

DOI: 10.7652/xjtuxb201504012

面向社交网络中多背景的信任评估模型

孟宪佳, 马建峰, 王一川, 卢笛

(西安电子科技大学计算机学院, 710071, 西安)

摘要: 针对信任由背景敏感性导致的在社交网络中难以有效评估用户间接信任值的问题,提出了一种面向多交互背景的间接信任评估模型(multiple-context trust evaluation, MCTE)。该模型利用相关性概念,通过对各背景下网络结构和用户信任关系的综合分析,建立覆盖在信任网络之上的相关性网络,进而利用交互背景的相关性计算跨背景用户的间接信任值。模型避免了多背景以及疏散网络中间接用户信任路径难以寻求,以及信任衰减对评估的影响,有针对性地为用户组建立相关网络,保证了预测的准确性及合理性。对真实社交网络的实验结果表明,MCTE模型不仅可以计算单一背景下用户的间接信任值,更适用于多交互背景下用户信任值的预测。与已有模型相比,评估准确度有较大的提高。

关键词: 社交网络;信任评估;多背景

中图分类号: TP309 **文献标志码:** A **文章编号:** 0253-987X(2015)04-0073-05

MCTE: a Trust Evaluation Model for Multiple Context in Social Networks

MENG Xianjia, MA Jianfeng, WANG Yichuan, LU Di

(School of Computer Science and Technology, Xidian University, Xi'an 710071, China)

Abstract: A novel multiple-context trust evaluation model, named MCTE, is proposed for social networks to deal with the problem that it is difficult to effectively evaluate indirect trust value between users due to the context sensitivity of trust. A relevance network on top of trust networks is established by taking advantages of the relevant concepts, and through a comprehensive analysis of network structures and user's trust relationship in each context. Then user's indirect trust across the context is calculated by using the relevance of context. The Model avoids the effect of the trust attenuation on evaluation and the problem that the trust path between indirect users is difficult to find in multiple context and sparse networks, so that relevance networks of user groups can be built, and the evaluation accuracy and reasonableness are ensured. Experimental results on real social networks shows that the MCTE model can compute the indirect trust value in one single context, and is suitable for the prediction of user's indirect trust in multiple context. Comparison with an existing model shows that the evaluation accuracy of the proposed model improves a lot.

Keywords: social networks; trust evaluation; multiple context

线上社交网络设计的初衷是让网络成员能够与更多的陌生成员进行交互,从而开拓新的商业、政

治、交际的可能性^[1-3],然而由于其开放性以及广泛的流行性,用户的隐私以及数据安全问题成为其达

收稿日期: 2014-10-15。 作者简介: 孟宪佳(1985—),男,博士生;马建峰(通信作者),男,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(61174146,61221063)。

网络出版时间: 2015-03-04

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20150304.1641.002.html>

到预期目的的重大阻碍之一^[4]。建立一个让社交网络成员公开、自由、安全分享信息的平台是必须要面对的问题^[5-6]。社交网络成员利用信任关系建立交互并处理复杂的系统和信息的不确定性所带来的各种威胁是解决这一问题的有效途径。在这种框架下,信任是社会成员之间所有交互的基础^[7],也是分享信息和开发新的社会关系的决定因素^[8]。

近年来,众多学者从事于社交网络中对未知用户的信任评估研究。Andersen 等提出了基于平均理念的评估方案^[9-10],这些方案适用于各种网络环境中,然而对于社交网络个体的特殊性不能有效地描述。Golbeck 等提出了基于信任传递性的评估模型^[11-12],这种模型利用信任的传递性建立节点间的信任路径对信任值进行评估,然而这些方法无法建立多背景下的信任路径。Hang 等提出了一种基于网络相似性的评估机制^[13],然而这种机制对于非对称的信任关系不能有效地区分。

对背景敏感是信任的重要属性^[14-15],如何综合利用已知背景下的信任信息,对各个背景下未知节点的信任值进行评估是上述方法不能有效解决的问题。本文提出一种基于网络相关性的信任评估模型(multiple-context trust evaluation, MCTE),这种模型利用相关性的概念计算网络节点,再利用相关性网络完善信任网络。最后,未知的信任值从完善后的信任网络得到。这种算法不仅适用于单背景下的陌生网络节点的信任评估,更在多背景环境下有着出色的表现,对于解决跨背景信任评估问题较现有算法有着较大的优势。通过对真实社交网络的数据进行测试,发现在准确度方面超过现有的算法,在节点的可计算度方面与传递法基本持平,略低于平均法,但是由于平均法的高误差,MCTE 方案在综合性能上超过现有的算法。

1 相关性网络

为了建立相关性网络,本文将共同信任以一种条件概率的形式定义,即假设存在用户 A、B、C,那么 A 与 B 共同信任 C 的概率为,在 B 信任 C 的条件下 A 同样信任 C 的概率。本文用 $T(A/B)$ 表示 A 对 B 的信任值, $T((A \cup B)/C)$ 表示 A、B 共同对 C 的信任值,那么在如图 1 所示的信任网络中,当 A、B 的独立性无法确定的情况下只能得到

$$T((A \cup B)/C) \neq T(A/C)T(B/C) \quad (1)$$

然而当 $T(B/C)=1$ 时可以得到

$$T((A \cup B)/C) \neq T(A/C) \quad (2)$$

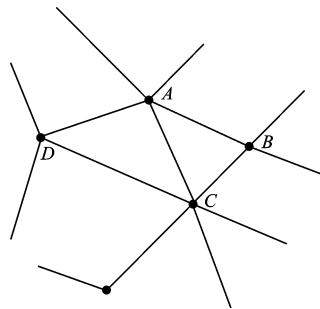


图1 简单信任网络

如上所述,当 B 完全信任 C 时,A 与 B 共同信任 C 的情况就相当于 A 单独信任 C 的情况。信任的准确量化对于社交网络用户来说往往是困难的,因此在实际应用中用户的信任常出现相信或不相信这种二元分布,即信任值为 1 的情况是大量存在的。

令 $T(A)$ 为 A 对任意与其相连的用户的信任值的期望,则

$$T(A) = E_X T(A/X) \quad (3)$$

在信任网络中,A 可能产生信任关系的用户未必都与其直接相连,对于和 A 不直接相连用户的信任值应视为一个缺省值而非 0。

由式(1)可得,在不能确定用户之间的独立性时同样满足

$$T(A \cup B) = E_X T((A \cup B)/X)$$

然而,在 B 对用户群 X 完全信任的情况下由式(2)可得

$$T(A \cup B) = E_X T(A/X) \quad (4)$$

令 $R(A/B)$ 表示用户 A 与用户 B 的相关度,它表示 A 在 B 完全信任某随机用户的条件下同样信任该用户的概率。A 与 B 的相关性可以在整个信任网络中传播。

$R(A/B)$ 的计算方法为

$$R(A/B) = E_X (T((A \cup B)/X) / T(B/X) = 1) = E_X T(A/X) \quad (5)$$

相关性不同于相似性。它们最直观的区别是相似性是一种对称的特性,也就是说假使 A 与 B 相似,那么 B 与 A 也相似,然而本文所述的相关性是非对称的,即 $R(A/B) \neq R(B/A)$ 。此外,在实际的信任网络中经常会见到用户 A、B 对某用户个体或群体拥有相同的信任值,不能由这种现象估计 A 与 B 的相关性,尽管其行为非常相似。

在定义了用户的相关性后,可以在信任网络之上构建相关性网络,从而扩展信任的传播性。由式(5)可得,当 $T(A/X) \neq \text{null}$ 且 $T(B/X)=1$ 时,

$R(A/B) = E_X T(A/X)$ 。可以将 $R(A/B)$ 视作当 $T(A/X) \neq \text{null}$ 且 $T(B/X) = 1$ 时的均值,从而由部分 $T(A/X)$ 观测值预测未知的 $T(A/X)$ 值。当用户样本非常庞大时,对任意用户 X 在 $T(B/X) = 1$ 时,可以近似地估计其信任值

$$T(A/X) \approx R(A/B)$$

进一步考虑到每一个至少通过某一用户群体 X 从而与 A 相关的用户 B ,可以得到

$$T(A/X) \approx E_B R(A/B) \quad (6)$$

图2所示为建立在信任网络之上的相关性网络,虚线代表相关性,实线代表信任值。

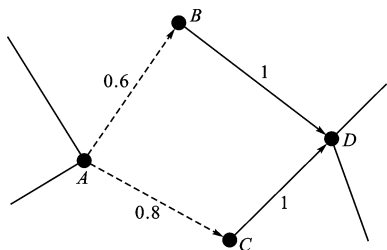


图2 相关性网络及底层信任网络

由式(6)可得用户 A 对 D 的信任值为

$$T(A/D) = (0.8 + 0.6)/2 = 0.7$$

相关性网络是建立在信任网络基础上的,因此可以依靠相关性网络由已知用户的信任关系估计未知用户的信任关系,将式(5)代入式(6)可得

$$R(A/X) \approx E_B R(A/B) = E_B E_Y (A/Y) \quad (7)$$

仅当 $T(B/Y) = 1$ 且 $T(B/X) = 1$

以上关于相关性网络的推导均建立在 A 与 B 的独立性不确定的基础上,但在实际的社交网络中,用户的独立性往往是很明显的。因此,在确定用户独立的基础上可以得到如下公式

$$E_X T((A \cup B)/X) = E_X T(A/X) T(B/X)$$

进一步由式(5)和式(6)可得

$$R(A/B) = E_X T(A/X) T(B/X) / E_X T(B/X) \quad (8)$$

$$T(A/X) \approx \sum_X R(A/B) T(B/X) / \sum_X T(B/X) \quad (9)$$

式(8)、式(9)在实际社交网络中的信任评估有着重要的作用。利用相关性的特点不仅可以对相同交互背景下的节点进行信任评估,同样适用于不同交互背景下的用户建立信任评估模型。

2 MCTE 模型

假设存在如图3所示的两种相互独立的交互背景 A 和 B ,在环境 A 中用户 C_1 对用户 C_2 的信任概

率为 $T_A(C_1/C_2)$,在环境 B 中用户 C_1 对用户 C_2 的信任概率为 $T_B(C_1/C_2)$,那么用户 C_1 在环境 A 与环境 B 中共同信任 C_2 的概率为

$$T_{AB}(C_1/C_2) = T_A(C_1/C_2) T_B(C_1/C_2)$$

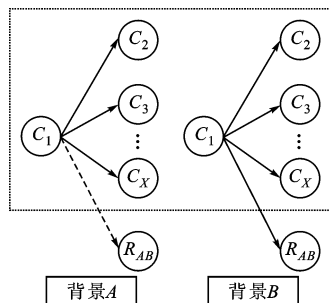


图3 双背景下的信任网络示例

在环境 A 中用户 C_1 对所有交互对象的信任的期望为 $E_X T_A(C_1/X)$,简记做 $T_A(C_1)$,同理在环境 B 中 C_1 对所有交互对象的信任的期望为 $T_B(C_1)$ 。这里在环境 A 与环境 B 中,用户 C_1 交互的对象集合为同一集合 $X = \{x_1, x_2, \dots, x_i\}$,且 x_i 不等于 C_1 。

用户 C_1 在环境 A 与环境 B 中同时信任集合 X 的概率为

$$T_{(A \cup B)}(C_1) = E_X (T_A(C_1/X) T_B(C_1/X))$$

根据贝叶斯公式

$$T_{A|B}(C_1) =$$

$$E_X (T_A(C_1/X) T_B(C_1/X)) / E_X T_B(C_1/X)$$

$T_{A|B}(C_1)$ 是环境 A 与环境 B 的相关性函数,为了更加明确地表示,将 $T_{A|B}(C_1)$ 记做 $R_{A|B}(C_1)$ 。再次通过贝叶斯公式可以求得如图1所示的在环境 A 中用户 C_1 对用户 $R_{(A \cup B)}$ 的信任值。具体方法如下

$$T_A(C_1/R_{(A \cup B)}) = R_{A|B}(C_1) \left(\sum_X T_B(C_1/X) + T_B(C_1/R_{(A \cup B)}) \right) / T_B(C_1/R_{(A \cup B)}) - \sum_X (T_A(C_1/X) T_B(C_1/X)) / T_B(C_1/R_{(A \cup B)})$$

上述计算方法是通过用户 C_1 在环境 B 中对 $R_{(A \cup B)}$ 的信任情况预测该用户在环境 A 中对 $R_{(A \cup B)}$ 的信任情况。在现实环境中,可能存在许多 C_1 与 $R_{(A \cup Y)}$ 交互的环境,单独依靠某一环境的信任状况作为参考不够全面,应该对这些环境中的信任情况作综合考虑。具体方法如下所述:假设存在 Y 个独立且与 C_1 拥有共同交互对象 $R_{(A \cup Y)}$ 的环境,由环境 Y 与环境 A 的相关性计算用户 C_1 对 $R_{(A \cup Y)}$ 的信任值 $T_A(C_1/R_{A \cup Y})$,再通过求 $T_A(C_1/R_{A \cup Y})$ 的期望,得到在环境 A 中用户 C_1 对 R 的信任值

$$T_A(C_1/R) = \sum_Y T_A(C_1/R_{A \cup Y}) / |Y| \quad (10)$$

式中: $|Y|$ 是不同环境的个数。

3 实验

本文提出的 MCTE 信任评估模型在实际社交网络中的实验主要面向两个问题: 第一是信任评估的准确度; 第二是在社交网络中用户的可评估度。所谓可评估度是指在有一定用户信任资料以及网络拓扑结构信息的情况下, 可以预测未知用户信任值的能力。

采用公开的 Advogato 数据集^[16]抽取出用户在不同背景下的交互以及评价记录, 分别建立相应的信任关系。该数据集的用户数量十分庞大, 考虑到实验环境等诸多方面的因素, 本文抽取 500 个活跃用户作为网络节点, 并以其存在的信任关系为边构建信任网络。时间方面采用的交互数据起始日期为 2012 年 9 月, 截止日期为 2013 年 6 月, 并分别抽取其中连续的 3 个月进行 3 次实验。为了与本文提出的信任模型比较, 对以上数据集还分别用两种现存的方案进行了信任评估。这两种方案分别是: ①平均法, 这种方法先找到所有除 A 外其他用户对 B 的信任值, 再求出这些值的均值作为 A 对 B 的信任值; ②传递法, 这种方法以信任的传递性为依据, 在 A 对 B 进行信任评估时, 首先找出在网络中所有从 A 到 B 的路径, 再由这些路径分别计算信任值, 最后综合这些信任值计算 A 对 B 的信任值。在综合各路径的信任值时现存的方案有多种, 有的采用均值的方法, 有的采用最大或最小值的方法, 有的对各路径进行分类, 并按照一定的规则对各路径的影响赋以权值, 从而进行综合评估。本文采用均值法处理各路径的信任值。至于其他的信任综合方法与均值法相比显示的优势远不及 MCTE 模型与其相比的优势。

首先在同一背景下以相关性网络方案与其他两种方案进行比较。图 4 描述了相关性网络方案与其他两种方案相比在准确度上的优势。图中横坐标 S 是边数与节点数的比值, 表示每个用户平均评价其他用户的数量。由图可以很明显地观察到, 相关性网络方案计算得到的信任值的准确度在全部覆盖率范围内保持在 0.85 ± 0.04 。在覆盖率小于 2 的区间, 准确度由 0.85 开始缓慢地减小。在覆盖率达到 4 附近时减小到最低, 这可能是由于覆盖率在较低的区间增长, 增加了许多干扰信息以使算法的准确度降低。当覆盖率达到 6 左右时, 准确度又回升到初始的 0.85, 相关性网络在准确度方面的优势较之

另外两种算法是非常明显的。平均法计算的信任准确度很低的原因是其对个体特殊性的不敏感, 从完全均衡的角度来度量用户的信任值, 忽略了信任的主观特性, 因此造成较大的误差。传递法在准确度上比平均法有了很大的提高, 其能够从特定用户出发, 找到一条由评估者到受评估者的路径, 然而和 MCTE 方法相比还存在一定的差距。

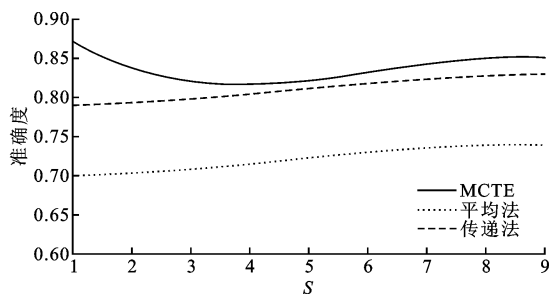


图4 3种评估方法在单背景下的准确度比较

图5显示了在可评估度方面相关性网络与其他两种方案的比较。可评估度定义为在所有的网络节点中可评估节点的百分比。从图中可以观察到, 平均法在可评估度上占有巨大的优势, MCTE 方案和传递法的结果有着近似的表现。在覆盖率等于 2 (平均每个用户只对两个用户进行过信任评估) 时, 仅有 8% 的可评估度。随着覆盖率达到 4, 可评估度上升到 31%。在覆盖率达到 6 时, 可评估度达到 45%, 并且上升趋势开始减慢。由图还可见, MCTE 方案的可评估度在覆盖率为 4 左右时上升趋势最明显。这一现象的原因是在现实网络中用户之间的连线往往出现高度的集群化现象, 而非完全随机的分布, 这同样也说明在稀疏网络中 MCTE 方案能够达到令人满意的效果。虽然平均法在可评估度上有着较大的优势, 但是由之前的实验可得其在准确度方面的表现较差。在现实的社交网络中, 信任评估的准确性较之网络的可评估度对用户来说往往更为重要。

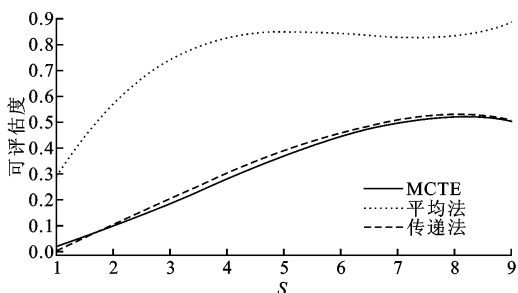


图5 3种评估方法在单背景下的可评估度比较

前文中对单背景 MCTE 方案与其他两种方案在准确性和可评估度方面做了比较, 下文对多背景

下信任值的准确性和可评估度做出比较。本文选取用户在2种和10种相互独立背景下的信任评估数据,如图6所示,随着背景的增加,MCTE模型的准确度也会提高。这是由于MCTE方案随着背景的增加可以选择更短并且相关性更强的路径对用户信任值进行评估。另外可以观察到,准确度随着背景的增长,每增加一个背景即增加 0.01 ± 0.005 的准确度,这也代表着在每一背景下的平均误差减少了大约0.01。在实际的社交网络中,交互背景的数量远不止几个或几十个,MCTE方案的使用将在更多背景的应用场合发挥更大的优势。

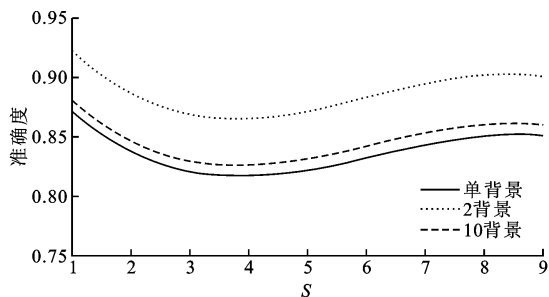


图6 信任值在多背景下准确度的比较

图7显示了在多背景下MCTE方案在可评估度方面的表现。与准确度类似,MCTE方案随着背景的增加可评估度也随之提高,其增加幅度大致在 0.01 ± 0.01 范围内。举例来说,从单背景到2背景,可评估度提高的均值为0.03,这代表着网络中增加了3%的节点可以评估,到10背景时,可评估度提高的均值达到0.09。在实际的社交网络中存在大量的交互背景,MCTE的性能将进一步提高。

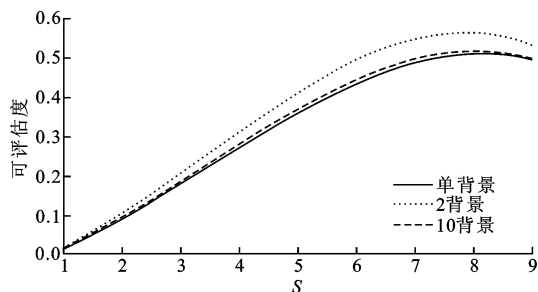


图7 信任值在多背景下可评估度的比较

4 结 论

社交网络中的间接信任评估是信任计算中的热点问题,由于信任对交互背景敏感的特性,本文在分析现存方案所存在的不足后,提出了基于相关性网络的信任评估模型MCTE,这种模型可以有效地克服现存的模型在以下两方面的缺点:仅考虑网络全

局性而忽略节点自身特性;仅通过简单的传递模型而忽略信任的衰减。MCTE利用相关性的概念在原有信任网络之上建立相关性网络,再由完备的相关性网络更新信任网络,从而得到未知网络成员的信任状况。这种方法不仅可以计算单背景下的未知用户的信任值,同样适用于多背景环境下。对真实社交网络的实验结果表明了MCTE模型的可用性,以及较现有模型的优势。

MCTE对于未知信任节点的评估强烈依赖于现存信任网络中信任关系的准确性,对于存在恶意或是不能给出准确信任信息的网络,模型的性能受到较大的限制。如何辨别恶意节点和优化现有信任关系的准确度是将来研究的方向之一。

参考文献:

- [1] SHERCHAN W, NEPAL S, PARIS C. A survey of trust in social networks [J]. ACM Computing Surveys, 2013, 45(4): 47-79.
- [2] BARNES J A. Class and committees in a Norwegian island parish [J]. Human Relations, 1954, 7(1): 39-58.
- [3] 张赛, 徐恪, 李海涛. 微博类社交网络中信息传播的测量与分析 [J]. 西安交通大学学报, 2013, 47(2): 124-130.
ZHANG Sai, XU Ke, LI Haitao. Measurement and analysis of information propagation in online social networks like microblog [J]. Journal of Xi'an Jiaotong University, 2013, 47(2): 124-130.
- [4] GROSS R, ACQUISTI A. Information revelation and privacy in online social networks [C]// Proceedings of the ACM Workshop on Privacy in the Electronic Society. New York, USA: ACM, 2005: 71-80.
- [5] MOLM L D, TAKAHASHI N, PETERSON G. Risk and trust in social exchange: an experimental test of a classical proposition [J]. American Journal of Sociology, 2000, 105(5): 1396-1427.
- [6] ZAPPEN J P, HARRISON T M, WATSON D. A new paradigm for designing e-government: Web 2.0 and experience design [C]// Proceedings of the International Conference on Digital Government Research. Berlin, Germany: Springer, 2008: 17-26.
- [7] ABDUL-RAHMAN A, HAILES S. Supporting trust in virtual communities [C]// Proceedings of the 33rd Hawaii International Conference on System Sciences. Piscataway, NJ, USA: IEEE, 2000: 1-9.

(下转第103页)