第35卷 第12期 计 算 机 工 程 2009年6月
Vol.35 No.12 Computer Engineering June 2009
·博士论文· 文章编号：1000—3428(2009)12—0032—02 文献标识码：A 中图分类号：TP391

基于自然语言的P2P信誉系统

邓子健[1]，来学嘉[1,2]，何大可[1]
(1. 西南交通大学信息安全与国家计算网格实验室，成都 610031；2. 上海交通大学计算机工程系，上海 200240)

摘 要： 提出一种基于CW的语义信誉系统，使买家可以使用自然语言评论卖家，网络上的节点能阅读针对卖家的评论。考虑交易额大小，以生成最终评论结果，给出价格因素对评论结果的影响。与使用数值描述节点可信度的传统信誉系统相比，该系统更符合人类的思维习惯。
关键词： 自然语言信誉；对等网络；信息安全

P2P Reputation System Based on Natural Language

DENG Zi-jian[1], LAI Xue-jia[1,2], HE Da-ke[1]
(1. Laboratory of Information Security and National Computing Grid, Southwest Jiaotong University, Chengdu 610031;
2. Department of Computer Engineering, Shanghai Jiaotong University, Shanghai 200240)

【Abstract】 This paper proposes a linguistic reputation system based on Computing with Words(CW). The buyer can comment the seller by using natural language, and nodes in the network can read the commentary for seller. This paper considers the weight of price of every tradeoff to obtain the last reputation result, and gives the effect of price factor on the reputation result. Compared with the traditional system using number to describe reliability of the node, this system is more tally with habit of human thought.
【Key words】 natural language reputation; P2P network; information security

1 概述

基于P2P网络的电子商务应用正快速发展，可信度是人们在网络上进行交流的基础。现有信誉系统用确定数字表示节点可信度，此类系统主要研究P2P的一般性应用。在人类社会中，通常用类似“这个人人品不错”的语言评价某人，而不是“这个人人品是0.7”。因此，本文提出一种应用在P2P网络上、基于CW(Computing with Words)的信誉系统，即自然语义信誉系统。

文献[1]方案通过搜集正负值数字来计算节点可信度，该可信度通过一个确定的数字表示。在eBay中，每次交易后，买方和卖方分别对其进行评论，系统搜集6个月内的评论以确定某个参与者的可信度。目前已有一些信誉系统采用基于语义变量的方法[2]，但基于模糊模型的方案存在评论者不便提供评论的问题。在此类系统中，模糊成员函数用来对不精确的输入进行归类，其评论仍然是一个具体数字。文献[3]提出基于CW的实用信誉系统，被使用在医疗系统中，患者可以通过手机联系终端数据库，以得到相关医生的评论。该方案的缺陷是通过手机查询方式获得节点评论在P2P网络上是不可行的，且没有解决评论的可靠性问题，在处理来自恶意节点的评论时该问题尤其严重。

2 CW基础知识

传统计算用来处理数字和符号，而人类则使用文字来交流。CW用文字代替数字，从而进行计算和推导。虽然CW很早就被提出，但模糊逻辑被提出后，CW发展为一种独特的方法论。CW的关键之一是融合了自然语言和计算模糊变量。文献[4-5]较详细地描述了CW。

3 系统设计

3.1 系统架构

本系统主要应用在半离散化P2P网络环境，该环境结合了中心化和完全离散化P2P网络的优点。该系统包括2类节点：(1)超级节点，又称为管理节点；(2)普通节点，通常由卖方节点和买方节点组成。系统假设超级节点已在网络中存在，它们在网络中是可信的，在组建P2P网络时，超级节点已相互认证。超级节点的主要工作是管理所属区域内的普通节点，转发来自其他区域的查询和回复消息。本文符号约定如下：PK表示公钥，SK表示私钥；$Cert_{SP_i}^{P_i}$表示管理节点SP_i给节点P_i颁布的证书；节点A对消息M的签名表示为$Sig_A(M)$。

3.2 系统工作流程

3.2.1 一次交易描述

在系统中的SP_i拥有公钥PK_{SP_i}和私钥SK_{SP_i}。当一个节点P_i加入网络时，该节点从它所在区域的SP_i获得证书$Cert_{SP_i}^{P_i}$和私钥SK_{SP_i}。一个买家P_b和一个卖家P_a进行一次交易，P_b对P_a给出一条基于自然语义的评论r。P_b向管理P_a的管理节点SP_a发送以下消息：$E_{PK_{SP_a}}(M \| Sig_{P_b}(M)) \| Cert_{SP_b}^{P_b}$。$M$表示$r \| T$，$T=Sig_{buyer}(E) \| Sig_{seller}(E) \| E$，$E$是本次交易的交易信息，包含卖方买方、交易额和时间戳，它被卖方和买方共同签名。

3.2.2 语义评论转化

SP_a需要将自然语义转化为CF(Canonical Form)格式。在方案描述中，由节点给出的自然语义评论基于以下格式：r = it is δ_i that X is ω_i，其中，δ_i是语义概率，如“likely”；X是卖家名字，该名字由该节点加入网络时被其管理的超级节点制定；ω_i可以是单个词语，如“trustable”，或复合词

基金项目： 国家自然科学基金资助项目(60573032)；华为科技基金资助项目(YJCB2007048IN)
作者简介： 邓子健(1982－)，男，博士研究生，主研方向：P2P数字版权系统；来学嘉、何大可，教授、博士生导师
收稿日期： 2009-01-29 **E-mail：** zijian.deng@gmail.com

语，如“not trustable”。方案定义的 CF 格式为

$r' = (ID_X \text{ is } G_i) \text{ is } \delta_i$

其中，ID_X 是约束变量；G_i 是约束模糊关系。SP_a 转化“X is ω_i”为“ID_X is G_i”。

ID_X 和 G_i 的定义如下：

$$ID_X = {}_{Arrributei2}RelationName_i[Attribute_{i1} = X] = {}_{Trustableworthiness}PEERID[PeerName] = {}_{Trustableworthiness}(PeerName)$$

$$G_i = {}_{Arrributej2}RelationName_j[Attribute_{j1} = \omega_i] = {}_{\mu}TRUSTABLEDEGREE[Trustableworthiness] = {}_{Trustableworthiness}(Trustableworthiness)$$

例如，若一条语义评论为 It is likely that Peer A is very trustable，则该语义评论的 CF 格式为

$r' = {}_{Trustableworthiness}(PeerA)$ is ${}_{\mu}(verytrustable)$ is likely

在 M 中的 E 包含买卖双方的姓名和交易额信息，若其他第三方节点得到此类信息，则评论者的匿名性无法得到保证。当 SP_a 接收到 P_b 向其发送的消息并成功验证后，SP_a 需要对消息进行转换。方案将具体金额划分为 3 类：小额 Pr_s，中额 Pr_m 和大额 Pr_h。在此基础上，SP_a 构造出新的 $M' = r \square \Pr$。

3.2.3　评论分发

当 SP_a 构造新的 M' 后，随机选择网络中的一个 SP，将 M'分发给该 SP，本系统存储 q 条语义评论，当系统评论大于 q 条 r' 时，系统保留最近的 q 条 r'。若某个节点想知道节点 A 的语义评论，则该节点向节点 A 的管理节点询问，以得到所有存放 M' 的 SP_a 地址信息，并获得相应的 r'。

3.2.4　全局语义评论计算

当一个节点得到他感兴趣的卖家的相关 M' 后，提取 r'，并开始根据 q 条 r' 计算其想得知的语义评论结果。

先简化限制关系 G_i，在语义评论中，G_i 的格式为 $G_i = mT_i = f(T_i)$，其中，m 为修饰词；T_i 为可信度。其简化规则如表 1 所示。图 1 描述了 Trust 模糊成员函数。

表 1　G_i 的简化规则

修饰词	$f(T_i)$
Not	$1-\mu_t(u)$
Very	$\mu_t^2(u)$
Not Very	$1-\mu_t^2(u)$
Very very	$\mu_t^4(u)$

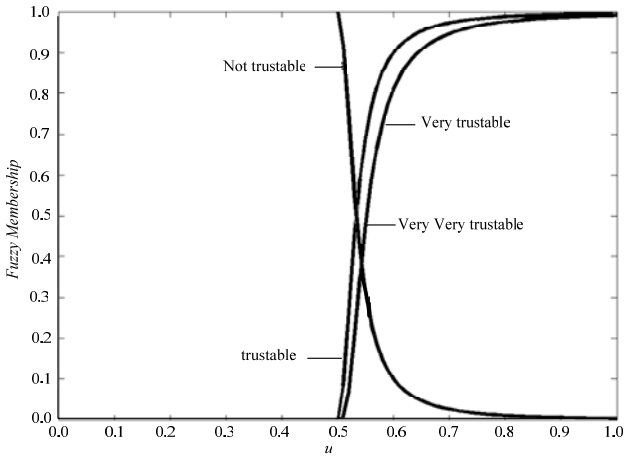

图 1　Trust 模糊成员函数

$\mu_{trustable}(u)$（简写为 $\mu_t(u)$）的定义如下：

$$\mu_{trustable}(u)\begin{cases}0 & u \in \left[0, \frac{1}{2}\right] \\ \frac{1}{1+20^{-2}(u-\frac{1}{2})^{-2}} & u \in \left[\frac{1}{2}, 1\right]\end{cases}$$

本文方案考虑了交易价格的因素，把具有高额支付的交易考虑到 r' 中。考虑价格因素后的 δ_i 如表 2 所示。

表 2　考虑价格因素后的 δ_i

修饰词	价格因素加入后的 δ_i
Not likely	Very not likely
likely	Very likely
Very likely	Very very likely

$\mu_{likely}(u)$ 的定义如下：

$$\mu_{likely}(u)\begin{cases}0 & u \in \left[0, \frac{1}{2}\right] \\ -4((u-1)^2+1) & u \in \left[\frac{1}{2}, 1\right]\end{cases}$$

图 2 描述了 likely 模糊成员函数。

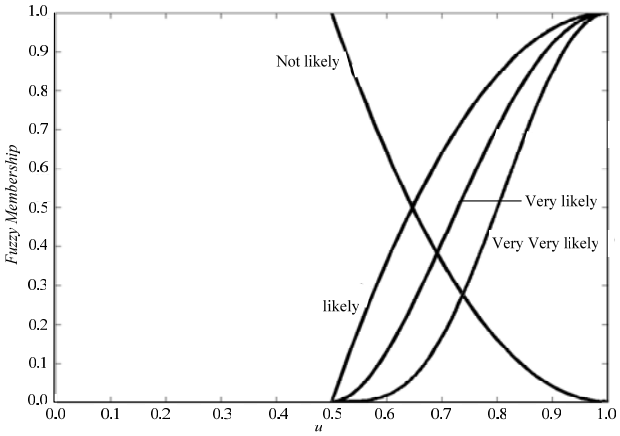

图 2　likely 模糊成员函数

用 P_i 代替该节点想得到的 ID_X 的模糊概率事件，可得

$$(ID_X) \text{ is } f(T_i) \text{ is } \delta_i \rightarrow P_i \text{ is } \delta_i \rightarrow prob\{ID_X \text{ is } G_i\} = \int_U \mu_{G_i}(u)p(u)\mathrm{d}u$$

其中，$p(u)$ 是 ID_X 在 U 上取值的概率密度；μ_{G_i} 是 G_i 的模糊成员函数。用 $\prod prob\{ID_X \text{ is } G_i\}$ 表示 ID_X is G_i 的概率密度的概率。根据概率理论可得

$$\prod_i(P) = \prod prob\{ID_X \text{ is } G_i\} = \mu_{\delta_i}(\int_U \mu_{G'_i}(u)p(u)\mathrm{d}u)$$

其中，μ_{δ_i} 是 δ_i' 的模糊成员函数。

利用极大约束原理，可得

$$\mu_{\delta'_i}(ID_X \text{ is } G') = \mu_{\delta'_i}(\int_U \mu_{G'_i}(u)p(u)\mathrm{d}u) = \max_P(\prod_1(P) \wedge \prod_2(P) \wedge \cdots \wedge \prod_n(P))$$

其中，$\mu_{\delta'_i}$ 和 $\mu_{G'_i}$ 分别是 δ'_i 和 G'_i 的模糊成员函数。参考语义概率的模糊成员变量，如 likely，最终可以得到 δ'_i。

4　系统分析

模拟一个包括 60 个超级节点和 400 个普通节点的 P2P 网络。系统中保留对某个卖家的 20 条语义评论，并有 10 个攻击节点对某个可信节点(假设为节点 A)进行合谋攻击，攻击节点评论该节点为不可信节点。系统每完成 2 000 笔交易时，询问该系统节点 A 的全局可信语义评论结果。存在 5 种状态：S1 为“It is very very likely that the Peer A is trustable”，S2 为

(下转第 37 页)