketch tree and event delivery tree. An event sketch tree can provide the stability for the upper event delivery tree. The event delivery tree can devote itself to swiftly match and deliver messages to the interested subscribers, without considering the churn in DHT.

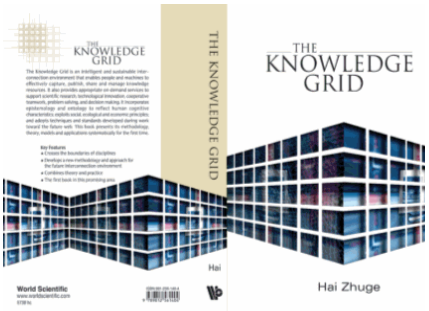
The initial simulation results indicate that XEvent can provide an efficient event notification system with quick recovery mechanisms and low Relative Delay Penalty and link stress compared with the SCRIBE [9] system.

REFERENCES

[1] A. Carzaniga, D. S. Rosenblum, and A. L. Wolf, "Content-based addressing and routing: A general model and its application," Department of Computer Science, University of Colorado, USA, Technical Report CU-CS-902-00, Jan. 2000.
[2] R. Chand and P. Felber, "A scalable protocol for content-based routing in overlay networks," in *IEEE International Symposium on Network Computing and Applications (NCA'03),* Cambridge, Massachusetts, Apr. 2003.
[3] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for Internet applications," in *Proceedings of ACM SIGCOMM'01*, San Diego, CA, Aug. 2001.
[4] A. Rowstron and P. Druschel, "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems," in *MiddleWare*, Germany, Nov. 2001.
[5] B. Y. Zhao, J. D. Kubiatowicz, and A. D. Joseph, "Tapestry: An infrastructure for fault-tolerant wide-area location and routing," University of California Berkeley, USA, Technical Report CSD-01-1141, 2001.
[6] M. Altinel and M. J. Franklin, "Efficient Filtering of XML Documents for Selective Dissemination of Information," in *Proceedings of the 26th International Conference on Very Large Data Bases,* Cairo, 2000, pp 53–64.
[7] A. K. Gupta and D. Suciu, "Stream processing of XPath queries with predicates," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, California, USA, 2003, pp. 419–430.
[8] Y. Diao and M. Franklin, "Query Processing for High-Volume XML Message Brokering," in *Proceedings of the 29th International Conference on Very Large Data Bases*, Berlin, Germany, 2003, pp. 261–272.
[9] M. Castro, P. Druschel, A. Kermarrec, and A. Rowstron, "SCRIBE: A large-scale and decentralized application-level multicast infrastructure," *IEEE Journal on Selected Areas in Communications,* vol. 20, no. 8, pp. 1489–1499, Oct. 2002.
[10] S. Q. Zhuang, B. Y. Zhao, A. D. Joseph, R. H. Katz, and J. Kubiatowicz, "Bayeux: An Architecture for Scalable and Fault-tolerant WideArea Data Dissemination," in *Proceedings of the 11th International Workshop on Network and Operating System Support for Digital Audio and Video,* USA, June 2001.
[11] P. R. Pietzuch and J. Bacon, "Peer-to-Peer Overlay Broker Networks in an Event-Based Middleware," in *Proceedings of the 2nd International Workshop on Distributed Event-based Systems,* USA, 2003.
[12] G. Koloniari, Y. Petrakis and E. Pitoura, "Content-Based Overlay Networks of XML Peers Based on Multi-Level Bloom Filters," in *International Workshop on Databases, Information Systems and Peer-to-Peer Computing,* Germany, 2003.
[13] Y. Chu, S. G. Rao, and H. Zhang, "A case for end system multicast (keynote address)," in *Proceedings of the 2000 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, USA, 2000.
[14] S. Rhea, D. Geels, T. Roscoe, and J. Kubiatowicz, "Handling Churn in a DHT," in *Proceedings of the USENIX Annual Technical Conference (USENIX),* Boston, Massachusetts, June 2004.
[15] Georgia tech internet topology model. http://www.cc.gatech.edu/fac/Ellen.Zegura/graphs.html
[16] http://www.cs.washington.edu/research/xmldatasets/
[17] http://www.alphaworks.ibm.com/tech/xmlgenerator

THE KNOWLEDGE GRID

By Hai Zhuge, World Scientific, 2004, ISBN: 981-256-140-4



Reviewed by Erlin Yao[1]

The recent booming interest in Knowledge Grid is just there. It can at least be revealed by the tens of thousands of related entries that Google.com returns based on these two buzzwords. If the ultimate goal of Grid computing is to facilitate distributed computing power access in a seemless manner, that of Knowledge Grid is to mobilize the distributed knowledge for effective reuse. “The Knowledge Grid”, written by professor Hai Zhuge -- the founder of the China Knowledge Grid Research Group (http://kg.ict.ac.cn), is by far the only academic monograph dedicated to this promising research area. This monograph, for the first time, provides a systematic presentation of the underlying methodologies, theories, models and applications of the Knowledge Grid.

Instead of being yet another book on one more form of Grid middleware, “The Knowledge Grid” features three distinct writing objectives: to bridge the boundaries of disciplines related to Knowledge Grid, to provide methodologies and approaches for developing future interconnection environment, and to integrate theory with practice. System architects, researchers and academics who are interested in knowledge-based systems and Grid computing should find this monograph a valuable and unique reference. This monograph also serves well an overview of the state of the art of the Knowledge Grid research, suitable particularly for researchers. If you want to find out if the Knowledge Grid is the killer methodology for solving knowledge sharing and integration problems or to see whether there are already better solutions on the horizon, “The Knowledge Grid”, altogether 264 pages, is a book written for providing some hints for the answers. However, if you are looking for technical details for building immediately Grid applications, you are better off with other books written around the Globus.

“The Knowledge Grid” is organized in four parts. The first part describes the methodology of the Knowledge Grid. The author systematically defines the Knowledge Grid, for the first time, as an intelligent and sustainable interconnection environment that enables people and machines to effectively capture, publish, share and manage knowledge resources. It also provides appropriate on-demand services to support scientific research, technological innovation, cooperative teamwork, problem-solving, and decision-making. It incorporates epistemology and ontology to reflect human cognitive characteristics; exploits social, ecological and economic principles; and adopts techniques and standards developed during work toward the future interconnection environment. The author also proposes the challenges and opportunities for Knowledge Grid research.

The second part of the book focuses on the Semantic Link Network and Resource Space Model — the two models proposed by the author for implementing a rich semantic layer for the next generation interconnection environment. The Semantic Link Network (SLN) model is a conceptual model to support the semantic overlay, where resources are not linked simply by hyperlinks but semantic links for enabling a particular form of semantic reasoning. The author starts from a metaphor of geographical mapping in organizing and locating resources, and then defines a Resource Space Model (RSM) as a semantic data model for uniformly, normally and effectively specifying and managing resources. RSM essentially organizes versatile resources according to a formally defined semantic normal form so that (1) one or a set of resources can be identified given a set of resource space coordinates, and (2) constraints, operations, and several other normal forms can be defined for the guarantee of the correctness of the corresponding deduction process. The combination of the Semantic Link Network and the Resource Space Model forms a unified semantic framework for resource management in the next generation interconnection environment, and the corresponding query language and programming environment will be needed as introduced in the book.

The third part of the monograph introduces the conceptual model on knowledge flow and peer-to-peer knowledge sharing. It discusses the computational model of knowledge intensity based on a number of fundamental knowledge flow principles and proposes the knowledge spiral model for revealing the intrinsic rules of knowledge flow. For the last part, it describes the properties of diverse scale–free networks covering random graph theories, small–world effects, the idea of preferential attachments, as well as the dynamic evolution of these networks. The goal is to explore an abstract live network — a possibly high overlay for the Knowledge Grid.

Unlike other books on Grids, “The Knowledge Grid” has no lengthy sample grid applications in it and the assessment of the potential of the Knowledge Grid is very much down to earth. The author considers the Knowledge Grid as a natural evolution of computing methods and technologies, instead of a “big bang” standard that can immediately make existing and competing standards obsolete. The co-existence of the current Web and the Knowledge Grid for years to come is what being predicted by the book author.

[1] Chinese Academy of Sciences, alin.yao@kg.ict.ac.cn

RELATED CONFERENCES & CALL FOR PAPERS

TCII Sponsored Conferences

WI'06
The 2006 IEEE/WIC/ACM International Conference on Web Intelligence
Hong Kong, China
December 18-22, 2006
http://www.comp.hkbu.edu.hk/iwi06/wi/

Web Intelligence (WI) has been recognized as a new direction for scientific research and development to explore the fundamental roles as well as practical impacts of Artificial Intelligence (AI) (e.g., knowledge representation, planning, knowledge discovery and data mining, intelligent agents, and social network intelligence) and advanced Information Technology (IT) (e.g., wireless networks, ubiquitous devices, social networks, wisdom Web, and data/knowledge grids) on the next generation of Web-empowered products, systems, services, and activities. It is one of the most important as well as promising IT research fields in the era of Web and agent intelligence.

The 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06) will be jointly held with the 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'06) and the 2006 IEEE International Conference on Data Mining (ICDM'06). The IEEE/WIC/ACM 2006 joint conferences are organized by Hong Kong Baptist University, and sponsored by IEEE Computer Society Technical Committee on Intelligent Informatics (TCII), Web Intelligence Consortium (WIC), and ACM-SIGART

Following the great successes of WI'01 held in Maebashi City, Japan , WI'03 held in Halifax, Canada, WI'04 held in Beijing, China and WI'05 in Compiegne University of Technology, France, WI'06 provides a leading international forum for researchers and practitioners (1) to present the state-of-the-art of WI technologies; (2) to examine performance characteristics of various approaches in Web-based intelligent information technology; and (3) to cross-fertilize ideas on the development of Web-based intelligent information systems among different domains. By idea-sharing and discussions on the underlying foundations and the enabling technologies of Web intelligence, WI'06 will capture current important developments of new models, new methodologies and new tools for building a variety of embodiments of Web-based intelligent information systems.

IAT'06
The 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology
Hong Kong, China
December 18-22, 2006
http://www.comp.hkbu.edu.hk/iwi06/iat/

The 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'06) will be jointly held with the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06) and the 2006 IEEE International Conference on Data Mining (ICDM'06). The IEEE/WIC/ACM 2006 joint conferences are organized by Hong Kong Baptist University, and sponsored by IEEE Computer Society Technical Committee on Intelligent Informatics (TCII), Web Intelligence Consortium (WIC), and ACM-SIGART. The upcoming meeting in this conference series follows the great success of IAT'99 held in Hong Kong in 1999, IAT'01 held in Maebashi City, Japan in 2001, IAT'03 held in Halifax, Canada, IAT'04 held in Beijing, China and IAT'05 held in Compiegne University of Technology, France.
IAT'06 provides a leading international forum to bring together researchers and practitioners from diverse fields, such as computer science, information technology, business, education, human factors, systems engineering, and robotics, to (1) examine the design principles and performance characteristics of various approaches in intelligent agent technology, and (2) increase the cross-fertilization of ideas on the development of autonomous agents and multi-agent systems among different domains. By encouraging idea-sharing and discussions on the underlying logical, cognitive, physical, and sociological foundations as well as the enabling technologies of intelligent agents, IAT '06 will foster the development of novel paradigms and advanced solutions in agent-based computing.

ICDM'06
The Sixth IEEE International Conference on Data Mining
Hong Kong, China
December 18-22, 2006
http://www.comp.hkbu.edu.hk/iwi06/icdm/

The 2006 IEEE International Conference on Data Mining (ICDM'06) will be jointly held with The 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06) and The 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'06). The IEEE/WIC/ACM 2006 joint conferences are organized by Hong Kong Baptist University, and sponsored by IEEE Computer Society Technical Committee on Intelligent Informatics (TCII), Web Intelligence Consortium (WIC), and ACM-SIGART. The upcoming meeting in this conference series follows the great success of ICDM'01 held in California, USA in 2001, ICDM'02 held in Maebashi City, Japan in 2002, ICDM'03 held in Florida, USA, ICDM'04 held in Brighton, UK and ICDM'05 held in Texas, USA.

The 2006 IEEE International Conference on Data Mining (ICDM'06) provides a premier forum for the dissemination of innovative, practical development experiences as well as original research results in data mining, spanning applications, algorithms, software and systems. The conference draws researchers and application developers from a wide range of data mining related areas such as statistics, machine learning, pattern recognition, databases and data warehousing, data visualization, knowledge-based systems and high performance computing. By promoting high quality and novel research findings, and innovative solutions to challenging data mining problems, the conference seeks to continuously advance the state of the art in data mining. As an important part of the conference, the workshops program will focus on new research challenges and initiatives, and the tutorials program will cover emerging data mining technologies and the latest developments in data mining.

Related Conferences

AAMAS'06
The Fifth International Joint Conference on Autonomous Agents and Multi-Agent Systems
Future University-Hakodate
May 8-12, 2006
http://www.fun.ac.jp/aamas2006/

AAMAS'06 encourages the submission of theoretical, experimental, methodological, and applications papers. Theory papers should make clear the significance and relevance of their results to the AAMAS community. Similarly, applied papers should make clear both their scientific and technical contributions, and are expected to demonstrate a thorough evaluation of their strengths and weaknesses in practice. Papers that address isolated agent capabilities (for example, planning or learning) are discouraged unless they are placed in the overall context of autonomous agent architectures or multiagent system organization and performance. A thorough evaluation is considered an essential component of any submission. Authors are also requested to make clear the implications of any theoretical and empirical results, as well as how their work relates to the state of the art in autonomous agents and multiagent systems research as evidenced in, for example, previous AAMAS conferences. All submissions will be rigorously peer reviewed and evaluated on the basis of the quality of their technical contribution, originality, soundness, significance, presentation, understanding of the state of the art, and overall quality.
In addition to conventional conference papers, AAMAS'06 will also include a demonstrations track for work focusing on implemented systems, software, or robot prototypes; and an industry track for descriptions of industrial applications of agents. The submission processes for the demonstration and industry tracks will be separate from the main paper submission process.

ISWC'06
The Fifth International Semantic Web Conference
Athens, GA, USA
November 5-9, 2006
http://iswc2005.semanticweb.org/

ISWC is a major international forum where visionary and state-of-the-art research of all aspects of the Semantic Web are presented. ISWC'06 follows the 1st International Semantic Web Conference (ISWC'02 which was held in Sardinia, Italy, 9-12 June 2002), the 2nd International Semantic Web Conference (ISWC'03 which was held in Florida, USA, 20 - 23 October 2003), 3rd International Semantic Web Conference (ISWC'04 which was held in Hiroshima, Japan, 7 - 11 November 2004), and 4th International Semantic Web Conference 2005 (ISWC'05 which was held in Galway, Ireland, 6 - 10 November, 2005)

ICTAI'06
The Eighteenth IEEE International Conference on Tools with Artificial Intelligence
Arlington, VA, USA
November 13-15, 2006
http://www.nvc.cs.vt.edu/ictai06/

The annual IEEE International Conference on Tools with Artificial Intelligence (ICTAI) provides a major international forum where the creation and exchange of ideas relating to artificial intelligence are fostered among academia, industry, and government agencies. The conference facilitates the cross-fertilization of these ideas and promotes their transfer into practical tools, for developing intelligent systems and pursuing artificial intelligence applications. The ICTAI encompasses all the technical aspects of specifying, developing, and evaluating the theoretical underpinnings and applied mechanisms of AI tools. A selection of the best papers in the conference will be published in a special issue of the International Journal on Artificial Intelligence Tools (IJAIT), special issues of other journals, or edited book.

CEC'06 & EEE'06
IEEE Joint Conference on E-Commerce Technology and Enterprise Computing E-Commerce and E-Services
San Francisco, California, USA
June 26-29, 2006
http://linux.ece.uci.edu/cec06/

The 8th IEEE Conference on E-Commerce Technology (CEC'06) and the 3rd IEEE Conference on Enterprise Computing, E-Commerce and E-Services (EEE'06) are the flagship annual conferences of the IEEE Computer Society Technical Committee on E-Commerce. The 2006 conference will be a joint event providing a platform for researchers interested in theory and practice of E-Commerce and Enterprise Computing.
The conference focuses on new technologies and methods geared towards business process innovation (i.e., optimizing existing or creating new business processes) for the purpose of optimizing business objectives. The joint programs of CEC'06 and EEE'06 will consist of tutorials, invited talks, paper presentations, and panel discussions. Submissions of high quality papers describing mature results or on-going work are invited. Technical contributions should be accompanied by a thorough evaluation of the results.

IEEE Computer Society
1730 Massachusetts Ave, NW
Washington, D.C. 20036-1903

Non-profit Org.
U.S. Postage
PAID
Silver Spring, MD
Permit 1398