

对等网络中一种基于概率的信任机制研究

王 平 邱 劲 邱玉辉

(西南大学计算机与信息科学学院 重庆 400715)

摘 要 P2P 网络由于其开放性和自治性,充斥着大量的不确定性因素。为了提高网络的整体可用性和服务质量,提出了一种基于 P2P 技术的信任机制,该机制使用 beta 概率分布来估价结点的信任,同时使用 Gossip 算法来传递聚集信誉。模拟实验表明,时间关联因子以及 Gossip 搜索机制的引入,能更有效地估价结点间的信任关系,减少网络拥塞,提高查询效率。

关键词 非结构化对等网络, Gossip 算法, beta 概率分布, 信任, 信誉

Trust Computing Based on Probability Theory in Unstructured P2P Network

WANG Ping QIU Jing QIU Yu-hui

(Institute of Computer and Information, Southwest University, Chongqing 400715, China)

Abstract Trust plays an important role in determining who and how to interact in open and dynamic P2P network. To this end, a trust mechanism for unstructured P2P network was proposed in which peers evaluate the trust based on beta probability distribution and aggregate the reputation with Gossip algorithm. The experimental data shows that our way can autonomously deal with the biased information and identify trust belief with low overhead and better efficiency.

Keywords Unstructured P2P networks, Gossip algorithm, Beta probability distribution, Trust, Reputation

随着互联网的飞速发展,各种信息资源海量膨胀。如何可扩展地支持海量数据信息共享、管理和应用,是目前互联网面临的严峻挑战。P2P 技术由于其自身的特点,在互联网中扮演了重要角色。然而,在 P2P 技术广泛应用于互联网的同时,也带来一系列有待进一步解决的问题,信任便是其中之一。

P2P 网络的开放性使得它不可避免地充斥着大量不可靠的服务和欺诈行为:交互中,自私的结点有可能违背其承诺;另一方面,由于环境的变化,结点经常需要和陌生的或者是交互很少的结点打交道。为了提高 P2P 网络的整体可用性和服务质量,信任机制的研究在 P2P 网络中扮演了重要角色。为此,人们提出了大量的信任评估机制:Aberer 和 Despotovic^[1]利用 P-Grid 提供的信息来估价结点在未来欺骗的可能性;Kamvar 等^[2]基于信任的传递性,提出了基于全局声誉的信任模型 EigenRep;Xiong 和 Liu 等^[3]提出了一种综合局部和全局声誉的信任机制 PeerTrust;Zhou Runfang 等^[4]利用 P2P 幂率分布特性来进行信任估价;Song 等^[5]则采用模糊逻辑推理规则来估价结点的全局声誉。

虽然研究者针对 P2P 网络中的信任问题做了大量研究,然而现有机制依然存在一系列问题有待进一步解决:Aberer 和 Despotovic 的方法没有考虑全局信息;FuzzyTrust 虽具有较高的恶意结点检测率,但不能解决针对信任机制的攻击;虽然 Kamvar, Xiong & Liu 以及 Zhou Runfang 等^[2-5]使用了不同的数学方法来传递和组合推荐信息,但是现有的信誉机制并

没有有效考虑 P2P 网络的可扩展性以及信任的时效性。为此,本文在现有研究的基础之上提出了一种基于非结构化 P2P 网络的信任机制。

本文第 1 节介绍了基于概率的信任评估机制;第 2 节对信誉的传递聚集机制进行了详细描述;第 3 节是实验及结果分析;最后是结论。

1 信任机制

我们的信任机制的主要思想是:对于 P2P 中任意一个结点,信任机制根据该结点执行过的交互情况和其他结点对其交互行为的评价,为该结点分配一个信任值(信任值的取值范围和高低由其应用领域所决定),使得其他结点可以根据其信任值做出是否与其交易的决策。本机制使用基于推荐信任 Gossip 算法进行消息传递聚集信誉,减少了消息冗余和网络拥塞;同时引入时间关联因子,能更有效地检测恶意结点的欺诈行为。

1.1 基本概念

在非结构化 P2P 中,假设在时间点 t , 结点 a 从 b 中对其资源 θ 进行下载, a 称为信任者, b 称为被信任者,交互的结果记为 $O_{a,b}(\theta)$ 。若 a 从 b 中成功获得资源 θ , 则令 $O_{a,b}(\theta) = 1$; 否则 $O_{a,b}(\theta) = 0$ 。从时间 t_1 到时间 t_2 这一时间段记为 $[t_1, t_2]$, a 与 b 基于 θ 进行的所有交互结果记为 $O_{a,b}^{[t_1, t_2]}(\theta)$ 。在某一时间点,特定的结点间只能完成一次交互。

在任意时间点 t , a 和 b 之间基于 θ 的交互历史记为 $E =$

到稿日期:2009-10-12 返修日期:2009-12-01 本文受 973 国家重点基金项目(2003CB31708)资助。

王 平(1979-),女,博士生,主要研究方向为分布式人工智能等, E-mail: wangping@swu.edu.cn; 邱 劲(1971-),男,讲师,主要研究方向为信息安全等; 邱玉辉(1938-),男,教授,博士生导师,主要研究方向为分布式人工智能等。

$\langle s_{a,b}^i, f_{a,b}^i \rangle$, 其中 $s_{a,b}^i$ 为交互成功的次数, $f_{a,b}^i$ 为交互失败的次数。被信任者 b 提供可靠或者不可靠资源的倾向性, 记为 $B_{a,b}(\theta) \in [0, 1]$ 。规定 $B_{a,b}(\theta)$ 为 a 从 b 中成功获得资源的概率, 即 $B_{a,b}(\theta) = p(O_{a,b}(\theta))$ 。

信任是一个非常复杂的主观概念, 这里引用 Gambetta^[6] 的定义, 把信任者 a 对被信任者 b 的信任 $T_{a,b}(\theta)$ 定义为信任者 a 对被信任者 b 在未来的交互中提供可靠资源 θ 的概率。

1.2 信任的估价

如果信任者 a 对被信任者 b 提供 θ 具有完整的知识, 则 b 提供可靠 θ 的概率可定义为 $B_{a,b}(\theta)$ 。然而, 要求 a 拥有 b 提供可靠 θ 的完整知识是不切实际的。根据贝叶斯概率统计方法, 可以利用 a 与 b 已有的交互历史事件来估价 $B_{a,b}(\theta)$ 的期望值。为此, α 关于 b 在时间点 t 提供可靠资源 θ 的信任 $T_{a,b}(\theta)$ 规定为在时间点 t , $B_{a,b}(\theta)$ 基于 $O_{a,b}^{[t_0, t]}$ 的期望值 (其中 t_0 为系统的初始时间):

$$T_{a,b}(\theta) = E[B_{a,b}(\theta) | O_{a,b}^{[t_0, t]}(\theta)] \quad (1)$$

这里, 使用 beta PDF (beta probability density function) 来计算 $E[B_{a,b}(\theta) | O_{a,b}^{[t_0, t]}(\theta)]$:

$$E[B_{a,b}(\theta)] = \frac{\alpha}{\alpha + \beta} \quad (2)$$

$$\alpha = s_{a,b}^{[t_0, t]} + 1, \beta = f_{a,b}^{[t_0, t]} + 1$$

1.2.1 信任的动态性: 时间关联因子

在估价 a 对 b 的信任值时, $E = \langle s_{a,b}^i, f_{a,b}^i \rangle$ 决定了 beta 函数中的 α 和 β 值。在标准 beta 概率分布函数中, a 和 b 的每一次交互结果都被赋予相同的权重, 但没有考虑 a 和 b 之间交互发生的时间。然而, 信任具有动态性, 信任作为信任者对被信任者未来行为的预测, 直观上两者之间发生较早的事件对信任值的影响应小于发生较晚的事件。为此, 引入时间关联因子 μ 来反映信任的这一特性。

$$\mu(o_i) = e^{-\frac{\Delta t(o_i)}{\lambda}} \quad (3)$$

其中, $\Delta t(o_i)$ 为当前时间与事件发生时间的间隔; λ 为衰减因子, 由其应用领域所决定。事件发生得越早 (即 $\Delta t(o_i)$ 越大), 领域事件变化越快 (即 λ 越小), 则 μ 越小, 即在信任估价时其重要性就越小。

假设 a 与 b 进行交互, 如果 b 提供了可靠资源 θ , 则有 $o = 1$; 否则 $o = 0$, 那么信任更新为:

$$\begin{aligned} \alpha &= \mu\alpha + o \\ \beta &= \mu\beta + (1 - o) \end{aligned} \quad (4)$$

假设 $\{o_1, \dots, o_n\}$ 为交互事件的结果序列, $\{t_1, \dots, t_n\}$ 为事件交互的时间序列 (这里时间开始点设为 0)。beta 分布中的参数 α, β 初始值均为 1。若 $o_0 = 1$, 则经过 n 次直接交互后:

$$s_{a,b}^n = o_n + \mu_{n-1}o_{n-1} + \dots + \mu_1o_1 + \mu_0 = \sum_{i=0}^n e^{-\frac{\Delta t(o_i)}{\lambda}} o_i \quad (5)$$

同理, 可得:

$$f_{a,b}^n = \sum_{i=0}^n e^{-\frac{\Delta t(o_i)}{\lambda}} (1 - o_i)$$

则有:

$$T_{a,b}^n(\theta) = \frac{\sum_{i=0}^n e^{-\frac{\Delta t(o_i)}{\lambda}} o_i + 1}{\sum_{i=0}^n e^{-\frac{\Delta t(o_i)}{\lambda}} o_i + \sum_{i=0}^n e^{-\frac{\Delta t(o_i)}{\lambda}} (1 - o_i) + 2} \quad (6)$$

注意, 即使结点没有进行交互, 但其信任值也随着时间的推移而发生衰减。

1.2.2 信任的可靠性: 信心

在 P2P 网络中, 结点间交互的次数越多, 则对未来交互结果估价的准确性越高。然而, 在 beta PDF 中, 并没有区分信任者是否对被信任者有足够的信息。Jøsang^[7] 虽然考虑了信任的不确定性, 但 Jøsang 在预测交互事件成功和失败的比例为 1:0 或 0:1 和 1:1 这两种不同的情形时, 得出的不确定性相同, 即没有考虑事件的冲突率 ζ ($\zeta = \frac{s+1}{s+f+2}$)。显然, 前者的不确定性要远小于后者。为此, 这里使用 $\langle s_{a,b}^i, f_{a,b}^i \rangle$ 来计算信心^[8]:

$$\begin{aligned} C(B_{a,b}(\theta) | \alpha, \beta) &= \frac{1}{2} \int_0^1 \frac{(B_{a,b}(\theta))^{\alpha-1} (1 - B_{a,b}(\theta))^{\beta-1}}{\int_0^1 U^{\alpha-1} (1 - U)^{\beta-1} dU} - 1 | dB_{a,b} \end{aligned} \quad (7)$$

其中, $\alpha = s_{a,b}^{[t_0, t]} + 1, \beta = f_{a,b}^{[t_0, t]} + 1$ 。信心的计算同时考虑了交互的次数和事件的冲突率 ζ 。

2 信誉

非结构化 P2P 网络中, 信誉在进行有效的信任决策中具有重要的作用^[9]。现有机制主要采用洪泛算法来传递聚集信息。洪泛算法虽具有良好的鲁棒性和扩展性, 但易产生大量冗余消息, 严重消耗资源, 造成网络拥塞。Gossip 算法作为一种典型的消息扩散算法而被广泛应用于分布式系统^[10, 11]。为此本文不仅考虑了结点间的直接交互, 同时采用基于推荐信任的 Gossip 算法 (G-R) 来传递消息聚集信誉。

2.1 推荐信任

在非结构化 P2P 网络中, 结点 a 没有足够的信息来估价目标结点 b 的信任时, a 将向其邻居查询 b 的交互信息。与传统的洪泛或随机相比, (G-R) 算法不再盲目路由, 而是引入反馈机制, 根据推荐信任值的高低进行优先转发。

这里按照上述信任计算机制来计算推荐信任:

$$TR_{a,w}(w \xrightarrow{\theta} b) = \frac{s_{a,w}^{[t_0, t]} + 1}{s_{a,w}^{[t_0, t]} + f_{a,w}^{[t_0, t]} + 2} \quad (8)$$

其中, a 为估价结点, b 为目标结点, w 为目标结点的见证者。如果 w 关于 b 的反馈信息与 a, b 的实际交互一致, 则 $s_{a,w} = 1$; 否则 $s_{a,w} = 0$ 。

2.2 基于推荐信任的 Gossip 搜索算法

Gossip 算法中, 每个收到消息的结点随机地选择 $fanout$ 个邻居结点进行传播。研究表明, 将概率优先机制引入路由策略可以提高搜索准确率^[12]。为此, 将概率优先机制引入 Gossip 算法, 即在查询目标结点的推荐信息的过程中, 每个结点根据反馈求取推荐信任, 并将所有邻居的推荐信任均一化为概率形式, 即转发概率, 然后选取转发概率最高的 $fanout$ 个结点进行消息传递, 从而控制本地信息流的传递, 提高整个路由效率。

P2P 网络中, 每个结点的信息如下。

- $Inf(i) = (ID_i, \theta_i, O_{i,X}^{[t_0, t]}(\theta), neighbor(i))$: ID_i 为结点 i 的 ID, θ_i 为结点 i 所拥有的资源, $O_{i,X}^{[t_0, t]}(\theta)$ 为结点 i 的直接交互历史, $neighbor(i)$ 为结点 i 的邻居结点集。

- $neighbor(i) = (ID_{j_k}, \theta_{j_k}, TR_{i,j_k}^i(\theta), k \in [1, n])$: ID_{j_k} 为结点 i 的邻居结点 j_k 的 ID, θ_{j_k} 为 j_k 所拥有的资源, $TR_{i,j_k}^i(\theta)$ 为 i 关于 j_k 的推荐信任。

假设 a 为源结点, b 为估价目标结点。在信任估价过程

中, a 首先检查 $O_{i,x}^{[0,t]}(\theta)$ 中是否有与 b 的直接交互。若有, 则直接估算; 否则, 执行如下操作:

1) 生成查询消息 $q(a, O_{i,b}^{[0,t]}(\theta))$, 初始化生存周期 TTL (这里 TTL 称为信息传播迭代的跳数。一个消息从一个结点转发到另一个结点的过程称为一跳。根据小世界网络模型, $TTL=6$)。

2) 检查 TTL 。若 $TTL=0$, 则消息无效, 丢弃; 若 $TTL>0$, 则根据邻居结点的推荐信任计算邻居结点的转发概率:

$$p_{i,j_k} = \frac{TR_{a,j_k}^c(\theta)}{\sum_{k=1}^n TR_{a,j_k}^c(\theta)} \quad (9)$$

3) 选取 p_{ij} 最高的 $fanout$ 个结点进行传播, 并令 $TTL=TTL-1$ 。

4) 邻居结点收到消息后, 检查其是否与目标结点有交互历史。如果有, 则返回相关信息给源结点; 否则, 将消息传递给邻居结点的邻居结点。重复 2)~4), 这样反复递归, 直到找到目标结点推荐信息, 或者 $TTL=0$, 消息无效被丢弃。具体算法详见文献[13]。

2.3 信誉的聚集

客观上, 参与评价的推荐结点数越多, 全局信誉的期望值就越接近真实值。为了防止少数恶意结点通过共谋来抬升彼此的信誉, 模型在聚集推荐信息的过程中, 不仅考虑了推荐结点的可信度、推荐结点对目标结点的信任, 而且考虑了推荐结点数。

假设 a 为源结点, c 为目标结点, w_i 为目标结点的推荐结点, m 为推荐结点数, 则 a 对 c 关于 θ 的全局信誉:

$$R_{a,c}^c(\theta) = e^{-m-1} \sum_{i=1}^m \left(\frac{TR_{a,w_i}^c(w_i \rightarrow c)}{\sum_{i=1}^m TR_{a,w_i}^c(w_i \rightarrow c)} \cdot T_{w_i,c}(\theta) \right) \quad (10)$$

3 实验与分析

为了验证本机制的准确性和有效性, 进行了如下模拟实验。

3.1 信任的动态性分析

实验 1 模拟结点突变行为。即结点先通过一段时间的合作交互并累积一定的信任后, 再进行连续的恶意攻击行为。结点进行了 100 次交互, 交互过程设定为 $0 \leq t \leq 50$ 交互成功, $51 \leq t \leq 100$ 交互失败。实验分别比较了 Jøsang 的信任模型以及引入时间关联因子的信任模型, 其中 λ 设置如下: $\lambda_1 =$

$$-\frac{5}{\ln(0.5)}, \lambda_2 = -\frac{5}{\ln(0.1)}。$$

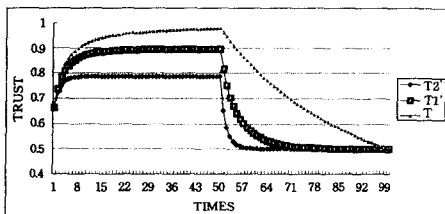


图 1 信任的动态性分析

实验结果如图 1 所示。X 轴为交互次数, Y 轴为信任值, 其中, T 为 Jøsang 模型数据, $T1'$ 表示衰减因子为 λ_1 的模型数据, $T2'$ 表示衰减因子为 λ_2 的模型数据。实验数据表明, 当结点行为发生突变时, 引入时间关联因子的模型都能较快地做出反应, 而 Jøsang 模型则反应较慢。即本机制在信任估价过

程中对结点的突变行为具有较好的适应能力, 能更好地检测出恶意结点的策略性欺骗。此外, 实验表明与领域相关的衰减因子 λ 越小, 事件衰减越快, 近期事件的比重越大, 则对结点的突变行为反应越灵敏。

3.2 信任的可靠性

信任的精确度 c 受交互的次数 t 和事件的冲突率 ζ 的影响。为了简化实验, 模拟了 $\zeta=0.5$ 时, c 与 t 的关系。实验结果如图 2 所示。

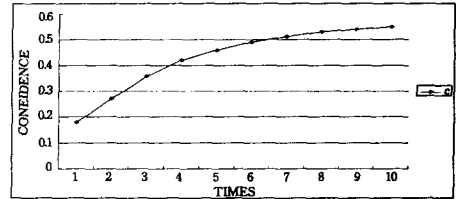


图 2 信任的可靠性分析

图 2 中 X 轴为交互次数, Y 轴为信心。实验数据表明, 在交互过程中, 当 ζ 为一定值时, 信任的精确性 c 随着交互次数的增加而增加。

3.3 查询的成功率及搜索代价

实验使用文献[11]的仿真网络环境, 其中相关参数设置如下: $fanout=3, TTL=6, \alpha_1=\alpha_2=1, \beta_1=0.2, \beta_2=0.1$ 。实验比较了传统的 Gossip 算法和基于概率优先的 Gossip 算法在推荐信息传递过程中的查询代价(一条消息访问过的结点数占整个网络结点数的百分比)和返回率(返回推荐数与网络中和目标结点有交互的结点数的比率)。理论上, 在消息的传递过程中, 消息访问的结点数越少, 查询的响应时间越短, 网络资源开销就越少。

实验结果如图 3 所示。其中 X 轴为返回率, Y 轴为查询代价。实验数据表明, 在推荐信息的传递聚集过程中, 本模型的查询开销要远小于传统的 Gossip 算法, 较好地解决了消息传递中的网络拥塞问题。

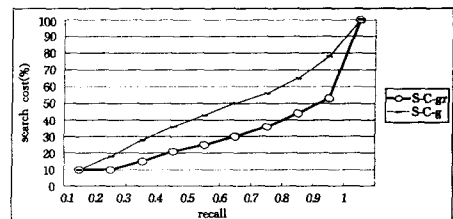


图 3 查询代价

结束语 信任是非结构 P2P 网络研究的一个热点问题。本文在总结现有的典型信任模型基础之上, 提出了一种基于 P2P 技术的概率信任机制。与现有机制相比, 本机制使用基于推荐信任的 Gossip 算法来传递消息, 降低了网络负载; 同时在估价信任关系时引入时间关联和衰减因子, 能更有效地检测恶意结点的欺诈行为。

参考文献

- [1] Karl A, Zoran D. Managing trust in a peer-2-peer information system[C]// Proceedings of the 10th International Conference on Information and Knowledge Management. New York, USA, 2001, 310-317
- [2] Kamvar S D, Schlosser M T. EigenRep: reputation management in p2p networks[C]// Lawrence S, ed. Proc. of the 12th Int'l

World Wide Web Conf. (WWW 2003). Budapest: ACM Press, 2003; 123-134

[3] Xiong L, Liu L. PeerTrust: supporting reputation-based trust for p2p electronic communities[J]. IEEE Trans. Knowledge and Data Engineering, 2004, 16(7): 843-857

[4] Zhou R, Hwang K. PowerTrust: A robust and scalable reputation system for trusted P2P computing[J]. IEEE Trans. Parallel and Distributed Systems, 2007, 18(4): 460-473

[5] Song S S, Hwang K, Zhou R F, et al. Trusted p2p transactions with fuzzy reputation aggregation[J]. IEEE Internet Computing, 2005, 9(6): 24-34

[6] Gambetta D. Can we trust trust [M]. Reprinted in electronic edition from Department of Sociology, University of Oxford; Basil Blackwell, 1988

[7] Jøsang A, Gray L, Kinatader M. Simplification and analysis of transitive trust networks[J]. Web Intelligence and Agent Systems Journal, 2006, 4(2): 139-161

[8] Wang Yonghong, Singh M P. Trust representation and aggregation in a distributed agent system[C]//Proceedings of the 21st National Conference on Artificial Intelligence (AAAI). Boston,

MA, USA, 2006

[9] Dong-Huynh T, Jennings N R, Shadbolt N R. Fire: An integrated trust and reputation model for openmulti-agent systems[J]. Journal of Autonomous Agents and Multi-Agent Systems, 2006, 13(2): 119-154

[10] Zhou R, Hwang K, Cai M. Gossip Trust for fast reputation aggregation in peer-to-peer networks[J]. IEEE TKDE, 2008, 20(9): 1282-1295

[11] Zhuge Hai, Chen Xue, et al. Trust-based probabilistic search with the view model of p2p networks[J]. Concurrency and Computation: Practice and Experience, 2006, 18(14): 1839-1855

[12] Menasce D A, Kanchanapli L. Probabilistic scalable p2p resource location services[J]. ACM SIGMETRICS Performance Evaluation Review, 2002(20): 48-58

[13] Wang Ping, Qiu Jing. A search algorithm based on referral trust in unstructured P2P systems[C]//Proceedings of the International Conference on ISECS. Nanchang, China, 2009: 644-649

[14] 陈学梁, 万晓瑜, 樊自甫. 无线接入网中 P2P 对网络业务性能的评估[J]. 重庆工学院学报: 自然科学版, 2006, 20(5): 92-96

(上接第 191 页)

关系词对具有同义关系。

4.2 实验结果和分析

图 3 给出了混合分类算法(HC)与贝叶斯分类算法(NB)和感知器分类算法(Perceptron)随训练集大小变化的曲线图。横坐标表示训练数据的大小,以千为单位。纵坐标是 F1 值。

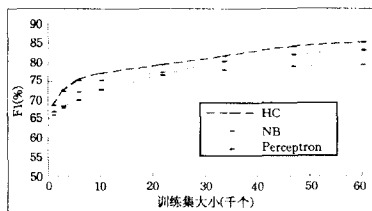


图 3 分类算法的比较结果

实验数据表明,本文提出的混合分类算法 HC 获得了较好的性能。NB 分类算法在训练数据量少的情况下,F1 值高于 Perceptron。但是随着训练数据的不断增加,Perceptron 分类算法的 F1 值高于 NB 分类算法。这是由于产生式分类算法需要已知训练数据的分布形式,然而实际情况往往无法得知训练数据的真正分布函数,因此 NB 的精度不高。但是,判定分类算法在足够的训练数据的前提下优于产生式分类算法。本文提出的混合分类算法,一部分特征值是通过训练数据估计而来,另一部分特征值是通过判定函数训练得到。因此,在训练数据较少的情况下 HC 曲线表现出和 NB 曲线一样快速上升的趋势。而在训练数据较多的情况下,NB 曲线基本保持不变,而 HC 曲线和 Perceptron 曲线仍呈上升趋势。

结束语 术语关系自动抽取逐渐成为当前热点研究方向,国内外学者做了许多有益尝试。其中,将关系抽取转化为分类问题,被认为是一种有效的抽取方法。然而,无论产生式分类算法还是判定分类算法,都有局限性。鉴此,本文提出了一种混合分类算法 HC。结合 Naive Bayes 算法和 Perceptron 算法的优点,一部分特征值是通过训练数据估计而来,另一部分特征值是通过判定函数训练得到。实验表明,在相同大小

的训练数据集下,本文提出的混合学习算法优于 Naive Bayes 算法和 Perceptron 算法。

参 考 文 献

[1] Boguraev B, Kennedy C. Applications of term identification technology: domain description and content characterization[J]. Natural Language Engineering, 1999, 5(1): 17-44

[2] Appelt D E. Introduction to Information Extraction[J]. AI Communications, 1999, 12(3): 161-172

[3] Aone C, Ramos M, Rees S. A large-scale relation and event extraction system[C] // Proceedings of the 6th Applied Natural Language Processing Conference, New York: ACM Press, 2000: 76-83

[4] Hearst M A. Automatic Acquisition of Hyponyms from Large Text Corpora[C] // 14th International Conference on Computational Linguistics, Nantes, France, 1992: 539-545

[5] Yu Hong, et al. Automatic Extraction of Gene and Protein Synonyms from MEDLINE and Journal Articles[C] // Proceedings of the American Medical Informatics Association 2002 Symposium (AMIA'2002), 2002: 919-923

[6] 刘克彬,李芳,刘磊,等. 基于核函数中文关系自动抽取系统的实现[J]. 计算机研究与发展, 2007, 44(8): 1406-1411

[7] Li Wenjie, et al. A Novel Feature-based Approach to Chinese Entity Relation Extraction[C] // Proceedings of ACL-08: HLT, Columbus, USA, 2008: 89-92

[8] Girju R, Badulescu A, Moldovan D. Learning Semantic Constraints for the Automatic Discovery of Part-Whole Relations [C] // Edmonton, Canada. Proceedings of HLT-NAACL, Edmonton, Canada, 2003: 80-87

[9] Fleischman M, Hovy E. Offline Strategies for Online Question Answering: Answering Questions Before They Are Asked[C] // Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics, Sapporo, Japan, 2003: 1-7

[10] Chang J T. Using Machine Learning to Extract Drug and Gene relationships from Text[D]. Stanford University, 2003

[11] 孙霞,郑庆华,王朝静,等. 一种基于生语料的领域词典生成方法[J]. 小型微型计算机系统, 2005, 26(6): 1088-1092