

基于可信理论的 Web 教育门户网站可用性评估模型

王 杨, 黄亚坤, 陈付龙, 赵传信

(安徽师范大学 数学计算机科学学院, 安徽 芜湖 241000)

摘 要: Web 教育网站的可信性和可用性是教育门户网站评价的重要内容. 从可信、可用角度, 提出了一种基于可信理论的 Web 教育门户网站可用性评估模型. 首先给出了 Web 教育资源的相关描述, 然后提出了基于可信理论的 Web 教育门户网站的评估模型(UEMBTT)和评估算法. 最后仿真实验表明, 该模型能够有效地对 Web 教育门户网站资源进行评估, 有助于提高 Web 教育网站可信性和可用性.

关键词: 可信理论; 教育门户网站; 信息熵; 可用性评估

中图分类号: G203

文献标识码: A

文章编号: 1000-7180(2013)03-0089-04

Usability Evaluation Model Based on Trusted Theory of the Web Educational Portal Web Site

WANG Yang, HUANG Ya-kun, CHEN Fu-long, ZHAO Chuan-xin

(College of Mathematics and Computer Science, Anhui Normal University, Wuhu 241000, China)

Abstract: Web credibility is an important factor to assess the availability of the portal. In this paper, we propose a usability evaluation model based on the credibility of the theory of web educational portal. Firstly, we give the description of the network of web educational resources. And we define the credibility of the evaluation mechanism from a trusted point of view. And the evaluation algorithm was presented. Finally, the simulation results show that this model can assess the resource of web educational portal effectively and it has a certain significance to improve the site availability.

Key words: credibility theory; education web portal; information entropy; credibility evaluation

1 引言

随着网络技术的发展, 高校门户网站不仅是向教育用户提供快捷的教育信息的载体, 同时也是展示高校特色和综合实力的一条重要途径^[1]. 传统的 Web 网站的可用性主要是指 Web 用户在网络环境下利用其完成具体任务时所达到的有效性、效率和满意度^[2].

近年来, 国内外的研究者对门户网站可用性研究工作主要有两类: 一是采用专家评价及问卷调查获取原始数据, 并构建评价指标体系^[3-4]; 二是通过收集网络用户对 Web 站点的访问数据进行分析, 即研究访问日志和网站结构, 利用可视化技术实现可用性研

究^[5-8]. 事实上, 随着网络教育资源存储量的指数级增长、网络教育资源污染现象(如冗余、空链接、虚假信息等的加剧以及网络黑客的攻击等问题的出现, 需要采用不同于传统的网站可用性评估方法对 Web 教育门户网站的安全可用性展开研究. 由于可信计算近年来一直是国际安全领域中的研究热点^[4], 本文基于可信理论提出一种 Web 教育门户网站可用性评估模型(Usability evaluation model based on Trusted theory of the web educational portal, UEMBTT), 对教育门户网站进行建模和评价.

2 相关概念的描述

2.1 Web 教育门户网站

收稿日期: 2012-06-20; **修回日期:** 2012-07-22

基金项目: 中国博士后基金(20100480701); 教育部人文社科青年基金项目(11YJC880119)

定义 1 主体(Subject). 是指 Web 教育门户网站中的教育资源实体, 这里用符号 s 表示.

定义 2 客体(Client). 指网络环境中对 Web 教育门户网站主体进行访问的对象实体, 用符号 c 表示.

定义 3 访问(Access). 是客体对象对主体信息执行的操作; 用 $\text{Access}(c, s)$ 表示客体对象对主体进行的一次访问.

定义 4 Web 教育门户网站(Web Education Portal Site). 指通过统一的协议对网络区域内的教育资源动态和安全进行管理的一组网络实体集合.

2.2 基于可信理论的 Web 教育门户网站可用性

对于 Web 教育门户网站可用性的研究, 本文主要从 Web 教育资源的直接可信度与 Web 用户的满意度两个方面进行观测.

定义 5 直接可信度.

假设 d_1 是主体资源链接的其他一级直接主体资源链接数目; d_2 主体资源链接的其他二级间接主体资源数目;

则直接可信度定义为对于该主体资源第 $n+1$ 个成功数的概率 $d(c, s)$, 该概率服从 Beta 分布, 据密度函数可以推导出:

$$d(c, s) = \begin{cases} \frac{d_1 + 1}{d_1 + d_2 + 2}, & \text{if } (d_1, d_2 > 0) \\ 0, & \text{else} \end{cases} \quad (1)$$

根据式(1), 当 d_1 为 0 时, 即该主体资源为孤立资源; d_2 为 0 时, 该资源为路由转发器, 这里无需进行可信度评估.

定义 6 满意度(Satisfaction). 当一次访问行为 $\text{Access}(c, s)$ 完成后提交的满意程度; 用符号 $A(c, s)$ 表示, 由于满意程度的主观性与不确定性, 我们提出基于信息熵理论的满意度函数.

假设满意评估指标 $e \in [0, 1]$ 之间的随机变量, $p(e)$ 表示取值为 e 的概率, 则构成如下函数:

$$A(c, s) = \begin{cases} 0, & \text{satisfactory} \\ 1, & \text{unsatisfactory} \\ -\int_e p(e) \log p(e) de, & \text{others} \end{cases} \quad (2)$$

这里采用信息熵的不确定度量方法来评估主体提供资源的质量, 0 表示本次访问中客体 c 对主体 s 完全不满意, 1 表示客体 c 对主体 s 完全满意, 值越大表示满意度越高.

由于恶意的客体对可信任的主体提供的服务仍

然可以做出错误反馈, 为了减少由错误反馈带来的影响, 引入反馈可信度的定义.

定义 7 反馈可信度. 客体对主体资源完成一次访问操作, 会对主体资源进行信息的评价反馈, 包括提供信息的真实性、准确度等, 用符号 CF 表示. 一般来说, 反馈可信度的影响因素有以下两种:

(1) 当前客体对主体访问的频繁次数, 访问的次数多, 其反馈可信度较高;

(2) 若前客体对主体的反馈信息与其他客体对主体反馈信息的相似度为一致性, 则一致性越高时, 反馈可信度越高.

反馈可信度直接反应主体历史访问的评价, 为了抑制恶意节点的攻击, 我们规定反馈可信度的评价建立在直接可信度与满意度评价的基础上. 下面给出 CF 的计算公式:

我们定义访问密度 Num 来表示客体对主体访问次数对反馈可信度的影响; 假设 N_c 表示客体对主体访问的次数, N_s 表示主体被访问的总次数, 则有:

$$\text{Num}(c, s) = \frac{N_c}{N_s} \quad (3)$$

我们引入一致性因子 $\text{Sim}(c, s)$ 的描述用当前客体一次访问主体的反馈信息 TS_c 与所有客体一次访问主体的反馈信息 TS_s 比较可得:

$$\text{Sim}(c, s) = \frac{TS_c}{TS_s} \quad (4)$$

进一步地, 我们定义反馈可信度 CF 计算公式如式(5):

$$CF(c, s) = \alpha \cdot \text{Num}(c, s) + (1 - \alpha) \cdot \text{Sim}(c, s) \quad (5)$$

式中, α 给出了两者之间的权重, 且当 $CF(c, s)$ 满足 $CF(c, s) \geq \frac{d(c, s) + A(c, s)}{2}$ 时, 定义的反馈可信度能够有效的进行评估.

定义 8 全局可信度.

对于教育网站中的任意主体资源节点 i 的全局可信度 AC_i 为:

$$\text{令 } p = \beta \cdot d_j(c, s) + (1 - \beta) \cdot A_j(c, s) \quad (6)$$

$$\text{则 } AC_i = \sum_{j \in n} (CF_j(c, s) \times p) \times T_j \quad (7)$$

其中, n 表示访问过主体资源 i 的所有客体集合. $A_j(c, s)$ 为客体 j 的满意度指标, $CF_j(c, s)$ 是客体 j 的反馈可信度, β 是影响因子, T_j 是活动周期, 引入 T_i 为了解决如下问题: (1) 资源的可信度与时间关联问题; (2) 信任衰减问题. 由此给出了活动周期 T_i 的公式:

$$T_i = \frac{1}{1 + \lambda \cdot \Delta t}$$

(8)

Δt 是可信度评估时间与被最近一次被访问时间差,为了更好的评估 Δt 的动态性 λ 为动态调节,可根据实际情况设置不同的值。

3 基于可信理论的 Web 教育门户网站可用性评估模型(UEMBTT)

当前的网络环境下,不同的教育门户网站通过一定的协议构成了开放的 Web 教育资源网。由于用户实体在访问客体资源时,两者各自所属域的不同对访问可信度评估的机制有一定的影响因素,下面给出可信度评估算法(Credibility Assessment Algorithm,CAA)。由于每个教育门户网站都是以某个实体作为中心,并进一步形成层次化的拓扑结构,由此我们提出信任代理的定义。

定义 9 信任代理(Trust Agent)信任代理负责本门户网中客体资源的可信度以及反馈可信度等信息的更新,用符号 TA 表示。

信任代理在一次访问行为 Access(*c*,*s*)过程中执行的动作主要有 2 个:(1)更新门户网中主体资源的不同信任度,并对陈旧的信息进行筛选;(2)控制恶意资源的攻击,保护门户网的安全性。

3.1 可信度评估算法(CAA)

在 Web 教育门户网站访问模型中,域中节点提供资源,客体在访问主体资源时,我们从直接可信度、满意度及反馈可信度来进行评判,若所属门户网不同,则对可信度评估同时还需要考虑门户网自身的信任度。基于自动机理论给出站点内和站点间访问示意图,如图 1 所示。

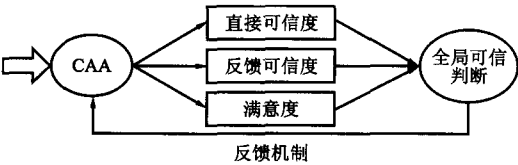


图 1 站点内访问示意图

客体资源在网间访问时,需要判断本门户网的信任度;复杂网络中,聚集系数用来衡量中心化程度,一般来说,当网络中心化程度越大,其可信度越大,所以本文采用聚集系数来作为门户网 *i* 的信任度 w_i ,计算方法如下

$$w_i = \frac{k_i(k_i - 1)}{2e_i}$$

(9)

其中, e_i 是全局网络中实际存在的边数, k_i 是门户网 *i* 的信任代理 TA_i 的度数, $0 < w_i < 1$,若 $w_i =$

1,则网络中任意两两直接均有连接。

引入自动机反馈机制有助于进行动态适应性的评价,下面给出可信度的算法描述。

3.2 算法的描述

算法:可信度评估算法(CAA)

Input:需进行访问的客体对象 C_i 与主体资源 S_j ;当前 Web 教育资源网络 *G* 中的相关信息。

Output:此次访问操作是否成功

Begin

$t \leftarrow \Delta t$; // 对 *t* 进行赋值

If($DC_i == DS_j$)// C_i 与 S_j 属于同一门户网

{ Access(C_i, S_j);

 If($AC(C_i, S_j) > w_i$)

 { Visit(C_i, S_j); // 完成访问

 Renew(DS_j, G); // 代理进行更新

 }

}

Else // C_i 与 S_j 不属于同一门户网

{ $w \leftarrow w_i$; // 对门户网的可信度判断

 If($w < \delta$)

 Break;

 Else

 Access(C_i, S_j);

}

End

算法中 δ 作为教育门户网可信度的判断标准,一般 δ 的取值为 Web 教育资源网络中所有门户网的可信度均值,当大于该值时,则定义为可信门户网站。

3.3 算法性能分析

CCA 主要思想是客体进行访问时,通过对主体资源从直接可信度、满意度和反馈可信进行评估。直接可信度是对当前主体的直接评估,算法时间复杂度为 $O(1)$;满意度计算则是从间接的角度,更全面的进行评估,算法的时间复杂度也可控制在 $O(1)$ 内。由于对主体的历史信息评估需要通过对所有与当前主体资源进行访问过的客体信息进行分析评估,因此,最后通过全局判断给出访问操作的结果。

4 仿真实验与结果分析

为了验证本文提出的模型和算法的性能,本文在 matlab 仿真软件中对模型中 CCA 算法进行仿真验证。模拟实验采取 3 个门户网站和 100 个用户,每个用户分别进行 10 次的交互,对进行可信度评估。

实验 1 有效性评估

假设 3 个教育服务门户网可信度分别为高、低

和随机网络,令初始信任度均为 0.5.规定所有用户为诚信用户.图 2 给出了门户网可信度的随时间步长变化的趋势图.

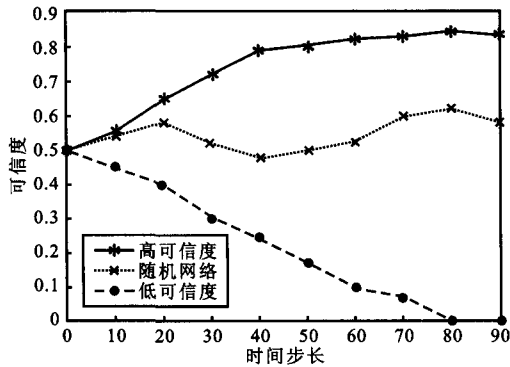


图 2 门户网可信度变化示意图

图 2 表明,随时间步长的变化,可信度高的门户网由于提供的服务质量好使得反馈信任度上升,网站的信任度也逐渐上升,可信度低的网门户网则相反逐渐降低直至为 0,没有价值存在;随机网络的信任度则是不稳定的进行变化,更加说明了 UEM-BCT 模型能够反应 Web 教育资源网络的有效性.

实验 2 调节因子的影响

为了对 UEMBCT 模型进行更好的评估,本实验仿真了全局可信度中直接可信度与满意度的比例影响,并给出了更新活动周期对门户网站可信度的影响,对模型的稳定性进行分析.实验结果如图 3 所示.

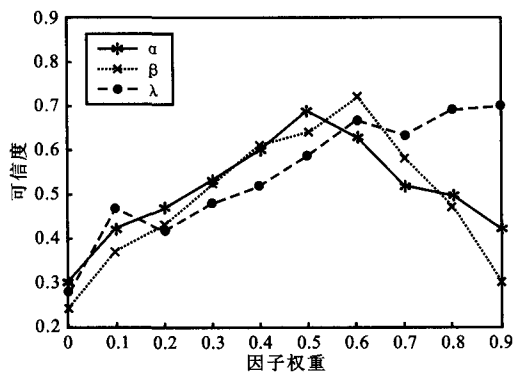


图 3 调节因子的影响

图 3 反应随着因子 α 增加,可信度呈现出先增加后降低的趋势,说明直接可信度与满意度控制在 0.5 范围可信度最高; β 增加对域可信度也呈现出先增加后降低的趋势,当在 0.6 时可以达到最大的可信度; λ 的增加使得活动周期减小,在一定范围内增大门户网站的可信度.

5 结束语

针对 Web 教育门户网站的可信度和可用性评估问题,本文从信任角度提出了一种 Web 教育门户网站可用性评估模型.本文提出的模型不仅有效地对 Web 教育资源进行评价,同时也能抑制恶意用户的攻击行为.我们下一步的工作主要是研究 Web 教育门户网站的可用性增强机制.

参考文献:

[1] 袁红. 基于网络内容分析的高校门户网站可用性测评[J]. 现代图书情报技术, 2010, (10): 70-75.

[2] 张越, 张硕文. 中国省会城市政府门户网站测评分析[J]. 山西财经大学学报, 2009, 31(1): 33-37.

[3] 刘爱华. 高校图书馆网站可用性评价研究[J]. 情报科学, 2008, (11): 1610-1610.

[4] 左双勇, 陈光喜, 丁勇, 等. 基于模糊决策理论的主观信任模型评估[J]. 微电子学与计算机, 2011, 28(9): 89-92.

[5] Munzner T, Burchard P. Visualizing the structure of the World Wide Web in 3D hyperbolic space [C]// Proceedings of VRML'95, San Diego, CA: IEEE, 1995.

[6] Shiozawa H, Matsushita Y. WWW visualization giving meanings to interactive manipulations [C]// Proceedings of Advances in Human Factors/Ergonomics Z1B (HCI International 97). [s. l.]: IEEE, 1997.

[7] Frecon E, Smith G. WebPath: a three dimensional Web history [C]// Proceedings of IEEE Symposium on Information Visualization (InfoVis 98), Chapel Hill, NC: IEEE, 1998.

[8] Chi Ed H, Pitkow J, Mackinlay J D, et al. Visualizing the evaluation of Web ecologies [C]// Proceedings of ACM SIGCHI'98, ACM Conference on Human Factors in Computing Systems. [s. l.]: ACM, 1998.

作者简介:



王 杨 男, (1971—), 教授. 研究方向为教育技术、机器学习、社会网络.