在线社交网络群体发现研究进展

潘 理* 吴 鹏 黄丹华
(上海交通大学电子信息与电气工程学院 上海 200240)
(信息内容分析技术国家工程实验室 上海 200240)

摘 要：群体是在线社交网络重要的中观组织。群体发现不仅有重要的理论意义，还推动了在线社交网络的应用与发展，有广泛的应用前景。该文总结论述了在线社交网络群体发现的研究进展。在分析群体形成机理的基础上定义在线社交网络群体，并介绍群体发现问题。根据挖掘群体时采用的不同特征，该文分别阐述基于个体属性特征的群体发现方法和综合属性与结构特征的群体发现方法。随后从特征选取和检测算法两个方面重点介绍了恶意行为群体的发现方法。最后，对群体发现进一步的研究方向进行展望。

关键词：在线社交网络；群体发现；恶意行为群体
中图分类号：TP393 **文献标识码**：A **文章编号**：1009-5896(2017)09-2097-11
DOI: 10.11999/JEIT161192

Reviews on Group Detection in Online Social Networks

PAN Li WU Peng HUANG Danhua
(*School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai* 200240*, China*)
(*National Engineering Laboratory for Information Content Analysis Technology, Shanghai* 200240*, China*)

Abstract: Groups are important mesoscopic organizations of Online Social Networks (OSNs). Group detection not only has important theoretical significance, but also has a wide range of applications. It promotes the application and development of online social networks. In this paper, group detection technology in online social networks is studied. Based on analyzing the formation mechanism of social groups, the online social network groups is defined and the group detection problem is introduced. According to different features adopted by group detection methods, the methods based on the attribute features only and those based on combination of attribute features and structure features are analyzed, respectively. Especially, it reviews the malicious behavior group detection methods by analyzing their feature selection mechanisms and detection models in detail. Finally, further research direction of group detection in online social networks is prospected.

Key words: Online Social Networks (OSN); Group detection; Malicious behavior group

1 引言

在线社交网络已经成为当今人们日常交流、信息发布与共享的重要平台。与此同时，用户在社交网络平台产生的大量数据为群体行为分析和数据挖掘带来了巨大的机遇和挑战。对在线社交网络进行群体挖掘，不仅可以从网络结构、节点属性等角度对在线社交网络进行全面理解，而且可以进一步在群体层面研究社交网络用户的行为模式和互动规律，实现群体行为的控制与引导[1]。

传统社交网络中的群体是从心理学、社会学、人类学等多角度因素，综合概括出来的一个较为抽象的概念。通常可以将群体理解为由两个或两个以上个体，因为一些相同的内在因素，如兴趣、目标、利益等，自发地或者有组织地聚集在一起形成的集合。集合内个体能够进行互动或信息传递共享，进而相互产生影响。在线社交网络群体以在线社交媒体平台为基础，是传统社交网络群体在互联网空间的映射[2]。在线社交网络群体具有的可计算特征也是传统社交网络群体特征在网络空间中的映射。群体的可计算特征包括用户在在线社交媒体平台上的交互关系和各种属性信息的统计特性。在线社交网络群体发现正是基于这些可计算特征实现的。在线社交网络的群体发现与传统的社区发现研究既有联系

收稿日期：2016-11-04；改回日期：2017-02-26；网络出版：2017-04-25
*通信作者：潘理 panli@sjtu.edu.cn
基金项目：国家自然科学基金(U1636105)，国家 973 计划项目(2013CB329603)
Foundation Items: The National Natural Science Foundation of China (U1636105), The National 973 Program of China (2013CB 329603)

也有区别，传统的社区发现[3–7]基于社区结构内聚的思想，主要考虑社交网络在结构上内聚的特征，即同一社区内的节点连接紧密，不同社区内的节点之间连接稀疏，而对社区形成的机理、目的、意义等关注较少。在线社交网络群体发现则更为关注个体在内容、行为等属性特征上的相似性。所以，仅仅基于结构连接稠密性的社区发现方法难以用于发现内容、行为等属性特征上相似的群体。从综合考虑个体间结构和属性的可计算特征的角度出发，在线社交网络群体发现通常可采用两类方法，即基于属性特征的群体发现方法以及综合属性和结构特征的群体发现方法。一方面，这样的在线社交网络群体发现结果将与应用的联系更为紧密，因此能够为信息检索、信息推荐、信息传播控制等诸多应用提供有力支撑。另一方面，当前在线社交网络中存在着一些特殊的恶意行为群体。这些群体因为特定的行为目的聚集，且具有特定的行为模式，如谣言群体，水军群体等。这些恶意行为群体通常为了自身利益而做出危害网络安全和其他网民利益的行为，并且随着在线社交网络的发展其危害日趋严重。对这些恶意行为群体的有效发现是进行合理管控的基础。因此，高效的群体发现方法对于促进在线社交网络服务以及网络空间治理均具有重要的研究意义和应用价值。

本文第 2 节首先介绍在线社交网络群体的形成机理及定义，并在此基础上引入群体发现问题；第 3 节和第 4 节分别阐述两类在线社交网络通用的群体发现方法，即基于个体属性特征的群体发现方法和综合属性与结构特征的群体发现方法；第 5 节总结论述恶意行为群体的发现方法；第 6 节对未来群体发现技术的发展方向进行展望；最后，第 7 节总结全文。

2 在线社交网络群体形成机理与群体发现

传统的社会群体是不同个体由于一定的社会关系或社会活动而产生互动并相互影响的集体。早期的社会学家、心理学家发现，这些社会群体的形成往往是多种因素共同作用的结果，这些因素从社会学的角度可以分为内部和外部因素[8]。外部因素主要包括亲缘关系、友缘关系等一系列社会关系，内部因素主要涉及与个体属性相关的利益、爱好等。由此可见，社会群体并不是个体的简单集合，群体成员通常相互依存、相互影响，并且有一定的分工。在线社交网络中关注的群体是以社交媒体为平台，基于一定的利益、爱好等目的，通过网络活动相互联系、相互影响的用户集合。这些群体中，传统的社会关系被淡化，用户之间通过建立好友关系或互动行为产生相应的连接，这些连接是群体形成的基础。根据连接的性质，在理论上可以将社交网络中的连接分为 4 类[9]：邻近连接、好友关系连接、互动连接和信息流连接。基于前两种连接，用户之间形成相应的关系网络；而基于后两种连接，用户之间能形成相应的交互网络。通常，关系网络中群体不仅在结构上内聚，在工作、学校等人口统计属性上也均质。相似地，交互网络中的群体通常具有内聚的结构和相似的行为属性。真实的在线社交网络中，不同类型的连接互相耦合并且同时存在[10]。比如地理上的邻近连接通常伴随着好友关系连接、互动连接以及信息流连接。因此，对社交网络中群体的研究需要全面考虑多种类型的属性信息和连接信息。

本文研究的群体与传统的“社区”有一定的联系但又相互区别。在线社交网络中的社区只要求结构上连接紧密，但未考虑网络其他的属性信息，因此挖掘出的社区中包含的个体在属性上未必相似。而在线社交网络中的群体种类繁多，部分类别的群体基于某些特定属性特征聚集，但在网络结构上不一定连接紧密，因此这些群体无法使用传统的社区发现方法挖掘。同时，网络结构是在线社交网络的基本组成部分，很多群体除了属性相似，在结构上也紧密连接，这些群体与传统的社区又有一定的联系。图 1 中用集合的形式表示出了群体、特殊行为群体及社区之间的关系。一方面，那些不仅结构内聚，而且在某些属性上均质，具有社会化语义特征的社区同时也是社交网络中的群体。而那些仅在结构上内聚而缺乏语义特征的社区则不是群体。另一方面，仅具有相似属性而连接松散的群体并不具有结构内聚的特征，因此也区别于社区。特殊群体是一类具有特殊行为模式的群体。部分具有结构内聚特征的特殊群体同时也是社区。综合群体的各项特征，这里我们将在线社交网络中的群体定义为以社交媒体为平台，属性相似性为基础，结构上连接紧密为重要可选特征的一类个体及其相关属性组成的集合。当然在不同的应用环境和应用目标下，群体以属性相似性和结构内聚性为本质具有各自特定的定义。这里给出社交网络中群体的一般化定义。

定义 1 给定一个社交网络 $\mathcal{S}=(\mathcal{V},\mathcal{E},\mathcal{F},\mathcal{P})$，其中 $\mathcal{V}$ 为所有用户节点的集合，$\mathcal{E}$ 为用户间所有连边的集合，$\mathcal{F}$ 表示用户的所有属性集，$\mathcal{P}$ 为社交网络平台，则社交网络中的群体定义为 $G=(V,E,F)$，其中，$V\subseteq\mathcal{V}$，$E\subseteq\mathcal{E}$，$F\subseteq\mathcal{F}$，且 V，E 满足结构内聚性条件 $\phi(V,E)$，V 在 F 上满足属性相似性条件 $\varphi(V,F)$。

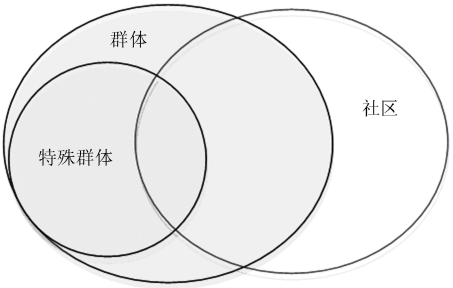

图 1 群体与社区的区别与联系

在线社交网络的群体发现就是针对群体的应用目标，选择群体在属性或结构上的可计算特征作为发现基准，挖掘具有特定属性或结构特征的用户子集的过程。大部分群体发现方法具有一般化的过程：首先基于属性和结构等方面的先验知识选取待发现群体的特征，然后基于选取的特征定义待发现群体，并建立描述群体聚集度的目标函数，最后设计算法求解该目标函数的最优化问题进行群体发现。该目标函数的最优化问题求解通常是 NP 难问题，因此需要设计相应的启发式算法进行近似优化求解。在这一研究过程中，选取哪些特征，如何基于特征定义群体聚集度目标函数的最优化问题，以及如何设计快速有效的启发式求解算法是需要解决的关键问题。特征选取是群体定义的关键，决定了发现的群体具有的性质。在线社交网络中的群体特征有很多表现形式，主要可以分为属性特征和结构特征两大类。一方面，根据群体的定义，群体内部个体总是在某些属性上相似，因此属性相似特征是群体的本质特征。另一方面，一些具有特定相似属性的用户会因互动频繁而产生紧密连接，因此群体通常还具有连接紧密的结构特征。针对选取的特征类型的不同，本文分别从基于个体属性特征和综合属性与结构特征两方面总结论述通用的群体发现方法。

3　基于个体属性特征的群体发现方法

在线社交网络个体的属性反映了个体的兴趣、身份、归属等信息。群体内个体通常具有某些相似的属性信息，因此群体具有个体属性相似的特征。聚类和分类是处理属性相似特征的两类常用的检测算法。聚类是一种无监督学习方法，特点是能够在没有训练样本时，基于个体属性相似性聚合群体，适用于发现一般化的属性相似的群体。分类是一种有监督学习方法，特点是需要提供大量训练样本和关键属性，适用于有目标地发现一些具有特殊属性特征的群体。下面分别总结论述基于属性聚类和基于属性分类的群体发现技术。

传统的基于属性聚类的群体发现方法用属性向量描述每个个体，以个体属性向量之间的相似性为基础，在属性空间中基于属性向量的分布划分群体。其中，基于相似度的划分算法和基于密度的网格算法是两类经典的属性聚类方法。基于相似度的划分算法[11−13]首先在属性空间中随机确定 K 个聚类中心，每个聚类中心代表一个群体。对每个聚类群体，通过最小化中心邻域个体属性向量到该群体聚类中心距离的标准差确定该群体的子空间。在每个子空间下计算个体属性向量到中心的距离，每个个体被划分到距其最近的中心。接下来每个群体的中心基于该群体内的个体分布重新设定，并基于新的中心重新分配个体到最近的群体。如此迭代直到收敛。基于相似度的划分算法通常只能发现“类圆形”群体聚类，且运行时间较长。而基于密度的网格算法[14−16]可以克服这些缺点。该类算法将属性空间用和坐标轴平行的网格划分成宽度相等的单元，那些至少包含一定数量个体的单元被认为是稠密的，每个群体由尽可能多的相邻稠密单元组成。由于稠密单元具有向下闭包性质，因此子空间群体可以用自底向上的方式快速检索。基于属性聚类的方法是一种无监督群体发现方法，该类方法基于个体属性值相似性自动聚类，无需提供训练集，但是这种方法难以发现具有特殊属性特征的群体。

基于属性分类的群体发现方法将个体的属性信息看作个体的特征，将已知类别标签的个体集合分成训练集和测试集。在训练集上执行分类器算法，生成分类模型。在测试集上执行分类模型，生成预测结果。根据预测结果，计算必要的评估指标。这些属性特征通常是基于大量用户数据的统计特征，例如 Zhang 等人[17]以大众点评上用户一天内的移动范围大小、一天内发布评论的数量、时间间隔等统计信息为特征，用分类模型支持向量机(Support Vector Machine, SVM)训练出一个分类器，将大众点评上的用户分为女巫(sybil)用户和非 sybil 用户，从而挖掘出 sybil 群体。相似地，Wang 等人[18]以用户的点击行为模式为特征，包括点击频率、间隔时间等，采用 SVM 对已知类别的用户数据集进行训练，挖掘人人网上的 sybil 群体。不同的分类算法在特定的应用环境下具有不同的效果，为了实现更精确的分类结果，有时需要对分类算法进行评估选择。为了发现 Twitter 上的垃圾邮件(spam)用户群体，McCord 等人[19]以用户和文本信息为特征，采用随机森林，朴素贝叶斯，支持向量机和 k-近邻 4 个分类方法，对 Twitter 上的用户进行分类。实验结果显示随机森林算法能更有效地挖掘 Twitter 上的

spam 群体。事实上，基于属性分类的群体发现方法可以采用大部分常见的有监督分类学习方法。基于属性分类的方法可以发现特殊类型的群体，然而这类方法需要提供类别已知的个体作为训练集。基于属性聚类和基于属性分类的群体发现方法的优点和缺点对比如表 1 所示。

4 综合属性与结构特征的群体发现方法

在线社交网络中很多群体除了属性相似，在结构上也具有紧密连接的特征。为了发现属性相似且结构内聚的群体结构，需要综合属性与结构特征设计群体发现方法。类似于基于属性特征的聚类和分类方法，综合性方法也分为发现一般化群体的聚类方法和发现特定目标群体的半监督方法。发现一般化群体的聚类方法不需要样本信息，仅基于个体属性相似性和连接紧密性聚合群体。在初始研究中[20−28]，研究者在能够获得的整个属性空间上考虑个体属性相似性，提出属性全空间聚类发现方法。后续研究[29−33]逐渐表明群体内个体属性通常只在部分属性上具有相似性，即在某个属性子空间上相似，因此在属性子空间上聚类进行群体发现更有效。另一方面，发现特定目标群体的半监督方法需要基于应用目标提供部分样本信息，来引导挖掘满足应用需要的目标群体。下面分别论述基于属性全空间聚类的群体发现方法，基于属性子空间聚类的群体发现方法和目标群体发现方法。

表 1 基于属性聚类和基于属性分类的群体发现方法对比

方法	简介	优点	缺点
基于属性聚类的群体发现方法[11−16]	无监督方法，将在某个属性子集上取值相似的个体聚集成群体	无需提供训练集，自动基于个体属性值相似性聚类	难以发现特殊类型的群体
基于属性分类的群体发现方法[17−19]	有监督方法，选取属性作为特征，在训练集上训练分类器，用训练过的分类器划分个体	能够发现特殊类型群体	需要提供训练集及关键的属性特征

综合属性和结构特征的属性全空间聚类方法[20−28]考虑个体在所有获得的属性上的均质性或相似性，要求群体内部个体在属性全空间下尽可能相似。如何将节点的属性信息加入到网络结构中是这一类方法需要解决的关键问题。一类较为直接的方法就是将节点的属性信息作为新的属性节点或者新的关联属性边加入到网络中。Zhou 等人[20−23]提出的结构属性聚类方法 SA-Cluster (Structural/Attribute-Cluster)将属性信息作为新的属性节点加入网络中建立增广网络，并基于邻域随机游走模型估计增广网络中节点之间的距离，基于该距离挖掘群体，将相距较短的节点划分到同一群体。Ruan 等人[24]基于关联属性边提出一种融合结构相似度和属性相似度的机制，并基于该机制设计群体发现算法 CODICIL(COmmunity Discovery Inferred from Content Information and Link-structure)。不同于 SA-Cluster，他们基于节点之间的属性相似度为每个节点创建 k 条关联属性边，属性边和原结构边共同组成联合边集。为了使算法适合大规模社交网络，Ruan 等人为每个节点采样少量重要的关联边。最后 CODICIL 在采样后的网络上快速划分群体。另一类属性全空间聚类方法是将属性和结构信息融合到一个统一的数学模型中，通过模型优化来实现群体的划分。Akoglu 等人[25]基于信息论从信息编码角度统一地建模属性和结构信息，提出综合发现方法 PICS(Parameter-free Identification of Cohesive Subgroups)。具体来说，PICS 对网络的邻接矩阵和属性矩阵进行二级编码，即将相关矩阵按群体进行分割，对矩阵的每一个子部分内部分别编码。当群体内部个体具有相似的连接模式和均质的属性时，矩阵的每一个子部分的 0-1 分布也相对均质，此时所需的编码长度较短。以矩阵编码长度为目标函数，初始时每个节点作为独立的群体，PICS 基于贪心思想迭代的扩展群体并交换矩阵中行列位置，使最终编码长度最小。基于属性全空间聚类的方法将属性和结构结合的思想引入到社交网络群体发现领域，但是随着获得的个体属性信息越来越多样化且属性维度越来越高，个体在属性全空间下的相似度越来越缺乏区分能力，因此全空间方法已难以满足群体发现的新要求。

不同于基于属性全空间聚类的方法，基于属性子空间的群体发现方法[29−37]为每个群体分配一个属性子集，要求群体内个体在属性子集内相似，即在属性子集组成的属性子空间上聚类群体。事实上，在线社交网络中的个体很难在所有属性上相似，每个群体内的个体仅在部分属性上相似，因此基于属性子空间的群体发现方法符合在线社交网络的特性。属性子空间聚类方法相对于属性全空间的最大难点就是属性子空间未知。因此，这一类方法需要在挖掘群体的同时选取最佳的属性子空间。

Günnemann 等人[31]提出的属性子空间群体发现方法，SSCG(Spectral Subspace Clustering for Graphs)为每个群体赋予一个独立的相关属性子集作为该群体的属性子空间，并在属性子空间下定义节点属性相似度，基于属性相似度为网络中的所有边赋予权值，进而在加权网络上定义挖掘群体的目标函数。通过最优化该目标函数挖掘群体及相应的子空间。Günnemann 等人[32]综合子空间聚类和稠密子图挖掘技术提出方法 GAMer(Graph & Attribute Miner)。他们通过要求每个群体在某个属性子集中取值相似来确定它的子空间，在属性子空间大小、连接密度和群体规模三方面提出群体需要满足的条件，最后使用集合枚举树有效枚举符合条件的群体。Huang 等人[33]基于网格聚类提出子空间群体发现方法，SCMAG(Subspace Clustering on Multi-valued Attributed Graph)。该方法首先将属性空间划分成网格，每个网格作为一个细胞，定义子空间熵度量网络个体属性在每个子空间的所有细胞中的分布情况，将熵较低的子空间作为感兴趣的子空间。在符合兴趣的子空间中，选取所有满足覆盖度条件和连接度条件的细胞，并将相邻的符合条件的细胞合并组成群体。大规模社交网络中存在大量多样的属性子空间，属性子空间的个数随着属性维度呈指数增长。另一方面，在真实应用中，并非所有属性子空间下的群体都是所需要的。因此，基于属性子空间聚类的群体发现方法在属性维度较大时不仅存在效率偏低的问题，而且缺乏面向具体应用的针对性。针对这一问题，研究者们提出面向应用的目标群体发现方法。

目标群体发现方法[38,39]基于具体应用，从特定的群体描述或者特定的样本个体出发，有目标的挖掘特定的群体。Pool 等人[38]提出挖掘具有特定描述的群体的方法 DCM(Description-driven Community Mining)。他们首先定义一个群体分数来衡量群体在结构上的内聚程度。然后定义一种群体描述语言描述群体，要求每个群体的描述尽可能简洁并且能够将描述的群体和网络其他个体区分开来。应用时，首先给定目标群体的描述，DCM 基于描述将网络个体进行分类，将满足描述的一类个体组成候选群体，对候选群体迭代调整群体组成和群体的描述，直到收敛，得到最符合给定描述的群体。Perozzi 等人[39]提出的方法 FocusCO (Focused Clustering and Outlier detection)仅挖掘符合用户兴趣的群体，要求挖掘的群体内的个体和用户提供的样本个体相似。FocusCO 的关键点在于通过一组用户提供的样本个体集推断用户的偏好，即用户感兴趣的属性子空间。他们将属性子空间推断问题建模为一个距离度量学习问题，通过最小化样本个体之间的距离度量学习子空间向量。基于推断到的属性子空间向量重新加权网络的边，并选出一组权重显著较大的边组成目标群体种子，最后迭代扩展这些种子得到目标群体。目标群体发现方法通常不需要划分整个网络，只需挖掘符合目标的群体，因此算法运行速度较快，且挖掘出的群体更能满足应用要求。综合属性与结构特征的 3 类群体发现方法的优点及缺点对比如表 2 所示。

类似于传统的社区发现方法，多数群体发现方法采用优化一个描述群体特征的指标函数的过程来发现最优的群体结构。在传统的社区发现方法中，描述社区特征的指标函数通常只考虑社区结构上连接的紧密性，如经典的模块度函数；而描述群体特征的指标函数不仅需要考虑结构上的内聚特征，还要考虑群体在特定属性上的均质特征。属性上的均质性和结构上的内聚性可以通过相似度、距离等指

表 2 综合属性与结构信息的 3 类群体发现方法对比

方 法	简 介	优 点	缺 点
基于属性全空间聚类的群体发现方法[20−28]	基于个体在所有属性的相似性及在结构上的内聚性划分群体	将属性和结构结合的思想引入到社交网络群体发现领域	不符合在线社交网络群体的特征，个体在属性全空间上缺乏区分力，运行时间长，不适用于大规模高维度在线社交网络
基于属性子空间聚类的群体发现方法[29−37]	基于个体在属性子集上的相似性及在结构上的内聚性划分群体	符合在线社交网络群体特征，个体在属性子空间上更具有区分度	运行时间较长，难以挖掘最符合应用目标的群体，缺乏针对性
目标群体发现方法[38,39]	基于个体在特定属性子集上的相似性及在结构上的内聚性挖掘局部群体	发现满足应用目标的群体，运行时间短，针对性强	难以为每个个体分配群体成员身份

标相互转化，因此多数情况下，群体的指标函数是一些经典的社区的指标函数的变形。SA-Cluster [20−23] 通过把属性信息作为属性节点加入网络中，来将群体内的属性均质性转化为结构内聚性。CODICIL[24] 基于节点之间的属性相似度为每个节点创建属性边，同样将属性均质性转化为结构内聚性。SSCG[31] 将节点之间的属性相似度作为权重赋予边，属性越相似则节点之间的边权重越大，从而将属性均质性转化为结构内聚性。基于属性均质性和结构内聚性的转换，传统的社区指标函数可以推广为群体的指标函数。

5 恶意行为群体发现

社交网络为人们信息的获取和发布以及日常的交互提供了一个非常便捷的平台。与此同时，在线社交网络的便捷性，匿名性等特征也给恶意行为群体的形成与发展提供了有利条件。恶意行为群体对社交网络造成的负面影响越来越不可忽视。在大量相关研究中，spam, sybil 和谣言是 3 类研究最为广泛且具有明确定义的恶意行为群体。

Spam 群体通常出现在邮件或网页中，他们通过创建大量的虚假账号和盗用正常用户的账号推荐一些链接，来诱导合法用户进入其推荐的恶意网站，以达到向合法用户发布广告、色情、钓鱼等恶意信息的目的[40,41]。相对于邮件 spam，在线社交媒体中的网页 spam 危害性更大[42]。

Sybil 是一种典型的恶意行为攻击方式，即利用社交网络中的少数节点控制多个虚假身份，从而利用这些身份控制或影响网络的大量正常节点来达到冗余备份的作用[43,44]。而 sybil 群体则指代那些被恶意操控进行协同攻击的虚假账户。

谣言被定义为一类真实性可疑且广泛传播的陈述，其看似可信其实难以证实，并常常引起公众怀疑和焦虑[45]。谣言群体则是传播这些陈述的用户群体。在线社交媒体平台的广泛流行为谣言的传播提供了一个非常有利的条件。因此在线社交网络中的谣言群体发现已成为近年来在线社交网络分析的一大热点。表 3 介绍了这 3 类恶意行为群体的研究现状。

前文论述了在线社交网络中一般化群体的通用发现方法，包括基于个体属性特征的群体发现方法和综合属性和结构信息的群体发现方法。恶意行为群体具有隐蔽性与反检测性，从一般化群体中鉴别出恶意行为群体需要挖掘与利用恶意行为群体不同于一般化群体的特征。下文分别从特征选取和检测算法两方面论述恶意行为群体发现的相关研究。

5.1 恶意行为群体的特征选取

目前对于恶意行为群体的描述通常基于两个方面的特征，即结构和属性上的特征。如何从纷杂的特征中选择最具代表性、最有区分度的特征是恶意群体检测的关键。下面分别综述恶意行为群体结构和属性方面的特征。

恶意行为群体的结构特征包括全局结构特征[46−50]和局部结构特征[51−53]。全局结构特征的计算利用网络全局结构信息。sybil 群体虽然能模仿正常用户的一般行为，但是难以和正常用户形成好友关系并产生频繁的互动[46]，即在对应的关系网络或交互网络中，这些恶意行为群体与正常用户连接就较

表 3 恶意行为群体研究现状

恶意行为群体类型	结构特征	内容属性特征	行为属性特征	进一步研究目标	应用场景
Spam 群体	入度、出度[53]、回复率、聚类系数[43]	发布内容的情感信息[54]、文本信息[55,56]	规律性重复行为[67]	已有的 spam 检测以结构属性为主，应进一步考虑其行为属性特征	邮件、网页、图片、好友请求、新闻中的链接
Sybil 群体	社区结构特性[49]	推荐系统中评分的统计特性[62,63,64]	移动地理位置范围[17]、点击行为模式[18]、日评论量、评论间隔等	sybil 用户之间不一定形成紧密连接，因此以网络结构为基础不能很好地检测出 sybil 群体，需要充分挖掘其行为上的属性特征	推荐网站上的评论，Facebook 等社交网站中的 sybil 攻击
谣言群体	传播行为扩散图中连通子图的结构特征[50]	内容中的矛盾信息[57]、特定短语[58]、询问语句等[59]	转发量、回复量[65]、时间突发性[66]等	能在谣言传播的初期检测出谣言群体并判断真伪。	Twitter、微博等在线服务网站上的帖子中传播的谣言、病毒式营销

为稀少。因此，可以将恶意行为群体与全局网络之间的割作为全局结构特征，用来检测 sybil 群体[47−49]。文献[50]根据用户的历史传播记录建立扩散图，并发现谣言信息的扩散图具有稀疏、零散的结构特征，因此根据扩散图中连通子图的大小和数量可以检测出谣言群体。而局部结构特征的计算则利用网络局部结构信息。Gan 等人[51]发现，spam 用户在入度、出度、回复率等结构特征上与正常用户不同。因此，可以基于用户的入度、出度和回复率等结构特征在社交网络中鉴别出 spam 群体。Boykin 等人[52]发现 sybil、spam 群体具有随机攻击特性，即与恶意行为群体相连的正常用户之间通常无连接关系。该特征反映为恶意行为群体的局部聚类系数较小，因此可以基于局部聚类系数检测出恶意行为群体。

在线社交网络中的用户属性特征描述了用户所具有的特性，包括人口统计属性、内容属性和行为属性等，其中内容属性和行为属性是恶意行为群体检测中常用的属性特征。用户的内容属性特征包括用户所发布、关注或转发内容的文本或语法的统计信息，是一类比较容易获得的属性特征，并且隐含了用户的话题、情感等多种信息。Spam 群体在内容属性上具有可识别的情感特征和文本统计特征[54−56]。Hu 等人[54]根据用户发布内容中隐含的情感信息分布差异，鉴别出这些用户中隐藏的 spam 用户。Shams 等人[55]对用户的一些文本信息进行统计，包括文字数字式词的出现频率、语法和拼写错误出现的频率等，建立检测模型，实现 spam 群体的检测。谣言群体的内容属性主要体现在所传播内容的话题特征、文本统计特征中[57−61]。早期的一些的研究通过转发内容中的矛盾信息[57]、特定的短　语[58]等文本统计特征发现谣言群体。Zhao 等人[59]证明了可以通过检索文本内容中出现的询问语句，再根据话题相关度进行聚类，排除包含询问语句的内容，对剩下的内容按与争议性话题相关度进行排名，最终检测出谣言群体。此外，考虑时间因素，文本的统计信息随时间的变化特征[60]可以用于对谣言群体的检测。在推荐系统中，sybil 群体的评论具有可鉴别的内容属性，即根据 sybil 群体对商品评分的统计特性可以检测出推荐系统中带有特定攻击目的的 sybil 群体[62−64]。

行为属性是用户行为方面的统计特征。基于不同的行为目的，恶意行为群体与正常用户有着不同的行为模式，因此，可以通过对恶意行为群体行为模式的理解，找出能够与正常用户行为模式相区别的行为属性。由于恶意行为群体的隐藏策略不断升级，导致其基于内容属性的检测结果不一定理想，因此许多关于恶意行为群体的研究中重点关注用户的行为属性。在对谣言群体的研究中，通过加入除文本内容信息之外的行为属性方面的统计信息，比如一个帖子的转发量、回复量[65]等，可以改善检测效果。除此之外，考虑谣言群体在行为的时间上具有突发性这一行为属性[66]，能够准确地检测出这些恶意行为群体。综合行为属性和地理位置因素，研究者们发现，正常用户的移动范围因受到地理位置的限制而局限于一个较小的区域，而 sybil 用户不会受到地理位置限制的假设。因此，能够利用用户的移动范围这一行为属性对 sybil 群体进行检测[17]。研究 spam 群体时，研究者们发现 spam 账户在不同的社交网络内容中的行为常常表现出一定的同步性，因此将用户行为的同步性、相似性作为行为属性可以检测出 spam 账户[67]。

由于恶意行为群体应对检测方法的策略不断提升，单纯地基于某一特征或某一类特征，未必能准确地检测出恶意行为群体。因此，采用多种特征相结合的方法来检测恶意行为群体越来越受到重视。

5.2 恶意行为群体的检测算法

在恶意行为群体的检测模型中，需要采用合适的检测算法来根据选取的特征集实现恶意群体的挖掘。这些算法主要分为两类，即有监督的分类算法和无监督的聚类算法。

分类方法是恶意行为群体检测中应用最多的一种检测算法，其一般过程在 3.1 节中已有叙述。已有的恶意行为群体检测模型中，涉及的分类算法包括决策树、支持向量机、随机森林、贝叶斯算法、K-近邻等。在建立恶意行为群体的检测模型时需要根据不同的应用场景和数据特征选择合适的分类算法。同时，为了提高精确度，模型中通常采用十折交叉验证法来评估。在对 spam 群体建立检测模型时，Gan 等人[51]从网页内容和网站链接中提取特征，使用决策树建立分类模型，模型中为了避免过度拟合采用了置信度阈值来剪枝。在对 sybil 群体建立检测模型时，Zhang 等人[17]建立 3 组不同的用户行为特征集，基于支持向量机在同一组数据集上进行 sybil 群体挖掘。根据 3 组特征集对应的检测效果，选出 sybil 群体的最优行为特征。在对谣言群体建立检测模型时，Ma 等人[68]将所有的群体特征分为基于内容、用户和扩散的特征，并采用线性支持向量机作为分类算法建立分类模型，检测出谣言群体。在分类算法的选择上，他们证明了在该检测环境下，由于随机森林无法获取复杂的传播信息，选取线性支持向量机作为分类算法检测效果更好。

恶意行为群体检测聚类算法的一般过程是根据

恶意行为群体和正常用户群体在所选特征集上差异较大，而同一恶意行为群体内的个体在所选特征集上差异较小这一特点，定义相似度指标，根据相似度逐点聚类。在这一类检测模型中，相似度指标或相应的距离指标的定义具有关键性意义。常用的衡量相似性的指标包括 Jaccord 相似性、相关系数、欧式距离等。在对 sybil 群体建立检测模型时，Wang 等人[18]将 Jaccord 相似度作为相似性指标，用聚类算法建立了基于用户行为序列相似性的检测模型，实现 sybil 群体的挖掘。相似地，Zhao 等人[59]考虑到与同一个谣言相关的帖子内容上具有极大的相似性，在谣言的检测模型中采用 Jaccord 相似度为相似性指标，用聚类算法来对内容相似的帖子进行聚类，挖掘出谣言群体。

6 进一步研究方向

在线社交网络属性多、复杂性高、群体结构多样，如何基于先验认识准确把握一般化群体结构和属性特征，进而准确地描述群体仍是需要解决的问题。在特征的选取方面，现有方法主要集中在一般化的结构连接紧密特征和人口统计属性相似特征，对用户的行为和情感等更加反应社交网络用户内在特性的特征关注较少，而这些特征往往是突发群体形成的关键因素。群体的应用非常广泛，在不同应用中往往需要发现有针对性的群体。目前对群体的定义主要基于一到两种特征，对群体的描述较为泛化，如何针对特定的应用环境从多个特征角度描述定义目标群体仍需进一步持续研究。快速的启发式群体发现算法一直是研究关键点，在线社交网络规模的巨大性要求进一步设计稳定有效的亚线性启发式算法。另外，当前大部分群体发现方法忽视了在线社交网络的动态演化特性，事实上个体的属性信息和网络关系结构都在不断的变化，因此群体结构也是动态演化的，如何快速准确发现不同时间段的群体并分析群体的演化规律及生命周期是进一步的研究方向。

在恶意群体检测研究中，一方面，不同类型的恶意行为群体具有不同的行为目的，表现不同的行为模式，因而通常具有不同的结构和属性特征。在进行特征选取时，要结合各种恶意行为群体的不同特点选取适当的特征。例如，Spam 群体主要通过散布大量链接来传播恶意内容，因此 Spam 群体的检测可以主要考虑了其发布的内容属性特征。而谣言群体的特点主要体现在其传播内容，传播广度和传播速度上，因此谣言群体在内容属性和行为属性上都体现出了可识别特征。另一方面，由于恶意行为群体应对检测方法的策略不断提升，单纯地基于某一特征或某一类特征，难以准确地检测出恶意行为群体。因此，需要采用多种特征相结合的方法来检测恶意行为群体。在特征选择方法上，目前对恶意行为群体特征的选择主要是基于先验知识或者使用监督学习的方法进行选择验证，需要有大量已知标签的训练集作为支撑。对于一些无法提供大量训练集的情况，传统基于训练集的有监督学习方法并不可行，而基于深度学习的特征自动选择方法将是进一步的研究方向。

7 结束语

群体是在线社交网络重要的中观结构。准确快速发现群体对在线社交网络分析、研究和应用有巨大的推动作用。本文首先介绍群体的形成机理，给出群体的一般化定义，并在此基础上介绍群体发现问题。然后基于群体的形成机理和定义分别阐述基于属性的群体发现方法和属性与结构信息结合的群体发现方法。接着总结介绍在线社交网络中恶意行为群体的鉴别研究。最后展望群体发现技术的发展方向，指出在线社交网络是一个大规模的复杂动态系统，还需要进一步研究更加适应社交网络复杂性和动态性的群体发现和应用方法。

参 考 文 献

[1] 方滨兴, 贾焰, 韩毅. 社交网络分析核心科学问题、研究现状及未来展望[J]. 中国科学院院刊, 2015, 30(2): 187–199. doi: 10.16418/j.issn.1000-3045.2015.02.007.
FANG Binxing, JIA Yan, and HAN Yi. Social network analysis—key research problems, related work, and future prospects[J]. *Bulletin of Chinese Academy of Sciences*, 2015, 30(2): 187–199. doi: 10.16418/j.issn.1000-3045.2015.02.007.

[2] 许进, 杨扬, 蒋飞, 等. 社交网络结构特性分析及建模研究进展[J]. 中国科学院院刊, 2015, 30(2): 216–228. doi: 10.16418/j.issn.1000-3045.2015.02.009.
XU Jin, YANG Yang, JIANG Fei, *et al.* Social network structure feature analysis and its modelling[J]. *Bulletin of Chinese Academy of Sciences*, 2015, 30(2): 216–228. doi: 10.16418/j.issn.1000-3045.2015.02.009.

[3] HE L, LU C-T, MA J, *et al.* Joint community and structural hole spanner detection via harmonic modularity[C]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016: 875–884.

[4] YING X, WANG C, WANG M, *et al.* CoDAR: Revealing the generalized procedure & recommending algorithms of community detection[C]. Proceedings of the 2016 International Conference on Management of Data, San

Francisco, CA, USA, 2016: 2181–2184.

[5] SHAHRIARI M, GUNASHEKAR S, DOMARUS M V, *et al*. Predictive analysis of temporal and overlapping community structures in social media[C]. Proceedings of the 25th International Conference Companion on World Wide Web, Geneva, Switzerland, 2016: 855–860.

[6] LIANG X, TANG J, and PAN L. A neighborhood vector propagation algorithm for community detection[C]. 2014 IEEE Global Communications Conference, Austin, TX, USA, 2014: 2923–2928.

[7] WU P and PAN L. Multi-objective community detection based on memetic algorithm[J]. *PloS One*, 2015, 10(5): e0126845. doi: 10.1371/journal.pone.0126845.

[8] CRANE R and SORNETTE D. Robust dynamic classes revealed by measuring the response function of a social system[J]. *Proceedings of the National Academy of Sciences,* of the United States of America, 2008, 105(41): 15649–15653. doi: 10.1073/pnas.0803685105.

[9] KANE G C, ALAVI M, LABIANCA G, *et al*. What's different about social media networks? A framework and research agenda[J]. *MIS Quarterly*, 2014, 38(1): 274–304.

[10] ATKIN R. Combinatorial Connectivities in Social Systems: An Application of Simplicial Complex Structures to the Study of Large Organizations[M]. Swiss, Birkhauser, 1977: 71–91.

[11] AGGARWAL C C, WOLF J L, YU P S, *et al*. Fast algorithms for projected clustering[C]. Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data, Philadelphia, Pennsylvania, USA, 1999: 61–72.

[12] WOO KG, LEE JH, KIM MH, *et al*. FINDIT: A fast and intelligent subspace clustering algorithm using dimension voting[J]. *Information and Software Technology*, 2004, 46(4): 255–271.doi: 10.1016/j.infsof.2003.07.003.

[13] YIP K P, CHEUNG D W, and NG M K. On discovery of extremely low-dimensional clusters using semi-supervised projected clustering[C]. 21st International Conference on Data Engineering (ICDE'05), Tokoyo, Japan, 2005: 329–340.

[14] AGRAWAL R, GEHRKE J, GUNOPULOS D, *et al*. Automatic subspace clustering of high dimensional data for data mining applications[C]. Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, Seattle, Washington, USA, 1998: 94–105.

[15] CHENG C-H, FU A W, and ZHANG Y. Entropy-based subspace clustering for mining numerical data[C]. Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, California, USA, 1999: 84–93.

[16] ASSENT I, KRIEGER R, *et al*. EDSC: Efficient density-based subspace clustering[C]. Proceedings of the 17th ACM Conference on Information and Knowledge Management, Napa Valley, California, USA, 2008: 1093–1102.

[17] ZHANG X, ZHENG H, LI X, *et al*. You are where you have been: Sybil detection via geo-location analysis in OSNs[C]. Global Communications Conference, Austin,TX, USA, 2014: 698–703.

[18] WANG G, KONOLIGE T, WILSON C, *et al*. You are how you click: Clickstream analysis for sybil detection[C]. Proceedings of the 22nd USENIX Conference on Security, Washington, DC, USA, 2013: 1–15.

[19] McCORD M and CHUAH M. Spam Detection on Twitter Using Traditional Classifiers[M]. In Autonomic and Trusted Computing. Springer, 2011: 175–186.

[20] ZHOU Y, CHENG H, and YU J X. Graph clustering based on structural/attribute similarities[J]. *Proceedings of the VLDB Endowment*, 2009, 2(1): 718–729. doi: 10.14778/1687627.1687709.

[21] ZHOU Y, CHENG H, and YU J X. Clustering large attributed graphs: An efficient incremental approach[C]. 2010 IEEE International Conference on Data Mining, Sydney, NSW, Australia, 2010: 689–698.

[22] CHENG H, ZHOU Y, and YU J X. Clustering large attributed graphs: A balance between structural and attribute similarities[J]. *ACM Transactions on Knowledge Discovery from Data*, 2011, 5(2): 1–33. doi: 10.1145/1921632.1921638.

[23] CHENG H, ZHOU Y, HUANG X, *et al*. Clustering large attributed information networks: An efficient incremental computing approach[J]. *Data Mining and Knowledge Discovery*, 2012, 25(3): 450–477. doi: 10.1007/s10618-012-0263-0.

[24] RUAN Y, FUHRY D, and PARTHASARATHY S. Efficient community detection in large networks using content and links[C]. Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 2013: 1089–1098.

[25] AKOGLU L, TONG H, MEEDER B, *et al*. PICS: Parameter-free identification of cohesive subgroups in large attributed graphs[C]. Proceedings of the 2012 SIAM International Conference on Data Mining, Anaheim, CA, USA, 2012: 439–450.

[26] XU Z, KE Y, WANG Y, *et al*. A model-based approach to attributed graph clustering[C]. Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, Scottsdale, Arizona, USA, 2012: 505–516.

[27] XU Z, KE Y, WANG Y, *et al*. GBAGC: A general Bayesian framework for attributed graph clustering[J]. *ACM Transactions on Knowledge Discovery from Data*, 2014, 9(1): 1–43. doi: 10.1145/2629616.

[28] WU P and PAN L. Multi-objective community detection

method by integrating users' behavior attributes[J]. *Neurocomputing*, 2016, 210. 13–25. doi: 10.1016/j.neucom.2015.11.128.

[29] SILVA A, WAGNER MEIRA J, and ZAKI M J. Mining attribute-structure correlated patterns in large attributed graphs[J]. *Proceedings of the VLDB Endowment*, 2012, 5(5): 466–477. doi: 10.14778/2140436.2140443.

[30] YANG J, MCAULEY J, and LESKOVEC J. Community detection in networks with node attributes[C]. 2013 IEEE 13th International Conference on Data Mining, Dallas, TX, USA, 2013: 1151–1156.

[31] GUNNEMANN S, FARBER I, RAUBACH S, *et al*. Spectral subspace clustering for graphs with feature vectors[C]. 2013 IEEE 13th International Conference on Data Mining, Dallas, TX, USA, 2013: 231–240.

[32] GUNNEMANN S, FARBER I, BODEN B, *et al*. GAMer: A synthesis of subspace clustering and dense subgraph mining[J]. *Knowledge and Information Systems*, 2014, 40(2): 243–278. doi: 10.1007/s10115-013-0640-z.

[33] HUANG X, CHENG H, and YU J X. Dense community detection in multi-valued attributed networks[J]. *Information Sciences*, 2015, 314: 77–99. doi: 10.1016/j.ins.2015.03.075.

[34] REVELLE M, DOMENICONI C, SWEENEY M, *et al*. Finding community topics and membership in graphs[C]. Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Porto, Portugal, 2015: 625–640.

[35] ATZMUELLER M, DOERFEL S, and MITZLAFF F. Description-oriented community detection using exhaustive subgroup discovery[J]. *Information Sciences*, 2016, 329. 965–984. doi: 10.1016/j.ins.2015.05.008.

[36] YIN H, HU Z, ZHOU X, *et al*. Discovering interpretable geo-social communities for user behavior prediction[C]. 2016 IEEE 32nd International Conference on Data Engineering (ICDE), Helsinki, Finland, 2016: 942–953.

[37] LIU L, XU L, WANGY Z, *et al*. Community detection based on structure and content: A content propagation perspective [C]. 2015 IEEE International Conference on Data Mining (ICDM), Atlantic City, NJ, USA, 2015: 271–280.

[38] POOL S, BONCHI F, and LEEUWEN M V. Description-driven community detection[J]. *ACM Transactions on Intelligent Systems and Technology*, 2014, 5(2): 28. doi: 10.1145/2517088.

[39] PEROZZI B, AKOGLU L, *et al*. Focused clustering and outlier detection in large attributed graphs[C]. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 2014: 1346–1355.

[40] THOMAS K, MCCOY D, GRIER C, *et al*. Trafficking fraudulent accounts: The role of the underground market in Twitter spam and abuse[C]. Proceedings of the 22nd USENIX Conference on Security, Washington, D.C., USA, 2013: 195–210.

[41] HUANG T-K, RAHMAN M S, MADHYASTHA H V, *et al*. An analysis of socware cascades in online social networks[C]. Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 2013: 619–630.

[42] ZHANG X, LI Z, ZHU S, *et al*. Detecting spam and promoting campaigns in Twitter[J]. *ACM Transactions on the Web*, 2016, 10(1): 1–28. doi: 10.1145/2846102.

[43] SINGH A, NGAN T W, DRUSCHEL P, *et al*. Eclipse attacks on overlay networks: Threats and defenses[C]. 25th IEEE International Conference on Computer Communications, Waikoloa, Hawaii, USA, 2006: 1–12.

[44] SIT E and MORRIS R. Security considerations for peer-to-peer distributed hash tables[C]. Revised Papers from the First International Workshop on Peer-to-Peer Systems, Springer-Verlag, 2002: 261–269.

[45] ZUBIAGA A, LIAKATA M, PROCTER R, *et al*. Towards detecting rumours in social media[C]. Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, Texas, USA, 2015: 35–41.

[46] 程晓涛, 刘彩霞, 刘树新. 基于关系图特征的微博水军发现方法[J]. 自动化学报, 2015, 41(9): 1533–1541.
CHENG Xiaotao, LIU Caixia, and LIU Shuxin. Graph-based features for identifying spammers in microblog networks[J]. *Acta Automatica Sinica*, 2015, 41(9): 1533–1541.

[47] VISWANATH B, MONDAL M, CLEMENT A, *et al*. Exploring the design space of social network-based sybil defenses[C]. 2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012), Bangalore, India, 2012: 1–8.

[48] VISWANATH B, POST A, GUMMADI K P, *et al*. An analysis of social network-based sybil defenses[J]. *ACM SIGCOMM Computer Communication Review*, 2011, 41(4): 363–374. doi: 10.1145/1851275.1851226.

[49] DANEZIS G and MITTAL P. SybilInfer: Detecting sybil nodes using social networks[C]. The Network and Distributed System Security Symposium, San Diego, CA, USA, 2009.

[50] KWON S, CHA M, JUNG K, *et al*. Prominent features of rumor propagation in online social media[C]. 2013 IEEE 13th International Conference on Data Mining (ICDM), Dallas, TX, USA, 2013: 1103–1108.

[51] GAN Q and SUEL T. Improving web spam classifiers using link structure[C]. Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web, Banff, Alberta, Canada, 2007: 17–20.

[52] BOYKIN P O and ROYCHOWDHURY V P. Leveraging

social networks to fight spam[J]. *Computer*, 2005, 38(4): 61–68. doi: 10.1109/MC.2005.132.

[53] FAKHRAEI S, FOULDS J, SHASHANKA M, *et al.* Collective spammer detection in evolving multi-relational social networks[C]. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, 2015: 1769–1778.

[54] HU X, TANG J, GAO H, *et al.* Social spammer detection with sentiment information[C]. Proceedings of the 2014 IEEE International Conference on Data Mining, Shenzhen, China, 2014: 180–189.

[55] SHAMS R and MERCER R E. Classifying spam emails using text and readability features[C]. 2013 IEEE 13th International Conference on Data Mining, Dallas, TX, USA, 2013: 657–666.

[56] SANDULESCU V and ESTER M. Detecting singleton review spammers using semantic similarity[C]. Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 2015: 971–976.

[57] CASTILLO C, MENDOZA M, and POBLETE B. Information credibility on Twitter[C]. Proceedings of the 20th International Conference on World Wide Web, New York, NY, USA, 2011: 675–684.

[58] ENNALS R, BYLER D, AGOSTA J M, *et al.* What is disputed on the web?[C]. Proceedings of the 4th Workshop on Information Credibility, Raleigh, North Carolina, USA, 2010: 67–74.

[59] ZHAO Z, RESNICK P, and MEI Q. Enquiring minds: Early detection of rumors in social media from enquiry posts[C]. Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 2015: 1395–1405.

[60] TAKAHASHI T and IGATA N. Rumor detection on Twitter[C]. Joint 6th International Conference on Soft Computing and Intelligent Systems and 13th International Symposium on Advanced Intelligent Systems, Kobe, Japan, 2012: 452–457.

[61] ZHOU X, CAO J, JIN Z, *et al.* Real-time news certification system on Sina Weibo[C]. Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 2015: 983–988.

[62] NOH G and KIM C K. RobuRec: Robust sybil attack defense in online recommender systems[C]. 2013 IEEE International Conference on Communications, Budapest, Hangary, 2013: 2001–2005.

[63] YANG Y, SUN Y, KAY S, *et al.* Securing rating aggregation systems using statistical detectors and trust[J]. *IEEE Transactions on Information Forensics & Security*, 2009, 4(4): 883–898. doi: 10.1109/TIFS.2009.2033741.

[64] YU H, SHI C, KAMINSKY M, *et al.* DSybil: Optimal sybil-resistance for recommendation systems[C]. 30th IEEE Symposium on Security and Privacy, Washington, DC, USA, 2009: 283–298.

[65] GUPTA A and KUMARAGURU P. Credibility ranking of tweets during high impact events[C]. Proceedings of the 1st Workshop on Privacy and Security in Online Social Media, Lyon, France, 2012: 2–8.

[66] GUPTA A, LAMBA H, and KUMARAGURU P. Prayforboston: Analyzing fake content on Twitter[C]. eCrime Researchers Summit, San Francisco, CA, USA, 2013: 1–12.

[67] CAO Q, YANG X, YU J, *et al.* Uncovering large groups of active malicious accounts in online social networks[C]. Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, Arizona, USA, 2014: 477–488.

[68] MA J, GAO W, WEI Z, *et al.* Detect rumors using time series of social context information on microblogging websites[C]. Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia, 2015: 1751–1754.

潘　理：　男，1974 年生，研究员，研究方向为网络安全管理、社交网络分析等.

吴　鹏：　男，1989 年生，博士生，研究方向为数据挖掘、社交网络分析等.

黄丹华：　女，1992 年生，博士生，研究方向为数据挖掘、社交网络分析等.