

## 基于启发式函数的多叉树防碰撞算法

丁治国<sup>1\*</sup>, 朱学永<sup>1</sup>, 雷迎科<sup>2</sup>, 王心灵<sup>1</sup>

(1. 合肥电子工程学院 网络信息管理中心, 合肥 230037; 2. 合肥电子工程学院 信息系, 合肥 230037)

(\* 通信作者电子邮箱 dhgzg@mail.ustc.edu.cn)

**摘要:**为克服传统二叉树防碰撞算法搜索效率低的缺点,提出了一种基于启发式函数的自适应多叉树防碰撞算法。新算法通过定义和计算启发式函数,有效地利用碰撞比特信息来估计节点内待识别标签的数量。新算法根据节点内的标签数量,可在不同节点和深度,自适应地调整搜索叉数,从而有效地提高了算法的搜索效率。理论分析和仿真实验证明:新算法克服了传统防碰撞算法的缺点,尤其在待识别标签数量较多场合,可有效地减少搜索和识别时间,提高射频识别系统的吞吐率。

**关键词:**射频识别;防碰撞算法;启发式函数;多叉树;吞吐率

**中图分类号:** TP18; TP301.6; TN911.23 **文献标志码:** A

### Multi-tree anti-collision algorithm based on heuristic function

DING Zhi-guo<sup>1\*</sup>, ZHU Xue-yong<sup>1</sup>, LEI Ying-ke<sup>2</sup>, WANG Xin-ling<sup>1</sup>

(1. Management Center of Network and Information, Electronic Engineering Institute of Hefei, Hefei Anhui 230037, China;

2. Department of Information, Electronic Engineering Institute of Hefei, Hefei Anhui 230037, China)

**Abstract:** In order to overcome the low efficiency of traditional binary-tree anti-collision algorithms, an adaptive multi-tree anti-collision algorithm based on heuristic function was presented in the paper. By defining the heuristic function which was computed by the number of collision bits, the new algorithm can estimate the number of tags in the branch effectively. Because the new algorithm can adjust the number of searching fork in different branches and depths dynamically, it improves the searching efficiency. The theoretical analyses and simulation results show that the new algorithm overcomes the deficiency of traditional algorithms. For the large number of tags in particular, it can reduce the searching and recognition time and increase the throughput of Radio Frequency Identification (RFID) system.

**Key words:** Radio Frequency Identification (RFID); anti-collision algorithm; heuristic function; multi-tree; throughput

## 0 引言

射频识别(Radio Frequency Identification, RFID)作为新一代非接触式自动识别技术在物流、销售、跟踪和定位等领域已得到广泛应用。随着有源标签的出现和RFID技术在高速移动物体中的应用,迫切需要读写器在有限时间内快速地识别大量标签。标签防碰撞算法就是要解决多个标签与读写器同时进行通信的问题。

传统的防碰撞算法基于时分多路技术,一般可分为两类:一类是基于时隙随机分配的ALOHA算法<sup>[1-4]</sup>,其特点是算法简单,便于实现,适用于低成本RFID系统。但由于该类算法存在“Tag starvation”问题,所以被称为可能性方法。另一类是基于二叉树的搜索算法<sup>[5-8]</sup>,该类算法比较复杂,识别时间较长,但不存在“Tag starvation”问题,故被称为确定性方法。该类方法往往对标签的设计要求较高,如增加随机数产生器、记位器或延迟器等,很难满足低端RFID系统的成本要求。因此,对于标签防碰撞算法的研究,依然是RFID领域炙手可热的研究课题之一。

本文提出了一种基于启发式函数的自适应多叉树防碰撞算法。新算法通过定义和计算启发式函数,有效地利用了碰撞比特信息来估计节点内待识别的标签数量。同时算法根据节点内标签数量,可在不同节点和深度,自适应地调整搜索叉

数,从而有效地减少了算法的搜索时间,大幅度提高了RFID系统的吞吐率。

## 1 现有防碰撞算法

随着有源标签的出现,RFID系统的通信距离显著增加,当读写器作用范围内有多个标签同时与读写器进行通信时,就不可避免地会相互产生干扰,即发生碰撞。标签防碰撞算法就是要提出相应的冲突消除策略,使读写器能依次完成对所有标签的识别和通信。

传统基于二叉树搜索的防碰撞算法利用曼彻斯特编码可以准确识别碰撞位这一特性,通过动态地修改查询码前缀,不断减少响应标签的数量,直至对唯一的标签进行识别。但值得注意的是,当待识别标签数量较多时,基于二叉树的防碰撞算法在搜索的初期会频频出现碰撞,搜索效率较低。如图1(a)所示,完成上述RFID系统内5个标签的识别共需要9个时隙,其中4个是碰撞时隙。

为此,文献[8]提出了一种基于四叉树的混合查询树算法。当读写器检测到发生碰撞时,该算法将响应标签分为4个分支,由于每个分支内响应标签数减少到了1/4,使得再次发生碰撞的概率降低。但随着搜索深度的增加,碰撞时隙内标签数量显著减少,会在减少碰撞时隙的同时产生大量空闲时隙。如图1(b)所示,完成上述所有标签的识别仍然需要9

收稿日期:2011-09-07;修回日期:2011-11-14。 基金项目:安徽省自然科学基金资助项目(090412055)。

作者简介:丁治国(1977-),男,安徽合肥人,讲师,博士,主要研究方向:射频识别、计算机网络; 朱学永(1958-),男,安徽舒城人,教授,主要研究方向:计算机网络、信息安全; 雷迎科(1976-),男,安徽安庆人,讲师,博士,主要研究方向:模式识别、机器学习; 王心灵(1975-),女,安徽颍上人,讲师,博士,主要研究方向:计算机网络、信息安全。

个时隙,即在减少2个碰撞时隙的同时增加了2个空闲时隙。

无论基于二叉树还是四叉树的防碰撞算法,在碰撞时隙,进行分叉的数量是固定的。如果算法能够根据搜索深度和标签数量自适应地调整搜索的叉数,就可以有效地提高算法的搜索效率。如图1(c)所示,自适应多叉树算法<sup>[9]</sup>仅需要7个时隙,就可以完成所有标签的识别。

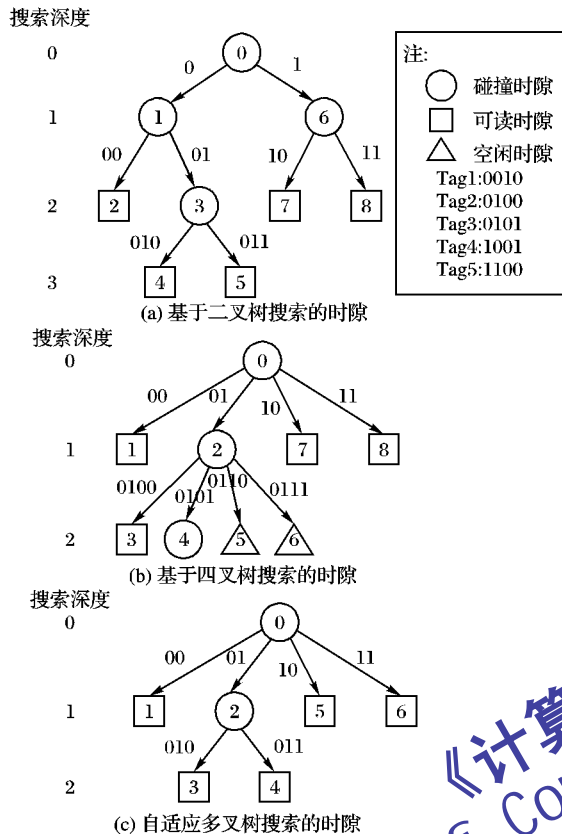


图1 时隙图

但自适应多叉树算法只能在四叉树或二叉树中选择搜索叉数,在标签数量较多的场合,识别效率仍然偏低。

## 2 基于启发式函数的多叉树防碰撞算法

在人工智能领域,为了减少搜索过程中的盲目性,一般采用启发式搜索方法。它利用所求解问题自身的某些特性,以指导搜索朝着最有利的方向进行。在RFID系统中,读写器要实现对所有标签的识别与通信,不能存在遗漏,因此防碰撞算法属于一种遍历性搜索方法。在防碰撞算法中,搜索叉数会直接影响到搜索效率和算法性能,因此可以借鉴启发式搜索的思想,利用搜索过程中获得的有用信息及时调整搜索叉数,以指导读写器利用最少的时隙数完成对所有标签的识别。

在启发式搜索中,用于评估节点从当前搜索状态 $x$ 到目标状态所付出的成本或代价的估计函数为启发式函数 $h(x)$ 。对于RFID系统,读写器可以根据标签响应中的碰撞比特信息来估计标签数量。因此,启发式函数定义为碰撞比特在标签响应中的比率:

$$h(x) = \frac{n_c(x)}{n} \quad (1)$$

其中: $x$ 为当前的碰撞时隙,读写器检测到发生碰撞的比特数为 $n_c(x)$ ,标签的响应长度为 $n$ 。由于 $0 \leq n_c(x) \leq n$ ,所以 $0 \leq h(x) \leq 1$ 。

首先分析启发式函数与时隙内发生碰撞的标签数量之间的关系。假设在当前时隙有 $N(x)$ 个标签的ID前缀符合查询码,从而发生碰撞,则启发式函数为:

$$h(x) = \frac{n \left[ 1 - \left( \frac{1}{2} \right)^{N(x)-1} \right]}{n} = 1 - \left( \frac{1}{2} \right)^{N(x)-1} \quad (2)$$

$$N(x) = 1 - \lg(1 - h(x)) \quad (3)$$

可见,利用启发式函数可以估计标签数量 $N(x)$ 。其中 $h(x)$ 越大,即碰撞比特所占的比例越大,则该时隙内未识别的标签数 $N(x)$ 较多,从当前状态到目标状态所需要付出的代价也越大。

假设某时隙内,标签数量为 $N(x)$ ,分配的叉数为 $L$ ,则 $m$ 个标签落在同一个分叉内的概率 $p(m)$ 满足二项式分布:

$$p(m) = \left[ \frac{m}{N(x)} \right] \left( \frac{1}{L} \right)^m \left( 1 - \frac{1}{L} \right)^{N(x)-m} \quad (4)$$

则可读时隙的概率 $p(1)$ 为:

$$p(1) = \frac{N(x)}{L} \left( 1 - \frac{1}{L} \right)^{N(x)-1} \quad (5)$$

系统的吞吐率为:

$$S = \frac{p(1)L}{L} = \frac{N(x)}{L} \left( 1 - \frac{1}{L} \right)^{N(x)-1} \quad (6)$$

希望吞吐率最大,所以对式(6)求 $1/L$ 导数:

$$\frac{dS}{d(1/L)} = N(x) \left( 1 - \frac{1}{L} \right)^{N(x)-1} - \frac{N(x)(N(x)-1)}{L} \left( 1 - \frac{1}{L} \right)^{N(x)-2} = 0 \quad (7)$$

解得:

$$L = N(x) \quad (8)$$

根据式(8)可知,要使系统吞吐率最大,分配叉数应与估计的标签数相等。但根据RFID系统的特点,分配的叉数应选择 $L=2^D$ , $D$ 为大于等于1的整数,根据式(3)和式(8)可得:

$$D = \left\lfloor \lg L \right\rfloor = \left\lfloor \lg N(x) \right\rfloor = \left\lfloor \lg \left[ 1 - \lg(1 - h(x)) \right] \right\rfloor \quad (9)$$

其中 $\lfloor x \rfloor$ 取小于 $x$ 的最大整数。但在实际的RFID系统中,标签ID是有限长度的。根据式(2)可知,随着标签数量的增加,发生碰撞的比特数会显著增加。当标签响应中每一比特都发生碰撞,即 $h(x) = 1$ 时,式(9)不成立。因此必须对 $h(x)$ 进行限制,令 $h(x) \leq (n-1)/n$ 。当ID长度 $n = 128$ 位时, $h(x) \leq 127/128$ ,根据式(9)可得, $D \leq 3$ 。即在这种RFID系统中,新算法会通过计算启发式函数的大小,自适应地选择二叉树、四叉树或八叉树进行搜索。

由于新算法是一种基于启发式函数的多叉树防碰撞算法(简称MHF算法)。如图2所示,算法步骤如下:

第1步 算法初始化,初始查询堆栈为空。

第2步 读写器读取堆栈中的查询码(当堆栈为空时,读写器发送零前缀的查询码)发送给标签。标签将查询码于自己的ID相比较,符合条件的标签响应。

第3步 读写器根据标签的响应状态进行相应的处理:

1)可读时隙。有且仅有一个标签响应,读写器完成对该标签的识别与通信。

2)空闲时隙。没有标签响应,无需识别。

3)碰撞时隙。根据式(1)计算启发式函数 $h(x)$ ,并根据式(9)计算搜索叉树,将标签响应的前 $D$ 个碰撞位,分别置为0或1,从而确定 $2^D$ 个新的查询码,并将其写入堆栈。例如:读写器的查询码为010,标签的响应为X0XX1时(其中X为发生碰撞的比特),如果计算得到的 $D = 1$ ,则进行二叉树搜索,新产生两个查询码0100和0101。如果计算得到 $D = 2$ ,则进行四叉树搜索,新产生4个查询码010000、010001、010100和010101。 $D = 3$ 时,依此类推。

第4步 读写器读取堆栈中的查询码,如果不为空,则返回到第2步继续搜索;否则,算法结束。

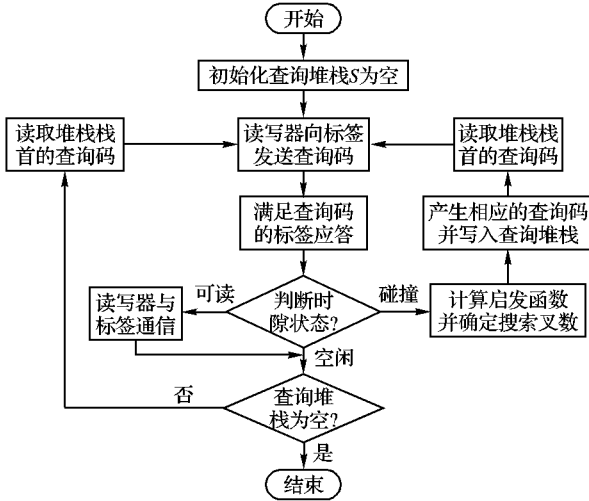


图2 基于启发式函数的多叉树防碰撞算法的流程

### 3 算法性能分析

通过时隙数和吞吐率的计算,对MHF算法的性能进行分析。

算法总的时隙数为不同叉数的时隙数之和:

$$T_{MHF} = \sum_{i=1}^D T_{2^i \text{ary}} \quad (10)$$

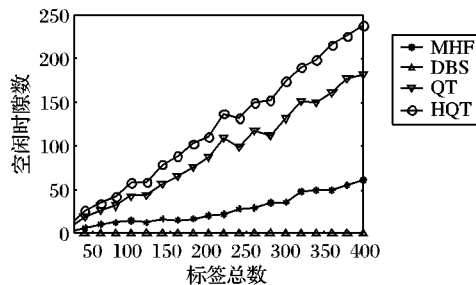
由于新算法的搜索叉数会随节点内标签数量、ID长度和碰撞比特等因素自适应地调整,因此很难精确计算时隙数。这里仅以标签长度为128位的RFID系统为例进行分析。

受式(9)的限制,  $1 \leq D \leq 3$ , 算法总时隙数为二叉树、四叉树和八叉树的时隙数之和,即:

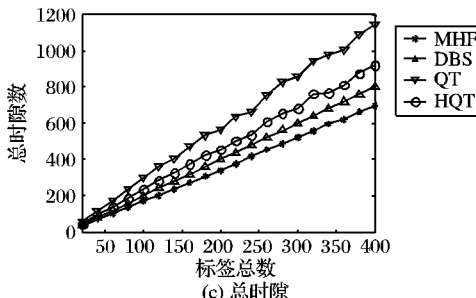
$$T_{MHF} = T_{2\text{ary}} + T_{4\text{ary}} + T_{8\text{ary}} \quad (11)$$

假设系统内有  $N$  个待识别标签,且标签的分布是均匀的。当搜索深度小于等于  $k$  时,子节点内标签数量满足  $N(x) \geq 8$  时,算法采用八叉树搜索,其中  $k = \lfloor \log_8 \frac{N}{8} \rfloor$ , 八叉树搜索的时隙数为:

$$T_{8\text{ary}} = \sum_{i=0}^k 8^i = \sum_{i=0}^{\lfloor \log_8 \frac{N}{8} \rfloor} 8^i \quad (12)$$



(a) 空闲时隙



(c) 总时隙

当搜索深度等于  $k+1$  时,子节点内标签数量应满足  $4 \leq N(x) \leq 8$ , 所以采用四叉树进行搜索,其时隙数为:

$$T_{4\text{ary}} = 4 \times 8^k = 4 \times 8^{\lfloor \log_8 \frac{N}{8} \rfloor + 1} \quad (13)$$

经过四叉树搜索后,每个节点内的标签数量平均减少为  $1 \leq N(x)/4 \leq 2$ , 当子节点内标签数量满足  $N(x) \leq 2$  时,使用二叉树搜索,故二叉树搜索的时隙数为:

$$T_{2\text{ary}} = 2T_{4\text{ary}} = 8^{\lfloor \log_8 \frac{N}{8} \rfloor + 1} \quad (14)$$

将式(12) ~ (14)代入式(11),可得:

$$T_{MHF} = T_{2\text{ary}} + T_{4\text{ary}} + T_{8\text{ary}} = 12 \times 8^{\lfloor \log_8 \frac{N}{8} \rfloor} + \sum_{i=0}^{\lfloor \log_8 \frac{N}{8} \rfloor} 8^i \quad (15)$$

即算法从初始状态搜索到某一个标签,一般经历的搜索深度为  $k+2$ , 其中深度从1到  $k$  为八叉树,  $k+1$  为四叉树,  $k+2$  为二叉树。

系统吞吐率为:

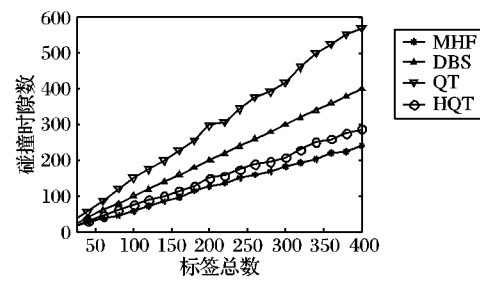
$$S_{MHF} = \frac{N}{T_{MHF}} = \frac{N}{12 \times 8^{\lfloor \log_8 \frac{N}{8} \rfloor} + \sum_{i=0}^{\lfloor \log_8 \frac{N}{8} \rfloor} 8^i} \quad (16)$$

### 4 仿真实验与分析

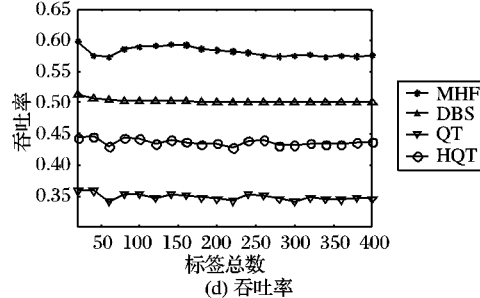
下面通过计算机仿真验证新算法的有效性,结果取100次实验的平均值。

图3分别为MHF、动态二叉树搜索<sup>[1]</sup> (Dynamic Binary Search, DBS)、查询树<sup>[7]</sup> (Query Tree, QT)和混合查询树<sup>[8]</sup> (Hybrid Query Tree, HQT)4种算法所需空闲时隙数、碰撞时隙数、总时隙数和吞吐率的比较。从仿真结果可以看到,QT和HQT算法分别基于二叉树和四叉树搜索,由于这两种算法没有利用碰撞比特信息,所以算法性能较差,适用于未采用曼彻斯特编码的RFID系统。其中HQT算法通过在标签中增加延迟器,有效地避免了一部分空闲时隙,所以性能优于QT算法。MHF和DBS算法都利用了碰撞比特的信息,其中DBS算法仅利用碰撞的首位信息,虽然该算法所需的空闲时隙最少(为0),但由于频繁出现碰撞,所以吞吐率较低。而MHF算法通过计算启发式函数,有效地利用所有的碰撞比特信息,所以具有最佳的搜索效率和性能。

图4分别为标签ID长度  $n$  为32, 64和128位对算法性能



(b) 碰撞时隙



(d) 吞吐率

图3 4种算法的性能比较



能的影响。标签 ID 越长,可利用的碰撞比特信息就越多,当标签 ID 长度为 128 时,算法会根据节点内标签的数量自适应地选择二叉树、四叉树和八叉树搜索进行搜索。而标签 ID 长度为 32 或 64 时,根据式(9), $1 \leq D \leq 2$ ,算法只能选择二叉树和四叉树,因此算法性能在标签长度为 128 位时,要优于 32 或 64 位,而在后两者情况下,算法的性能相近。

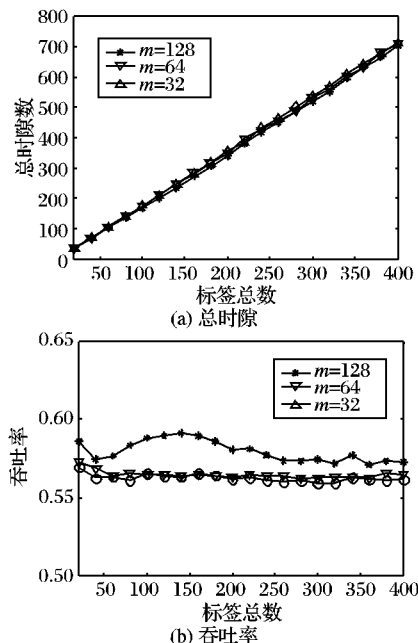


图4 标签 ID 长度对算法性能的影响

值得注意的是,在新算法中无论是启发式函数的计算还是搜索叉数的确定都是在读写器端实现的,不需要增加标签的硬件成本,因此新算法具有较高的实用性。

## 5 结语

本文提出了一种基于启发式函数的自适应多叉树防碰撞算法。该算法利用曼彻斯特编码可以识别碰撞位的特性,通

过定义和计算启发式函数,以估计节点内待识别标签的数量,可在不同搜索节点和深度,自适应地调整搜索叉数,从而有效地提高了算法的搜索效率。新算法克服了传统防碰撞算法的缺点,尤其在待识别标签数量较多场合,可有效地减少搜索和识别时间,提高 RFID 系统的吞吐率。

### 参考文献:

- [1] FINKENZELLER K. RFID Handbook: fundamentals and applications in contactless smart cards and identification[M]. Hoboken: John Wiley & Sons, 2003.
- [2] HWANG T - W, LEE B - G, KIM Y - S. Improved anti - collision scheme for high speed identification in RFID system [C]// Proceedings of First International Conference on Innovative Computing, Information and Control. Piscataway, NJ: IEEE Press, 2006: 449 - 452.
- [3] KIM J G. A divide-and-conquer technique for throughput enhancement of RFID anti-collision protocol [J]. IEEE Communications Letters, 2008, 12(6): 474 - 476.
- [4] EOM J B, LEE T J, RIETMAN R. An efficient framed - slotted ALOHA algorithm with pilot frame and binary selection for anti-collision of RFID tags [J]. IEEE Communications Letters, 2008, 12(11): 861 - 863.
- [5] JIHOON M, WONJUN L, SRIVASTAVA J. Adaptive binary splitting for efficient RFID tag anti-collision [J]. IEEE Communications Letters, 2006, 10(3): 144 - 146.
- [6] LAI Y - C, LIN C - C. A pair-resolution blocking algorithm on adaptive binary splitting for RFID tag identification [J]. IEEE Communications Letters, 2008, 12(6): 432 - 434.
- [7] CHOI J H, LEE D, LEE H. Query tree-based reservation for efficient RFID tag anti-collision [J]. IEEE Communications Letters, 2007, 11(1): 85 - 87.
- [8] RYU J, LEE H, SEOK Y. A hybrid query tree protocol for tag collision arbitration in RFID systems [C]// ICC'07: IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2007: 5981 - 5986.
- [9] 丁治国,朱学永,郭立,等. 自适应多叉树防碰撞算法研究[J]. 自动化学报, 2010, 36(2): 237 - 241.

(上接第 660 页)

荐算法,最后通过实验验证了该算法能够有效地提高推荐精度。下一步将考虑用户评分数据对云模型推荐精度的影响,如评分用户多少、两个用户共同评分的多少对云模型相似度的影响。此外,考虑如何综合利用各种推荐系统,解决推荐系统中的冷启动问题。

### 参考文献:

- [1] 刘建国,周涛,汪秉宏. 个性化推荐系统的研究进展[J]. 自然科学进展, 2009, 19(1): 1 - 15.
- [2] 孙小华. 协同过滤系统的稀疏性与冷启动问题研究[D]. 杭州: 浙江大学, 2005.
- [3] 张光卫,李德毅,李鹏. 基于云模型的协同过滤推荐算法[J]. 软件学报, 2007, 18(10): 2403 - 2411.
- [4] 邓爱林,朱扬勇,施伯乐. 基于项目评分预测的协同过滤推荐算法[J]. 软件学报, 2003, 14(9): 1621 - 1628.
- [5] 孙小华,陈洪,孔繁胜. 在协同过滤中结合奇异值分解与最近邻方法[J]. 计算机应用研究, 2006, 23(9): 206 - 208.
- [6] 余志虎,戚玉峰. 一种基于云模型数据填充的算法[J]. 计算机技术与发展, 2010, 20(12): 35 - 37.
- [7] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based collaborative filtering recommendation algorithms [C]// Proceedings of the

10th International World Wide Web Conference. New York: ACM Press, 2001: 285 - 295.

- [8] 白丽君,张永奎,陈鑫卿. 协作过滤研究概述[J]. 电脑开发与应用, 2002, 15(11): 2 - 3.
- [9] 李德毅,刘常昱. 论正态云模型的普适性[J]. 中国工程科学, 2004, 6(8): 28 - 34.
- [10] 李德毅,刘常昱,杜鹃,等. 不确定性人工智能[J]. 软件学报, 2004, 15(11): 1583 - 1594.
- [11] 张新香,刘腾红. 利用云模型改进基于项目的协同过滤推荐算法[J]. 图书情报工作, 2009, 53(1): 117 - 120.
- [12] 汪静,印鉴. 一种优化的 Item-based 协同过滤推荐算法[J]. 小型微型计算机系统, 2010, 31(12): 2338 - 2342.
- [13] 徐翔,王照法. 协同过滤算法中的相似度优化方法[J]. 计算机工程, 2010, 36(6): 52 - 57.
- [14] BREESE J, HECKERMAN D, KADIE C. Empirical analysis of predictive algorithms for collaborative filtering [C]// Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence. Madison: Morgan Kaufmann, 1998: 43 - 52.
- [15] KONSTAN J, MILLER B, MALTZ D, et al. Grouplens: Applying collaborative filtering to usenet news [J]. Communications of the ACM, 1997, 40(3): 77 - 87.