MINING SOCIAL NETWORKS FOR ANOMALIES: METHODS AND CHALLENGES

Byung Wook Kwon

CONTENTS

CONTENTS

1. Abstract

- Online social networks have received a dramatic increase of interest in the last decade due to the growth of Internet and Web 2.0.

- they have become primary targets for malicious users who attempted to perform illegal activities and cause harm to other users.

- Anomaly detection in social networks refers to the problem of identifying the strange and unexpected behavior of users by exploring the patterns hidden in the networks.

- In this paper, we provide a comprehensive review of a large set of methods for mining social networks for anomalies by providing a multi-level taxonomy to categorize the existing techniques based on the nature of input network.

2. Background

2.1 Social Network Analysis

- A social network is a social structure that represents the relationships or interactions among social entities such as friends.
- Social networks are typically modeled by using graphs.
- Social Network Analysis (Wasserman and Faust, 1994) is the field of study which investigates the properties of social networks.
- SNA helps us to explore the relationships between individuals who are connected through social networks, and provides understanding about the inherent patterns that are embedded in these social graphs (Scott, 2011).

2. Background

2.2 Anomaly detection

- An anomaly is a data item whose behavior varies significantly from the regular pattern of the rest of the data.
- Anomaly detection or outlier detection is a key problem in data mining and is defined as the problem of detecting anomalous patterns in a dataset.
- Whereas other data mining algorithms such as classification, clustering, and frequent pattern analysis find popular patterns from the dataset, anomaly detection identifies relatively small set of objects that differ from the normal behavior of the larger number of data items in the dataset (Dang et al., 2014).

2. Background

2.3 Graph-based anomaly detection

- Graph-based anomaly detection is the problem of finding anomalies from data that are represented as graphs, by using graph mining techniques. Graph mining, which has been a popular area of research in recent years, is the novel approach for extracting useful knowledge from graph data by using techniques from fields such as machine learning, data mining, statistics, pattern recognition, and graph theory (Jiang, 2011).

- Substantial research works have been directed towards graph mining because of its various applications in a multitude of practical domains including bioinformatics, chemical compound analysis, program flow structures, computer networks, and online social networks.

3. Application domains of anomaly detection in social networks

- **3.1 Fraud detection in online social networks**

- Online social networks have become new targets for cybercrime, and malicious users attempt to perform illegal activities such as cyber attacks, bullying, fraudulent information dissemination, organized crimes, and even terrorist attack planning on these systems (Yu et al., 2015).

- The fraudulent activities of users in online social networks necessitate the importance of digital forensics in social networks arena.

- Anomaly detection can provide inherent and invaluable information for detecting criminal activities in social networks for Social Network Forensics (SNF) (Keyvanpour et al., 2014).

3. Application domains of anomaly detection in social networks

- **3.2 Fraud detection in online social networks**

- An insider threat is a security threat to an organization from individuals within the organization. The threat can be fraud, theft of sensitive information, and destruction or compromise of hardware and software resources.

- Network-based anomaly detection measures can be used for identifying insider threats. The work presented in Eberle et al.

3. Application domains of anomaly detection in social networks

- **3.3 Fraud detection in online social networks**

- Nowadays, it has become a common practice for customers to read the opinions or reviews written by other customers before deciding to purchase a product.

- These reviews are also used as feedbacks by manufacturers to identify the problems of their products. However, in review websites such as amazon.com, fraudsters write fake reviews or bogus reviews to defame or boost the reputation of manufacturers and vendors, and to mislead the customers.

4. Related work

- Anomaly detection is a broad field that has been well-researched in a wide variety of application domains. Several extensive survey articles and books (Hodge and Austin, 2004; Chandola et al., 2009; Aggarwal, 2013) have been published on non-network based anomaly detection problem.

- They mainly focus on independent and identically distributed and multidimensional data objects. Hodge and Austin (2004) provide a comprehensive review of outlier detection approaches formulated in machine learning and statistical domain.

5. Different aspects of anomaly detection in social networks

Table 1
Related survey articles and their area of focus.

Article	Area of focus
Hodge and Austin (2004)	Outlier detection approaches in machine learning and statistics
Chandola et al. (2009)	General, non-network based anomaly detection
Aggarwal (2013)	General, non-network based anomaly detection
Agyemang et al. (2006)	Outlier detection techniques for numeric and symbolic data
Chandola et al. (2012)	Anomaly detection in discrete sequences
Zimek et al. (2012)	Unsupervised outlier detection in high-dimensional numerical data
Gupta et al. (2014a)	Outlier detection for temporal data
Schubert et al. (2014)	Local outlier detection with respect to spatial, video, and network data
Patcha and Park (2007)	Anomaly detection on computer networks
Gogoi et al. (2011)	Outlier detection methods in network anomaly identification
Bhuyan et al. (2014)	Anomaly detection on computer networks
Ahmed et al. (2016)	Anomaly detection on computer networks
Zhang et al. (2010)	Outlier detection techniques for wireless sensor networks
Akoglu et al. (2015)	General graph-based anomaly detection techniques
Ranshous et al. (2015)	Anomaly detection in dynamic networks
Savage et al. (2014)	Anomaly detection in online social networks

5. Different aspects of anomaly detection in social networks

- **5.1. Nature of input networks**
- In the case of social network anomaly detection, the input networks that are used by various methods can be categorized as static/dynamic or attributed/unattributed, based on the type of analysis performed for anomaly detection.

- **5.1.1. Static versus dynamic networks**
- the static network is able to represent only a single snapshot of the social network data at any instant of time. However, social networks are constantly evolving, leading to dynamic networks.

- Dynamic networks from different application domains evolve differently over time which leads to application-specific anomalies and approaches. For slowly evolving networks such as bibliographic networks, snapshot analysis can be employed, by analyzing the state of the networks at two different instants of time

5. Different aspects of anomaly detection in social networks

- **5.1.2. Unattributed versus attributed networks**
- In unattributed or unlabeled networks, the attributes or labels associated with the individuals or their interactions such as any details about the type of interaction, the age of the individuals involved in the interaction, or the duration of the interaction are not considered for anomaly detection.

- **5.2 Types of anomailes in social networks**
- Another important characteristic of an anomaly detection technique is the type of anomalies and the manner in which the anomalies are reported.
- The detection methods un-cover anomalous units such as nodes, edges, subgraphs, and/or events in the networks. The methods either assign an anomaly score to each unit, or classify each unit as normal or anomalous.

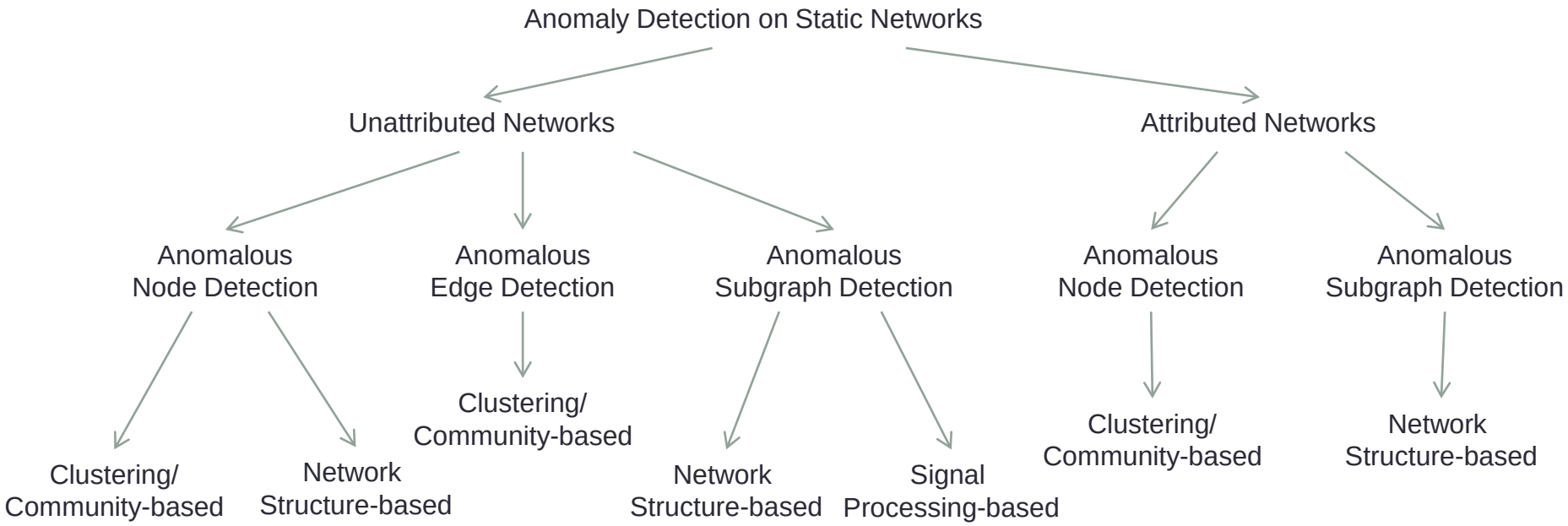
5. Different aspects of anomaly detection in social networks

- **5.2.2. Anomalous edges**
- If we need to identify unusual or irregular interactions among the users in the network, we may view a subset of edges as anomalous.
- The edge weight can correspond to the number of messages exchanged between individuals in the network.

- **5.2.3. Anomalous subgraphs**
- Anomalous subgraph detection aims at finding the subnetworks such that the pattern of interaction among the nodes in the subnetwork is irregular compared to the other nodes in the network.

- **5.2.4. Events**
- Events occur exclusively in dynamic networks and are the discrete time steps at which the social network is substantially different from the preceding and the succeeding networks in the dynamic network sequence.

6. Anomaly detection in static social networks



- Fig. 1. Taxonomy of the survey on static social networks. Methods are categorized based on the nature of input network, type of anomalies they identify, and the underlying approach.

6. Anomaly detection in static social networks

- **6.1 Static unattributed networks**
- Anomalies in static unattributed social networks are identified by analyzing the interaction between the users or structure of the network at any instant of time, as the details about the individuals as well as their interactions are either ignored or not available.

- In this section, first we categorize the anomaly detection methods based on the underlying approach, and then we provide a review of the various anomaly detection techniques that are developed for static unattributed networks.

6. Anomaly detection in static social networks

- (a) **Clustering/Community-based approaches:** The community-based approaches for anomaly detection in static networks find densely connected nodes in the network as clusters or communities, and identify the nodes or edges that inter-connect different communities.
- (b) **Network structure-based approaches:** This group of approaches exploit the given network structure or different shapes of to-pology to compute graph-specific feature space and to detect anomalous nodes (Akoglu et al., 2010; Hassanzadeh and Nayak, 2013a,b).
- (c) **Signal processing-based approaches:** Signal processing on graphs is an emerging field that extends the traditional signal processing for graphs and treat comparatively small anom-alous subgraphs as signals embedded in a much bigger network taken as the background noise.

6. Anomaly detection in static social networks

- **6.1.1. Anomalous node detection**
- This section reviews the anomaly detection methods that have been developed for spotting anomalous nodes in static unattributed networks.

- (a) **Clustering/Community-based approaches:** One of the earliest works on community-based anomaly detection by Sun et al.(2005) addresses the problems of neighborhood formation and spotting anomalous nodes in bipartite networks.
- The authors proposed an algo-rithm to find the community or neighborhood of each node in the bipartite graph using random walks with restart and graph parti-tioning, and used this algorithm to detect anomalous nodes in the network.

6. Anomaly detection in static social networks

- (b) **Network structure-based approaches:** Network structure-based approach: Akoglu et al. (2010) pro-posed a network structure-based technique called OddBall to discover anomalies such as near-clique, near star, heavy vicinity, and dominant edge patterns from large, weighted networks.

- The Hassanzadeh et al. (2012) proposed a frame-work based on egonet's features for identifying anomalous nodes in social networks.

- The framework aims to find the common be-havior obeyed by majority of the nodes by computing graph the-oretic properties of a node's egonet such as number of nodes, number of edges, the average betweenness centrality, and com-munity cohesiveness of the node's super-egonet. It then models the relationships between these metrics by using distribution models such as linear and power laws.

6. Anomaly detection in static social networks

- **6.1.2. Anomalous edge detetion**
- The existing anomalous edge detection methods on static un-attributed networks are based on clustering/community-based approaches. This section provides an overview of the different approaches developed for detecting anomalous edges from a static unattributed network.

- **Clustering/Community-based approach:** Chakrabarti (2004) developed AUTOPART, a parameter-free, scalable, and iterative approach, to detect anomalous edges and node groups.

- AUTOPART automatically partitions the network into clusters by re-organizing the rows and columns of the adjacency matrix. The edges that do not belong to any cluster as well as the edges that interconnect different clusters represent anomalies.

6. Anomaly detection in static social networks

- **6.1.3. Anomalous subgraph detection**
- The algorithms developed for detecting anomalous subgraphs from a static unattributed network are based on network structure-based and signal processing-based approaches.

- **Network structure-based approach:** In e-mail spam and viral marketing in social networks, the fraudulent user creates a set of fake identities and uses these identities to interact with a large random set of innocent users.

- The Shrivastava et al. (2008) defined Random Link Attacks (RLA) to model such collaborative malicious activities and proposed algorithms to mine subgraphs satisfying the RLA property. RLA is detected in two steps:

6. Anomaly detection in static social networks

- the first step is identifying the suspect nodes that are possibly part of the attacker cluster by conducting two tests on each individual nodes in the network: clustering test and neighborhood independence test.
- As the innocent users are chosen at random, they are improbable to be interconnected, forming a star-like pattern in the network.
- In order to detect the suspect nodes in the network, the triangles in each egonet are counted, with a lower triangle count indicating an attacker.
- In the next step, the attack set is identified by growing the neighborhood of the suspect nodes.

6. Anomaly detection in static social networks

- **Signal processing-based approach:** The signal processing-based framework proposed in Miller et al. (2010) uses L1 properties of the eigenvectors of the network's modularity matrix to determine the presence of an anomalous subgraph.

- The framework aims at determining whether an observed network was generated by a given random process or if there is other behavior that deviates from the mode.
  - H0 : The network is the large background graph without anomalous subgraphs (noise).
  - H1 : The network is the large background graph with anomalies (signalþnoise).

- In Miller et al. (2011), the detection framework and the related algorithms proposed in Miller et al. (2010) are applied for the specific problem of detecting threat subgraphs embedded in social networks.

6. Anomaly detection in static social networks

- **6.2. Static attributed network**
- Anomalies in static attributed networks are determined by analyzing the network topology as well as the attributes associated with nodes and/or edges.

- The anomaly detection methods developed for static attributed networks mainly identify anomalous nodes and subgraphs.

- In this section, first we categorize the methods based on the underlying approach they follow, and then provide an overview of the methods that have been proposed for detecting anomalies on static attributed networks.

6. Anomaly detection in static social networks

- Types of approaches: The anomaly detection approaches that have been reported for static attributed networks can be categorized into two groups: clustering/community-based and network structure-based approaches.

- **Clustering/Community-based approach:** This type of anomaly detection methods identifies the subset of nodes within the context of communities such that their characteristics differ significantly from other members of the same community.

- The community-based anomaly detection methods proposed in Gao et al. (2010) and Muller et al. (2013) integrate attribute graph clustering and outlier detection in a single algorithm.

6. Anomaly detection in static social networks

- **Network Structure-based approach:** This class of anomaly detection methods aims to spot subgraphs or substructures in the network that are unusual with respect to the structure as well as labels or attributes of the network. Most of these methods make use of the SUBDUE (Holder et al., 1994) system, which is based on MDL principle, for discovering frequent substructures within networks.

6. Anomaly detection in static social networks

- **6.2.1. Anomalous node detection**
- The detection methods that have been reported for anomalous nodes are based on clustering / community-based approach. This section reviews the different methods for identifying anomalous nodes.

- Gao et al. (2010) introduced the concept of community outliers, and developed a unified framework for detecting outliers and for discovering communities. They proposed community outlier detection algorithm (CODA) that combines both community detection and outlier detection in a probabilistic formulation based on Hidden Markov Random Fields (HMRF), by combining information from both network structure and node attributes.

6. Anomaly detection in static social networks

- GoutRank (Muller et al., 2013) is an approach for ranking network nodes according to their degree of deviation in both node attributes and network structure. The approach focuses on the detection of complex outliers that deviate with respect to a subgraph of highly connected nodes. Whereas the individual outlier is highly similar to other nodes in the subgraph, it deviates significantly with respect to a subset of significant attributes called subspaces.

- Yang et al. (2015) developed a framework based on bipartite graph and co-clustering for detecting anomalous users and messages in microblogging. A bipartite graph is constructed to represent homogeneous and heterogeneous interactions between users and messages.

6. Anomaly detection in static social networks

- **6.2.2. Anomalous subgraph detection**
- The anomaly detection methods that aim to identify anomalous subgraphs are based on network structure-based approach

- Noble and Cook (2003) introduced two methods for identifying unusual patterns in a network with categorical labels using the SUBDUE system.
- The first method identifies specific, unusual substructures within a network. In the second method, anomalous subgraphs are detected by partitioning the network into distinct, separate subgraphs and then comparing each of them against the other subgraphs for unusual occurrences.
- The main idea behind the two methods is that subgraphs containing frequent substructures are generally less anomalous than subgraphs with few frequent substructures.

6. Anomaly detection in static social networks

- The GBAD algorithms (Eberle and Holder, 2007) operate on unweighted networks with discrete node and edge labels, and cannot incorporate continuous labels.

- To overcome this problem, Davis et al. (2011) presented YAGADA (Yet Another Graphbased Anomaly Detection Algorithm), an algorithm to search attributed networks for anomalies using both structural information and numeric labels.

- If the numeric values in the network are normal, the algorithm discretizes the values with the same constant categorical label. If the values are abnormal, they are assigned an anomalous score. When the network is subsequently searched for frequent substructures, the nodes with the same constant value are incorporated into frequent patterns.

6. Anomaly detection in static social networks

Table 2
Summary and comparison of articles on anomaly detection in static social networks.

Research paper	Network		Anomaly			Approach used	Reporting of anomalies
	Unattributed	Attributed	Node	Edge	Subgraph		
Sun et al. (2005)	✓		✓			Clustering/community-based	Normality scores
Xu et al. (2007)	✓		✓			Clustering/community-based	Binary labels
Sun et al. (2010)	✓		✓			Clustering/community-based	Binary labels
Akoglu et al. (2010)	✓		✓			Network structure-based	Anomaly scores
Hassanzadeh et al. (2012)	✓		✓			Network structure-based	Anomaly scores
Hassanzadeh and Nayak (2013b)	✓		✓			Network structure-based	Binary labels
Hassanzadeh and Nayak (2013a)	✓		✓			Network structure-based	Binary labels
Chakrabarti (2004)	✓		✓	✓		Clustering/community-based	Binary labels
Tong and Lin (2011)	✓		✓	✓		Clustering/community-based	Binary labels
Shrivastava et al. (2008)	✓				✓	Network structure-based	Node subsets
Miller et al. (2010)	✓				✓	Signal processing-based	Node subsets
Miller et al. (2011)	✓				✓	Signal processing-based	Node subsets
Miller et al. (2015)	✓				✓	Signal processing-based	Node subsets
Gao et al. (2010)		✓	✓			Clustering/community-based	Binary labels
Muller et al. (2013)		✓	✓			Clustering/community-based	Anomaly scores
Yang et al. (2015)		✓	✓			Clustering/community-based	Node subsets and messages
Noble and Cook (2003)		✓			✓	Network-structure based	Anomaly scores
Eberle and Holder (2007)		✓			✓	Network-structure based	Anomaly scores
Davis et al. (2011)		✓			✓	Network-structure based	Anomaly scores
Gupta et al. (2014b)		✓			✓	Network-structure based	Anomaly scores

7. Anomaly detection in dynamic social networks

- As dynamic social networks are constantly undergoing alterations to their structure and/or attributes, the major tasks in identifying anomalies are to detect change-points or events in time at which the majority of the nodes or edges deviate from their normal behavior, and to identify the particular parts of the network that are responsible for the change-point. When considering the dynamic nature of networks, new types of community-based anomalies such as formation of new communities, and splitting up and disappearance of existing communities are also possible (Chen et al., 2012d). This section provides a review of the existing methods for identifying anomalies in dynamic social networks. The taxonomy of the survey is shown in Fig. 2.

7. Anomaly detection in dynamic social networks

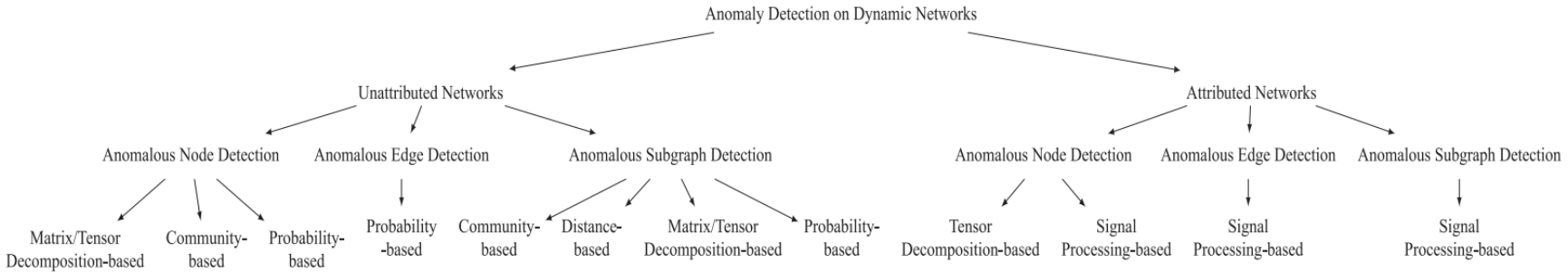


Fig. 2. Taxonomy of the survey on dynamic social networks. Methods are categorized based on the nature of input network, type of anomalies they identify, and the underlying approach.

7. Anomaly detection in dynamic social networks

- **7.2 Dynamic unattributed networks**
- For detecting anomalies in dynamic unattributed networks, only the alterations in structure of the network in subsequent time steps are considered; the attributes associated with nodes or edges are not considered.
- (a) **Matrix/Tensor decomposition-based approaches:** The decomposition-based methods identify anomalies by decomposing the adjacency matrix or tensor representation of the dynamic networks and by analyzing the eigenvectors, eigenvalues, or singular values appropriately.
- Tensors are generalizations of matrices. A dynamic network is represented as a third order tensor in which the first two dimensions denote an adjacency matrix and the third dimension denotes the dynamics of the network.

7. Anomaly detection in dynamic social networks

- (b) **Community-based approaches**: Even though the communities in a network contract, expand, merge, split, appear, vanish, or re-appear after a time period, majority of the nodes within a community follow similar evolution trends.

- (c) **Probability-based approaches**: Probability-based approaches generally build models of normal behavior of the network-based on probability theory, probabilistic distributions, and scan statistics (Heard et al., 2010; Priebe et al., 2005; Pandit et al., 2007).

- Several previous graph snapshots are used for building a model for normal behavior, and every new incoming network instance is compared with this model. Any deviations from this model are reported as anomalies.

7. Anomaly detection in dynamic social networks

- **7.2.1. Anomalous node detection**
- This section reviews the various anomaly detection methods that have been developed for identifying anomalous nodes in the dynamic unattributed networks.

- Sun et al. (2007) employed matrix decomposition for anomaly detection and proposed Compact Matrix Decomposition (CMD) to calculate sparse low rank matrix approximations.
- The low rank decompositions, such as Singular Value Decomposition (SVD) and CUR, reveal hidden variables and associated patterns from high dimensional data.
- CMD is computationally efficient and requires less space compared to SVD and CUR. In order to detect anomalies using CMD, the low-rank approximations of input networks are used to summarize the dynamic networks.

7. Anomaly detection in dynamic social networks

- The method proposed in Priebe et al. (2005) applies scan statistics on disjoint one-week windows of the ENRON who-e-mails-whom network to find network instances that have remarkably high interactions compared to the past.

- The number of edges in the neighborhood of each node is taken as the local statistics.

- The network instance is flagged as anomalous, if the scan statistics is above a threshold value, and the nodes and edges that contribute most to the change are deemed as anomalies.

7. Anomaly detection in dynamic social networks

- Another method that aims to find anomalous nodes based on probabilistic models is by Pandit et al .(2007), who proposed an approach called NetProbe to detect fraudsters in online auction networks.

- NetProbe represents auction users and their transactions as a Markov Random Field to identify the subgraphs of fraudsters in the network, and uses belief propagation to predict which users are likely to commit frauds in the future.

- It classifies the users of the auction network as fraudster, accomplice, or honest. The interaction between fraudsters and accomplices form bipartite cores, and fraudsters can be uncovered by discovering those cores.

7. Anomaly detection in dynamic social networks

- **7.2.2. Animalous edge detection**
- Anomalous edge detection methods aim to find patterns of interaction among individuals that deviate from the usual pattern of interaction in the network. The anomalous edge detection methods on dynamic unattributed networks are based on probability-based approach.

- Probability-based approach: Anomalous edges in a dynamic network can be discovered by applying link prediction techniques. Future interactions can be predicted through link prediction and the interactions that are very unlikely to occur are deemed as anomalous. Huang and Zeng (2006) developed an anomalous e-mail detection framework by applying a link prediction based formulation. The probability that an interaction occurs between two individuals is estimated using expectation maximization, and is used for assigning likelihood scores to each interaction afterwards.

7. Anomaly detection in dynamic social networks

- **7.2.2. Animalous subgraph detection**
- The methods that detect anomalous subgraphs in dynamic unattributed networks are based on community-based, distance-based, matrix/tensor decomposition-based, and probability-based 224 P.V. Bindu, P.S. Thilagam / Journal of Network and Computer Applications 68 (2016) 213–229 approaches. This section discusses the different anomalous subgraph detection methods that have been developed for dynamic unattributed networks.

- **Community-based approach**: When considering the dynamic nature of networks, new types of community-based anomalies such as formation of new communities, and splitting up and disappearance of existing communities are also possible.

7. Anomaly detection in dynamic social networks

- **Distance-based approach**: Mongiovi et al. (2013a) developed an iterative formulation called NetSpot for spotting and summarizing anomalous subgraphs called Significant Anomalous Regions (SAR) in weighted dynamic networks.

- **Matrix/Tensor decomposition-based approach**: The matrix/tensor decomposition-based methods for anomalous node detection discussed in Section 7.2.1 can detect anomalous subgraphs as well.

- **Probability-based approach**: Thompson and Eliassi Rad (2009) put forth a probability-based algorithm to find anomalous subgraphs from a time-evolving network. The recent behavior pattern of the network is modeled as a cumulative network that summarizes all past edges but gives more weightage to recent edges by using an exponential decay model.

7. Anomaly detection in dynamic social networks

- **7.3. Dynamic attributed networks**
- Types of approaches: The methods that have been proposed for detecting anomalies from dynamic attributed networks can be categorized into three: tensor decomposition-based, probability based, and signal processing-based approaches.

- (a) Tensor decomposition-based approaches: As discussed in Section 7.2.1, the decomposition-based anomaly detection methods use matrices or tensors to represent the dynamic networks, and exploit their decomposition to detect the anomalies. The method proposed by Papalexakis et al. (2012) uses tensor decomposition to detect anomalies in dynamic attributed networks.

7. Anomaly detection in dynamic social networks

- (b) Probability-based approaches: Based on probability theory, the approaches presented in Heard et al. (2010) build models of normal behavior of dynamic networks by analyzing several previous network instances. Every new incoming network instance is compared with this normal model, and any deviations are flagged as anomalous.

- (c) Signal processing-based approaches: As discussed in Section 6.1.3, signal processing on graphs treat comparatively smaller anomalous subgraphs as signals embedded in a much bigger network taken as the background noise.

7. Anomaly detection in dynamic social networks

- **7.3.1. Anomalous node, edge, and subgraph detection**
- Tensor decomposition-based approach: Papalexakis et al. (2012) presented a fast and parallelizable approach called ParCube, for speeding up sparse tensor decompositions by using random sampling techniques.

- Probability-based approach: The probability-based method by Heard et al. (2010), discussed in Section 7.2.1, can detect anomalous nodes, edges, and subgraphs in dynamic networks with categorical attributes as well.

- Signal processing-based approach: The proposed framework by Miller et al. (2013) exploits signal processing for graphs to find anomalous nodes and subgraphs from dynamic attributed networks.

7. Anomaly detection in dynamic social networks

Table 3
Summary and comparison of articles on anomaly detection in dynamic social networks.

Research paper	Network		Anomaly				Approach used	Reporting of anomalies
	Unattributed	Attributed	Node	Edge	Subgraph	Event		
Sun et al. (2007)	✓		✓		✓	✓	Matrix decomposition	Graph encoding cost for each time step
Akoglu and Faloutsos (2010)	✓		✓			✓	Matrix decomposition	*z*-scores
Yu et al. (2013)	✓		✓				Matrix decomposition	Node subsets
Sun et al. (2006)	✓		✓		✓	✓	Tensor decomposition	Reconstruction error
Kolda and Sun (2008)	✓		✓		✓	✓	Tensor decomposition	Reconstruction error
Koutra et al. (2012)	✓	✓	✓		✓	✓	Tensor decomposition	Tensor factors at each time step
Gupta et al. (2012b)	✓		✓				Community-based	Community memberships
Gupta et al. (2012a)	✓		✓				Community-based	Anomaly scores at time step
Ji et al. (2013)	✓		✓				Community-based	Anomaly scores at each time step
Rossi et al. (2013)	✓		✓				Community-based	Role memberships
Araujo et al. (2014)	✓	✓	✓				Community-based	Tensor factors at each time step
Priebe et al. (2005)	✓		✓	✓			Probability-based	Scan statistics
Huang and Zeng (2006)	✓			✓		✓	Probability-based	Likelihood scores
Pandit et al. (2007)	✓		✓				Probability-based	Node labels
Aggarwal et al. (2011)	✓			✓		✓	Probability-based	Likelihood at each time step
Heard et al. (2010)	✓	✓	✓	✓	✓	✓	Probability-based	*p*-values
Chen et al. (2012d)	✓				✓		Community-based	Community outliers and time stamps
Mongiovi et al. (2013a)	✓				✓		Distance-based	Anomaly scores
Mongiovi et al. (2013b)	✓				✓		Distance-based	Anomaly scores
Thompson and Eliassi-Rad (2009)	✓				✓		Probability-based	Anomaly scores
Papalexakis et al. (2012)		✓	✓			✓	Tensor decomposition	Tensor factors for each time step
Miller et al. (2012)	✓		✓			✓	Signal processing	Node subsets
Miller et al. (2013)		✓	✓		✓	✓	Signal processing	Node subsets

9. Conclusion

- In this paper, we have presented a structured review of the various methods proposed for anomaly detection in social networks represented as graphs.
- Mining social networks for anomalies is a challenging and computationally intensive task due to the huge size of the network and its dynamic nature.
- When selecting an appropriate algorithm, one has to consider different aspects of the application such as the type of the network being examined and the types of anomalies to be detected.
- This comprehensive review provides a better understanding of the various techniques that have been developed for mining social networks for anomalies.