An Exhaustive Review on State-of-the-art Techniques for Anomaly Detection on Attributed Networks

Wasim Khan[a], Mohammad Haroon[b]

[a]Integral University, Department of Computer Science and Engg., Lucknow/India (ORCID: 0000-0003-2311-1451)
[b]Integral University, Department of Computer Science and Engg., Lucknow/India (ORCID: 0000-0001-7967-7302)
[a]khanwasim051@gmail.com

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Social Network sites are one of the most prominent websites used in almost every facet of life. A social network reflects relationships among social entities, including acquaintances, co-workers, or co-authors. With the extreme success of Social networks, the misuse has also been escalated and unlocked the way for various illegal behavior and security threats. Social Network anomaly detection has become a critical topic to be explored by researchers. The nature of the input data is a significant aspect of an anomaly detection technique. In the field of anomaly detection in the social network, input networks can be classified as static or dynamic, Attributed or unattributed. The number of nodes and the connections between the nodes in static networks would not change with time. Attribute networks are pervasive in various domains and constitute a vital element of modern technological architecture, where node attributes support the topological structure in data exploration. While social networks build up over time, evaluating them as if they had been static is very beneficial. Numerous studies have been done for Anomaly detection, but to the best of our knowledge, research in static attributed anomaly detection has been very limited. This review tries to portray earlier research on detecting anomalies for social static attributed networks and thoroughly discusses state-of-the-art embedding approaches

Keywords: Anomaly Detection, Attributed Networks, Online Social Networks, Static Attributed Network.

1. Introduction

In the last decade, Social networking sites have generated growing interest proliferation due to rapid internet development. The social platforms encourage users to connect with others quickly irrespective of geographical positions and therefore got enormous coverage. It helps members remain in contact with family and friends, connect with older friends, and build a strong relationship with others based on mutual characteristics (Bindu & Thilagam, 2016) (Anand et al., 2017). A significant percentage of the population is currently accessing online social media such as Twitter, LinkedIn, Facebook, etc., and this has a considerable effect on almost all areas (Elghanuni et al., 2019). In comparison, its growing success and unrestricted use have resulted in the widespread abuse of social networks and becoming a primary target for attackers (Kaur & Singh, 2016). When a user's behavior fits standard trends, we define its use, as usual, else unusual (Hassanzadeh et al., 2012). Unexpected and irregular behavior in social media shows different practices from those in the same network. Various researchers have defined this terminology differently as an anomaly, an exception, or an outlier. Identifying anomalies or outliers is the central issue in data mining and has been described as an abnormal pattern detection task from the dataset. Although various anomaly detection methods are available for different problems, this sector is too young and fast expanding.

Two approaches are mainly used for social network anomaly detection: Behavior-based and Graph-based. Behavior-based methods deal with nodes' behavioral features such as the quantity and intensity of texts, the nature of exchanged information, the comments and likes on an article, and interaction span. The task of detecting anomalies on datasets that are depicted as graphs is Graph-based anomaly detection. It is possible to represent an online social network as a graph whereby the nodes represent individuals. The edges reflect the connections between nodes using various ties such as friendship, association, family, etc. (Hassanzadeh et al., 2012)

There are two social network classification forms based on input data type: Static vs. Dynamic and Attributed vs. Unattributed (Wasim et al., 2021). A static social network can be represented as just one snapshot of the data at any moment because the number of nodes and links among the nodes are not modified in static networks across time. Dynamic networks experience distinct transformation, e.g., node or edge addition or removal and node or edge attribute change. Users' characteristics or communication, like information about the communication type, users' date of birth, or communication period, are not used to identify anomalies in unattributed networks. In Attributed networks, features are associated with users or links among them and considered to identify anomalies, including the network structure [Figure 1].

Graphs not only act as organized data repositories but are indeed crucial in the advanced machine learning task. Graph-based identification of anomalies on social platforms is a new study area with a solid foundation in standard graph theory and communication patterns between users and groups. Many detailed survey studies have been carried out on Anomaly detection techniques and Graph-based anomaly detection techniques (Bindu &

Thilagam, 2016) (Anand et al., 2017) (Kaur & Singh, 2016). In this paper, we present a thorough and extensive review of the research work conducted in the field of Graph-based Anomaly detection, primarily focusing on the embedding-based methods for the attributed static social network.

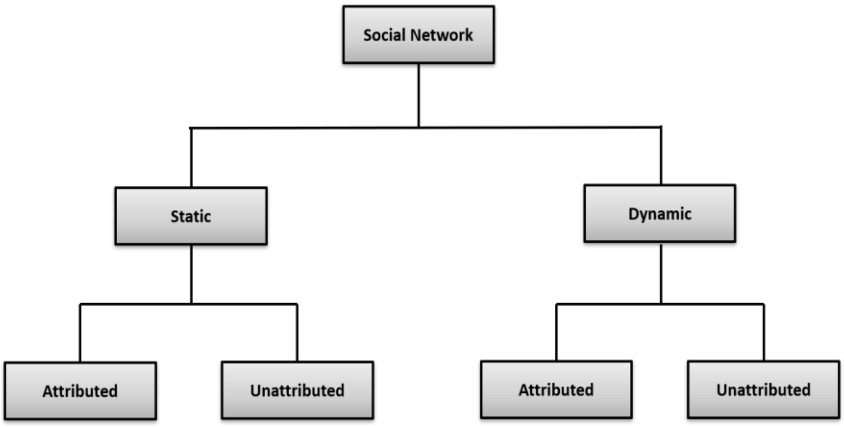

Figure 1. Social Network Anomaly Detection

1.1. Motivation

This review will be the first review of network embedding techniques for anomaly detection on attributed social networks to the best of our knowledge. Usually, nodes' abnormal behavior on attributed networks is not limited to the shared interactions with network structure only. Still, their node features (i.e., attributes) are also evaluated, and there is a growing intense trend in finding anomalies on these networks. Existing methods are hard to be implemented directly on attributed networks due to the heterogeneous nature and volume of real-world data. Current Anomaly detection techniques generally target either topological structure or nodal attributes alone (Liu et al., 2017). Also, Attribute networks typically include high-voluminous data instances and high-dimensional functionality. Raising concerns in identifying anomalies on attributed networks compelled us to propose this review work, analyzing different factors with state-of-the-art alternatives to offer users an overview of the various published studies in network embedding-based anomaly detection. It allows us to learn these strategies in solving future Anomaly detection problems and explore options in the days ahead.

1.2. Review contribution

So far, numerous review works have been published on various aspects of Social Network Anomaly detection. Our paper's contribution is as follows.

- Identifying the key aspects associated with the problem of anomaly detection in attributed static social networks.
- Providing a comprehensive review of state-of-the-art for anomaly detection in the static attributed social network.
- A couple of datasets and evaluation parameters have been explored for anomaly detection on the static attributed social network.
- Exploring the research challenges and open problems in the area of anomaly detection in static attributed social networks.

The rest of this paper is organized as follows. In Section 2, we provide the details of Network embedding. This section briefly describes the details of the embedding approach and its relation to anomaly detection. In Section 3, we identified the classification for different anomaly detection methods on static attributed networks. Section 4 discusses benchmarked datasets for anomaly detection on attributed networks. Evaluation parameters used for measuring the performance of anomaly detection approaches are presented in section 5. Section 6 discusses state-of-the-art approaches for anomaly detection in static attributed social networks. Finally, Section 7 discusses the challenges and the conclusion is presented in Section 8.

2. Network Embedding

The emergence of improved computational capability combined with both the widespread availability and high vitality of graph-structured datasets has led to a rise in network embedding research. Transformation of a graph to a low dimensional space in which its structural details and attributes are retained to the fullest extent is the main reason behind this pattern (Lerique et al., 2019). The objective is to identify unnoticed or complicated network properties either explicitly or through passing the learned representations to a downstream deduction network. Recent studies aimed to create embedded architectures that evaluate node representations through the methods of deep learning. For example, DeepWalk calculates metrics on node occurrences with a disjointed, random path by analyzing the input graph (Perozzi, Al-Rfou, et al., 2014). It endorses the SkipGram method to increase the likelihood of node neighborhood detection provided its embedding.

In recent years, several methods have been proposed that learn the representations which encode spatial graph features. These methods use the concept of learning the mapping of nodes embedding or whole graph embedding into the Low-dimensional vector points in embedding space Rd. This mapping is used effectively so that structural connections in this learned space represent the original graph structure. After that, the learned embeddings are utilized as input features for deep learning or machine learning tasks.

There are two major activities to learn graph representation. Initially, an encoder is used to every node in the graph is mapped to a node embedding that is a low-dimensional vector in latent space. The next step is to use a decoder to rebuild information from each node's neighborhood in the actual graph. The encoder is formally a process of mapping nodes $v \epsilon V$ to vector embeddings $Z_v \epsilon R^d$. As an input, the encoder uses node I.D.s to create the embedding of nodes.

$ENCD: V \rightarrow R^d$

In other words,

$ENCD(v) = Z[v]$

$Z \epsilon R^{|V| \times d}$ is embedding vectors matrix for all nodes, and Z[v] implies the row of Z that matches to node v. The decoder's role is to recreate graphical details from the node embedding produced by the encoder. For instance, provided the embedding z_u of a node u, the decoder would try to anticipate the neighbors set N(u) of node u [Figure 2].

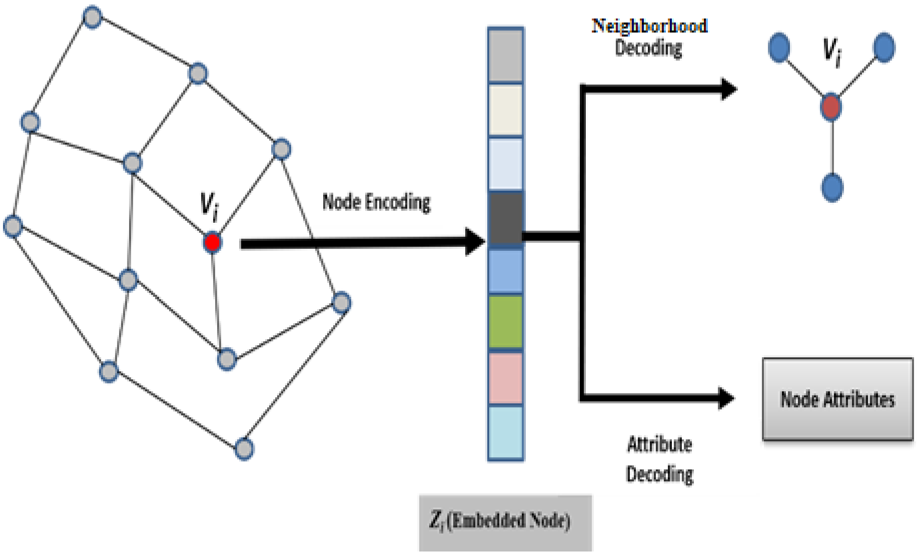

Figure 2. Node Embedding

2.1. Network embedding and anomaly detection

Due to network architecture's sparsity on the attributed networks and nonlinearity of node structure and attributes, node embedding methods based on deep learning for Anomaly detection are popular these days. Deep learning-based node embedding methods for Social network anomaly detection, generally, first of all, transforms the attributed network to the embedding representations as an encoding method. Then node attributes and topological structure are rebuilt with the use of respective decoding methods. After encoder and decoder processes using deep learning, nodes' reconstruction errors are then exploited to detect anomalous nodes on attributed networks.

DEFINITION: An attributed network G= (V, E, X) has (1) Node set V= {v1, v2, v3,…. vn} where |V|=n; (2) Edge set E, where |E|=m and (3) node attributes X ϵ $R^{n \times d}$ where the kth row vector x_k ϵ R^d (k=1, 2,…., n) is the kth node attribute information.

PROBLEM: Attributed network anomaly detection: Suppose numbers of nodes are N, G is an attributed network, A is the adjacency matrix, and X is a matrix for attribute information for n nodes. The objective is to evaluate all the nodes by their level of abnormality in such a way as to place the nodes in high-ranking positions, which vary spectacularly from the majority of reference nodes (Ding et al., 2019).

3. Methods Classification

Several methods for anomaly detection have been published recently in the literature (Pei et al., 2020). Many anomaly detection algorithms pursue an unsupervised manner because the costs for collecting training data become exorbitant. They can be divided into four categories: Community analysis, residual analysis, subspace selection, and deep learning methodologies [Figure 3].

Methods of community analysis identify anomalies by observing the existing node abnormality to other nodes within the same community (Perozzi & Akoglu, 2016) (Perozzi & Akoglu, 2018) (Gaot et al., 2010). Residual analysis approaches systematically simulate residual information using matrix factorization to recreate the attributed network(J. Li et al., 2017)(Peng et al., 2018). Subspace selection methods initially explore attributes subspace and then find abnormalities in the subspace that has already been learned (Perozzi, Akoglu, et al., 2014)(Sánchez et al., 2013)(Muller et al., 2013).

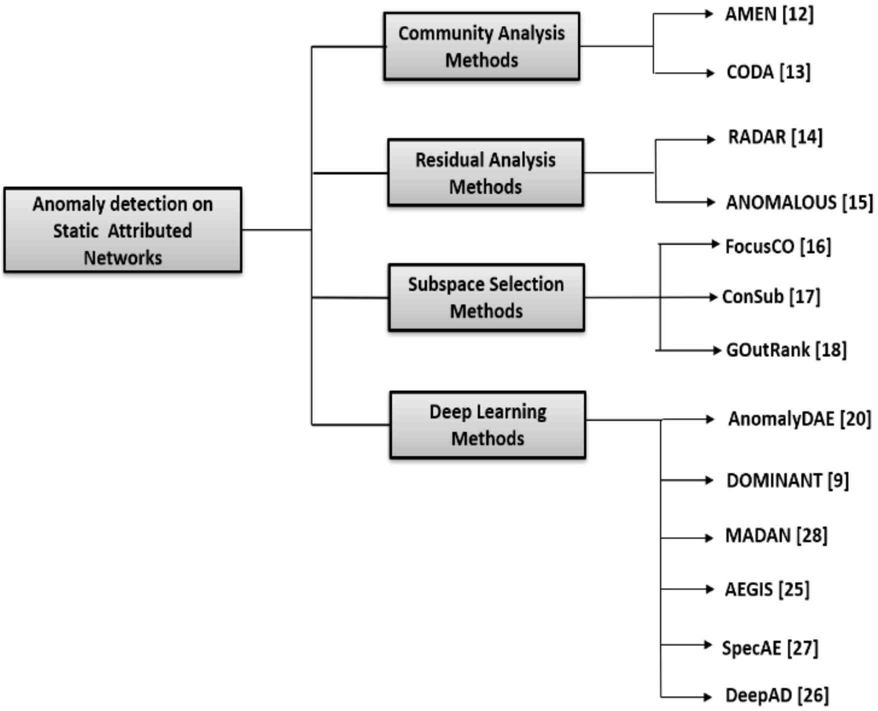

Figure 3. Methods Classification

While the algorithms mentioned above had their fair proportion of performance, these approaches are either subject to excessive overhead computing related to subspace selection and shallow learning processes or overlook the casual relationship between nodes and attributes through learning only node representations. But to catch attribute and structure anomalies, the interaction between these two separate modality inputs is essential for the anomaly detection task. Deep learning approaches are currently used to resolve the mentioned issues. These approaches employ deep neural models to explore network nonlinearity and unsupervised or supervised identification of anomalies. Deep learning embedding-based methods encapsulate the interactions between the node attribute and the network structure for decent quality embeddings.

4. Attributed Datasets

This section discusses benchmarked datasets for anomaly detection on attributed networks. The statistics of datasets are shown in Table 1.

- BlogCatalog: It is a forum for blog sharing where bloggers will follow one another to create a social network. Users are assigned several tags that portray them and their articles; these tags are known as node attributes (J. Li et al., 2015)(Fan et al., 2020) (Huang et al., 2018).
- Flickr: Flickr is a website that allows uploading photos. Similarly, BlogCatalog has allowed users to follow each other and form a social network. Users are allowed to follow each other, thus creating a social network. Their tagged interests specify node attributes on users (J. Li et al., 2015).
- ACM: This is also an attributed network from the academic sector. It is indeed a citation network, so each article is considered a node, and links are the citation ties between the various papers. Each article's attributes are centered on the article's abstract (Fan et al., 2020) (Huang et al., 2018).
- Disney and Books: These are datasets from co-purchase networks that are obtained from Amazon. These contain attributes describing online products' characteristics, such as reviews, cost, or even more (Muller et al., 2013).
- Enron: Enron is an email network with edges that represent the email transfer between individuals. Each node includes several attributes that describe the message's metadata, such as the length of the text, average number of receivers, etc. This dataset has been widely used as a spam detection benchmark (Metsis et al., 2006).
- PolBlog: It is an online blog citation network. The words that emerged in each blog using a predefined word dictionary are used as attributes of each node (Perozzi, Akoglu, et al., 2014).
- Twitter, Facebook, and Google+: Each comprises an ego network set that includes ground-truth social groups. Facebook's edges reflect undirected connection, whereas Google+ and Twitter indicate directed "follower" connection. Google+ and Facebook attributes include user profile data, including employer information, position, and university information. The Twitter attributes are taken from each user's usernames and hashtags (Perozzi & Akoglu, 2018).
- Cora, Citeseer, and Pubmed: These three datasets are citation networks where node represents publication and the edges represent the citations among them. Each publication includes several keywords and these keywords are considered as attributes (Liang et al., 2018)(Yang et al., 2016).

Table 1. Statistics of Attributed Datasets.

Dataset	**Nodes(V)**	**Edges(E)**	**Attributes(A)**
BlogCatalog	5,196	171,743	8,189
Flickr	7,575	239,738	12,047
ACM	16,484	71,980	8,337
Disney	124	334	28
Books	1,418	3,695	28
Enron	13,533	176,987	20
PolBlog	362	1,288	44,839
Twitter	81,306	1,768,149	1-2271
Facebook	4,039	88,234	42-576
Google+	107,614	13,673,453	1-4122
Cora	2,780	5,278	1,433
Citeseer	3,327	4,732	3,703
PubMed	19,717	44,338	500

5. Evaluation Parameters

This section highlights the various evaluation metrics to measure the performance of anomaly detection techniques on attributed networks.

- The ROC curve, a commonly used anomaly detection metric, is a plot of a true-positive result (Anomalous node is identified as an anomaly) against a false-positive result (Normal node is recognized as an anomaly), based on the facts and findings.
- The AUC value represents the area under the Receiver Operating Characteristic (ROC) curve, in which a randomly selected irregular node will be ranked higher than that of a normal node. As the AUC is higher, the more the model distinguishes between normal and abnormal nodes.
- While each anomaly detection technique produces an anomaly ranking list, Precision@K is used to calculate the percentage of true anomalies identified in its highest-ranking K nodes by a particular detection procedure.
- Recall@K measure evaluates the percentage of true anomalies found in the total number of ground truth anomalies by a unique detection technique.

6. Existing Key Methods

Several papers related to the embedding approach for Anomaly detection in static attributed networks have been published during the last decade. The objective of this section is to provide a detailed overview of some of the baseline approaches. Table 2 summarizes various state-of-the-art approaches.

Kaize Ding et al. introduced a novel framework, AEGIS, that can detect anomalies on recently identified nodes without extra training and attain better analytical results. The proposed framework comprises two stages of learning. Graph differentiation layers (GDN) are used to develop an autoencoder network at the first stage to learn the representations of attributed network nodes. The second step targets the generative adversarial system by using the established node representations of anomalies, reliably representing regular data distribution. The framework AEGIS exceeds all the critical approaches to all three real-world datasets. The experiment on datasets also indicates that although this technique is usually built for inductive identification of anomalies, it can also obtain better efficiency in the transductive environment (Ding et al., 2020).

Considering the dynamic interactions between the topological structure and node attributes, Haoyi Fan et al. proposed a deep combined representation learning method to detect anomalies through a dual autoencoder. Complex interactions between the structure of a network and the node's attribute are observed, and both structural and attribute observations assess anomalies. It learns the embedding for both nodes and attributes together in a spatial domain. The Attention mechanism is also used in the structure encoder to understand the significance between a node and its peers to effectively access structure patterns that are essential for detecting anomalies. Eventually, irregularities could be observed by evaluating node regeneration errors from the structure and attribute viewpoints both. Experiments were performed on real-world datasets. Blog-Catalog, Flickr, and ACM. Experiment results indicate that the proposed approach, AnomalyDAE vastly outweighs all benchmarks and works best on several datasets than existing state-of-the-art approaches (Fan et al., 2020).

A novel combined embedding method, DeepAD, was proposed by Dali Zhu et al. to exploit the high non-linear behavior in both attributes and network structure and find irregularities based on reconstruction errors. GCN imposes local neighborhood links inherently on every layer of encoding as routes to add neighborhood embeddings. First-order proximity can be considered the controlled information to limit the node pairs' proximity in latent representations. Motivated by the concept of Laplacian eigenmaps, a penalty function is added to limit the local vicinity when nodes close to one another are traced away from each other. The higher-order proximity relates to the correlation between the neighborhood details of two nodes. The higher order of neighborhood data is embedded with the convolutions iteration. Cross-model proximity is maintained by enhancing their estimated probability of communication (Zhu et al., 2020).

Kaize Ding et al. emphasized resolving the issues associated with existing approaches like the Non-linear nature of Data, Network sparsity, and non-ability of handling complicated relationships. They used a graph convolution network (GCN) to smoothly form the network structure and attributes for mapping node embedding, followed by a deep autoencoder utilizing this embedding to recreate the actual data. The cohesion between GCN and autoencoder helps to detect abnormalities by calculating node regeneration errors. Substantial experiments on datasets proved their proposed method's usefulness (Ding et al., 2019).

Yuening Li et al. in proposed a Spectral autoencoder-based anomaly detection framework for attributed networks to tackle the problem created due to the homophily hypothesis and over-smoothing. They tried to explore local anomalies (Community anomalies) and global anomalies both for the attributed networks. They used a Spectral autoencoder to embed topological relation and node attributes into an embedded space together. They selected an autoencoder to learn both embedding and reconstruction errors on all nodal attributes to identify global anomalies. They introduced a novel graph convolution encoder and decoder networks to learn the node's local representations depending on each node's neighborhood. After calculating the reconstruction errors, each node's suspect level is measured by evaluating the energy of its embedding representation in the Gaussian mixture model. Experimental tests on real-world datasets with five standard methods assessed the proposed model's efficiency and concluded that SpecAE retains exceptional results (Y. Li et al., 2019).

Leonardo Gutierrez-Gomez et al. suggested a novel MADAN approach (Multiscale Anomaly Detection in Attributed Networks) that offers a good rating and analysis process anomalous nodes at all levels in a network within their context. A highly concentrated node implies a node that is quite distinct in its features from its peers with the help of the proposed gaussian weighting function. Finally, a rating is given to classify potential anomalies using the concentration to measure the extent of divergence of a particular node at a specified point in time and relation to its context. Experiments performed on synthetic and real-world datasets demonstrate that the proposed approach helps identify ground truth anomalies and recognize new anomalies, including their contexts (Gutiérrez-Gómez et al., 2019).

Zhen Peng et al. developed ANOMALOUS, a joint modeling method to detect anomalies on attributed networks by streamlining the anomaly detection and attribute selection. Instead of considering both as two distinct steps, utilize them as a whole. It involves extracting specific instances from the representation features deeply connected to the CUR decomposition-based topological structure. Finally, it tests each instance's normality using the residual analysis approach. The success of the proposed method is demonstrated by observations on both synthetic and real-world data sets (Peng et al., 2018).

Residual analysis-based Anomaly detection was introduced by Jundong Liy et al. Residual analysis is used to differentiate the original data from the predicted data by calculating the residuals and identifying anomalies. Instances containing high residual error are much prone to being anomalies, as their patterns would not match with many other references. This algorithm gives a useful way to identify anomalies globally and relies not on peculiar characteristics of anomalies. The AUC values yielded with this method prove that the RADAR is markedly good than other baseline approaches (J. Li et al., 2017).

Most current approaches cannot be tailored to users' specific needs since they do not regard the user preferences eccentricity. Therefore, Peng et al. proposed a novel framework, ALARM, that integrates user interests in identifying anomalies and addressing heterogeneous attribute features simultaneously using multiple graph encoders and a well-designed, self-learning, and user-friendly aggregator. Synthetic dataset experiments reveal that although the dataset includes no natural multi-view attributes, the proposed method still delivers the desired outcomes by dividing the attribute set arbitrarily (Peng et al., 2020).

FocusCO is an approach of seeking focused communities and anomalies in massive attributed datasets. The proposed approach first cuts off clusters of identical, strongly connected nodes and shows cohesiveness in focus attributes, which subsets. Then focused anomalous nodes are detected adhering to a topological structure-based focused cluster but indicate absolute deviations from focus attributes. An effective algorithm is offered, incorporating the user's focus attributes and simultaneously extracting focused clusters and anomalies (Perozzi, Akoglu, et al., 2014).

For graphs containing many numeric attributes, Sanchez et al. proposed ConOut, a new local outlier ranking model that specifies each object's subgraph and its statistically significant attribute subset locally. This filtering of the context makes a high distinction between an outlier and standard objects. They measure the outliers rating by integrating both the topological structure and attribute value variation from this context. The context selection enables local anomalies detection that is not possible provided the global partition or the whole graph. Unlike subspace selection approaches, the running time of ConOut is much smaller (Sánchez et al., 2014).

GoutRank is a method for scoring the nodes according to their level of variance in both topological structure and attributes. The approach aims to detect complicated outliers that diverge concerning a strongly connected subgraph. While the specific outlier in the subgraph is practically identical to other nodes, but It strongly differs from a subset of relevant attributes called subspaces. This approach uses the technique of subspace analysis and graph clustering for preprocessing to produce the ranking of outliers. In attributed networks, Outlier scoring functions can find better quality outliers (Muller et al., 2013).

Gao et al. pioneered community anomalies and created a standardized system for anomaly detection and community exploration. They presented a community outlier detection algorithm (CODA), a deterministic formulation that incorporates both community and anomaly detection based on Hidden Markov Random Fields (HMRF). The topological data is encoded employing the HMRF model as structural restrictions on the hidden states. In contrast, the node data is configured as a combination of multinomial or Gaussian distributions. Extensive experiments indicate that the proposed CODA approach continually exceeds the reference algorithms on synthetic data. It also detects significant community anomalies from the DBLP dataset (Gaot et al., 2010).

Table 2. Summary of the State-of-the-art approaches.

Author/Study/Ref	Method	Working	Features
"Inductive Anomaly Detection on Attributed Networks," Kaize Ding, Jundong Li, Nitin Agarwal, and Huan Liu, 2020, (Ding et al., 2020)	Deep Learning Method (Graph Neural network with generative adversarial Learning employed).	• First of all, an autoencoder is used to learn the anomaly aware representations of the node. After that, it trains a generative adversarial network to facilitate the generalizability of recently added data. • The generator seeks to produce probable descriptive anomalies, whereas the discriminator attempts to discover a boundary of decisions that differentiates possible anomalies from standard data. The generator helps enhance the discriminator's ability to distinguish standard data by collecting probable descriptive anomalies.	• This approach can be used in both inductive and transductive environments to implement anomaly detection. • The proposed method obtained an outstanding Anomaly Detection efficiency under inductive conditions above other standard techniques and revealed its potential to perceive anomalies on recently introduced data except retraining.
"ANOMALYDAE: Dual Autoencoder for Anomaly Detection on Attributed Networks", Haoyi Fan, Fengbin Zhang, Zuoyong Li, 2020, (Fan et al., 2020)	Deep Learning Method (Dual Autoencoder with attention mechanism employed).	• It learns the embedding for both nodes and attributes together in a spatial domain. • The Attention mechanism is also used in the structure encoder to understand the significance between a node and its peers to access structure patterns that are essential for detecting anomalies effectively. • Eventually, irregularities could be observed by evaluating node regeneration errors from the structure and attribute viewpoints both.	• The results indicate that the proposed approach outperforms robust anomaly detection benchmarks with an enormous-margin because encoders based on Graph convolution networks can acquire different node embeddings by collating neighbor characteristics but could not accumulate complicated topological structures and node attributes connections together.

"DeepAD: A Joint Embedding Approach for Anomaly Detection on Attributed Networks", Dali Zhu, Yuchen Mal, and Yinlong Liu., 2020, (Zhu et al., 2020)	Deep Learning Method (Stacking GCN Layers).	• Graph Convolution Network (GCN) is used to embed the attributes and topological structure. It acquires the attributed input network's representations through the loss function minimization until the objective function's convergence. • Stochastic Gradient descent is used to tailor the parameter's value, and node abnormalities are ranked then after a few iterations.	• In the original networks, the proposed method collectively conserves first order, high order, and cross-modal similarity to achieve a coherent comparative representation.
"Deep Anomaly Detection on Attributed Networks", Kaize Ding, Jundong, Rohit Bhanushali, Huan Liu, 2019, (Ding et al., 2019)	Deep Learning Method (Graph Convolution Network followed by a Deep Autoencoder).	• GCN simplifies the Network Sparsity problem by detecting the data nonlinearity and dynamic relationship on attributed networks. • A deep autoencoder framework is presented for Anomaly Detection by recreating the actual network with embedded nodes.	• They fixed the drawbacks of current approaches by managing node relationships with several non-linear layers of transformation. • Experiments evaluated the proposed model against popular Anomaly detection methods like LOF, SCAN, AMEN, RADAR, and ANOMALOUS. Results clearly showed the proposed model's supremacy over other methods.
"SpecAE: Spectral AutoEncoder for Anomaly Detection in Attributed Networks", Yuening Li, Xiao Huang, Jundong Li, Mengnan Du, and Na Zou, 2019, (Y. Li et al., 2019)	Deep Learning Method (Graph Convolution and Deconvolution followed by Laplacian Sharpening and density estimation).	• Attributed Networks are mapped to the tailored space by using the spectral convolution autoencoder and deconvolution decoder to identify the local and global anomalies. • Then Laplacian sharpening is employed to intensify the deviations between anomaly representations. • Finally, the density evaluation model is coupled with the observed representations and reconstruction errors to find the anomalies.	• The proposed approach addresses constraints of the homophily hypothesis and over-smoothing of node representation issues.

"Multi-scale Anomaly Detection on Attributed Networks," Leonardo Gutierrez-Gomez, Alexandre Bovet, Jean-Charles Delvenne, 2019, (Gutiérrez-Gómez et al., 2019)	Deep Learning Method (Gaussian weighting Function followed by the heat kernel as filtering operator with the Markov stability).	• The widely dispersed signals of impulses identify anomalous nodes through each vertex of the graph after smoothing. Heat kernel used for filtration helps us leverage the Markov stability connection to locate the network's outlier node background.	• The proposed approach is reliable and scalable for complex systems due to the exponential Chebychev measurements of the Laplacian graph. • It also offers a smoother community identification technique using a continuous-time Markov stability system as an additional advantage.
"ANOMALOUS: A Joint Modeling Approach for Anomaly Detection on Attributed Networks", Zhen Peng, Minnan Luo, Jundong Li, Huan Liu, Qinghua Zheng, 2018, (Peng et al., 2018)	Residual Analysis Method (CUR Decomposition and Residual analysis as a whole).	• The proposed method ensures that contrary to choosing attributes or instances alone, the compilation from both instance and attribute can enhance efficiency because they select indicative instances and attributes jointly to identify anomalies that can ease the adverse impact of irrelevant or redundant attributes.	• After conducting tests on various real-world datasets, it is finally concluded that the suggested methodology is quite resilient for growing dimensions and an increasing number of insignificant attributes. • The primary reason seems to be that it will flush away noisy data and concurrently choose instances and representative attributes to detect anomalies.
"Radar: Residual Analysis for Anomaly Detection in Attributed Networks," Jundong Liy, Harsh Daniy, Xia Huz, Huan Liuy, 2017, (J. Li et al., 2017)	Residual Analysis	• Networks with interconnected instances, associations among instances often make the residual process model complex. The integration of attributed information modeling and network modeling must be combined to achieve consistency between them.	• They have primarily evaluated that (a) how to classify the residual details of attributes for finding anomalies while no previous information regarding anomalies is available; (b) How to use cohesiveness between topological information and calculated residuals to generally classifying anomalies.
"Accelerated Local Anomaly Detection via Resolving Attributed Networks", Ninghao Liu, Xiao Huang and Xia Hu, 2017, (Liu et al., 2017)	Hybrid Method (Joint Embedding of Attributed data and Network structure followed by the novel parallel algorithm for accelerated optimization).	• Actual features are first translated into a low dimensional vector. • Nonnegative matrix factorization (NMF) based clustering is then applied to accomplish the node clustering and evaluate each cluster's data distribution. • After that, Cosine Similarity is used to compute each node's	• They investigated the problem of detecting local anomalies effectively and efficiently in attributed networks. They created SGD based small-batch process to speed up the optimization process and manage massive datasets in real-life scenarios.

		normality score. A ranked list is created with the ascending order of the nodes concerning their ratings. The higher-ranking nodes are known as anomalies.	
"Scalable Anomaly Ranking of Attributed Neighborhoods," Bryan Perozzi and Leman Akoglu, 2016, (Perozzi & Akoglu, 2016)	Community Analysis Method (Proposed metric Normality Calculation followed by an optimization approach).	• After calculating the normality metric that analyses and measures the neighborhoods' quality, an optimization method is used to derive weight vectors of attributes in a specific neighborhood dynamically, aiming to define a subspace that maximizes the normality of the neighborhood. • Their normality levels then sort Neighborhoods, while those with the least numbers are considered abnormal.	• Proposed metric normality is peculiar to its notion of "acquittal" of the boundary's edges, that hardly any of the current metrics demonstrated. Instead, they overlooked the boundary. • Normality uses topological structure and node attributes simultaneously, and its definition extrapolates several node attribute graphs. • Experimental studies on real-world datasets illustrate this metric's effectiveness for the neighborhoods' quality rating, surpassing existing metrics and methods. It also works as a useful tool to analyze the structure-attribute association.
"A Deep Multi-View Framework for Anomaly Detection on Attributed Networks", Zhen Peng, Minnan Luo, Jundong Li, Luguo Xue, and Qinghua Zheng, 2015, (Peng et al., 2020)	Deep Learning Method (Multiple graph encoders and a well-designed aggregator).	• Initially, provided attributed network is encoded by using several GNNs to address multi-view attributes. • After that, the hidden embeddings obtained from these multi-view attributes are filtered with an adequately designed aggregator to create coherent representations and then decoded from the aspects of attributes and structure both. • After encoding and decoding processes, measured reconstruction errors are finally used for anomaly detection.	• This approach revisited the issue of multi-view anomaly detection and dealt effectively with the separation and affiliation of many views. • The proposed ALARM model's quality deterioration is comparatively low than the other baseline methods due to an aggregator's deployment. Aggregator combines multiple encoders output with learning parameters that provide extra scope for adjustment to fit into the deep structure.
"Focused Clustering and Outlier Detection in Large Attributed Graphs", Bryan Perozzi,	Subspace Selection Method (Identifies the focus by using	• First, they defined the node attributes' significance weights that render the exemplary nodes	• The proposed method's main feature is to imply the users' interest by extracting focus attributes with the collection of

Leman Akoglu, Patricia Iglesias Sánchez and Emmanuel Müller, 2014, (Perozzi, Akoglu, et al., 2014)	inverse Mahalanobis distance, extracts focused clusters, and detects outliers with conductance measure).	identical. • After assessing attribute weights, focused clusters are extracted. • Anomalous nodes are also detected from each cluster simultaneously with identifying the clusters in a unified manner.	exemplar nodes selected by the user. • This approach correlates well with the graph's size since user-interest-based local clustering is done instead of a global graph clustering.
"Local Context Selection for Outlier Ranking in Graphs with Multiple Numeric Node Attributes", Patricia Iglesias Sánchez, Emmanuel Müller, Oretta Irmler, Klemens Böhm, 2014, (Sánchez et al., 2014)	Community Analysis Method	• First of all, the Proposed method calculates the context of the graph and attributes locally for each database object. • Then, it chooses a set of appropriate attributes for each measured context. A statistical analysis assesses attribute significance that compares each attribute's local and global deviation. • Finally, the outliers are detected based on the high distinction between the outliers and their local context.	• The proposed approach may find contextual anomalies that other anomaly detection models neglect. • Overall, this approach measures the anomaly rating of excellent quality correlated well with the number of attributes. • It matches the value with performance when linking graph structure to the attribute data. •
"Ranking Outlier Nodes in Subspaces of Attributed Graphs", Emmanuel Muller, Patricia Iglesias Sanchez , Yvonne Mulle, Klemens Bohm, 2013, (Muller et al., 2013)	Subspace selection method (Graph clustering techniques and subspace analysis).	• GOutRank first select subspace clusters with the help of techniques used in previous works (Moser et al., 2009) (Günnemann et al., 2010). • Then objects are scored in these various subspace clusters using scoring functions. • Finally, Outliers are detected after examining their participation in either large, high dimensional, or densely connected subgraphs.	• The proposed approach deals with two issues: Difficulty with selecting subgraphs with corresponding subspaces and the node scoring problem in multiple such subspaces. • GOutRank utilizes the inherent strength of graph clustering and also its connectedness to the subspace analysis. • This approach has been developed for complicated outliers, which differ significantly from a subclass of relevant attributes, mostly concerning a localized subgraph. So, it addresses the problem of finding hidden anomalies also.

"On Community Outliers and their Efficient Detection in Information Networks", Jing Gao, Feng Liang, Wei Fan, Chi Wang, Yizhou Sun, and Jiawei Han, 2010, (Gaot et al., 2010)	Community Analysis Method	• They have presumed that Normal objects are chosen randomly by K Communities and outliers. A combination of multinomial or Gaussian distributions is used to model each object's attributes, whereas relations are used to estimate previous distributions on hidden labels. • After that, ICM (Iteration Conditional Modes) and E.M. (Expectation-Maximization) based robust algorithm is proposed to train parameter values and deduce the community anomaly detection model's hidden categories.	• CODA is the very first research in which community anomalies are detected by concurrently examining data and links. • Although the algorithm's computing cost is constant throughout the number of nodes, a proper initial setup of the clusters is vital to the proposed method's success.

7. Challenges for Anomaly Detection on Attributed Networks

Some of the challenges that are found after the review of state-of-the-art approaches are as follows:

- Attributed networks are generally non-linear and sparse, and it is a primary challenge for anomaly detection on these networks due to network sparsity and data nonlinearity (Ding et al., 2019).
- When using subspace selection methods, a challenging issue is choosing congruent subspaces to model the relationship between attribute values and topological structure (Sánchez et al., 2013).
- Standard anomaly detection approaches rely on spatial information or graph models, but many of today's databases contain multi-dimensional quantitative data and entity relationships in attributed graphs. The main challenge is how to work with these various types of data in a unified way (Gutiérrez-Gómez et al., 2019).
- Anomalous nodes are often context-dependent, so identifying anomalies is naturally related to community detection. Still, there is a severe challenge to perform multiscale community detection where edge weights attach the associated nodes (Gutiérrez-Gómez et al., 2019).
- The attributed networks in the real world often include high-dimensional features and massive data instances. With the substantial data scales and heterogeneous attributed networks, current anomaly detection approaches cannot be implemented directly (Liu et al., 2017).
- Despite the methodological success of the homophily hypothesis, specific unpopulated attributes may exist that do not fulfill the theory of homophily, and the presence of these Structurally insignificant attributes may cause harmful effects on the detection of anomalies (Peng et al., 2018).
- Until now, in attributed networks, limited attention is being paid to inductive anomaly detection (Ding et al., 2020).
- In attributed networks, various types of anomalies frequently blend, all of which are impossible to classify if we do not have previous data awareness (J. Li et al., 2017).
- The smoothing methods generally applied to attributed networks do not fit with the anomaly detection process. The node representations could be over smoothed and render abnormal nodes less distinctive than the remaining nodes within the same community (Y. Li et al., 2019).

8. Conclusion

The large size of the attributed networks and their complex nature render anomaly detection in the static attributed network an arduous and computationally intensive process. It has recently been increasingly crucial for anomalous nodes to be detected within attributed networks, with widespread deployments in different high-impact fields, such as social networks, cybersecurity, and healthcare. The review paper assembles and arranges the state-of-the-art techniques into various categories and describes the complementary approaches for detecting anomalies in Static social attributed networks. Since it was tough to cover every method in this paper, the best attempts have been devoted to cover most. We also explored the different challenges for future studies in this area. Despite considerable efforts in anomaly detection, several shortcomings exist that could be resolved and addressed in the future. This more in-depth review offers a better understanding of the critical approaches designed for anomaly detection in Static social attributed networks.

Acknowledgement

This work is acknowledged under Integral University manuscript No IU/R&D/2021-MCN0001123

References

1. Anand, K., Kumar, J., & Anand, K. (2017). Anomaly detection in online social network: A survey. Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2017, Icicct, 456–459. https://doi.org/10.1109/ICICCT.2017.7975239
2. Bindu, P. V., & Thilagam, P. S. (2016). Mining social networks for anomalies: Methods and challenges. Journal of Network and Computer Applications, 68, 213–229. https://doi.org/10.1016/j.jnca.2016.02.021
3. Ding, K., Li, J., Agarwal, N., & Liu, H. (2020). Inductive anomaly detection on attributed networks. IJCAI International Joint Conference on Artificial Intelligence, 2021-Janua(1), 1288–1294. https://doi.org/10.24963/ijcai.2020/179
4. Ding, K., Li, J., Bhanushali, R., & Liu, H. (2019). Deep anomaly detection on attributed networks. SIAM International Conference on Data Mining, SDM 2019, 2, 594–602. https://doi.org/10.1137/1.9781611975673.67
5. Elghanuni, R. H., Ali, M. A. M., & Swidan, M. B. (2019). An Overview of Anomaly Detection for Online Social Network. ICSGRC 2019 - 2019 IEEE 10th Control and System Graduate Research Colloquium, Proceeding, August, 172–177. https://doi.org/10.1109/ICSGRC.2019.8837066
6. Fan, H., Zhang, F., & Li, Z. (2020, February 10). AnomalyDAE: Dual autoencoder for anomaly detection on attributed networks. http://arxiv.org/abs/2002.03665
7. Gaot, J., Liangt, F., Fan, W., Wangt, C., Sunt, Y., & Hant, J. (2010). On community outliers and their efficient detection in information networks. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 813–822. https://doi.org/10.1145/1835804.1835907
8. Günnemann, S., Färber, I., Boden, B., & Seidl, T. (2010). Subspace clustering meets dense subgraph mining: A synthesis of two paradigms. Proceedings - IEEE International Conference on Data Mining, ICDM, 845–850. https://doi.org/10.1109/ICDM.2010.95
9. Gutiérrez-Gómez, L., Bovet, A., & Delvenne, J. C. (2019). Multi-scale anomaly detection on attributed networks. ArXiv. https://doi.org/10.1609/aaai.v34i01.5409
10. Haroon, M. (n.d.). " A Systematic Study on Aspects and Methods to Detect Anomalies in Online Social Networks ." 50(2), 1–8.
11. Hassanzadeh, R., Nayak, R., & Stebila, D. (2012). Analyzing the effectiveness of graph metrics for anomaly detection in online social networks. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7651 LNCS, 624–630. https://doi.org/10.1007/978-3-642-35063-4_45
12. Huang, X., Song, Q., Li, J., & Hu, X. (2018). Exploring expert cognition for atributed network embedding. WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining, 2018-Febua, 270–278. https://doi.org/10.1145/3159652.3159655
13. Kaur, R., & Singh, S. (2016). A survey of data mining and social network analysis based anomaly detection techniques. Egyptian Informatics Journal, 17(2), 199–216. https://doi.org/10.1016/j.eij.2015.11.004

14. Lerique, S., Abitbo, J. L., & Karsai, M. (2019). Joint embedding of structure and features via graph convolutional networks. ArXiv.
15. Li, J., Dani, H., Hu, X., & Liu, H. (2017). Radar: Residual analysis for anomaly detection in attributed networks. IJCAI International Joint Conference on Artificial Intelligence, 0, 2152–2158. https://doi.org/10.24963/ijcai.2017/299
16. Li, J., Hu, X., Tang, J., & Liu, H. (2015). Unsupervised streaming feature selection in social media. International Conference on Information and Knowledge Management, Proceedings, 19-23-Oct-, 1041–1050. https://doi.org/10.1145/2806416.2806501
17. Li, Y., Huang, X., Li, J., Du, M., & Zou, N. (2019). SpecAE. 2233–2236. https://doi.org/10.1145/3357384.3358074
18. Liang, J., Jacobs, P., Sun, J., & Parthasarathy, S. (2018). Semi-supervised embedding in attributed networks with outliers. SIAM International Conference on Data Mining, SDM 2018, 153–161. https://doi.org/10.1137/1.9781611975321.18
19. Liu, N., Huang, X., & Hu, X. (2017). Accelerated local anomaly detection via resolving attributed networks. IJCAI International Joint Conference on Artificial Intelligence, 0, 2337–2343. https://doi.org/10.24963/ijcai.2017/325
20. Metsis, V., Androutsopoulos, I., & Paliouras, G. (2006). Spam filtering with Naive Bayes - Which Naive Bayes? 3rd Conference on Email and Anti-Spam - Proceedings, CEAS 2006.
21. Moser, F., Colak, R., Rafiey, A., & Ester, M. (2009). Mining cohesive patterns from graphs with feature vectors. Society for Industrial and Applied Mathematics - 9th SIAM International Conference on Data Mining 2009, Proceedings in Applied Mathematics, 2, 589–600. https://doi.org/10.1137/1.9781611972795.51
22. Muller, E., Sanchez, P. I., Mulle, Y., & Bohm, K. (2013). Ranking outlier nodes in subspaces of attributed graphs. Proceedings - International Conference on Data Engineering, 216–222. https://doi.org/10.1109/ICDEW.2013.6547453
23. Pei, Y., Huang, T., van Ipenburg, W., & Pechenizkiy, M. (2020). ResGCN: Attention-based deep residual modeling for anomaly detection on attributed networks. ArXiv, 1–22.
24. Peng, Z., Luo, M., Li, J., Liu, H., & Zheng, Q. (2018). Anomalous: A joint modeling approach for anomaly detection on attributed networks. IJCAI International Joint Conference on Artificial Intelligence, 2018-July, 3513–3519. https://doi.org/10.24963/ijcai.2018/488
25. Peng, Z., Luo, M., Li, J., Xue, L., & Zheng, Q. (2020). A Deep Multi-View Framework for Anomaly Detection on Attributed Networks. IEEE Transactions on Knowledge and Data Engineering, 14(8), 1–1. https://doi.org/10.1109/tkde.2020.3015098
26. Perozzi, B., & Akoglu, L. (2016). Scalable anomaly ranking of attributed neighborhoods. 16th SIAM International Conference on Data Mining 2016, SDM 2016, 207–215. https://doi.org/10.1137/1.9781611974348.24
27. Perozzi, B., & Akoglu, L. (2018). Discovering Communities and Anomalies in Attributed Graphs: Interactive Visual Exploration and Summarization. ACM Transactions on Knowledge Discovery from Data, 12(2), 1–40. https://doi.org/10.1145/3139241
28. Perozzi, B., Akoglu, L., Iglesias Sánchez, P., & Müller, E. (2014). Focused clustering and outlier detection in large attributed graphs. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1346–1355. https://doi.org/10.1145/2623330.2623682
29. Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). DeepWalk: Online learning of social representations. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 701–710. https://doi.org/10.1145/2623330.2623732
30. Sánchez, P. I., Müller, E., Irmler, O., & Böhm, K. (2014). Local context selection for outlier ranking in graphs with multiple numeric node attributes. ACM International Conference Proceeding Series. https://doi.org/10.1145/2618243.2618266
31. Sánchez, P. I., Müller, E., Laforet, F., Keller, F., & Böhm, K. (2013). Statistical selection of congruent subspaces for mining attributed graphs. Proceedings - IEEE International Conference on Data Mining, ICDM, 647–656. https://doi.org/10.1109/ICDM.2013.88

32. Yang, Z., Cohen, W. W., & Salakhutdinov, R. (2016). Revisiting semi-supervised learning with graph embeddings. 33rd International Conference on Machine Learning, ICML 2016, 1, 86–94.
33. Zhu, D., Ma, Y., & Liu, Y. (2020). DeepAD: A joint embedding approach for anomaly detection on attributed networks. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 12138 LNCS. Springer International Publishing. https://doi.org/10.1007/978-3-030-50417-5_22