

基于生物记忆原理的智能伸展树模型

黄光球, 周 娇

(西安建筑科技大学管理学院, 西安 710055)

摘要: 针对普通伸展树节点移动方式单调、灵活性和实时性差的缺点, 提出一种基于生物记忆原理的智能伸展树模型。该模型利用短时记忆元不断衰减和更新的特点及短时记忆元与长时记忆、综合记忆的关系, 建立相应的数学模型, 借助节点访问度反映节点向前移动的步幅大小。实验结果表明, 智能伸展树具有良好的实时性和动态性, 具有一定的应用价值。

关键词: 记忆原理; 智能伸展树; 遗忘; 查找

Intellectualized Stretching Tree Model Based on Biological Memory Principle

HUANG Guang-qiu, ZHOU Jiao

(School of Management, Xi'an University of Architecture and Technology, Xi'an 710055)

【Abstract】 Aiming at the shortness of single node moving mode, poor flexibility and real-time in common stretching tree, this paper proposes a model of intellectualized stretching tree based on biological memory principles. The model establishes corresponding math model through the short memory cell's characteristics of decreasing and updating and the relationship of short memory, long memory and comprehensive memory. The node accessing range reflects the step number of node forward moving. Experimental result shows that intellectualized stretching tree can track the latest development in real time, dynamic character and has application value to some degree.

【Key words】 memory principle; intellectualized stretching tree; forgetting; search

1 概述

伸展树在每次访问树节点时, 将对树执行根本的调整。经常访问的节点频繁地被提升而成为根节点或它们不会离根位置太远; 不活跃的节点慢慢地被推得离根越来越远。伸展树中节点的搜索访问与搜索引擎具有相似的理论基础。高级搜索主要包括 2 个方面: 快速搜索和智能搜索。目前, 常见的搜索模型主要包括以下几个方面: 文献[1-2]将人工智能技术运用到信息检索中, 由于每次的检索都依照已有的规则, 因此缺乏灵活性; 文献[3]提出了一种快速字符串匹配搜索算法, 该算法缺乏对以前的字符串匹配信息的充分应用。文献[4]在扩展查询方面有了很大改进, 对已查询的信息却没有进行合理的归纳总结。另外, 其他实现搜索引擎的方法(轮廓查询技术、均值以及本体等)都具有类似的缺点: 系统移植性不好, 将具体搜索信息抽象成知识很困难, 搜索范围受已知知识的限制等。

针对上述问题, 本文提出一种基于生物记忆原理的智能伸展树模型, 大大提高了节点信息的访问效率。

2 基于生物记忆原理的智能伸展树模型

2.1 生物记忆原理

现代认知心理学认为, 记忆是通过识记、保持、再现等方式, 在人的头脑中积累和保存个体经验的心理过程^[5]。记忆包括瞬时记忆、短时记忆、长时记忆 3 个步骤。瞬时记忆通过一种或几种感官进入瞬时记忆存储器的是物质刺激的直接映像或“原始信息”, 其存储容量是感官的生理限度^[6]。长时记忆是短时记忆长期刺激的结果, 同时又会随着时间的推移而衰减。遗忘是指由于新事件的加入而将旧事件挤出记忆

库的方式。

2.2 基于记忆原理的智能伸展树模型工作原理

智能伸展树将记忆原理运用到伸展树的节点访问中, 其工作原理如图 1 所示。

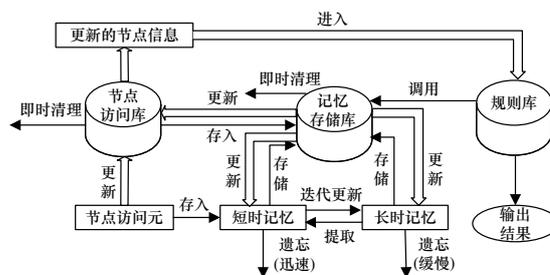


图 1 基于记忆原理的智能伸展树模型

节点访问元进入该系统时, 将节点访问信息存入节点访问库, 同时进入短时记忆。短时记忆元进入记忆存储库。短时记忆的迭代更新和自身的遗忘过程形成了对某节点信息的长时记忆, 此时该节点的长时记忆也进入记忆存储库。在短时记忆与长时记忆的共同作用下, 形成了对节点访问信息的综合记忆。记忆存储库中形成的长时记忆值和综合记忆值进入节点访问库, 更新节点信息。为了节省系统的空间, 节点

基金项目: 陕西省自然科学基金资助项目(2007E217); 陕西省教育厅专项基金资助项目(09JK524)

作者简介: 黄光球(1964—), 男, 教授、博士, 主研方向: 信息检索, 网络安全; 周 娇, 硕士研究生

收稿日期: 2009-12-30 **E-mail:** huangnan93@sohu.com

访问库和记忆存储库都有容量限制。更新后的节点信息进入规则库后，按照一定的规则进行节点的移动。

2.3 基于记忆原理的智能伸展树定义

定义1 智能伸展树(Intellectualized Stretching Tree, IST树)是一种基于生物记忆原理的伸展树，其节点具有记忆功能。在智能伸展树中，节点的记忆值是通过多个访问该节点的短时记忆元的遗忘、即时更新以及长时记忆的迭代更新来确定的。

定义2 节点访问度(Node Accessing Range, NAR)表示某时刻访问该节点向前移动的步数。访问越频繁，节点访问度增加幅度越大，记忆力越深刻。节点访问度越大，向前移动的步数越多，即移动速度越快；反之，移动速度越慢。

定义3 节点访问元(Node Accessing Unit, NAU)有如下6元组组成：

$$NAU = \{ID, T, M_L, M_C, NAR, tag\}$$

其中， ID 为访问节点号； T 为访问时间； M_L 为长时节点记忆值； M_C 为综合节点记忆值； NAR 为节点访问度； tag 为标志，若此刻访问该节点时，则 $tag=1$ ，记忆元既衰减又进行记忆值的更新，否则 $tag=0$ ，即记忆元只进行衰减。

定义4 短时记忆元(Memory Cell for Short, MCs)是记忆系统中传递的最小记忆单元。智能伸展树中的每个节点都包括若干个短时记忆元，但短时记忆元的数目不一定相同，它是根据对这个节点的访问次数来决定的。其结构如下：

$$MC_s = \{ID, T, time\}$$

其中， ID 为节点号； T 为节点访问时间； $time$ 表示访问该节点的次数。

定义5 规则库(Rules Database, RD)是规定节点伸展规则的库。规则库的结构定义如下：

$$RD = \{\max(M_C), \min(M_C), span, NAR\}$$

其中， $\max(M_C)$ 表示综合记忆的最大值； $\min(M_C)$ 表示综合记忆的最小值； $span$ 指分界的跨度值，这个值可以根据实际情况进行调整。

定义6 短时记忆集是有容量限制的集合，当访问某节点的短时记忆元超过一定数量时，则应进行即时清理，丢弃那些最旧的记忆元。设短时记忆集为 G_s ，容量限制为 L_s ，当前访问某个固定节点的短时记忆元个数为 n ， MC_s^i 为访问某个固定节点的第 i 个短时记忆元。

$$G_s = \{MC_s^1, MC_s^2, \dots, MC_s^i, \dots, MC_s^n\}, n \leq L_s$$

2.4 基于记忆原理的智能伸展树的性质

性质1 智能伸展树采用中序遍历的方式进行遍历。当搜索智能伸展树中的某个节点时，采用从左到右、自上而下的搜索方式，即高度越小，越靠左的节点离根节点越近。

性质2 当访问智能伸展树中的某个节点时，按照指定的伸展规则向上移动。而伸展规则主要根据节点访问度而变化。

3 智能伸展树节点记忆的遗忘和更新数学模型

3.1 节点记忆的遗忘数学模型

智能伸展树中对每个节点的记忆水平随着时间的推移而逐渐减弱，根据艾宾浩斯遗忘曲线可得衰减函数：

$$M_n = M_{n-1} e^{-a(t_2-t_1)}$$

其中， M_n 表示此刻的节点记忆值； M_{n-1} 表示上次该节点的记忆值； t_2-t_1 (即 Δt) 为这次节点访问时间与上次访问时间的时间差； a 为衰减调整变量。

根据实际情况，该衰减函数应满足条件：

$$\lim_{t_2-t_1 \rightarrow 0} M_{n-1} e^{-a(t_2-t_1)} = M_{n-1}, \quad \lim_{t_2-t_1 \rightarrow \infty} M_{n-1} e^{-a(t_2-t_1)} = 0$$

3.2 节点记忆的更新数学模型

在智能伸展树中，随着时间的推移，每个节点的记忆值都在实时地发生变化。对于访问频率大、时间间隔大的节点，其节点记忆值增加幅度逐渐升高；对于访问频率小、访问间隔大的节点，其节点记忆值逐渐减小，直至被遗忘。

某节点的每个短时记忆元的更新都满足遗忘规律，短时记忆值用 M_s 表示，短时记忆元的更新数学模型如下：

$$M_s(T-t') = N e^{-\varepsilon(T-t')} \quad (1)$$

其中， N 为节点记忆常量； ε 表示短时调整常量； T 表示当前时刻； t' 表示访问该节点的短时记忆元的时刻；短时记忆元的初始值为 N 。

令 $x = \sum_{i=1}^n M_s(T-t_n)$ ，则长时记忆的更新模型如下：

$$M_L = M_0 e^{-\eta(T-t_0)} + f(x)$$

其中， $f(x)$ 称为长时记忆迭代增长速度函数。在不同条件下， $f(x)$ 可有多种表示形式，用于控制迭代增长速度。由于函数 $f(x) = \lg \sqrt{1+x}$ 具有很好的平稳性，本文采用该函数作为长时记忆迭代增长速度函数。因此，其更新模型表示如下：

$$M_L = M_0 e^{-\eta(T-t_0)} + \lg \sqrt{1 + \sum_{i=1}^n M_s(T-t_n)} \quad (2)$$

其中， M_0 为长时记忆的初始值； η 为长时调整常量； t_0 为某节点第一个记忆元的访问时间； t_n 为访问某节点的第 n 个短时记忆元的时刻。

节点的综合记忆值包括访问该节点的所有短时记忆元的衰减更新以及长时记忆的迭代更新。其数学模型为

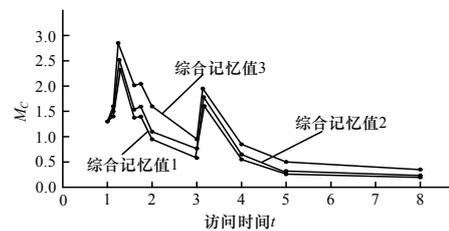
$$M_c = \sum_{i=1}^n M_s(T-t_n) + M_L \quad (3)$$

由式(1)~式(3)可知，最终的节点综合记忆值更新模型为

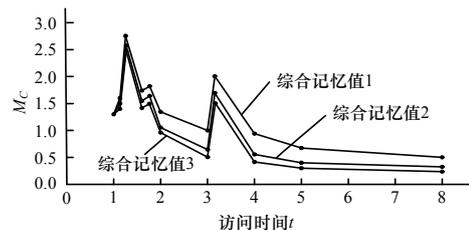
$$M_c = N \sum_{i=1}^n e^{-\varepsilon(T-t_n)} + M_0 e^{-\eta(T-t_0)} + \lg \sqrt{1 + N \sum_{i=1}^n e^{-\varepsilon(T-t_n)}} \quad (4)$$

3.3 模型参数与记忆值的变化关系

智能伸展树能够根据每个节点的记忆值来确定节点访问度，模型中的参数与记忆值的变化关系如下：(1)长时调整变量 η 不变，短时调整变量 ε_1 、 ε_2 和 ε_3 满足关系： $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$ 。(2)短时调整变量 ε 不变，长时调整变量 η_1 、 η_2 和 η_3 满足关系： $\eta_1 < \eta_2 < \eta_3$ ，记忆值变化情况如图2所示。



(a) η 不变时 ε 增加时的综合记忆值变化情况



(b) ε 不变时 η 增加时的综合记忆值变化情况

图2 模型参数与记忆值的变化关系

4 基于记忆原理的智能伸展树模型实现

4.1 智能伸展树模型实现流程及记忆库的更新

智能伸展树实现流程如图3所示。记忆库的更新和清理在整个记忆系统中十分重要，包括所有节点的短记忆元和长时记忆的衰减更新过程以及记忆库中记忆元的清理工作。这些都是整个记忆模型能较好地完成指定任务的重要保证，对记忆库系统更新及清理的关系如图4所示。

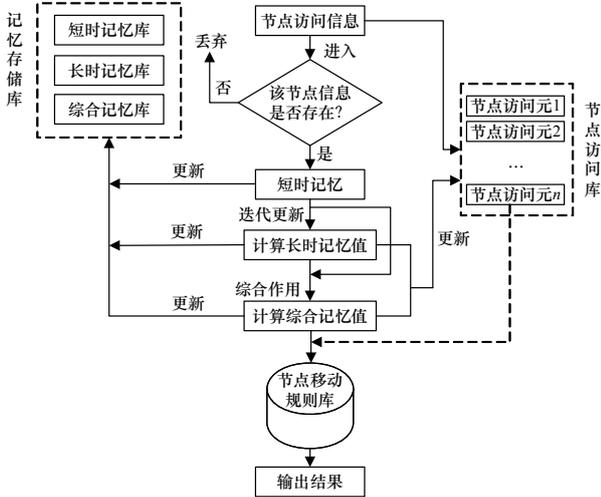


图3 智能伸展树模型实现流程

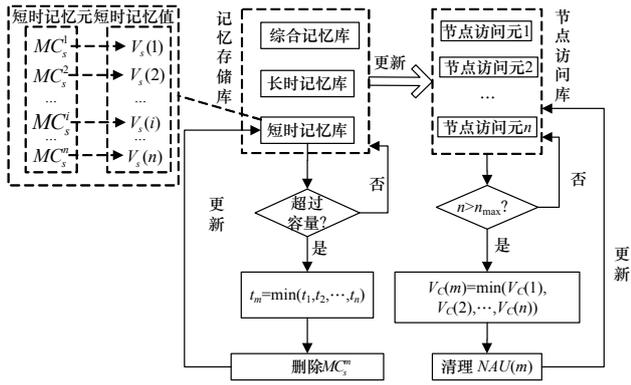


图4 两大记忆库更新及清理关系

4.2 智能伸展树模型算法实现步骤

根据以上基于记忆原理的智能伸展树模型的流程描述，现将实现算法描述如下：

Step1 初始化 $M_L, M_C, N, \varepsilon, \eta, M_0$ 。

Step2 将节点访问元输入系统，记录其节点访问信息，更新 MC_S 。如果 $n < L_S$ ，则执行 Step3；否则，执行 Step5。

Step3 将 MC_S 作为输入，按照式(2)计算并更新该节点的长时记忆值，即 M_L 。

Step4 利用式(4)计算并更新综合记忆值，即 M_C 。如果某节点综合记忆值超过某一固定值 N' ，则该值不再衰减，即使以后不再出现该节点，其记忆强度依然不变。

Step5 更新清理记忆库，包括以下 2 种记忆库的更新清理工作：

Step5.1 记忆存储库。

将访问该节点时间最早、最旧的短记忆元删除，重新计算长时记忆及综合记忆值，返回 Step3；

Step5.2 节点访问库。

由系统选择综合记忆值最小的一个节点记忆元，然后删

除，返回 Step3。

Step6 重复执行 Step2~Step5，直到节点访问元输入完毕。

Step7 将计算出的智能伸展树中所有节点的综合记忆值输入规则库，按照特定的规则进行节点的伸展操作，得到最新的伸展树。

4.3 智能伸展树与普通伸展树的对比

普通伸展树节点伸展方式单调，每访问节点一次，向上伸展一步(如 zig-zig, zig-zag)，有很大的局限性。而智能伸展树的性质决定了其伸展步数最少为一步，当频繁访问某节点时，节点的记忆性使得记忆强度增加，节点访问度提高，此时的伸展步数相应增加，该节点到达根节点的时间缩短；最差的情况是某个节点一直未被访问，这时该节点的伸展情况与普通伸展树相同。

进一步讨论，可以将节点当作不同的信息，当搜索某个信息的频率高时，该信息向前移动的步数增加，更容易搜索到；那些很少搜索或基本不搜索的信息逐渐移到数据库的后面。这种方式使得每个信息具有记忆性，大大提高搜索效率。

5 实例研究

5.1 实验数据描述

假设智能伸展树如图5所示，观察节点访问时的记忆值变化情况。根据经验，取实验中的参数值： $N=0.2, \varepsilon=4, \eta=0.5, M_0=1, N'=2.9$ 。当 $M_C \in (0,1)$ 时， $NAR=1$ ；当 $M_C \in (1,2)$ 时， $NAR=2$ ；依次类推。节点访问格式为：(节点访问号，访问时间 m)。

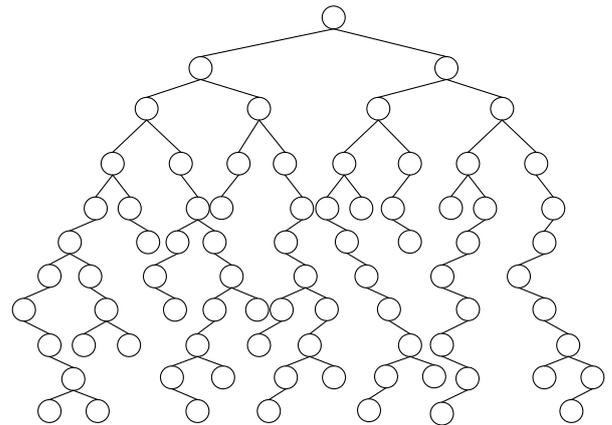


图5 智能伸展树

节点访问情况如下：(3,1.00) (24,1.01) (66,1.02) (76,1.03) (24,1.04) (3,1.11) (66,1.12) (66,1.13) (60,1.14) (60,1.15) (60,1.16) (60,1.17) (60,1.18) (60,1.19) (3,1.20) (3,1.21) (3,1.22) (3,1.23) (3,1.24) (3,1.25) (3,1.30)

5.2 智能伸展树节点访问结果分析

按照上述参数输入程序后所得结果绘制成图形，得到 4 个节点的长时记忆值和综合记忆值的变化趋势如图6、图7所示。

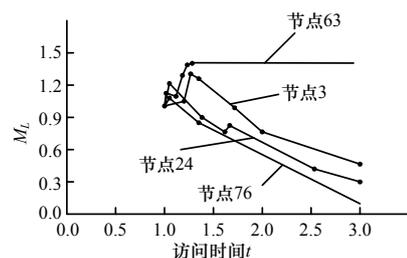


图6 长时记忆曲线

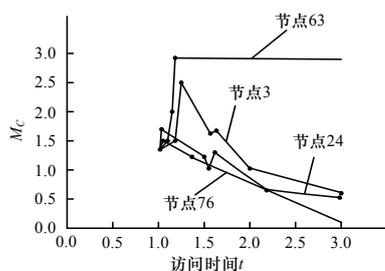


图7 综合记忆曲线

对于曲线趋势的解释如下:

(1)由于访问节点3比较频繁,其记忆值普遍较高;访问节点24的频率比较均匀,但访问间隔时间较长,整体呈下降趋势;在某段时间内访问节点63相当频繁,其综合记忆值超过2.9,因此记忆值不再衰减;节点76只被访问了2次就不再访问,在记忆衰减作用下,曲线下降趋势很快。

(2)从长时记忆曲线来看,总体变化较平缓。当访问该节点比较集中时,曲线上升幅度较大;反之,当访问该节点间隔时间较长时,由于短时记忆元遗忘值较大,多个短时记忆元迭代起来,使得长时记忆曲线有下降趋势。

(3)综合记忆曲线比长时记忆曲线变化幅度大,但总体变化趋势与长时记忆基本相同。当访问节点的频率较大时,一方面由于访问该节点的所有短时记忆元遗忘相对减缓;另一方面由于长时记忆的作用,使得综合记忆曲线上升趋势明显;反之,下降趋势明显。

5.3 智能伸展树与普通伸展树效率对比

根据本实验所得节点访问量,将智能伸展树与普通伸展树进行效率对比,如表1所示。

表1 智能伸展树与普通伸展树效率对比

节点号	普通伸展树	智能伸展树
节点3	5(zag-zig,zig-zag,zig-zig,zig-zig,zig-zig)	3((zag-zig,zig-zag),(zig-zig,zig-zig),zig-zig)
节点24	5(zig-zag,zig-zig,zag-zag,zag-zag,zig-zig)	3((zig-zag,zig-zig),(zag-zag,zag-zag),zig-zig)
节点63	3(zig-zig,zig-zag,zag-zag)	2((zig-zig,zig-zag),zag-zag)
节点76	5(zag-zig,zag-zag,zig-zig,zag-zag,zag-zag)	3((zag-zig,zag-zag),(zig-zig,zag-zag),zag-zag)

由表1可以看出智能伸展树的节点向根节点移动的速度比普通伸展数要快将近2倍,随着伸展树层数的增加,智能伸展树的优势将更加明显。

6 结束语

基于生物记忆原理的智能伸展树模型将记忆原理与伸展树融合在一起,为用户提供了一种具有实时性的节点移动方式。该模型提高了伸展树的节点访问效率,有效节省系统开销,同时也为搜索技术提供了一个新的思路。

参考文献

- [1] 邱均平,余以胜.基于知识库系统的智能搜索引擎研究[J].情报科学,2006,24(3):413-416.
- [2] 张玉红,王华,蒋一峰,等.基于知识库系统的中文智能搜索引擎[J].计算机应用,2008,28(S2):61-64.
- [3] 谭强,沈雁.一种快速字符串匹配搜索算法[J].计算机应用,2003,23(1):202-204.
- [4] 梁晓辉,崔杜武.基于对等网络的智能信息搜索模型[J].计算机工程,2008,34(2):166-168.
- [5] 陈永明,罗永东.现代认知心理学[M].北京:团结出版社,1989.
- [6] 邓立平.关于短时记忆存储库容量的甄别[J].云南师范大学学报,1997,29(3):62-63.

编辑 索书志

(上接第163页)

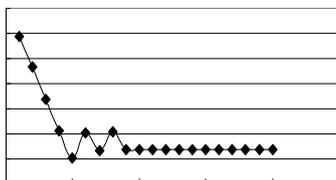


图1 少量虚假信息时算法的比较

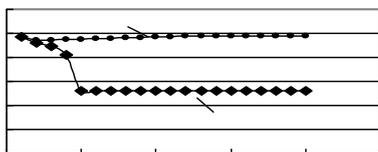


图2 大量虚假信息时算法的比较

可以看出,在少量的恶意推荐信息下2种算法表现是一样的(图1),都能够过滤掉恶意的推荐信息。但在大量的恶意推荐时(图2),改进的算法对信任值的影响很小,信任值基本上没有太大的变化,而改进后的算法可以较好地过滤掉虚假推荐。

5 结束语

信任模型本身的安全问题最终决定了模型的可用性。本文重点关注虚假推荐问题,考虑节点推荐的信任值的标准差,进而得到虚假过滤算法,能够有效地抑制大规模节点提供的

虚假推荐信息,提高信任模型的安全性。

参考文献

- [1] 林欣,赵铁柱.基于P2P网络的信任和信誉评估模型[J].计算机工程,2009,35(2):133-135.
- [2] Dellarocas C. Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior[C]//Proc. of ACM Conference on Electronic Commerce. [S. l.]: ACM Press, 2000: 150-157.
- [3] Audun J. A Survey of Trust and Reputation Systems for Online Service Provision[J]. Decision Support Systems, 2007, 43(2): 618-644.
- [4] Chen M, Singh J. Computing and Using Reputations for Internet Ratings[C]//Proc. of the 3rd ACM Conference on Electronic Commerce. [S. l.]: ACM Press, 2001.
- [5] Buchegger S, Boudec J Y L. A Robust Reputation System for Mobile Ad-hoc Networks[C]//Proc. of the 2nd Workshop on the Economics of Peer-to-Peer Systems. Lausanne, Switzerland: [s. n.], 2003.
- [6] Jigar P, Luke T W T. A Probabilistic Trust Model for Handling Inaccurate Reputation Sources[C]//Proc. of the 3rd International Conference on Trust Management. Paris, France: [s. n.], 2005: 193-209.
- [7] Mui L, Mohtashemi M, Szolovits C A P, et al. Ratings in Distributed Systems: A Bayesian Approach[C]//Proc. of the Workshop on Information Technologies and Systems. [S. l.]: MIT Press, 2001.

编辑 索书志

