Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 7, No. 1, March 2019, pp. 37~50
ISSN: 2089-3272, DOI: 10.11591/ijeei.v7i1.773

Network anomaly detection research: a survey

Kurniabudi[1], Benni Purnama[2], Sharipuddin[3], Darmawijoyo[4], Deris Stiawan[5], Samsuryadi[6], Ahmad Heryanto[7], Rahmat Budiarto[8]
[1,2,3]STIKOM Dinamika Bangsa, Indonesia
[4,5,6,7]Faculty of Computer Science, Universitas Sriwijaya, Indonesia
[8]Albaha University, Saudi Arabia

Article Info

Article historys:

Received Oct 19, 2018
Revised Jan 09, 2019
Accepted Jan 21, 2019

Keywords:

Anomaly identity
Data type
Intrusion detection system
Network anomaly detection

ABSTRACT

Data analysis to identifying attacks/anomalies is a crucial task in anomaly detection and network anomaly detection itself is an important issue in network security. Researchers have developed methods and algorithms for the improvement of the anomaly detection system. At the same time, survey papers on anomaly detection researches are available. Nevertheless, this paper attempts to analyze futher and to provide alternative taxonomy on anomaly detection researches focusing on methods, types of anomalies, data repositories, outlier identity and the most used data type. In addition, this paper summarizes information on application network categories of the existing studies.

Corresponding Author:

Deris Stiawan,
Faculty of Computer Science,
Universitas Sriwijaya,
Palembang, Indonesia.
Email: deris@unsri.ac.id

1. INTRODUCTION

Anomaly (also known as an outlier) detection is an important issue in information security as defined in [1] and [2]. Anomaly and misuse detection are alternative approaches used to recognize intrusions [3] and as part of the Intrusion Detection System (IDS). IDS consists of three major groups: Signature-based Detection (SD), Anomaly-based Detection (AD) and Stateful Protocol Analysis (SPA) [4]. It is important for administrators to recognize anomalies on the network that can help in managing and troubleshooting security issues [5].

Researches on network anomaly detection have been done for quite long time. Up to now, anomaly detection research is still widely progressing. As mentioned in [6], besides it is an important research area it also has dynamic issues. The research topics of anomaly detection are too diverse. Starting from discussing and proposing models [7], [8] to frameworks [9], [10], and to research concerning methods [11], [12]. Moreover, the evaluation techniques and evaluation approaches of anomaly detection become more important, because they affect the accuracy of the identification. The evaluation and validation approaches used by researchers vary. Researchers use experiments [13], [14], simulations [15], or both approaches [16].

This paper attempts to analyze and to provide alternative taxonomy on anomaly detection researches focusing on methods, types of anomalies, data repositories, outlier identity and the most used data type. In addition, this paper summarizes information on application network categories of the existing studies.

This article is structured as follows. Section II provides information on preliminary studies and relevant researches, Section III describes the research methodology, section IV discusses the observation results of the survey study and Section V concludes the survey study results and provides future research plan.

2. PRELIMINARY STUDIES AND RELEVANT RESEARCHES

Anomaly detection can be interpreted as a detector of unexpected events, patterns, and behaviors, and deviates from the normal concept [17]. Anomaly detection method firstly defines the normal system behavior profile then any deviation from the profile will be marked as an anomaly [18]. Multiple devices connected to the network introduce new challenges in anomaly detection. Researchers have developed various methods, frameworks, techniques and algorithms in order to produce automatic and reliable anomaly detection.

Researches on anomaly detection have been spread and carried out in different aspects. Especially in networking, this concept appears along with IDS research. As mentioned by Abduvaliyev *et al.* [19] anomaly detection is one of three main techniques that can be used in IDS. This technique identifies wether the network traffic considered as normal or abnormal. By implementing this concept the IDS is expected to be able to detect new or unknown anomalies/attacks. Oreilly *et al.* [20] study on detecting anomalies in a non-stationary environment of wireless sensor networks. Authors in [6] discuss a variety of anomalous detections based on methods, systems and tools. While Al-Musawi *et al.* [21] present a grouping of important anomaly detection techniques for identifying traffic anomalies. The first two articles focus more on wireless sensor networks. The following article only focuses on methods, systems and tools. The last article focuses on anomaly detection on Border Gateway Protocol (BGP).

Extensive survey studies have also been carried out, however the studies are too diverse. Each researcher uses a different approach and focus topics. For example, Zhang *et a*l. [1] evaluate and compare the existing outlier detection techniques specifically developed for wireless sensor networks (WSNs). Gogoi *et al.* [22] provide a comprehensive up-to-date survey on outlier detection methods. While Marnerides & Mauthe [23] disscus the dimensions of theoretical methodologies and traffic features. Bhuyan *et al.* [6] present a structured and comprehensive survey on anomaly-based network intrusion detection and Weller-Fahi *et al.* [24] present a taxonomy of network anomaly detection. Patcha *et a*l. [25] and Garcia-Teodore *et a*l. [26] present existing solutions and latest technological trends of network anomaly detection. Table 1 summarizes the discussion topics covered by this paper and other existing survey studies.

Table 1. Comparing our survey with existing survey

Discussion Topics	Researchers							This paper
	Chandola et al., 2009 [7]	Zhang et al., 2010 [1]	Gogoi et al.,2011 [22]	Marnerides & Mauthe, 2014, [23]	Bhuyan et al., 2014 [6]	Weller-Fahy et al., 2015 [24]	M.Ahmed et al.,2016 [27]	Kurniabudi et al., 2019
Detection Technique/ Method	√	√	√	√	√	√	√	√
Type of Anomaly/ Outlier	√	√	√	√	√	√	√	√
Type of Attack	-	-	-	-	-	√	√	√
Output anomaly/ outlier	√	√	-	-	-	-	√	√
Data Repositaries	√	-	√	√	√	√	√	√
Data Types	√	√	√	-	√	√	-	√
Anomaly/ Outlier identity	-	√	-	-	-	-	-	√
Research Challenge	√	√	√	-	√	-	√	√
Categorize Network	-	-	-	-	-	-	-	√
Evaluation method	-	-	-	-	-	-	-	√

Authors of this paper believe, in addition to the knowledge of methods, systems, techniques, and types of anomalies, researchers need also to gain an understanding of the anomaly detection research trends related to the research area, data sources, evaluation methods and performance measures used. In particular, this survey differs from the existing surveys in the following:

1) In contrast to [17], this survey paper focuses on anomaly detection in network system.
2) In contrast to [1], this survey paper is not restricted to WSN, instead, this survey paper presents and compares anomaly detection methods on various network applications.

3) Similar to [22], this paper provides an up-to-date survey on the anomaly detection methods in difference ways. It presents the trend of the most used detection techniques by researches and their application domains.
4) In contrats to [6], this paper provides evaluation method that has been used in network anomaly detection.
5) In contrast to [24], this paper presents detection method, type of anomaly, and data type. In addition, it includes anomaly identity, output of anomaly, and application network categories.
6) Similar to [27], this paper compares the types of anomaly and types of attacks. In addition, it includes information about the types of networks and domains where detection is performed.

3. METHODOLOGY

This survey study uses Systematic Literature Review (SLR) [28], [29]. In this study the authors do a similar article search through the IEEE Xplore and ScienceDirect. The search process uses two keywords. The first keyword is "network anomaly detection"; the second keyword is "network outlier detection". To link the first and the second keywords the Boolean "OR" is used and focus the search on Full text and Metadata only. The search focuses on articles published in Journals & Conferences. For the year they are published, the range is determined from 2007 to 2017. Based on these criteria, the search results 329 articles. Further filtering is done through abstract search, to ensure that the article matches the topics covered and obtained 34 articles.

4. DISCUSSION

This section discusses anomaly detection studies, methods or techniques, types of anomaly and attack, output, data repositories, data types, anomaly identity, evaluation methods and research challenges in anomaly detection research.

4.1. Anomaly detection studies

Network anomaly detection is an extensive and widely studied research topic. Table 2 Shows summary of the existing anomaly detection researches. Each researcher utilizes different techniques to solve problems in network anomaly detection. Authors in [30] use Path Computation Element (PCE) Anomaly Detector (PAD) to detect malicious utilization in computing services. Research in [31] proposes an Integrated Anomaly Detection System (ADS). This system combines host-based anomaly detection and network-based anomaly detection for detecting Cyber intrusions to substations of power grid. Authors in [32] introduce a model of traffic matrix estimation and anomaly detection. This model is capable for detecting and correcting data errors. Other reseachers focus on anomalous behavior, such as researchers in [33] and [34] discuss anomalous behavior detection models on cloud computing networks. The models have ability to detect and predict anomalies in real-time. Study in [35] proposes anomaly detection techniques to deal with long-term anomalies, while study in [36] deals with different times. Research in [37] and [38] each introduces detecting anomalous traffic, and special anomaly detection of wireless sensor networks, respectively. Study in [39] proposes a principal component analysis (PCA)-based anomaly method whilst study in [40] uses a PCA sparse. Authors in [41] propose a toolkit for analyzing and detecting anomalous behavior on the Internet. The works in [42]-[44] propose an anomalous mobile agent-based detection scheme, diagnostic and detection of occurrences on a network-wide, and PCA-based distributed-anomaly detection scheme, respectively. Study in [45] introduces an anomaly mobile anomaly detection mechanisms based on information entropy. In [46], the researchers propose anomaly detection for big data mobile networks. Lastly, authors in [47] propose an algorithm for the detection of anomalies on large-scale networks. Based on the observation on the existing studies, the current anomaly detection research trends focus on models, methods, schemes, and algorithms to create a reliable anomaly detection system.

4.2. Network Categories and Application Domain

In this paper, the category of the application network is based on the information about the environment and application domains applied to the anomaly detection. The category considers also the type of traffic or data used as follows: 1) Smart network; includes smart control system in smart city, home, and industries; 2) Large scale network; inculdes Internet Service Provider (ISP), Multi-Protocol Label Switching (MPLS), backbone network and cloud computing. 3) Wireless Sensor Network. 4) Mobile networks, and 5) Conventional nework; includes computer network and TCP network. Table 2 presents the statistics on anomaly detection works that have been carried out in each network category.

Table 2. Number of article by network category

Nework Category	Number of Article
Smart City/ Home/ Industrial	5
Large-scale Network	10
Wireless Sensor Network	11
Mobile	1
Conventional Network	7

4.3. Anomaly detection methods

Methods used by researchers are also evolving and diverse enough. This sub-section presents the methods used in anomaly detection research. The observations of survey papers conclude several similar methods. As an example, authors in [1] identify statistical, nearest neightbor, clustering, classification, spectral-decomposition as anomaly detection methods. Whereas authors in [23] identify statistics, digital signal processing, information theory as anomaly detection methods. Survey by [8] identify statistical, classification-based, clustering and outlier-based, soft computing, knowledge-based and combination-learners. Research in [27] identify a statistical, information theory, clustering and classification. Thus, it can be concluded that anomaly detection methods used by researchers are clustering, classification, statistical, information theory, nearest neighbor, spectral-decomposition, soft computing, knowledge-based, digital signal processing and combination-learner. Figure 1 shows the summary of detection methods concluded from surveys.

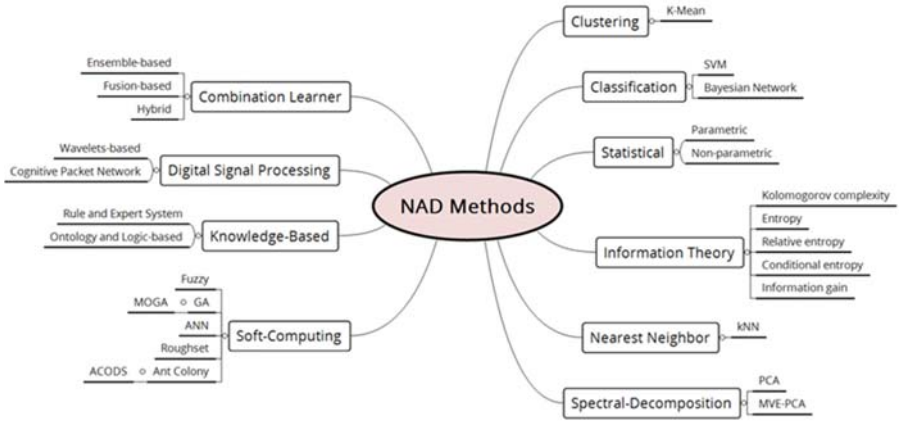

Figure 1. Network anomaly detection methods.

The following are a brief description of each of the method:

a) **Clustering,** a method for grouping a number of similar objects into groups called clusters so that objects in the same cluster share similarities with each other than objects found in other clusters [6]. Researchers in [2] use clustering algorithm in preprocessing step to clustering sensor data into normal cluster and outlier cluster. K-means is the most common algorithm for clustering, usually combined with another technique for outlier detection on data stream [48]. Researchers in [49] combine K-Means and Iterative Dichotomiser 3 (ID3) method for anomaly detection, resulting in high accuracy. In order to detect anomaly using K-means, firstly need to set nomal clusters, anomalous clusters, and suitable similarity measures. Secondly perform an offline preprocessing phase [50].
b) **Classification,** it starts with learning a set of instances data (*training*) and classify an unseen instance into one of the learned (normal/outlier) class (*testing*) [1]. A classification method identifies membership of a set of categories of observations, based on a set of training data that contains observations of categories whose memberships are known [6]. Researchers in [51] use a Support Vector Machines (SVM) classifier to detect network anomaly traffic. One SVM class is most widely used for anomaly detection and it is used to effectively separate normal and anomalous data from the features space learned [52].
c) **Statistical**, it is the earliest method, which is used for outlier detection problems. Based on how the probability model is built the statistical-based techiques are categorized into parametric and non-parametric [1]. The statistical method approach is based on the development of probabilistic data models as well as the use of mathematical methods from applied statistics and probability theory [22]. Statistical is a method of mathematical scheme that uses temporal characteristics, events and trends to create process

profiles and capture specific dynamics (eg network anomalies) relying heavily on statistical methods [23]. Statistically, anomalies are observations that are suspected either partially or entirely irrelevant as they are not generated by a stochastic model assumed [6].

d) **Information theory,** this method analyzes content information using information theory such as: Kolomogorov complexity, entropy, relative entropy, etc. to explain dataset charateristic [17], and involves information quantification [23]. Information theory uses one of the following measurements: entropy, conditional entropy, relative entropy, relative conditional entropy, or information gain [27]. Researchers in [53] propose the Method of Entropy Spaces (MES), which useful to detecting anomalous traffic. Having done evaluation in a real scenario, the proposed method achieved good performance in detecting anomalies. Meanwhile researchers in [45] use information entropy in anomaly detection mechanism for mobile payment application. The proposed mechanism can improve system stability and reduce false alarm.

e) **Nearest neighbor,** this method measures similarity or distance of data instance to differentiate data instance [17]. The most common method is the use of an approach to analyze sample data with respect to its nearest neighbors in the data mining community and machine learning [1]. The suitable nearest neighbor algorithm for anomaly detection is *k-nearest neighbor* (k-NN). k-NN calculates the nearest neighbors of a record using a suitable distance calculation metric such as Euclidean distance or Mahalanobis distance [22]. Fawzy *et al.* [2] propose outlier detection approach by using nearest neighbor. The experiment results show that the method achieves high accuracy rate for identifying outlier. Chorppath *et al.* [54] compare three machine learning techniques, which are SVM, Naive Bayes and k-NN. Performance measurements show k-NN technique has a lowest true positive rate (TPR) and highest false positive rate (FPR) among the three methods.

f) **Spectral-decomposition,** this method uses a combination of attributes that captures most of the variability in the data in order to find the approximate data [17]. Spectral method aims to find the normal behavior mode in the data by using the principle component [1]. In the early step of outlier detection, spectral decomposition-based approach uses PCA to reduce dimensionality [55]. Similar to [55], Zolotukhin *et al.* [56], use PCA to reduce dimensionality of feature vectors corresponding with web resources. PCA is the most common method used for analysis high-dimensionality data [40]. In Oreilly *et al.* [44] the Minimum Vollume Elliptical PCA (MVE-PCA) is introduced. This method shows superior performance from a classic PCA. Experiment results show that the computational complexity of distributed MVE-PCA is lower than centralized MVE-PCA.

g) **Soft computing,** Soft computing is usually thought of as encompassing methods such as genetic algorithms, artificial neural networks, fuzzy sets, rough sets, ant colony algorithm and artificial immune system 6]. Authors in [57] employ Multi-Objective Genetic Algorithm (MOGA) to detect anomalies from large data sets by analyzing subspaces, where in high-dimensional space context, subspace anomalies concerned as anomalies. Authors in [58] combine genetic algorithm and fuzzy logic. Firstly, the Genetic Algorithm is used to generate digital signature of network segment by using flow analysis. Then, Fuzzy Logic is applied to detect anomaly on instances data. The proposed method achieves 96.53% accuracy and 0.56% false positive rate. In [59] using modification of ant colony optimation metaheuristic that called Ant Colony Optimization for Digital Signature (ACODS) is compared with the PCA for Digital Signature (PCAD). The result from Normalised Mean Square Error (NMSE) correlation coefficient of the methods present similar result. This soft computing method not only works well in detecting anomalies, but is also used for feature selection, such as in [60] that use rough set theory for feature selection.

h) **Knowledge-based,** In this method, the network or host event is checked and matched with predefined rules or attack patterns. The goal is to identify known attacks in common mode so that handling the actual event becomes easier [6]. Samples In Alipour *et al.* [14] build an online model to detect anomalies. This model identifies abnormal actifities by monitoring n-gram of state transition in real traffic sessions. Any state transition violation considered as an abnormal activity.

i) **Digital signal processing,** this method is used to represent network traffic into the form of signal components that can be processed dependently [23]. Typically, a signal is converted from the time (or space) domain into the frequency domain, e.g., by means of a Fourier transform. There are two signal-processing-based approaches: wavelet-based approach and cognitive packet network (CPN)-based approach [22].

j) **Combination learner,** this method use several techniques simultaneously or combined to improve the accuracy of the anomaly detection system. Combination learner inculdes: ensemble based, fusion based and hybrid [6]. In ensemble-based technique multiple model can be combined to classify data instances. The same algorithms can be applied to different dataset or/and same dataset and can be trained with different algorithms [61]. Researchers in [62], propose an ensemble of five binary classifiers to detect anomalies from wireless sensor network. Each classifier uses vary algorithms, from simple average

computing to complex algorithms such as neural or Artrifical Neural and Fuzzy Inference System (ANFIS) network. The experimental results show the efficiency of the ensemble method. In paper [50], a heterogeneous set of local online learning classifier was developed to automatically recognize anomaly in data without any prior knowledge. Then, by using ensemble-based method a multiple and diverse individual classifier will be combined. Fisher's method or median is used to agregate the individual classifier that applied in parallel for same data. Experiment results confirm that this ensemble method improves the anomaly detection accuracy. While authors in [63], propose a complex combination of anomaly detector with unsupervised (mean, max, rank BFS, mean rank) and supervised (SVM-perf, TopPush, RankBoost, and Acc@Top) methods. All of these methods are compared with two existing anomaly detection systems which are Net-Flow and HTTP network anomaly detection. The experimental results show that the proposed method outperforms the prior methods with significant accuracy.

Gogoi *et al*. [22] and Comput et al [64] categorize anomaly detection methods as supervised and unsupervised methods. The following are brief descriptions of both methods.

a. **Supervised Method**, requeires pre-labeled data, tagged as normal or abnormal. Usually train the data with normal pattern and try to detect attack with comfirmity normal pattern. This method can detect known attack [22], uses prior knowledge to build a normal profile [64] and generally labeled data is needed [65].
b. **Unsupervised Method**, does not need a pre-labeled data set, can detect unknown attack [22], with non prior knowledge of data [64], however, use some measurement criteria to identify outliers [1]. In unsupervised (or cluster) method the data point that separated from normal will be considered as anomaly [66].

Other than supervised and unsupervised methods, a pre-defined data anomaly detection method can be be defined as semi-supervised method [1], [6]. A large amount of unlabeled data, used together with pre-labeled data to build better classifiers is practiced in semi-supervised method [9]. Semi-supervised method assumes the training data has only labeled instances for normal class. The use of labels for anomaly class is not required. They much easier compared to supervised approach [6], [67]. An example of semi-supervised method is proposed in [68] that presents a semi-supervised statistical method. This method is then compared with Naive Bayes method, resulting the proposed method overcome Naive Bayes in detection capabilities.

Table 3. Comparison the method types of anomaly and attacks

Author Name	Methods	Algorithm	Pros and contras
Gharbaoui et al.,2013 [30]	Statistical	Sequential Hypothesis Testing (SHT)	Detection capabilities achieve a good performance, Reduce false alarm, Tested on realtime situation is needed
Nguyen & Roughan, 2013 [37]	Statistical	Hidden Markov Model (HMM)	Low computation and communication overheads, Suitable for adoption by ISPs, Only on small size data
Fernandes et al., 2016 [60]	Statistical	PCA + Ant Colony	Able to detect anomalous behavior, Computation very complex
Parwez et al.,2017 [46]	Clustering	k-means and hierarchical clustering + neural-network	Low complexity in k-means clustering, Better performance with hierarchical clustering, Hierarchical clustering facing space complexity for large data set
Zhu, C. et al., 2015 [9]	Classification	Bayesian Network	Very effective to detect anomaly, Requires user interference (expert) to adapt changed probabilities
Z. Zhang et al., 2016 [45]	Information theory	Information entropy + Neural Network back propagation	Can improve the stability of the system, Dynamica-lly adjust to traffic change, lower false alarm rate, Entropy value is detected to be too sensitive
Shabtai et al., 2010 [69]	Knowledge based	knowledge-based temporal abstraction (KBTA)	Support misuse detection and anomaly detection, KBTA was adapted for mobile devices that are limited in resources (i.e., CPU, memory, battery).
Usman et al., 2015 [70]	Soft computing	Fuzzy Logic	High accuracy in detect cross-layer anomalies, Low energy consumption, Initial domain knowledge is needed, Unreliable to transmit mobile agent (in poor communication)
Alipour, H. et al.,2015 [14]	Supervised	n-grams	System can detect difference attack, High detection rate, Low false alarm, Work with pre-labeled data
Dromard et al., 2017 [5]	Unsupervised	grid clustering algorithm	High performance in detection rate, Low false alarm, The speed must be improve
Lu & Ghorbani, 2009 [71]	Digital signal processing	Wavelet analysis	High-detection rates, Not accurate for real large-scale Wifi traffic
Fawzy et al., 2013 [2]	Nearest neighbor	k-nearest neighbor (kNN)	Able to detect outlier values, Can classify noisy data or interesting event, not tested in larger dataset
Oreilly et al., 2016 [45]	Spectral-decomposition	PCA	Able to reduce dimensionality, Computational is complex
Erfani et al.,2016 [50]	Combination-leaner	Deep belief networks (DBNs) + one-class SVM	Efficient, accurate and scalable anomaly detection. Able to implement with large-scale and high-dimensional domain, Tested on sensor network datasets only, so no guarantee for other domain

Networks with a variety of applications and equipment generate huge amounts of data, both in number and type. This is related to data dimensions. As a general knowledge, dimensionality is one of the problems in anomaly detection[17], [57]. Table 3 compares the methods used by researchers in solving problems in anomaly detection. So many used by researcher in anomaly detection studies, result in different pros and contras. From the best knowledge of the authors of this survey, the most popular issues in anomaly detection include and not limited to detection capabilities such as detection rate and false alarm. Another issue that related to detection capabilities is dimensional reduction and the computational complexity. Some researchers concern about computational time and scalability. This survey concludes some methods have achieved high performance in detection capability, however, there is consequency such as high false alarm, computational complexity, computational times and scalability. To overcome the problems, some researchers have proposed methods for dimensional reduction such as Juvonen *et al.* [72], Erfani *et al.* [52] and.Wei *et al.* [73]. Whereas Zhang *et al* [74] propose an algorithm with efficient computational time, and Sommer & Paxson use machine learning technique to improve the accuracy of network intrusion detection [75].

4.4. Types of Anomaly

The definition of anomalous types by researchers is associated with the network charateristic and area. This concept depends on complexity source of traffic, such as researchers in [22] define the type of anomaly based on [17], as point anomaly, contextual anomaly and collective anomaly. Researchers in [27] map the type of anomaly with type of attacks. In [9], the researchers perform behavioral anomaly detection in smart asisted living environtment. The authors separate point anomaly to three types as: spatial anomaly, timing anomaly, and duration anomaly. They also detect contextual anomalies that are defined as sequence anomaly. While researchers in [1] categorize the sources of outliers into two types: an error and an event.

The authors of this arcticle compare the type of anomaly detected by the existing works and in the category of which, anomaly detection network is done, as presented in Table 4. The authors of this paper discover that in smart, large-scale network and WSN, anomaly detection is performed to recognize collective anomaly, which is usually in the form of a DOS attack.

Table 4. Comparing Types of anomalies and Network Category

Author	Type of Anomaly						Network Category				
	Point	Collective	Contextual	Error	Event	Others	Smart	Large Scale	WSN	Mobile	Conventional
[11]	-	√	-	-	-	-	-	-	-	-	√
[30]	-	√	-	-	-	-	-	√	-	-	-
[7]	-	√	-	-	-	-	-	-	-	-	√
[13]	-	-	-	-	√	-	√	-	-	-	-
[37]	√	-	-	-		-	-	√	-	-	-
[40]	-	-	-	-	√	-	-	-	-	-	√
[76]	-	-	-	√	-	-	-	-	√	-	-
[31]	-	√	-	-	-	-	√	-	-	-	-
[14]	-	√	-	-	-	-	-	-	√	-	-
[9]	√	√	-	-	-	-	-	-	√	-	-
[8]	-	√	-	-	-	-	√	-	-	-	-
[15]	-	√	-	-	-	-	√	-	-	-	-
[42]	-	-	-	-	-	A,B	-	-	√	-	-
[70]	-	-	-	√	-	-	√	-		-	-
[35]	-	-	-	-	√	-	-	-	√	-	-
[39]	-	√	-	-	-	-	-	√	-	-	-
[32]	-	-	-	-	-	A,B	-	√	-	-	-
[43]	-	-	-	-	√	-	-	√	-	-	-
[10]	-	-	-	-	-	-	-	√	-	-	-
[45]	-	-	-	√	-	-	-	-	-	√	-
[33]	-	-	-	-	-	B	-	√	-	-	-
[47]	-	√	-	-	√	-	-	√	-	-	-
[12]	-	√	-	-	-	B	-	√	-	-	-
[32]	-	-	-	-	√	-	-	√	-	-	-
[46]	-	-	-	-	-	B	-	-	√	-	-
[5]	-	√	-	√	-	-	-	√	-	-	-

A=Attack, B=Behaviour

Findings of this work show that very limited numbers of researchers who detect anomaly points and even no one has done contextual anomalies. Indeed, this result has not been confirmed, however, some researchers did not mention the types of anomaly that were detected. Some researchers only provide information that they recognize the attack as an anomaly, some recognize the network traffic behavior

[33],other recognize the control and data planes [12] or wireless sensor network [46]. Even some researchers recognize both attack and behavior [42], [32]. Overall, this survey paper has confirmed that collective anomaly is the most popular research type of anomaly detection and the most researchers have resolved. Thus, the authors of this paper map out the types and sources of anomalies used by the researchers in the existing works as in Figure 2.

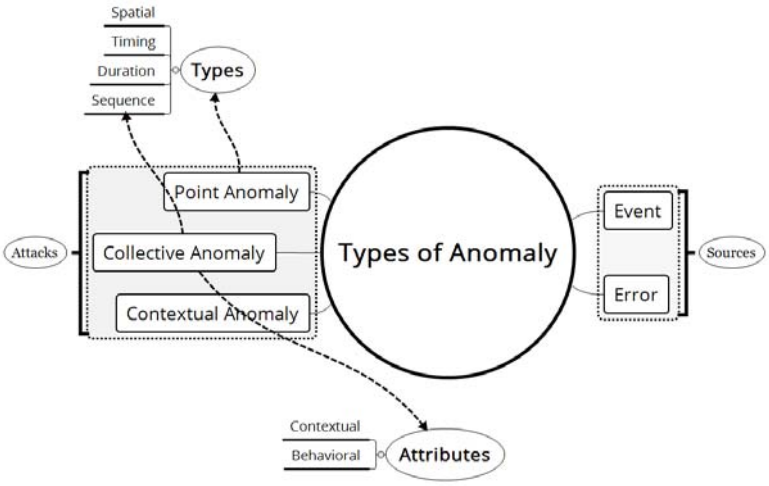

Figure 2. Mapping types of anomaly

4.5. Output of Anomalies

Anomaly detection technique capabilities can be seen based on test or measurement values. Usually these results are indicated in two types of scores or labels [17] [27].

a) Scores, are test results that provide anomalous scores on the data, depending on the extent to which the data is considered anomalous. This technique generates an anomaly ranking.
b) Labels, test results outcomes as a label in the data as normal or anomalous.

A score, which is a value that combine (i) distance or deviation with reference to a set of profiles or signatures, (ii) influence of the majority in its neighborhood, and (iii) distinct dominance of the relevant subspace [6]. Usually labeling techniques depends on (i) the size of groups generated by an unsupervised technique, (ii) the compactness of the group(s), (iii) majority voting based on the outputs given by multiple indices, or (iv) distinct dominance of the subset of features [6].

Whereas in wireless networks, anomalies are distinguished by scalar and outlier score[1] as follows.

a) Using zero-one classification measurement in scalar scale, which classifies each data into the normal class or outlier class.
b) While in outlier score technique, each measurement result gives a score. Score is based on measurement level.

4.6. Data Repositories

Anomaly detection is an important part of data analysis and is useful for recognizing network intrusion]. In order for the analysis to work properly, it must be supported by reliable data. In anomaly detection studies on the network, the type of traffic data used may vary. The more complex the dataset, the techniques used will have more challenges [17]. Researchers in [11] use dataset of the Los Angeles Network Data Exchange and Repository (LANDER), researchers in [76] use the dataset of the USA Army Research Laboratory (ARL), researchers in [35]use datasets from the Intel Berkeley Research Lab, researchers in [59] use the Abilene network dataset and researchers in [33] use the KDD-99 dataset. Researchers use topologies that are designed to meet the research needs, for example: a topology of 2 domains, 28 nodes, 55 bidirectional links and each link provides 2.5 Gbps bandwidth [30]. Authors in [70] build a topology that represents a minimalist smart home.

While other researchers use data sources captured from traffic in the specific network. Such as in [14] research, they capture wireless traffic from the ECE department at the University of Arizona. Researchers in [12] use traffic data captured from King Saud University network infrastructure. The authors of this article observe that there are three types of data usually used in network anomaly detection research. First, using data that captured directly from the real network. Second, using publicly available dataset, and the third, using data that captured from topologies specifically designed for testing, often called testbed topologies. Figure 3 plots the data presented in Table 5, showing 52% of researchers use the publicly available dataset as traffic data for analysis purposes. 35% of researchers use testbed topology, and 13% use data captured directly from the network.

Table 5. The comparison of source of data vs evaluation method used by researchers

Author(s)	Source of data			Evaluation Method	Comments
	Captured	Testbed	Dataset		
Tartakovsky et al.,2013 [11]	-	-	√	Experimental	Los Angeles Network Data Exchange and Repository (LANDER) project; flow data captured by Merit Network Inc.
Gharbaoui et al.,2013 [30]	-	√	-	Simulation	Simulated topology: 2 domains, 28 nodes, 55 bidirectional links. Each link direction supports a bitrate of 2.5 Gbps
Rahmé et al.,2013 [7]	-	-	√	Simulation & Experimental	Traffic collected from RENATER network
Difallah et al.,2013 [13]	√			Experimental	Data from Water Distribution Network
Nguyen & Roughan, 2013 [37]	√	-	-	Simulation	Using sample data from multiple ISP
Jiang et al., 2013 [40]	-	-	√	Experimental	world dataset from: financial markets, wireless sensor networks, and machinery operating condition
Zhang, R. et al. 2013 [76]	-	-	√	Simulation & real measurement	Data set from USA Army Research Laboratory (ARL)
Hong, J. et al. 2014 [31]		√		Simulation	WSU cyber security testbed
Alipour, H. et al.,2015[14]	√	--	-	Experimental	Data set: 2 channel wifi captured from Wireless traffic of ECE department at the University of Arizona
Zhu, C. et al., 2015 [9]	-	√		Experimental	Data from weareble Sensor
Zhou, C., 2015 [8]	-	√	-	simulation and real-time	Designed data set
Ntalampiras, S. 2105 [15]		√	-	Simulation & Experimental	IEEE 30-busmodel
Chen, P. et al. 2015 [42]	-	-	√	Simulation	data set: real-world industrial environment
Usman et al.,2015 [70]	-	√	-	Experimental	testbed that representing a minimal working smart home sensor network (two sensor nodes and a laptop node)
Xie, M. et al. [36]	-	-	√	Experimental	Data set: Intel Berkeley Research Lab
Kumarage, H. et al. [38]	-	-	√	Experimental	Data set: Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) and data set: Intel Berkeley Research Laboratory
Wang, J. et al., [36]	-	√	-	Simulation	Traffic from simulation network
Ding, M., & Tian, H. [39]	-	-	√	Experimental	Data set: Abilene network
Dong, W., el al. [16]	-	√	-	Simulation & Experimental	Data from real-world sensor network system
Lutu, A., et al. 2016 [41]	-	√	-	Experimental	Unique dataset from internet
Zhang, Q., & Chu, T. 2016 [32]	-	-	√	Simulation	TUIDS data set: Network Security Laboratory at Tezpur University
Zhang, Y., et al. 2016 [43]	-	√	-	Experimental	data collected multidomain network: data collected from DOE lab and ESnet hub sites
Mardani, M., & Giannakis, G. B. [10]	-	√	-	Simulation	synthetic and real Internet data
Zhang, Z. et al., [36]	-	-	√	Simulation	Data set: KDDCUP99
Ye, X. et al. 2016 [33]	-	-	√	Experimental	Data set: KDD-99
Kasai, H., et al. 2016 [47]	-	-	√	Experimental	Synthetic and real-world (Abilene Network Dataset)
AsSadhan, B. Et al 2017[12]	√	-	-	Experimental	data capture from Internet traffic at King Saud University network
Böse, B., [34]	-	-	√	Experimental	Data set: DARPA ADAMS
Parwez et al., [46]	-	-	√	Experimental	CDR (Call Detail Record) dataset
Dromard et al., [5]	-	-	√	Experimental	ONTS dataset, theMAWILab network traces

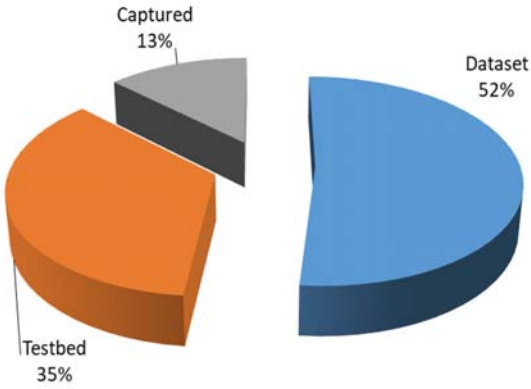

Figure 3. Data used in anomaly detection research

4.7. Data Type

The main aspect in anomaly detection researches is the type of input data. Input data can be a set of attributes (often known as variables, characteristics, features, fields or dimensions). The attributes can be of different types such as binary, categorical or continuous [17], [22]. The type of input data determines the detection method that can be used to analyze the data. Each data instance may consist of only one attribute (univariate) or multiple attributes (multivariate) [6]. The technique of detecting outliers on sensor data usually considers the following two aspects[1]:

1) Attributes; An outlier in univariate data with a single attribute can be easily detected if the single attribute is anomalous with respect to that attribute of other data. The sensor node equipped with multiple sensors and also certain correlations may exist among attributes of sensor data. In this case outlier detection method for WSNs should be able to analyze multivariate data;
2) Correlations; defines dependencies: (i) dependencies among the attributes of the sensor node, and (ii) dependency of sensor node readings on history and neighboring node readings].

4.8. Outlier/ Anomaly Identity

Generally the outlier detection method does not distinguish between errors and events, tend to regard the outlier as an error. This fact results in the loss of important information hidden from an event. Thus, indentify outlier source and distinction between errors, events and malicious attacks is one of the challenges in detecting outliers in WSNs. This survey work concludes that error and event as a type of anomalies, however are also considered as source of anomaly. As shown in Table 4, the researchers identify error [45], [70], [76] and event[43], [47], [34]. Research in [76] identifies error sensor in WSN, while researchers in [5], identifies network errors or failures in large-scale networks by evaluating traffic flow. On the other hand, researchers in [40] identify anomaly in data stream by simulations of some abnormal events such as box removal and replacement, rotation, and flipping. Whereas researchers in [34] identify suspicious activities in real time by evaluate an event session in data stream. Lastly, researchers in [43] detect anomaly event and leverage Q-statistic event correlation analysis in large scale network.

4.9. Evaluation Method

Evaluation and validation are among the one of the important stages in every study and researchers use different approaches to do so. Researchers use experiments to evaluate the proposed works. For example, experiments to evaluate anomaly detection system on smart city infrastructure network [77], experiments to verify the framework [9], and experiments to analyze perfSONAR performance in detecting occurrence, experiments on calculation of accuracy of normal and abnormal data points [38]. The use of testbed such as, test Joint Sparse PCA Algorithms [40], monitor traffic and test system performance detection [14]. Then the use of simuilation, such as evaluating the ability of PAD detect malicious traffic [30], evaluating Hidden Markov Model (HMM) to detect SSH burce-force attack [37], validating the integrated ADS method [31], simulatingTraffic Matrix estimation and anomaly detection [32], testing and validating the sliding mode method on real data traffic [7]. Figure 4 illustrates the statistics of the use of evaluation and validation method used in network anomaly detection. It shows 58% of the researchers use the experimental approach, 26% utilze the simulation, while 10% use experiments and 6% use other approaches.

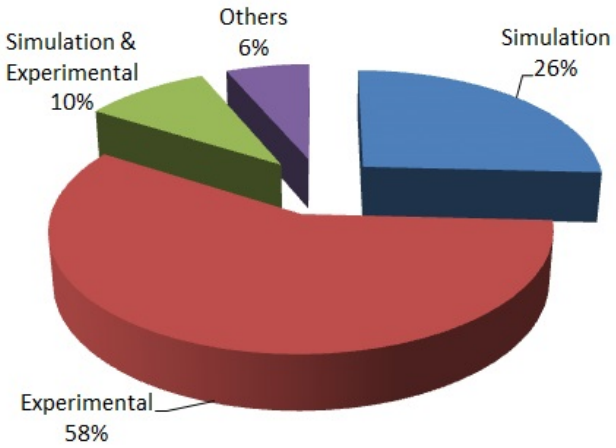

Figure 4. Evaluation and validation methods used in the network anomaly detection researches.

4.10. Open Issues and Research Challenges

As a general knowledge, the main issues of network anomaly detection include detection capabilities [71], [78], [79], high dimensionality of data [53], [57], [80] computational complexity and computational times [72], [81], [82]. The detection capability is related with detection rate [14], [52], [66] and false alarm [83]. As mentioned in [1], the challenge in traditional outlier detection is how to achieve high detection rate and low false alarm at the same time. Many researches have been carried out to build anomaly detection systems with high detection rate [28]. However, more issues come with the rapid network development. More complex network will produce heterogeneous and huge volume of data such as the Internet of Things (IoT), this complexity becomes challenges in anomaly detection. In IoT many sensors and devices with different protocols interconnected and produce data stream and result in high dimensionality data. The dimensionality related to the size of data traffic. Heterogeneous of traffic becomes challenge in data analysis. Data captured from IoT network must be extracted with specific technique to become readable information. Since many protocols have contributed to data stream, spesific method is needed to read this difference of data. Thus, more challenges in extracting data. On the other hand, to analyze the data with huge volume, the high capabilities and intelligent algorithms are needed which in turn result in computational complexity. Reseacher must take into consideration on how to select significant and important features from the extracted feature, this so called dimentional reduction. The work becomes a challenge, because an unknown feature that relevant to detection of anomalous traffic and now known as an attack. Outcome of the survey done in this work shows that most detection is successfully done as off-line. Thus, it becomes a challenge to build a real-time network anomaly detection.

5. CONCLUSION

In this survey study the authors have reviewed articles on network anomaly detection collected from IEEE Explorer and ScienceDirect. As a general knowledge, anomaly detection research field is very wide and dynamic. The survey study summarized current anomaly detection research trends and focus on models, methods, schemes, algorithms to create a reliable anomaly detection system. The study found out that current network anomaly detection has been done on network category of smart network, large scale network, wireless sensor network, mobile networks and conventional nework, include computer network and Transmission Control Protocol (TCP) network. The study concluded that the most popular issues in anomaly detection include the high dimensionality of data, detection capabilities, complexity of computational, and computational times. Although each researcher uses different terminology to measure performance, however, the goal is a same, i.e.: to build a reliable anomaly detection system. The network anomaly detection must achieve high performance with high accuracy on detection rate and low false alarm at the same time. Further, modern network anomaly detection should have ability for real-time detection and automatic profile update. The survey study showed 52% of researchers use publically available benchmark dataset as traffic data. Study also showed that 58% of researchers used experimental approach in evaluating or validating the proposed works. Taking into account the current research trends and network developments, future research is still highly likely to address anomaly on large-scale networks, which generate a variety of traffic types, and real-time observations.

REFERENCES

[1] N. Zhang, Meratnia and P. Havinga, "Outlier Detection Techniques for Wireless Sensor Networks: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[2] A. Fawzy, H. M. O. Mokhtar, and O. Hegazy, "Outliers detection and classification in wireless sensor networks," *Egypt. Informatics J.*, vol. 14, no. 2, pp. 157–164, 2013.

[3] F. Seredynski and P. Bouvry, "Anomaly detection in TCP/IP networks using immune systems paradigm," *Comput. Commun.*, vol. 30, no. 4, pp. 740–749, 2007.

[4] H.-J. Liao, C.-H. Richard Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 16–24, 2013.

[5] J. Dromard, G. Roudiere, and P. Owezarski, "Online and Scalable Unsupervised Network Anomaly Detection Method," *IEEE Trans. Netw. Serv. Manag.*, vol. 14, no. 1, pp. 34–47, 2017.

[6] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," *Ieee Commun. Surv. tutorials*, vol. 16, no. 1, pp. 303–336, 2014.

[7] S. Rahme, Y. Labit, F. Gouaisbaut, and T. Floquet, "Sliding modes for anomaly observation in TCP networks: From theory to practice," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 3, pp. 1031–1038, 2013.

[8] C. Zhou *et al.*, "Design and analysis of multimodel-based anomaly intrusion detection systems in industrial process automation," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 45, no. 10, pp. 1345–1360, 2015.

[9] C. Zhu, W. Sheng, and M. Liu, "Wearable sensor-based behavioral anomaly detection in smart assisted living systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1225–1234, 2015.

[10] M. Mardani and G. B. Giannakis, "Estimating traffic and anomaly maps via network tomography," *Biol. Cybern.*, vol. 24, no. 3, pp. 1533–1547, 2016.

[11] A. G. Tartakovsky, A. S. Polunchenko, and G. Sokolov, "Efficient computer network anomaly detection by changepoint detection methods," *IEEE J. Sel. Top. Signal Process.*, vol. 7, no. 1, pp. 4–11, 2013.

[12] B. AsSadhan, K. Zeb, J. Al-Muhtadi, and S. Alshebeili, "Anomaly Detection Based on LRD Behavior Analysis of Decomposed Control and Data Planes Network Traffic using SOSS and FARIMA Models," *IEEE Access*, 2017.

[13] D. E. Difallah, P. Cudre-Mauroux, and S. A. McKenna, "Scalable anomaly detection for smart city infrastructure networks," *IEEE Internet Comput.*, vol. 17, no. 6, pp. 39–47, 2013.

[14] H. Alipour, Y. B. Al-Nashif, P. Satam, and S. Hariri, "Wireless anomaly detection based on IEEE 802.11 behavior analysis," *IEEE Trans. Inf. forensics Secur.*, vol. 10, no. 10, pp. 2158–2170, 2015.

[15] S. Ntalampiras, "Detection of integrity attacks in cyber-physical critical infrastructures using ensemble modeling," *IEEE Trans. Ind. Informatics*, vol. 11, no. 1, pp. 104–111, 2015.

[16] W. Dong, L. Luo, C. Chen, J. Bu, X. Liu, and Y. Liu, "Post-Deployment Anomaly Detection and Diagnosis in Networked Embedded Systems by Program Profiling and Symptom Mining," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 12, pp. 3588–3601, 2016.

[17] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, 2009.

[18] R. Goel, A. Sardana, and R. C. Joshi, "Parallel Misuse and Anomaly Detection Model," vol. 14, no. 4, pp. 211–222, 2012.

[19] A. Abduvaliyev, A.-S. K. Pathan, J. Zhou, R. Roman, and W.-C. Wong, "On the vital areas of intrusion detection systems in wireless sensor networks," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 3, pp. 1223–1237, 2013.

[20] C. OReilly, A. Gluhak, M. Imran, and S. Rajasegarar, "Anomaly Detection in Wireless Sensor Networks in a Non-Stationary Environment," *Ieeexplore.Ieee.Org*, vol. 16, no. 3, pp. 1–20, 2013.

[21] B. Al-Musawi, P. Branch, and G. Armitage, "BGP Anomaly Detection Techniques: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 1, pp. 377–396, 2017.

[22] P. Gogoi, D. K. Bhattacharyya, B. Borah, and J. K. Kalita, "A survey of outlier detection methods in network anomaly identification," *Comput. J.*, vol. 54, no. 4, pp. 570–588, 2011.

[23] A. K. Marnerides and A. Mauthe, "Traffic Anomaly Diagnosis in Internet Backbone Networks : A Survey," *Comput. NETWORKS*, no. August, 2014.

[24] D. J. Weller-Fahy, B. J. Borghetti, and A. A. Sodemann, "A Survey of Distance and Similarity Measures Used Within Network Intrusion Anomaly Detection," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 1, pp. 70–91, 2015.

[25] A. Patcha and J.M Park. An overview of anomaly detection techniques: existing solutions and latest technological trends. Computer Networks vol. 51: pp. 3448–70, 2007;.

[26] P. García-Teodoro, E. Díaz-VerdejoJesús, G. Maciá-Fernández E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges", Computers & Security, vol. 28, no. 1-2, pp. 18-28, 2009. DOI: 10.1016/j.cose.2008.08.003.

[27] M. Ahmed, A. Naser Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *J. Netw. Comput. Appl.*, vol. 60, pp. 19–31, 2016.

[28] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3," *Engineering*, vol. 45, no. 4ve, p. 1051, 2007.

[29] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009.

[30] M. Gharbaoui, F. Paolucci, A. Giorgetti, B. Martini, and P. Castoldi, "Effective statistical detection of smart confidentiality attacks in multi-domain networks," *IEEE Trans. Netw. Serv. Manag.*, vol. 10, no. 4, pp. 383–397, 2013.

[31] J. Hong, C.-C. Liu, and M. Govindarasu, "Integrated anomaly detection for cyber security of the substations," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1643–1653, 2014.
[32] Q. Zhang and T. Chu, "Structure regularized traffic monitoring for traffic matrix estimation and anomaly detection by link-load measurements," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 12, pp. 2797–2807, 2016.
[33] X. Ye *et al.*, "An anomalous behavior detection model in cloud computing," *Tsinghua Sci. Technol.*, vol. 21, no. 3, pp. 322–332, 2016.
[34] B. Bose, B. Avasarala, S. Tirthapura, Y. Y. Chung, and D. Steiner, "Detecting Insider Threats Using RADISH: A System for Real-Time Anomaly Detection in Heterogeneous Data Streams," *IEEE Syst. J.*, vol. 11, no. 2, pp. 471–482, 2017.
[35] M. Xie, J. Hu, and S. Guo, "Segment-based anomaly detection with approximated sample covariance matrix in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 2, pp. 574–583, 2015.
[36] J. Wang and I. C. Paschalidis, "Statistical traffic anomaly detection in time-varying communication networks," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 2, pp. 100–111, 2015.
[37] H. X. Nguyen and M. Roughan, "Multi-observer privacy-preserving hidden markov models," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6010–6019, 2013.
[38] H. Kumarage, I. Khalil, and Z. Tari, "Granular evaluation of anomalies in wireless sensor networks using dynamic data partitioning with an entropy criteria," *IEEE Trans. Comput.*, vol. 64, no. 9, pp. 2573–2585, 2015.
[39] M. Ding and H. Tian, "PCA-based network Traffic anomaly detection," *Tsinghua Sci. Technol.*, vol. 21, no. 5, pp. 500–509, 2016.
[40] R. Jiang, H. Fei, and J. Huan, "A family of joint sparse PCA algorithms for anomaly localization in network data streams," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 11, pp. 2421–2433, 2013.
[41] A. Lutu, M. Bagnulo, C. Pelsser, O. Maennel, and J. Cid-Sueiro, "The BGP visibility toolkit: Detecting anomalous internet routing behavior," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 1237–1250, 2016.
[42] P.-Y. Chen, S. Yang, and J. A. McCann, "Distributed real-time anomaly detection in networked industrial sensing systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3832–3842, 2015.
[43] Y. Zhang, S. Debroy, and P. Calyam, "Network-wide Anomaly Event Detection and Diagnosis with perfSONAR," *IEEE Trans. Netw. Serv. Manag.*, vol. 13, no. 3, pp. 666–680, 2016.
[44] C. Oreilly, A. Gluhak, and M. A. Imran, "Distributed Anomaly Detection Using Minimum Volume Elliptical Principal Component Analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 9, pp. 2320–2333, 2016.
[45] Z. Zhang, X. Wang, and L. Sun, "Mobile payment anomaly detection mechanism based on information entropy," *IET Networks*, vol. 5, no. 1, pp. 1–7, 2016.
[46] M. S. Parwez, D. Rawat, and M. Garuba, "Big Data Analytics for User Activity Analysis and User Anomaly Detection in Mobile Wireless Network," *IEEE Trans. Ind. Informatics*, vol. 3203, no. c, pp. 1–1, 2017.
[47] H. Kasai, W. Kellerer, and M. Kleinsteuber, "Network Volume Anomaly Detection and Identification in Large-Scale Networks Based on Online Time-Structured Traffic Tensor Tracking," *IEEE Trans. Netw. Serv. Manag.*, vol. 13, no. 3, pp. 636–650, 2016.
[48] P. Chauhan and M. Shukla, "A review on outlier detection techniques on data stream by using different approaches of K-Means algorithm," *2015 Int. Conf. Adv. Comput. Eng. Appl.*, pp. 580–585, 2015.
[49] S. R. Gaddam, V. V Phoha, S. Member, and K. S. Balagani, "K-Means + ID3 : A Novel Method for Supervised Anomaly Detection by Cascading K-Means Clustering and ID3 Decision Tree Learning Methods," vol. 19, no. 3, pp. 345–354, 2007.
[50] H. H. W. J. Bosman, G. Iacca, A. Tejada, H. J. W??rtche, and A. Liotta, "Ensembles of incremental learners to detect anomalies in ad hoc sensor networks," *Ad Hoc Networks*, vol. 35, pp. 14–36, 2015.
[51] G. Yan, "Network anomaly traffic detection method based on support vector machine," pp. 2–5, 2016.
[52] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognit.*, vol. 58, pp. 121–134, 2016.
[53] R. Zempoaltecatl-Piedras, P. Velarde-Alvarado, and D. Torres-Roman, "Entropy and Flow-based Approach for Anomalous Traffic Filtering," *Procedia Technol.*, vol. 7, no. 33, pp. 360–369, 2013.
[54] A. K. Chorppath, T. Alpcan, and H. Boche, "Bayesian Mechanisms and Detection Methods for Wireless Network with Malicious Users," *IEEE Trans. Mob. Comput.*, vol. 15, no. 10, pp. 2452–2465, 2016.
[55] D. S. Shukla, A. C. Pandey, and A. Kulhari, "Outlier detection: A survey on techniques of WSNs involving event and error based outliers," *Proc. Int. Conf. Innov. Appl. Comput. Intell. Power, Energy Control. with Their Impact Humanit. CIPECH 2014*, no. November, pp. 113–116, 2014.
[56] M. Zolotukhin, T. Hämäläinen, T. Kokkonen, and J. Siltanen, "Analysis of HTTP requests for anomaly detection of web attacks," *Proc. - 2014 World Ubiquitous Sci. Congr. 2014 IEEE 12th Int. Conf. Dependable, Auton. Secur. Comput. DASC 2014*, pp. 406–411, 2014.
[57] J. Zhang, H. Li, Q. Gao, H. Wang, and Y. Luo, "Detecting anomalies from big network traffic data using an adaptive detection approach," *Inf. Sci. (Ny).*, vol. 318, no. August, pp. 91–110, 2015.
[58] A. H. Hamamoto, L. F. Carvalho, L. D. H. Sampaio, T. Abrão, and M. L. Proença, "Network Anomaly Detection System using Genetic Algorithm and Fuzzy Logic," *Expert Syst. Appl.*, vol. 92, pp. 390–402, 2018.
[59] G. Fernandes, L. F. Carvalho, J. J. P. C. Rodrigues, and M. L. Proença, "Network anomaly detection using IP flows with Principal Component Analysis and Ant Colony Optimization," *J. Netw. Comput. Appl.*, vol. 64, pp. 1–11, 2016.
[60] D. J. Weller-fahy, B. J. Borghetti, and A. A. Sodemann, "Within Network Intrusion Anomaly Detection," vol. 17, no. 1, pp. 70–91, 2015.
[61] G. Folino and P. Sabatino, "Ensemble based collaborative and distributed intrusion detection systems: A survey," *Journal of Network and Computer Applications*, vol. 66. Elsevier, pp. 1–16, 2016.

[62] D. Curiac and C. Volosencu, "Expert Systems with Applications Ensemble based sensing anomaly detection in wireless sensor networks," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9087–9096, 2012.
[63] M. Grill and T. Pevný, "Learning combination of anomaly detectors for security domain," *Comput. Networks*, vol. 0, pp. 1–9, 2016.
[64] J. P. D. Comput, H. Kumarage, I. Khalil, Z. Tari, and A. Zomaya, "Distributed anomaly detection for industrial wireless sensor networks based on fuzzy data modelling," *J. Parallel Distrib. Comput.*, vol. 73, no. 6, pp. 790–806, 2013.
[65] N. Stakhanova, S. Basu, and J. Wong, "On the symbiosis of specification-based and anomaly-based detection," *Comput. Secur.*, vol. 29, no. 2, pp. 253–268, 2010.
[66] S. Rajasegarar, C. Leckie, and M. Palaniswami, "Hyperspherical cluster based distributed anomaly detection in wireless sensor networks," *J. Parallel Distrib. Comput.*, 2013.
[67] O. Osanaiye, K. R. Choo, and M. Dlodlo, "Distributed Denial of Service (DDoS) Resilience in Cloud: Review and Conceptual Cloud DDoS Mitigation Framework," *J. Netw. Comput. Appl.*, 2016.
[68] N. B. Aissa and M. Guerroumi, "Semi-supervised Statistical Approach for Network Anomaly Detection," *Procedia Comput. Sci.*, vol. 83, no. Fams, pp. 1090–1095, 2016.
[69] A. Shabtai, U. Kanonov, and Y. Elovici, "Intrusion detection for mobile devices using the knowledge-based, temporal abstraction method," *J. Syst. Softw.*, vol. 83, no. 8, pp. 1524–1537, 2010.
[70] M. Usman, V. Muthukkumarasamy, and X.-W. Wu, "Mobile agent-based cross-layer anomaly detection in smart home sensor networks using fuzzy logic," *IEEE Trans. Consum. Electron.*, vol. 61, no. 2, pp. 197–205, 2015.
[71] W. Lu and A. A. Ghorbani, "Network anomaly detection based on wavelet analysis," *EURASIP J. Adv. Signal Process.*, vol. 2009, 2009.
[72] A. Juvonen and T. Hamalainen, "An Efficient Network Log Anomaly Detection System Using Random Projection Dimensionality Reduction," *2014 6th Int. Conf. New Technol. Mobil. Secur.*, pp. 1–5, 2014.
[73] J. Wei, M. Meng, J. Wang, Q. Ma, and X. Wang, "Neurocomputing Adaptive semi-supervised dimensionality reduction with sparse representation using pairwise constraints," *Neurocomputing*, vol. 177, pp. 564–571, 2016.
[74] J. Zhang, Y. Xiang, Y. Wang, W. Zhou, Y. Xiang, and Y. Guan, "Network traffic classification using correlation information," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 104–117, 2013.
[75] R. Sommer and V. Paxson , "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection", Proceedings of the IEEE Symposium on Security and Privacy, Oakland, California, pp. 305-316, 2010.
[76] R. Zhang, P. Ji, D. Mylaraswamy, M. Srivastava, and S. Zahedi, "Cooperative sensor anomaly detection using global information," *Tsinghua Sci. Technol.*, vol. 18, no. 3, pp. 209–219, 2013.
[77] D. E. Difallah, P. Cudre-Mauroux, and S. A. McKenna, "Scalable Anomaly Detection for Smart City Infrastructure Networks," *IEEE Internet Comput.*, vol. 17, no. 6, pp. 39–47, 2013.
[78] W. Khreich, S. S. Murtaza, A. Hamou-lhadj, and C. Talhi, "Combining Heterogeneous Anomaly Detectors for Improved Software Security," *J. Syst. Softw.*, 2017.
[79] C. Zhou *et al.*, "Design and analysis of multimodel-based anomaly intrusion detection systems in industrial process automation," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 45, no. 10, pp. 1345–1360, 2015.
[80] A. Juvonen, T. Sipola, and T. Hämäläinen, "t," *Comput. Networks*, vol. 91, pp. 46–56, 2015.
[81] A. Satoh, Y. Nakamura, and T. Ikenaga, "A flow-based detection method for stealthy dictionary attacks against Secure Shell," *J. Inf. Secur. Appl.*, vol. 21, pp. 31–41, 2015.
[82] V. Jyothsna and V. V. Rama Prasad, "FCAAIS: Anomaly based network intrusion detection through feature correlation analysis and association impact scale," *ICT Express*, vol. 2, no. 3, pp. 103–116, 2016.
[83] G. Fernandes, J. J. P. C. Rodrigues, and M. L. Proença, "Autonomous profile-based anomaly detection system using principal component analysis and flow analysis," *Appl. Soft Comput.*, vol. 34, pp. 513–525, 2015.