基于共同交易行为的在线交易欺诈

摘要

1 绪论

1.1 背景

在线交易欺诈现象普遍存在，但是在线交易平台并没能提出系统性的解决方法，用于在交易过程中识别欺诈者，它们只是利用用户的交易行为和个人信息来判断用户的可信性。典型的方式就是基于反馈累积的信誉评价机制，如国外的ebay和国内的淘宝，但这种机制很容易通过互刷来进行欺诈。从某种意义上讲，这种机制反而促进了在线交易的欺诈行为。所以，我们现在面临的研究问题就是，在拥有在线交易平台的用户信息和他们之间的交易记录的前提下，如何检测出具有潜在欺诈行为的用户，如何去发现那些欺诈团伙？

1.2 国内外研究现状

1.2.1 基于统计的方法

2009，D-S证据理论法 [5]

1.2.2 基于数据挖掘的方法

2009，基于时间分段模型的在线交易早期欺诈检测机制[6]，考虑到欺诈者行为对时间的依赖性，将欺诈者的生命周期划分为潜伏期和作案期，将用户的交易记录根据累积的评分进行分段，分成80%、85%、90%、95%和100%五段[7]，然后利用Weka3.6中的分类回归树对各个时间段分别进行训练，最终使用AdaBoost [8]集成算法将弱分类器构建成强分类器。

早在2005年D.H. Chau就提出从在线交易用户的个人资料和交易记录中抽取出一些重要的特征属性，包括最近15天、30天买卖物品的均价和方差，总共包含17个特征属性。然后，利用这些特征属性和获取的真实欺诈用户名单对C5.0决策树进行训练，最终使用训练模型来发现潜在的欺诈用户[9]。

在2006年D.H. Chau他们又提出了2LFS (2-Level Fraud Spotting) 方法用于发现欺诈者，并且还分析了欺诈者使用的典型欺诈行为[10]。2LFS方法由两层构成：第一层为用户级别本证，利用【7】中提出的17个特征属性和决策树分类算法得到用户的类别标识（诚实者、欺诈者）；第二层为图级别特征，将第一层获得的用户类别标识作为交易图中用户的初始类别标识，并利用马尔科夫随机场模型和置信传播算法进行迭代计算。在对图级别特征的计算过程之中，引入了同盟者用户类别标识，并且还针对同盟者和欺诈者的联合欺诈行为提出了近似Bipartite Cores现象。【13】

1.2.3 基于社交网络分析的方法

2007，Y.H. Ku等人提出利用SNA (Social Network Analysis)和决策树分类算法来发现现在的欺诈者。【15】

2008，Wang使用SNA中的K-core值和中心加权算法，用于发现在线交易中的欺诈团伙【16】

2012，S.J. Lin等人提出结合排名思想和SNA来发现在线交易中的欺诈团伙。【17】

2 相关知识及研究

2.1社交网络分析

社交网络分析通过研究网络间的关系，有助于把个体间关系、微观网络与大规模的社会系统的宏观结构结合起来。从社交网络的角度出发，用户在社交网络中的相互作用可以表示为基于关系的一种模型或规则，而基于这种关系的有规律模式反映了社交网络的结构，这种结构的量化分析是社交网络分析的出发点。社交网络分析强调的是行动者彼此之间存在的关系，研究人员对这些关系建立分析模型，研究这些关系给社交网络整体结构或者网络内部的个体带来的影响。

2.1.1 PageRank

2.1.2 TrustRank

TrustRank是一个基于链接分析的技术，用于半自动化的垃圾网页检测。因为搜索引擎在计算网页排名(PageRank)的时候，非常依赖网页之间的链接，而且链接的质量越来越显得重要。但是，一些垃圾网页可能误导搜索引擎，获得远远高于他们本身应得的排名。

在计算网页的TrustRank值时，充分利用了优质网页它们的近似独立性，也就是优质的网页经常会链接至一些优质的网页，而它们很少会只想垃圾网页。

对于TrustRank算法，首先会在全部的网页中选择大小为L(L << N)的种子集，然后，人类专家会对种子集中的页面进行评估，给种子集中鉴定为合法的页面赋予TrustRank值为1，而其它的页面TruestRank值为0，并对网页的静态得分进行归一化。最后，通过制定的网页Trust值传递衰减过程，对网页的TrustRank值进行迭代计算。（类似PageRank）

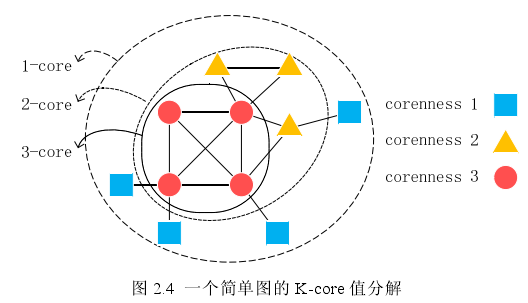
2.1.3 BadRank

网页BadRank值的计算思想与TrustRank值的计算恰恰相反，它是通过种子集中垃圾页面来发现更多垃圾页面的一种方法。

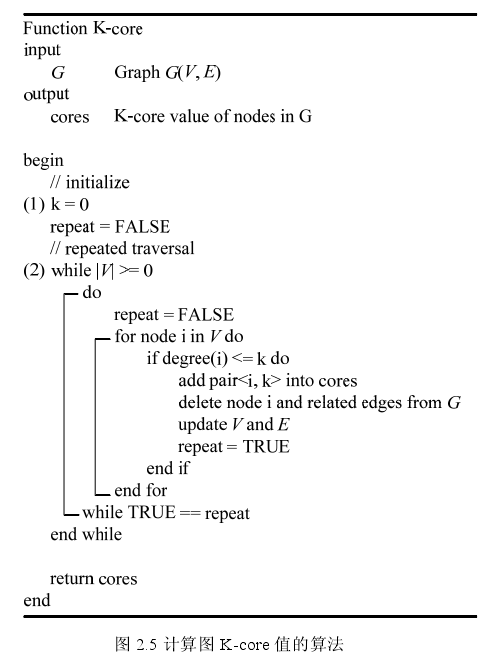
这种方法基于的假设性原理是：指向垃圾页面的页面通常也是垃圾页面。所以，对于BadRank算法，首先，将网页中的所有链接反向，得到的是一个反向的网页链接图。然后，选择一些页面作为种子集，并人工的对这些种子集中的页面进行评估，给种子集中鉴定为垃圾的页面赋予BadRank值为1，其余为0，然后进行归一化处理，最后，在网页的反向链接图上，采用BadRank算法计算网页的BadRank值，该值越高表明网页是垃圾页面的概率越大。

2.1.4 K-core

K-core：在图G的一个连通子图中所有节点的度数至少为k，则这个连通子图中所有节点的K-core值大于等于k，连通子图的形成是通过重复的从图G中删除顶点度数小于k的节点以及该节点相连接的边。

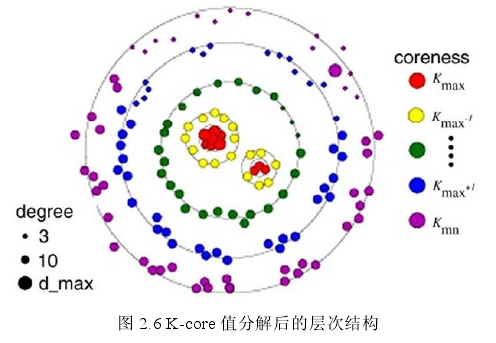


一个节点的K-core值为k，当且仅当该节点属于K-core值为k的连通子图，但却不属于任何K-core值为k+1的连通子图。顶点K-core值中最大的k被定义为图G的K-core值。



顶点集V应该按度数从小到大排列。

顶点K-core值可以用来描述顶点在图G中所处的位置，顶点的K-core值越大，表明该顶点离图G的核心越近，反之顶点越靠近图G的边缘。

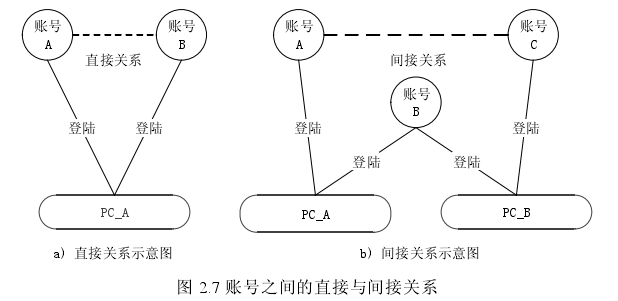


需要注意的是，利用K-core将图中顶点进行分层的方法，它并不能够很好的表示图中结点的差异性，其主要原因有两个方面：其一，按照K-core的思想，属于同一层次的顶点之间因为K-core值相等，所以它们之间没有差异性，但实际情况并非如此，同一层次的顶点在度数之间的差异可能会很大；其二，相对于图的规模很大时，图的K-core值经常会比较小，利用K-core值来区分顶点的粒度较粗，该指标对结点影响力的辨析能力差强人意。

2.2 团伙欺诈的发展过程

2.2.1 多用户抱团方式

一个用户通过注册多个账号来给特定的账号提升信誉度。但是这种提升信誉的方式很容易被发现。下面以IP地址检测为例：



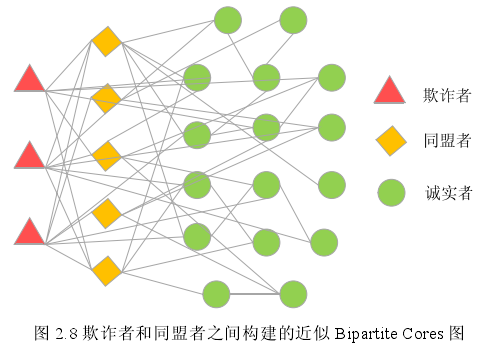
对于这种多个账号之间形成的直接或间接的关系，可以称之为抱团现象。这种方式在后期并不会被用户采取，因为

1) 交易平台可以采取很简单的社交网络分析技术(K-core)，利用紧密连通子图来发现行为异常的团伙，或者通过分析用户间的多层间接关系来挖掘他们之间的异常行为。

2) 若团伙中的一名成员被发现，则交易平台可以很容易的发现团伙中的其它成员。

2.2.2 同盟者团伙方式

在这种团伙欺诈方式中，欺诈者们会注册多个账号，并将这些账号分成两类（欺诈者和同盟者），这些账号一般不会再同一台电脑上登录，这样就可以避免交易平台对账号的多层间接分析。其中，欺诈者账号会被用来进行最终的欺诈交易，同盟者会通过与诚实者之间的交易来提高自己的信誉度，同盟者的行为完全类似于诚实者，同时这些同盟者会与欺诈者进行交易，用来提高欺诈者的信誉度，最终会在欺诈者、同盟者和诚实者之间形成近似的Bipartite Cores图。



这样，当欺诈者被交易平台发现时，由于同盟者的行为类似于正常用户，所以他们并不会被发现，这些同盟者账号可以在团伙欺诈中被重复利用。

2.2.3 平台欺诈方式