sensors

MDPI

Review

Cyber-Attack Prediction Based on Network Intrusion Detection Systems for Alert Correlation Techniques: A Survey

Hashim Albasheer [1,2], Maheyzah Md Siraj [1,*], Azath Mubarakali [2], Omer Elsier Tayfour [2], Sayeed Salih [3], Mosab Hamdan [4], Suleman Khan [5], Anazida Zainal [1] and Sameer Kamarudeen [2]

1 School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Skudai Johor 81310, Malaysia; hthashim2@graduate.utm.my (H.A.); anazida@utm.my (A.Z.)
2 College of Computer Science, King Khalid University, Abha 61421, Saudi Arabia; aabdurrahman@kku.edu.sa (A.M.); oalser@kku.edu.sa (O.E.T.); sameer@kku.edu.sa (S.K.)
3 College of Computer and Information Sciences, King Saud University, Riyadh 11461, Saudi Arabia; salih.sayd@gmail.com
4 Department of Computer Science, University of São Paulo, São Paulo 13566-590, Brazil; mosab.hamdan@ieee.org
5 School of Psychology and Computer Science, University of Central Lancashire, Preston PR1 2HE, UK; skhan92@uclan.ac.uk
* Correspondence: maheyzah@utm.my

Abstract: Network Intrusion Detection Systems (NIDS) are designed to safeguard the security needs of enterprise networks against cyber-attacks. However, NIDS networks suffer from several limitations, such as generating a high volume of low-quality alerts. Moreover, 99% of the alerts produced by NIDSs are false positives. As well, the prediction of future actions of an attacker is one of the most important goals here. The study has reviewed the state-of-the-art cyber-attack prediction based on NIDS Intrusion Alert, its models, and limitations. The taxonomy of intrusion alert correlation (AC) is introduced, which includes similarity-based, statistical-based, knowledge-based, and hybrid-based approaches. Moreover, the classification of alert correlation components was also introduced. Alert Correlation Datasets and future research directions are highlighted. The AC receives raw alerts to identify the association between different alerts, linking each alert to its related contextual information and predicting a forthcoming alert/attack. It provides a timely, concise, and high-level view of the network security situation. This review can serve as a benchmark for researchers and industries for Network Intrusion Detection Systems' future progress and development.

Keywords: intrusion detection; alerts correlation; attacks prediction; machine learning

Citation: Albasheer, H.; Md Siraj, M.; Mubarakali, A.; Elsier Tayfour, O.; Salih, S.; Hamdan, M.; Khan, S.; Zainal, A.; Kamarudeen, S. Cyber-Attack Prediction Based on Network Intrusion Detection Systems for Alert Correlation Techniques: A Survey. *Sensors* **2022**, *22*, 1494. https://doi.org/10.3390/s22041494

Academic Editor: Luis Velasco

Received: 22 December 2021
Accepted: 29 January 2022
Published: 15 February 2022

1. Introduction

Network Intrusion Detection Systems (NIDs) are rapidly becoming ubiquitous due to sophisticated risk associated with network attacks. The Network Intrusion Detection Systems (NIDSs) are designed to safeguard the security needs of enterprise networks against cyber-attacks. NIDS networks suffer from several limitations, such as the generation of a high volume of low-quality alerts. Moreover, 99% of the alerts, produced by NIDSs, are false positives [1,2]. The Alert Correlation (AC), which is known as IDS post-processing, has been proposed to overcome these limitations. The Intrusion Alert Prediction can assist early caution and prevention to avoid the attacks from escalating and damaging the network as a proactive approach [3].

An intrusion can be characterized as a framework transgression of the security policy that refers to the security components, which are committed to identify infringement of framework security. Intrusive activities are not the same as from normal system activities and the abnormal system activities. Intrusion Detection Systems (IDSs) do not replace other security strategies, for example, verification and access control preventions methods, as

they are integral for the existing security. Distinctive advances approaches are utilized as a part of IDS, including data mining [4], machine learning [5], hidden Markov models [6], honeypot [7], genetic algorithms [8], and fuzzy [9], deep learning [10]. In addition, different types of IDSs have additionally been created, which incorporate the Network Intrusion Detection Systems (NIDS), Host-Based IDS, Stack-Based IDS, Protocol-Based IDS (PIDS), and Graph-Based IDS [11,12]. The technique, like packet sniffing, is used by NIDS, which helps in examining the collected network information. In addition, it tries to find access to a computer network that is unauthorized. A commonplace NIDS incorporates various sensors to observe packet traffic, administration capacities, and at least one administration reassures for human interface.

Table 1 shows the comparison of proposed survey with the existing survey articles, such as [13–16] focus on alert correlation techniques or a mapping among framework techniques and components. No articles comprehensively reviewed cyber-attack prediction based on intrusion alert correlation techniques, considering the intrusion alert dataset. In addition, the development of alert correlation systems has been such that several different systems have been proposed in the meantime, and so there is a need for an update. The taxonomy of alert correlation approaches and components is presented in this paper. In view of the discussion on prior surveys, this article focuses on the following:

- State of the art intrusion alert prediction methods.
- Presenting a classification and comparison of alert correlation approaches.
- State of the art intrusion alert datasets, which are not considered in the existing surveys.

Table 1. Comparison of this survey and existing surveys. (✔: Topic is covered, ☒ the topic is not covered).

Survey (Year)	Alert Correlation Approaches	Alert Prediction Methods	Dataset Issue
Sadoddin (2006) [15]	✔	☒	☒
Mirheidari (2013) [16]	✔	☒	☒
Salah (2013) [14]	✔	✔	☒
Yu Beng (2014) [13]	✔	☒	☒
Proposed Survey	✔	✔	✔

The survey is organized as follows: The state-of-the-art intrusion alert prediction models are presented in Section 2. In Section 3, we present an alert correlation taxonomy. A generalized component in intrusion alert correlation models are presented in Section 4. In Section 5, presented a number of datasets have been used for testing Alert Correlation research. Some discussion and future research direction are reviewed in Section 6. Finally, we conclude the survey in Section 7.

2. State-of-the-Art Intrusion Alert Prediction Models

NIDS technology plays an essential role in protecting communication networks from cybercrime. However, these technologies are not viable in predicting future attacks. It produces much of alerts when attack activities/intrusions have effectively occurred. A proactive approach is to predict and lead conceivable attacks to prevent damage. However, network intrusion prediction is still an active investigation.

This section presents current methodologies for prediction algorithms, which considered as applicable to predicting intrusion alerts. Their main advantages and disadvantages are summarized in Table 2.

2.1. Plan Library

The plan is utilized for predicting an attacker's behavior [17,18]. defined how a plan library of specific attacks to predict an attack plan. The security expert is required to accumulate the plan library manually. However, this can be time-consuming, and it is not always able to reply to a new form of attack variants. The complexity of plan matching

increases because of variation of missing actions in an attack sequence. Therefore, the plan library expected to be updated regularly to ensure it meets the new attack sequence.

2.2. Network Attack Graph

The process of utilization of a network attack in the form of a graph is undertaken to discover the available security vulnerabilities and find all possible attack sequences [19]. The network attack graph is made of correlating intrusion alerts that are in accordance with source and destination Internet Protocol (IP) addresses. The anticipated next alert is determined through predictability scores resulting after the attack graph. It provides the graphical stream relating to the attack sequence to the network administrator. In order to enhance the prediction, some alerts need to be removed manually because the generated graph includes low probable alerts in attack sequence. However, the method cannot discover out-of-sequence attacks.

2.3. Sequence Pattern Mining

The technique, known as sequence pattern mining, decreases the exertion to advance pattern rules [20]. A historic attack sequence considered to be vulnerable several recent attack strategies, while making sure it utilizes the resultant database. The authors of [21] observed that many of the attacks are normally completed within a specific time span and put forward in an incremental mining algorithm to differentiate sequential attack patterns over the separated time window. The database is rationalized within a smaller period afterward the arrival of a new form of attack procedure. After the fundamental regulation generation, the performance of subsequent updates is speedier as the amount of new alert sequence then decreases.

2.4. Machine Learning and Data Mining

The machine learning procedures are utilized to learn past behavior of the attackers. Afterward, the knowledge is utilized to anticipate the subsequent stages of a future attack in real time [22,23]. The authors of [24] presented a prediction algorithm, known as Nexat. It comprises three operational stages, which incorporate the information extraction stage, the preparation stage, and the expectation stage. At run time, it utilizes the trained database and weighted probability to anticipate the next alert. The historical database that is large enough is required to find the next fit to the historical data; it cannot predict the new attacks.

2.5. Relational Time Series

Another way to process intrusion alerts is carried out by arranging them in relational time series (RTS) form [25]. The IDS generates certain security alerts in sequential order by time of arrival. The relational percepts are formed by these alerts sequence. Every percept is characterized as $pi = r(c_1, c_2 \ldots c_m)$, where r represents the predicate and $c_{i\in(1\ldots m)}$ are constants that denote objects. For security alerts, r is the ready type/identity and $c_{i\in(1\ldots m)}$ refers to an entity, for example, source or destination IP. The main aim of the technique is to predict the steps of the future attacker, which can be observed as future NIDS alerts.

2.6. Bayesian Network

The Bayesian network-based approach has been proposed for learning an attack strategy to correlate alert and predict conceivable forthcoming attacks in an online system [26,27]. The general architecture of this intrusion alert prediction model includes two segments. The offline attack pattern recognition is utilized to remove attack action patterns automatically, which create correlation rules by examining alert using Bayesian networks. The online alert correlation is associated with alerts in three steps that include alert fusion, attack thread reconstruction, and attack scenario reconstruction.

2.7. Deep Learning

Recently, Deep Learning techniques have made great progress in many domains, not limited to the field of network security (NIDS). These techniques/approaches aim to observe/encode/predict/classify patterns or pattern sequences by learning a good feature representation. Some of the most successful deep learning methods involve artificial neural networks, such as stacked auto-encoder (SAE), convolutional neural networks (CNN), recurrent neural networks (RNN), deep belief networks (DBN), and reinforcement learning.

Studies show that deep learning completely betters than traditional methods in most of areas. The most important advantage of deep learning is replacing handcrafted features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction [28,29].

Works in [3,30,31] proposed a deep learning method for intrusion alert prediction.

Table 2. Summarizes intrusion alert prediction methods.

Methods	Advantages	Disadvantages	References
Deep Learning	Useful in feature learning.	There is a confusion of how to adopt deep learning in alert correlation applications.	[3,30,31]
Bayesian Network	High accuracy for detecting well-known attacks.	Use complex correlation techniques to discover root-case or attack scenario.	[6,26,27]
Relational Time Series	Performing well in alert reduction, clustering, aggregation, and pattern-matching.	Difficult aspects of selecting representation methods when determining a compatible and adequate similarity measure.	[22,25]
Machine Learning and Data Mining	Performed well in static networks, used to represent a well-defined attack scenarios.	Scalability is the big issue.	[5,19,22,23]
Sequence Pattern Mining	These methods can detect known and unknown attack scenarios.	Complex algorithms.	[20,21]
Network Attack Graph	High accuracy for known attacks.	Complex correlation algorithms.	[19,27]
Plan Library	Correlation result is easy to understand. Discovers all attack scenario variants.	Knowledge build by expert, or predefined scenarios. Scalability problem.	[17,18,32]

3. Taxonomy

NIDS post-processing has been studied for over 10 years to overcome the limitation of NIDS, particularly a high volume of low-quality alerts. In Figure 1, we present a taxonomy of existing alert correlation approaches, which can be categorized as similarity-based [33], statistical-based [34], and knowledge-based alert correlation [16,35]. Regardless of these approaches, a hybrid-based approach has also been shown [36,37]. The similarity-based approach focusses on addressing the issue of enhancing the nature of alerts attributes. The statistical-based approaches are managing the issue of perceiving the attack scenario, in view of the statistical or causal relationship between alerts; knowledge-based approaches are dealing with attack definition considering ready importance. Additionally, a hybrid-based approach can exploit the qualities of three correlation approaches.

3.1. Similarity-Based Approach

The similarity-based approach is defined as a measure to find the similarity between two alerts or alert clusters. This approach clusters similar alerts in time to reduce the number of alerts and increase its ability to discover the known attacks. The approach can discriminate between false positive, redundant, and invalid alerts, whereas the types of attacks do not need to be defined. The similarity between alerts can be modeled based on simple rules, hierarchal rules, and automatic generation using machine learning:

1. Similarity-based on Simple Rules—The Simple Rule similarity-based algorithm defines a simple rule to measure the similarity between alert attributes, which can be aggregated together [38–40].
2. Similarity-based on Hierarchal Rules—The Hierarchal Rule similarity-based algorithm defined similarity measures in a generalized hierarchical concept. It has been defined as a similarity that works on the basis of root cause analysis to detect root causes

in networks [41,42]. The root of cause of similar alerts are same to the root cause of hierarchy.

3. Similarity-based on Machine Learning—The similarity measures are generated automatically. The pre-requisite for supervised algorithms characterizes the existence of clustered alerts, which is set by the machine-learning algorithms. Algorithms, without this requirement (unsupervised), require training to measure the similarity between alerts. A neural network used to perform clustering based on alert reconstruction error [5,27,43,44]. It is a one-step process of clustering and decision-making based on cluster statistics. The online learning capability applies false and true alert patterns on the basis of labeled data.

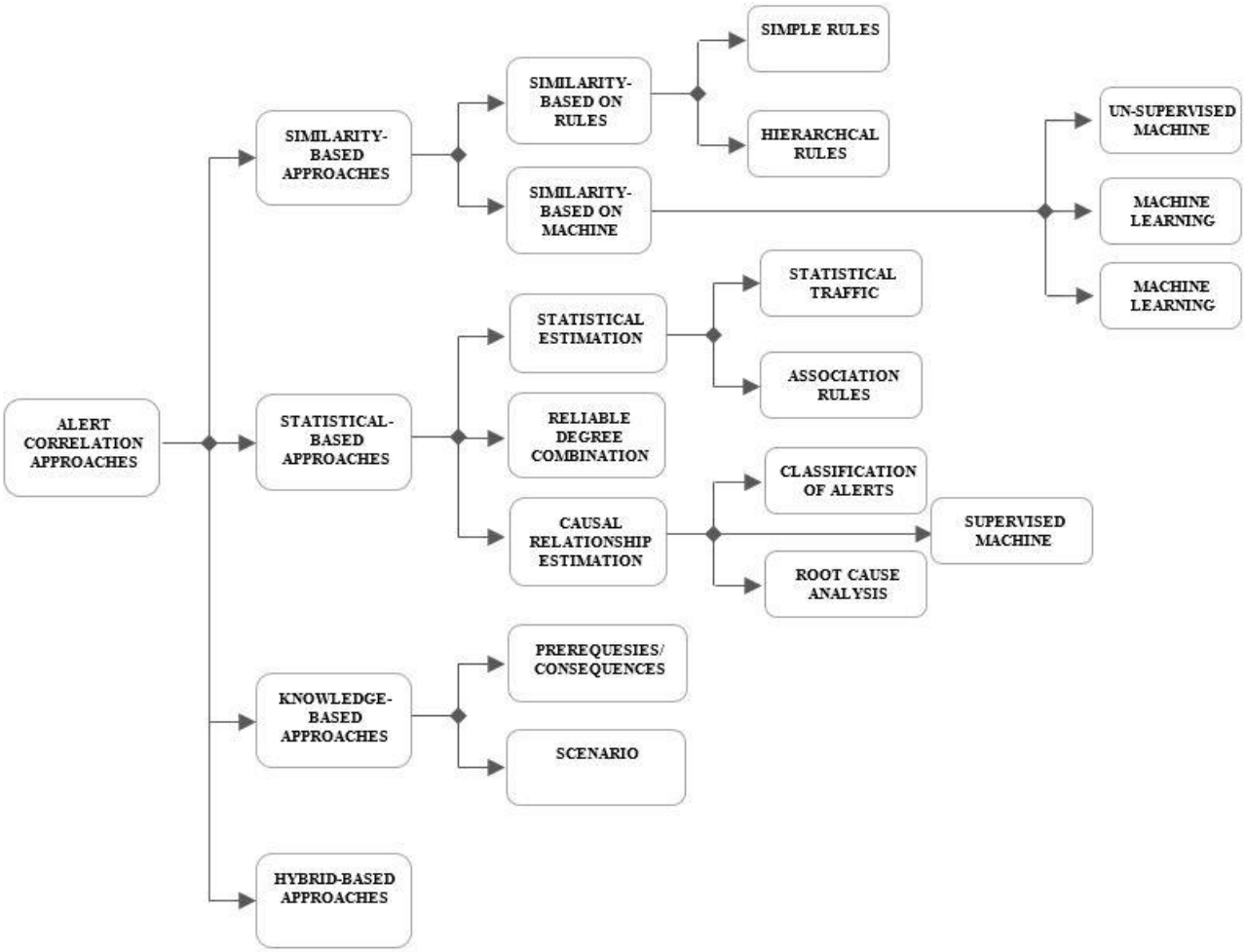

Figure 1. Taxonomy of alert correlation approaches.

3.2. Knowledge-Based Approach

Existing works of this approaches are based on the knowledgebase of attack definitions, which has been divided into two components:

1. Scenario—The principle utilization of scenario-based or pre-defined scenario approaches is to predict the multi-step attacks, as a real attack consists of a sequence of steps [27,45]. An attack scenario is defined by its relating steps or stages, which are required to succeed. Different attack scenario modeling languages are presented, such as [19,27,45–47], However the main idea in all of them is to define the stages and preconditions of an attack, as well as its goals based on the rules stored in the knowledge base, which specifies that the stages of the attack are necessary to achieve success. The knowledge rules are constructs either by machine learning or manually by experts.

However, the correlation rules in machine learning approaches are obtained using the training stage and labeled data.

2. *Prerequisites/Consequences*—These methodologies observe and control implications of alerts and existing knowledge in the network and then detect/predict the security event. If some post-conditions of the first alert match some pre-conditions of the next alert, then the two alerts can be logically correlated. However, a complete scenario description is not required here. One of the first algorithms to use background knowledge has been proposed by [48]. They used provides/requires a model for describing causal relationships between alerts using JIGSAW language. However, it can only detect known attacks. Attack scenarios are modeled in terms of concept and capabilities. Concepts are an abstraction of attacks, and functionality is a mandatory and provided condition associated with each attack concept. The correlation task runs when a match is found between the conditions of two temporally ordered alerts. Therefore, each alert received is modeled into a concept with the relevant required and provided capabilities. It does not represent the attack scenario as a set of states, but as a set of concepts and capabilities.

Several efforts based on this model have been proposed, and they use different definitions and epistemological representations [22,49,50].

3.3. Statistical-Based Approach

The basic idea of these approaches is that relevant attacks have similar statistical features, and a proper classification can be found by detecting these similarities. These approaches do not need any context knowledge. They store the causal relationship between different events and use previous statistical analysis of data to analyze the frequency of occurrence during system training, and then generate attack steps. This knowledge is used for correlating different attack steps after learning these relationships and being confirmed by the supervisor.

1. Statistical traffic estimation—the main application of this estimation is to detect alerts, which are regularly repeated to find their repetition pattern. The patterns of occurred alerts are recognized, and the repetition pattern is derived based on the statistical data of each alert and detects the dissimilarity with these patterns in the future [51,52].
2. Causal relation estimation—it estimates causal relationships between alerts, which predicts next alert occurrence and detects attacks. The alert sequence (pattern) is used for detecting false cases. Several works have been introduced to analyze the causal relationships between alerts [50,53,54].
3. Reliability degree combination—the Reliability Degree Combination is responsible for combining reliability with the similar alerts [55,56]. Changing the reliability to alerts is proposed based on equivalent alert repetitions. This combination aims to change the priority or severity of an alert based on the other resources. The presented algorithms require a large amount of labeled previous data for generating the probability models.

3.4. Hybrid-Based Approach

The hybrid-based approach tries to take advantage of each of the three correlational approaches. Work [57,58] proposed a hybrid model, which provides alert correlation based on statistics, similarity, and knowledge correlation approaches. Their main goal is to enhance the detection and prediction of alerts and recognize attack scenarios.

3.5. Comparison of Alert Correlation Approaches

In this subsection, we provide a comparison of existing approaches, which has been outlined in Table 3. The capability metrics for evaluating alert correlation techniques are defined in the following section.

Table 3. A comparison of existing approaches (Capable = ✔, incapable = ☒).

Approach	Pre-Knowledge or Rule	Alert Reduction	Reducing False Alerts	Alert Prioritization	Extract Attack Scenario	Predict Next Alert	Construct and Predict Attack
Similarity-based	✔	✔	✔	☒	✔	☒	☒
Statistical-based	✔	✔	✔	✔	✔	✔	☒
Knowledge-based	✔	☒	☒	✔	✔	☒	☒
Hybrid-based	✔	✔	✔	✔	✔	☒	☒

4. Generalized Components in Intrusion Alert Correlation Models

There is much need for Alert Correlation (AC) to ensure the improvement on the quality of alerts that have been made available by NIDS. The prediction of sophisticated attack scenarios requires the development of an effective, efficient, and accurate alert correlation/prediction model. The AC model consists of several tasks that include normalization, reduction, severity/prioritization, attack detection, and prediction to present a viewpoint of network security situations. Generalized components in intrusion alert correlation/prediction models has been shown in Figure 2. The following sub-sections briefly explain each of them.

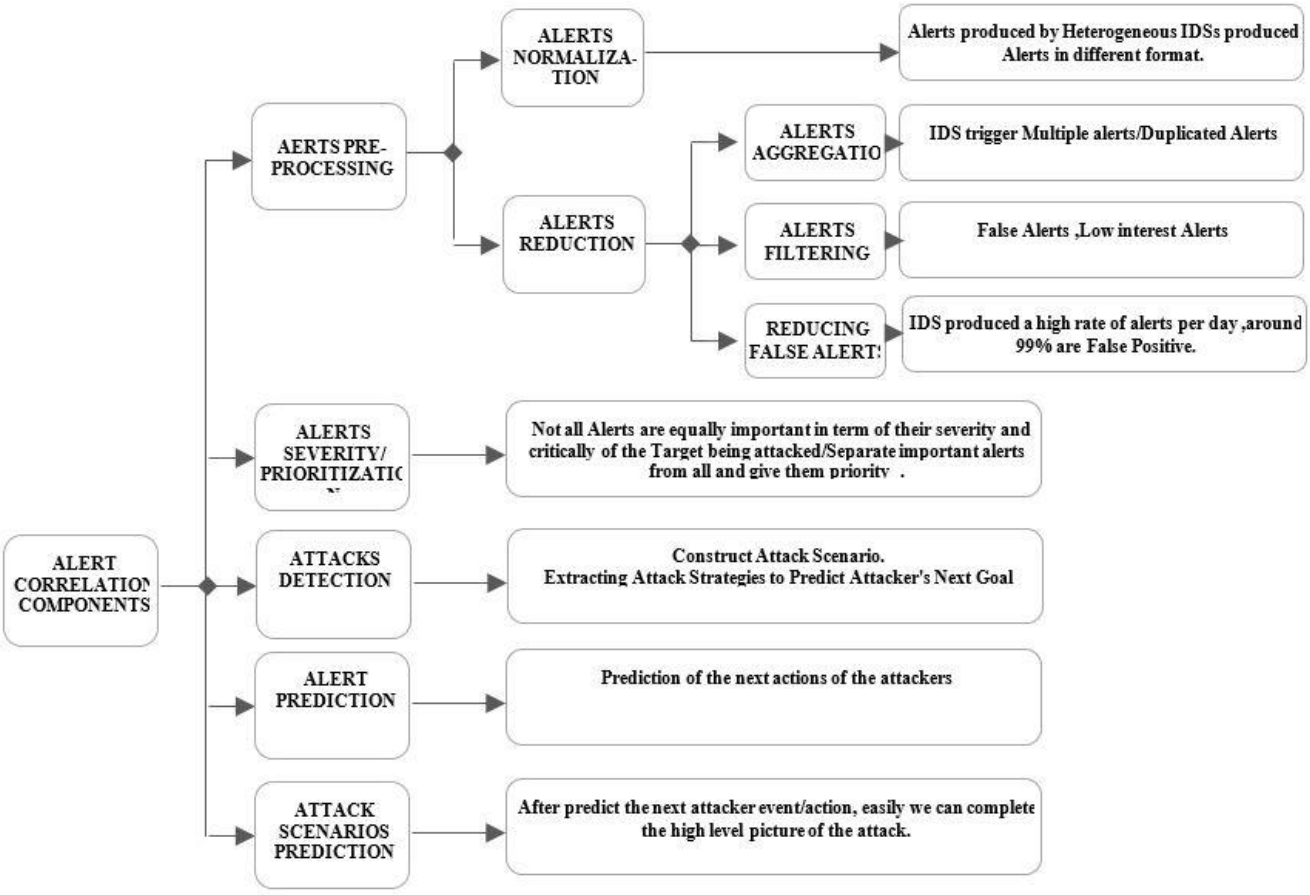

Figure 2. Mapping of alerts correlation components with research problems.

4.1. Alert Normalization/Formatting

Among the preprocessing tasks studied by AC, formatting alerts can be considered as an important initial task. Currently, most organizations implement different types of NIDS (heterogeneous NIDS), so they generate alerts in different data formats. Alert normalization is the process of converting different alert data formats from multiple intrusion sensors into appropriate and acceptable standard formats for other correlation components. The

authors of [59] tended to the issue of formatting and standardizing the unformatted alerts. Their work has motivated to design a format called Intrusion Detection Message Exchange Format (IDMEF), which can be received in a wide range of IDS.

Another alert format based on IDMEF is called IDEA (intrusion detection extensible alert) [60]. IDEA uses the latest JSON format instead of XML and adds alert classification. Figure 3 shows an example of alert in IDEA format.

```
{
    "Format": "IDEA0",
    "ID": "3ad275e3-559a-45c0-8299-6807148ce157",
    "DetectTime": "2019-03-22T10:12:56Z",
    "Category": ["Recon.Scanning"],
    "ConnCount": 633,
    "Description": "Ping scan",
    "Source": [
       {
         "IP4": ["11.12.13.14"],
         "Proto": ["icmp"]
       }
    ],
    "Target": [
       {
         "Proto": ["icmp"],
         "IP4": ["10.10.10.0/24"]
       }
    ],
    "Node": [
       "Name": "cz.cesnet.tarpit",
       "SW": ["LaBrea"],
       "Type": ["Connection", "Tarpit"]
    ]
}
```

Figure 3. Example of an IDEA alert in JSON format.

4.2. Alert Reduction

NIDS can produce thousands of alerts per day. This flow contains repeated/redundant and low interesting alerts. About 99% of the alerts are false positive [61]. To reduce the alerts, we categorized the related works into two groups: aggregation and filtration.

1. Alert aggregation:

Alerts belonging to the same attack may be generated by the same NIDS or different NIDS. They are recognized by a similar IP addresses source and destination and merged with repeated/redundant alerts. This case increases the dimensionality of created alerts. In fact, redundant alerts are usually false positives.

Aggregation is utilized to group redundant/repeated alerts and represent to as a hyper-alert or single meta-alert. Alerts are clustered (or aggregated) based on attributes/features similarities [59] using predefined rules operator/function. A new cluster is created to represent aggregate clusters, and a new global alert is created to represent the new cluster [62,63]. Clustering is practical because a similarity search between alerts can be performed automatically on many alerts

Another effort attempts to reduce unnecessary alerts by validating alerts with vulnerabilities assessment and then aggregate alerts [64,65]. Kavousi et al. [26] aggregate alerts that have been caused for same attack stage. Zomlot [66] used a support vector matrix (SVM) to reduce alerts, and the non-interested alerts are not removed; they claim it will be helpful to link true attacks. However, reducing redundant alerts does not truly eliminate false positives. Therefore, the next problem is to filter (remove) and verify false positive alerts.

2. Alert filtration—technically, false positives alerts are normally caused by runtime limitations, specificity of detection signature, and environment dependency.

False positives alerts need to be verified and filtered (removed) to ensure effective alert correlation. In [67], a clustering algorithm based on the XML distance measure used to aggregate alerts in clusters is implemented. Each XML file represents a sequence of alerts for a network session. In [68], the authors verify and filter false positive alerts using machine learning, whereas [69] have adopted a fuzzy-based classifier to generate a fuzzy rules classifier to differentiate between alerts as true or false positives using knowledge, while [70] have used genetic based fuzzy classifier to generate the classification rules.

4.3. Alert Severity/Prioritization

Not all intrusion alerts generated by NIDS are equally important to the severity and target of the attack [71], so some critical alerts need to be separated and prioritized from the rest of the alert set.

Work by [72] categorized alerts severity into three types: high, medium, and low, the type of severity based on NIDS signature files; while Alsubhi [55] proposed a technique based on fuzzy logic for scoring and prioritizing alerts. Their method evaluates alerts according to several criteria and used fuzzy logic inference mechanism prioritize alerts.

4.4. Attack Prediction and Construction

As is known, the prediction of future actions of an attacker is one of the most important goals here. NIDS technologies play a vital role in protecting communication networks from cybercrime. However, these techniques are not very effective in predicting future action of the attacker. Even worse, it generates alerts when attack/intrusion activities occur. A predictive approach is to anticipate and implement potential attacks to prevent damage. Accordingly, the future attack step can be derived after detecting a few steps of the attack in progress. However, anticipating the next actions of attackers is a difficult and important task. Predicting an attack future action can help intrusion prevention systems to function properly. Details about intrusion alert prediction are given in Section 2.

5. Intrusion Alerts Datasets

A major challenge for the research community is to obtain suitable datasets for evaluating various research designs in the domains of NIDS [73,74]. A complex and new case of intrusions, new bugs, new security issues, and vulnerabilities are growing every day. The evaluation process of alert correlation research shares the same challenges of NIDS research, as AC is a complementary system to NIDS. One of these challenges is the unavailability of enough datasets. However, a number of datasets have been used for testing Alert Correlation researches, such as KDD 99 [75], DARPA 2000 [76,77], Koyto2006+ [78,79], and Defcon [77], and, recently, alert sharing platform (SABU) [80] datasets. There are verities of security issue concerning datasets for evaluating alert correlation. The authors of [73] listed some issue regarding the use of datasets, which include data privacy, obtaining approval from the data owner, the scope of evaluative datasets, a documentation problem, understanding datasets, data labeling, availability of evaluative datasets, and discrepancies in evaluative datasets.

1. Data privacy—realistic data are not allowed to be shared among users due to security policies, sensitivities of realistic data, lack of trust, and risk of disclosing digital information.
2. Getting approval from data owner—some data owners require approval to obtain access to the dataset, and this approval process is frequently delayed.
3. The scope of evaluative datasets—most publicly available datasets become out of date and unsuitable for making strong scientific claims because of variability in network segments.
4. Different research objectives—the aims and methods of studies are factors that influence the choice of datasets.
5. Data labeling—some available datasets are manually labeled datasets, while some are packet traces without identifiers, which influences the validity of the datasets.
6. Source of evaluating datasets—there are three conceivable approaches to making more reasonable datasets for assessing the alert correlation researches to limit the

effect of challenges. The datasets can be extracted from any of the general population (for example, DARPA), local area systems, and genuine systems (Koyto2006+), as shown in Figure 4.

7. Alert sharing platform (SABU) dataset: This dataset was recently published as the first intrusion alert dataset; it contains intrusion alerts collected from heterogeneous intrusion detection systems in different organizations [80]. Alerts are formatted in the (IDEA) format and classified using the eCSIRT.net taxonomy. The IDEA format, as shown in Figure 4, is a descriptive data model that uses a modern JSON structure. Nearly 12 million alerts have been collected in three organization from different data sources (34 IDS, honeypots, and other data sources). This dataset has been used in significant research to test alert correlation and prediction models [3,30,81,82].
8. DARPA 2000 datasets: The DARPA-2000 dataset is among the notable datasets, which are utilized to assess alert correlation research. The datasets are made by the Lincoln Lab, Massachusetts Institute of Technology (MIT) USA [76]. This dataset comprises two multi-stage scenarios, called LLDDOS 1.0 (Scenario One) and LLDDOS 2.0.2 (Scenario Two). Both attack scenarios contain a progression of attack activities to start distributed denial of service assaults (DDoS). A significant number of investigations have been made into this dataset to test the alerts correlation frameworks [20,26,44,57].
9. KDD: The KDD 99 Dataset is created on basis of the data captured in DARPA'98 IDS evaluation program and used for the evaluation of Alerts Correlation researches [83]. KDD 99 network traffic consists of 41 features and is labeled as either an attack activity or normal activity. The attacks are categories as user to root attack (U2R), denial of service attack (DoS), remote to local attack (R2L), and probing attack.
10. Datasets from Real Networks: The second approach is to deal with more sensible datasets from genuine networks, which are a great source of real events. Such datasets, as a rule, require endorsement from administrators of the networks.
11. KYOTO2006+: The Kyoto 2006+ is datasets that have been used for evaluating Alerts Correlation researches [79]. These datasets were created from real network traffic from Nov 2006 to Aug 2009, without human intervention at the University of Kyoto, Japan. They consist of 24 statistical features where 14 conventional features are extracted based on KDD99 datasets and 10 additional features. These feature assist in investigating the happenings within the network.

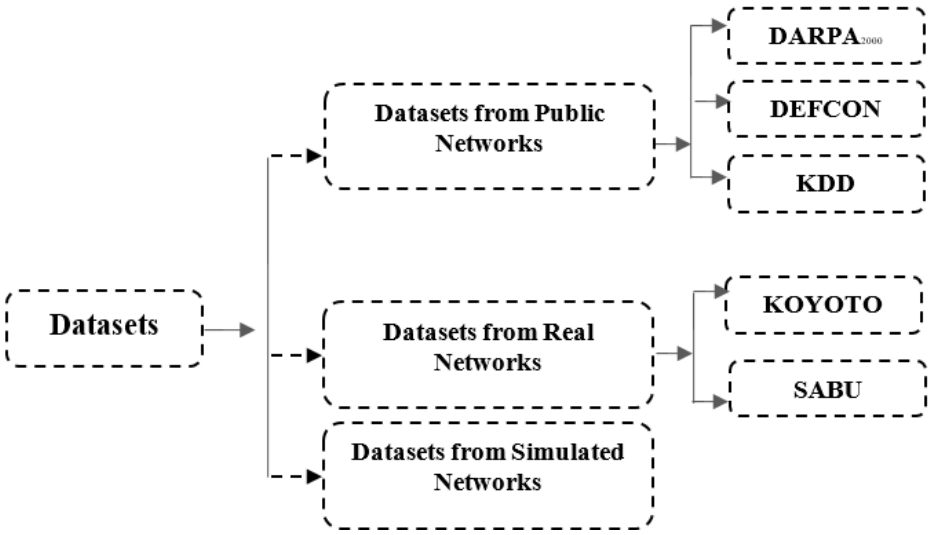

Figure 4. The source of datasets.

Datasets from Simulated Networks: The third approach is to set up local area networks to mimic a few attacks in line. There are several researchers in this research domain that

applied this approach [84]. The comparison of different types of attacks and methods is shown in Table 4.

Table 4. Comparison of different types of attacks and methods.

Used Algorithms	Objective(s)	Accuracy (%)	Datasets	Type of Attacks	References
K-means + KNN	IDS	93.55	KDD-Cup'99	DoS, User to Root (U2R), Remote to Local (R2L) and Probing Attack	[85]
SVM + KNN + PSO	IDS	88.44	KDD-Cup 99	DoS, User to Root (U2R), Remote to Local (R2L) and Probing Attack	[86]
HC + SVM	IDS	95.72, 69.8	KDD Cup 1999, 1998 DARPA	DoS, User to Root (U2R), Remote to Local (R2L) and Probing Attack	[87,88]
RF + AODE	IDS	90.51	Kyoto	Various attacks against honeypots	[89]
FL + ES	Network forensics	91.5	DARPA 2000	DDoS, DARPA attacks	[90]
FL + GA	IDS	94.6	DARPA-KDD99	DoS, User to Root (U2R), Remote to Local (R2L) and Probing Attack	[91]

SVM: support vector machine; KNN: K-nearest neighbors; PSO: particle swarm optimization; AODE: average one-dependence estimator; HC: hierarchical clustering; AODE: average one-dependence estimator; RF: random forest; FL: fuzzy logic; ES: expert system; GA: genetic algorithm; IDS: intrusion detection system.

6. Discussion and Future Research Direction

Predicting the next activities of the attacker is an imperative and difficult task [13,20,81,82]. Prediction encourages intrusion frameworks to respond appropriately before the network is compromised and gives the chances to overcome the benefits of the attacker. However, existing works on conquering the limitations of NIDSs (in term of delivering a high volume of low-quality alarms) do not manage alert correlation [40] and attacker prediction [24] as a proactive approach. The present expectation methods exhibited on intrusion alerts prediction have constrained capacity to anticipate attack situations that have never been experienced. Various researchers exploring AC have recommended that analyzing intrusion alarms created by NIDS gives a brief and a high-level figure when attack exercises have occurred in a timely manner [19,24,26,38,49,71].

AC is a complex multi-organized change process, and the majority of the existing procedure experiences complex relationship rules definition [26] that constrains the capacities of recognizing new attack situations. The capacities are restricted because of hard-coded area information, which is precisely predefined, depending on the learning of human expert. Prediction in known conditions can be unraveled utilizing information base framework or Bayesian derivation with a foreordained structure. The probabilistic and Markov methodologies will function admirably when the environment is obscure yet exceptionally tedious. At the point when nature is obscure yet stationary with restricted social and question assortment, Bayesian and Markov will take some opportunity to learn. It is significant and reasonable to propose an Intrusion Alerts Correlation Prediction Model, which incorporates an expectation of the assailant's subsequent stage and the aggressor's new situation. Intrusion detection alert can be expressed as a relational time series (RTS). The NIDS creates cautions as malicious exercises that arrive in time succession.

7. Conclusions

The trends and solutions have been highlighted according to the issues and advantages of intrusion alert prediction. The growing interest in the internet, communication networks, telecommunications alternative, and cybercrimes issues have increased the necessity of robust predictive mechanisms. Advanced and modern attacks are faced that exploit the emerging services, as simple attacks are no longer used. The Network Intrusion Detection Systems (NIDSs) are useful if the detected results are reviewed and analyzed to derive security for the current system. NIDS operations generate a high volume of low-quality alerts. However, 99% of the alerts produced by NIDSs, are a false positive. The study has presented an overview of past and recent works in the field of NIDSs and alert correlation. Several NIDS approaches have been discussed that provide taxonomy

to show the types of NIDS. The limitation of NIDSs that motivate the alert correlation research has been summarized. Alert-correlation-related works have also been discussed. Different approaches including similarity-based, knowledge-based, statistical-based, and hybrid-based approaches have also been discussed. The study concludes that the existing intrusion alert prediction systems have limited prediction capabilities.

Author Contributions: Conceptualization, H.A., M.M.S. and A.Z.; methodology, H.A.; data curation, H.A., S.S., O.E.T., M.H., S.K. (Sameer Kamarudeen) and S.K. (Suleman Khan); validation, H.A.; formal analysis, H.A., S.S. and M.H.; writing—original draft preparation, H.A. and M.M.S.; writing—review and editing, S.S., S.K. (Suleman Khan), A.M, O.E.T., M.H., S.K. (Sameer Kamarudeen) and H.A.; project administration, H.A.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors also extend their appreciation to the Deanship of Scientific Research at King Khalid University for supporting this work through Research Groups, under grant number (R.G.P. 2/164/42).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bhatti, D.G.; Virparia, P.V. Soft Computing-Based Intrusion Detection System with Reduced False Positive Rate. In *Design and Analysis of Security Protocol for Communication*; Wiley Online Library: Hoboken, NJ, USA, 2020; pp. 109–139.
2. Thudumu, S.; Branch, P.; Jin, J.; Singh, J.J. A comprehensive survey of anomaly detection techniques for high dimensional big data. *J. Big Data* **2020**, *7*, 1–30. [CrossRef]
3. Ansari, M.S.; Bartos, V.; Lee, B. Shallow and Deep Learning Approaches for Network Intrusion Alert Prediction. *Procedia Comput. Sci.* **2020**, *171*, 644–653. [CrossRef]
4. Puthran, S.; Shah, K. Intrusion detection using data mining. *Int. J. Comput. Intell. Stud.* **2020**, *9*, 292–306. [CrossRef]
5. Ayub, M.A.; Johnson, W.A.; Talbert, D.A.; Siraj, A. Model evasion attack on intrusion detection systems using adversarial machine learning. In Proceedings of the 2020 54th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 18–20 March 2020; pp. 1–6.
6. Kalnoor, G.; Gowri Shankar, S. A Model-Based System for Intrusion Detection Using Novel Technique-Hidden Markov Bayesian in Wireless Sensor Network. In *Information and Communication Technology for Competitive Strategies (ICTCS 2020)*; Springer: Singapore, 2022; pp. 43–53.
7. Negi, P.S.; Garg, A.; Lal, R. Intrusion detection and prevention using honeypot network for cloud security. In Proceedings of the 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 29–31 January 2020; pp. 129–132.
8. Jain, V.; Agrawal, M. Applying genetic algorithm in intrusion detection system of iot applications. In Proceedings of the 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, 15–17 June 2020; pp. 284–287.
9. Masdari, M.; Khezri, H. A survey and taxonomy of the fuzzy signature-based intrusion detection systems. *Appl. Soft Comput.* **2020**, *92*, 106301. [CrossRef]
10. Gamage, S.; Samarabandu, J. Deep learning methods in network intrusion detection: A survey and an objective comparison. *J. Netw. Comput. Appl.* **2020**, *169*, 102767. [CrossRef]
11. Chou, D.; Jiang, M. A Survey on Data-driven Network Intrusion Detection. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–36. [CrossRef]
12. Ahmed, N. *Intrusion Detection System: A Survey and Taxonomy*; HAL Open Science: Lyon, France, 2021.
13. Yu Beng, L.; Ramadass, S.; Manickam, S.; Soo Fun, T. A survey of intrusion alert correlation and its design considerations. *IETE Tech. Rev.* **2014**, *31*, 233–240. [CrossRef]
14. Salah, S.; Maciá-Fernández, G.; Díaz-Verdejo, J.E. A model-based survey of alert correlation techniques. *Comput. Netw.* **2013**, *57*, 1289–1317. [CrossRef]
15. Sadoddin, R.; Ghorbani, A. Alert correlation survey: Framework and techniques. In Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap between PST Technologies and Business Services, Markham, ON, Canada, 30 October–1 November 2006; pp. 1–10.
16. Mirheidari, S.A.; Arshad, S.; Jalili, R. Alert correlation algorithms: A survey and taxonomy. In Proceedings of the International Symposium on Cyberspace Safety and Security, Zhangjiajie, China, 13–15 November 2013; pp. 183–197.
17. Li, T.; Liu, Y.; Liu, Y.; Xiao, Y.; Nguyen, N.A. Attack plan recognition using hidden Markov and probabilistic inference. *Comput. Secur.* **2020**, *97*, 101974. [CrossRef]
18. Geib, C.W.; Goldman, R.P. Plan recognition in intrusion detection systems. In Proceedings of the Proceedings DARPA Information Survivability Conference and Exposition II. DISCEX'01, Anaheim, CA, USA, 12–14 June 2001; pp. 46–55.
19. Hu, H.; Liu, J.; Zhang, Y.; Liu, Y.; Xu, X.; Huang, J. Attack scenario reconstruction approach using attack graph and alert data mining. *J. Inf. Secur. Appl.* **2020**, *54*, 102522. [CrossRef]

20. Zhang, K.; Luo, S.; Xin, Y.; Zhu, H.; Chen, Y. Online Mining Intrusion Patterns from IDS Alerts. *Appl. Sci.* **2020**, *10*, 2983. [CrossRef]
21. Zhang, A.-F.; Li, Z.-T.; Li, D.; Wang, L. Discovering novel multistage attack patterns in alert streams. In Proceedings of the 2007 International Conference on Networking, Architecture, and Storage (NAS 2007), Guilin, China, 29–31 July 2007; pp. 115–121.
22. Mahdavi, E.; Fanian, A.; Amini, F. A real-time alert correlation method based on code-books for intrusion detection systems. *Comput. Secur.* **2020**, *89*, 101661. [CrossRef]
23. Cheng, Q.; Wu, C.; Zhou, S. Discovering Attack Scenarios via Intrusion Alert Correlation Using Graph Convolutional Networks. *IEEE Commun. Lett.* **2021**, *25*, 1564–1567. [CrossRef]
24. Cipriano, C.; Zand, A.; Houmansadr, A.; Kruegel, C.; Vigna, G. Nexat: A history-based approach to predict attacker actions. In Proceedings of the 27th Annual Computer Security Applications Conference, Orlando, FL, USA, 5–9 December 2011; pp. 383–392.
25. Tan, T.K.; Darken, C.J. Learning and prediction of relational time series. *Comput. Math. Organ. Theory* **2015**, *21*, 210–241. [CrossRef]
26. Kavousi, F.; Akbari, B. A Bayesian network-based approach for learning attack strategies from intrusion alerts. *Secur. Commun. Netw.* **2014**, *7*, 833–853. [CrossRef]
27. Liu, J.; Liu, B.; Zhang, R.; Wang, C. Multi-step Attack Scenarios Mining Based on Neural Network and Bayesian Network Attack Graph. In Proceedings of the International Conference on Artificial Intelligence and Security, New York, NY, USA, 26–28 July 2019; pp. 62–74.
28. Wang, Z. The applications of deep learning on traffic identification. *BlackHat USA* **2015**, *24*, 1–10.
29. Chandra, B.; Sharma, R.K. Exploring autoencoders for unsupervised feature selection. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–16 July 2015; pp. 1–6.
30. Ansari, M.S.; Bartoš, V.; Lee, B. GRU-based deep learning approach for network intrusion alert prediction. *Future Gener. Comput. Syst.* **2022**, *128*, 235–247. [CrossRef]
31. Bartos, V.; Zadnik, M.; Habib, S.M.; Vasilomanolakis, E. Network entity characterization and attack prediction. *Future Gener. Comput. Syst.* **2019**, *97*, 674–686. [CrossRef]
32. Chintabathina, S.; Villacis, J.; Walker, J.J.; Gomez, H.R. Plan recognition in intrusion detection systems using logic programming. In Proceedings of the 2012 IEEE Conference on Technologies for Homeland Security (HST), Waltham, MA, USA, 13–15 November 2012; pp. 609–613.
33. Wu, M.; Moon, Y.B. Alert correlation for detecting cyber-manufacturing attacks and intrusions. *J. Comput. Inf. Sci. Eng.* **2020**, *20*, 011004. [CrossRef]
34. Shin, Y.; Lim, C.; Park, M.; Cho, S.; Han, I.; Oh, H.; Lee, K. Alert correlation using diamond model for cyber threat intelligence. In Proceedings of the European Conference on Cyber Warfare and Security, Coimbra, Portugal, 4–5 July 2019; Academic Conferences International Limited: Oxfordshire, UK, 2019; pp. 444–450.
35. Wang, W.; Jiang, R.; Jia, Y.; Li, A.; Chen, Y. KGBIAC: Knowledge graph based intelligent alert correlation framework. In Proceedings of the International Symposium on Cyberspace Safety and Security, Xi'an, China, 23–25 October 2017; pp. 523–530.
36. Siraj, M.M.; Albasheer, H.H.T.; Din, M.M. Towards predictive real-time multi-sensors intrusion alert correlation framework. *Indian J. Sci. Technol.* **2015**, *8*, 1. [CrossRef]
37. Siraj, M.M.; Maarof, M.A.; Hashim, S.Z.M. Intelligent alert clustering model for network intrusion analysis. *Int. J. Adv. Soft Comput. Appl.* **2009**, *1*, 1–16.
38. Cuppens, F. Managing alerts in a multi-intrusion detection environment. In Proceedings of the Seventeenth Annual Computer Security Applications Conference, New Orleans, LA, USA, 10–14 December 2001; pp. 22–31.
39. Valdes, A.; Skinner, K. Probabilistic alert correlation. In Proceedings of the International Workshop on Recent Advances in Intrusion Detection, Davis, CA, USA, 10–12 October 2001; pp. 54–68.
40. Elshoush, H.T.; Osman, I.M. Intrusion alert correlation framework: An innovative approach. In *IAENG Transactions on Engineering Technologies*; Springer: Dordrecht, The Netherlands, 2013; pp. 405–420.
41. Julisch, K. Clustering intrusion detection alarms to support root cause analysis. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **2003**, *6*, 443–471. [CrossRef]
42. Al-Mamory, S.O.; Zhang, H. Ids alerts correlation using grammar-based approach. *J. Comput. Virol.* **2009**, *5*, 271–282. [CrossRef]
43. Dain, O.; Cunningham, R.K. Fusing a heterogeneous alert stream into scenarios. In *Applications of Data Mining in Computer Security*; Springer: Boston, MA, USA, 2002; pp. 103–122.
44. Smith, R.; Japkowicz, N.; Dondo, M.; Mason, P. Using unsupervised learning for network alert correlation. In Proceedings of the Conference of the Canadian Society for Computational Studies of Intelligence, Windsor, ON, Canada, 28–30 May 2008; pp. 308–319.
45. Cheung, S.; Lindqvist, U.; Fong, M.W. Modeling multistep cyber attacks for scenario recognition. In Proceedings of the Proceedings DARPA Information Survivability Conference and Exposition, Washington, DC, USA, 22–24 April 2003; pp. 284–292.
46. Kovačević, I.; Groš, S.; Slovenec, K. Systematic Review and Quantitative Comparison of Cyberattack Scenario Detection and Projection. *Electronics* **2020**, *9*, 1722. [CrossRef]
47. Zali, Z.; Hashemi, M.R.; Saidi, H. Real-time attack scenario detection via intrusion detection alert correlation. In Proceedings of the 2012 9th International ISC Conference on Information Security and Cryptology, Tabriz, Iran, 13–14 September 2012; pp. 95–102.
48. Templeton, S.J.; Levitt, K. A requires/provides model for computer attacks. In Proceedings of the 2000 Workshop on New Security Paradigms, Ballycotton, Ireland, 18–21 September 2000; pp. 31–38.

49. Ning, P.; Cui, Y.; Reeves, D.S.; Xu, D. Techniques and tools for analyzing intrusion alerts. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **2004**, *7*, 274–318. [CrossRef]
50. Zhu, B.; Ghorbani, A.A. Alert correlation for extracting attack strategies. *Int. J. Netw. Secur.* **2006**, *3*, 244–258.
51. Viinikka, J.; Debar, H.; Me, L.; Lehikoinen, A.; Tarvainen, M. Processing intrusion detection alert aggregates with time series modeling. *Inf. Fusion* **2009**, *10*, 312–324. [CrossRef]
52. Melo, R.V.; de Macedo, D.D.; Kreutz, D.; De Benedictis, A.; Fiorenza, M.M. ISM-AC: An immune security model based on alert correlation and software-defined networking. *Int. J. Inf. Secur.* **2021**, 1–15. [CrossRef]
53. Ning, P.; Xu, D.; Healey, C.G.; Amant, R.S. Building Attack Scenarios through Integration of Complementary Alert Correlation Method. In Proceedings of the NDSS, San Diego, CA, USA, 5 February 2004; pp. 97–111.
54. Yang, J.; Zhang, Q.; Jiang, X.; Chen, S.; Yang, F. Poirot: Causal Correlation Aided Semantic Analysis for Advanced Persistent Threat Detection. *IEEE Trans. Dependable Secur. Comput.* **2021**. [CrossRef]
55. Alsubhi, K.; Al-Shaer, E.; Boutaba, R. Alert prioritization in intrusion detection systems. In Proceedings of the NOMS 2008—2008 IEEE Network Operations and Management Symposium, Salvador, Brazil, 7–11 April 2008; pp. 33–40.
56. Asharf, J.; Moustafa, N.; Khurshid, H.; Debie, E.; Haider, W.; Wahab, A. A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. *Electronics* **2020**, *9*, 1177. [CrossRef]
57. Siraj, M.M.; Maarof, M.A.; Hashim, S.Z.M. Classifying security alerts from multiple sensors based on hybrid approach. In Proceedings of the International Conference on Informatics & Applications, Kuala Terengganu, Malaysia, 3–5 June 2012; pp. 174–181.
58. Chadza, T.; Kyriakopoulos, K.G.; Lambotharan, S. Analysis of hidden Markov model learning algorithms for the detection and prediction of multi-stage network attacks. *Future Gener. Comput. Syst.* **2020**, *108*, 636–649. [CrossRef]
59. Debar, H.; Wespi, A. Aggregation and correlation of intrusion-detection alerts. In Proceedings of the International Workshop on Recent Advances in Intrusion Detection, Davis, CA, USA, 10–12 October 2001; pp. 85–103.
60. Kácha, P. Idea: Security event taxonomy mapping. In Proceedings of the 18th International Conference on Circuits, Systems, Communications and Computers, Santorini Island, Greece, 17–21 July 2014.
61. Roschke, S.; Cheng, F.; Meinel, C. A new alert correlation algorithm based on attack graph. In *Computational Intelligence in Security for Information Systems*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 58–67.
62. Man, D.; Yang, W.; Wang, W.; Xuan, S. An alert aggregation algorithm based on iterative self-organization. *Procedia Eng.* **2012**, *29*, 3033–3038. [CrossRef]
63. Spathoulas, G.P.; Katsikas, S.K. Enhancing IDS performance through comprehensive alert post-processing. *Comput. Secur.* **2013**, *37*, 176–196. [CrossRef]
64. Nguyen, T.H.; Luo, J.; Njogu, H.W. An efficient approach to reduce alerts generated by multiple IDS products. *Int. J. Netw. Manag.* **2014**, *24*, 153–180. [CrossRef]
65. Sadighian, A.; Fernandez, J.M.; Lemay, A.; Zargar, S.T. Ontids: A highly flexible context-aware and ontology-based alert correlation framework. In Proceedings of the International Symposium on Foundations and Practice of Security, La Rochelle, France, 21–22 October 2013; pp. 161–177.
66. Zomlot, L.; Chandran, S.; Caragea, D.; Ou, X. Aiding intrusion analysis using machine learning. In Proceedings of the 2013 12th International Conference on Machine Learning and Applications, Miami, FL, USA, 4–7 December 2013; pp. 40–47.
67. Long, J.; Schwartz, D.; Stoecklin, S. Distinguishing false from true alerts in snort by data mining patterns of alerts. In Proceedings of the Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security, Orlando, FL, USA, 18 April 2006; International Society for Optics and Photonics: Bellingham, WA, USA, 2006; p. 62410B.
68. Maggi, F.; Zanero, S. On the use of different statistical tests for alert correlation–short paper. In Proceedings of the International Workshop on Recent Advances in Intrusion Detection, Gold Goast, Australia, 5–7 September 2007; pp. 167–177.
69. Huang, C.-J.; Hu, K.-W.; Cheng, H.; Chang, T.-K.; Luo, Y.-C.; Lien, Y.-J. Application of type-2 fuzzy logic to rule-based intrusion alert correlation detection. *Int. J. Innov. Comput. Inf. Control* **2012**, *8*, 2865–2874.
70. Hassan, M.M.M.; Baruah, H.K. Fuzzy classifier for ids alerts using genetic algorithm. *Int. J. Res.* **2014**, *2*, 228–238.
71. Ghorbani, A.A.; Lu, W.; Tavallaee, M. Alert management and correlation. In *Network Intrusion Detection and Prevention*; Springer: Boston, MA, USA, 2010; pp. 129–160.
72. Siraj, M.M.; Maarof, M.A.; Hashim, S.Z.M. A hybrid intelligent approach for automated alert clustering and filtering in intrusion alert analysis. *Int. J. Comput. Theory Eng.* **2009**, *1*, 539. [CrossRef]
73. Nehinbe, J.O. A critical evaluation of datasets for investigating IDSs and IPSs researches. In Proceedings of the 2011 IEEE 10th International Conference on Cybernetic Intelligent Systems (CIS), London, UK, 1–2 September 2011; pp. 92–97.
74. Thakkar, A.; Lohiya, R. A review of the advancement in intrusion detection datasets. *Procedia Comput. Sci.* **2020**, *167*, 636–645. [CrossRef]
75. UCI KDD University of California. KDD Cup 99 Dataset. 1999. Available online: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html (accessed on 15 October 2021).
76. MIT. MIT Lincoln Laboratory, D.I.D.E. DARPA Intrusion Detection. 2000. Available online: https://archive.ll.mit.edu/ideval/data/2000data.html (accessed on 15 October 2021).
77. Lippmann, R.; Haines, J.W.; Fried, D.J.; Korba, J.; Das, K. The 1999 DARPA off-line intrusion detection evaluation. *Comput. Netw.* **2000**, *34*, 579–595. [CrossRef]

78. Kyoto University's. Traffic Data from Kyoto University's Honeypots. 2006. Available online: https://www.takakura.com/Kyoto_data/ (accessed on 15 October 2021).
79. Song, J.; Takakura, H.; Okabe, Y.; Eto, M.; Inoue, D.; Nakao, K. Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation. In Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security, Salzburg, Austria, 10 April 2011; pp. 29–36.
80. Husák, M.; Žádník, M.; Bartoš, V.; Sokol, P. Dataset of intrusion detection alerts from a sharing platform. *Data Brief* **2020**, *33*, 106530. [CrossRef]
81. Pekarčík, P.; Gajdoš, A.; Sokol, P. Forecasting Security Alerts Based on Time Series. In Proceedings of the International Conference on Hybrid Artificial Intelligence Systems, Gijón, Spain, 11–13 November 2020; pp. 546–557.
82. Husák, M.; Čeleda, P. Predictions of Network Attacks in Collaborative Environment. In Proceedings of the NOMS 2020—2020 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 20–24 April 2020; pp. 1–6.
83. Sallay, H.; Ammar, A.; Saad, M.B.; Bourouis, S. A real time adaptive intrusion detection alert classifier for high speed networks. In Proceedings of the 2013 IEEE 12th International Symposium on Network Computing and Applications, Cambridge, MA, USA, 22–24 August 2013; pp. 73–80.
84. Lyons, K.B. *A Recommender System in the Cyber Defense Domain*; AFIT Scholar: Dayton, OH, USA, 2014.
85. Tsai, C.-F.; Lin, C.-Y. A triangle area based nearest neighbors approach to intrusion detection. *Pattern Recognit.* **2010**, *43*, 222–229. [CrossRef]
86. Aburomman, A.A.; Reaz, M.B.I. A novel SVM-kNN-PSO ensemble method for intrusion detection system. *Appl. Soft Comput.* **2016**, *38*, 360–372. [CrossRef]
87. Horng, S.-J.; Su, M.-Y.; Chen, Y.-H.; Kao, T.-W.; Chen, R.-J.; Lai, J.-L.; Perkasa, C.D. A novel intrusion detection system based on hierarchical clustering and support vector machines. *Expert Syst. Appl.* **2011**, *38*, 306–313. [CrossRef]
88. Khan, L.; Awad, M.; Thuraisingham, B. A new intrusion detection system using support vector machines and hierarchical clustering. *VLDB J.* **2007**, *16*, 507–521. [CrossRef]
89. Jabbar, M.; Aluvalu, R. RFAODE: A novel ensemble intrusion detection system. *Procedia Comput. Sci.* **2017**, *115*, 226–234. [CrossRef]
90. Liao, N.; Tian, S.; Wang, T. Network forensics based on fuzzy logic and expert system. *Comput. Commun.* **2009**, *32*, 1881–1892. [CrossRef]
91. Chadha, K.; Jain, S. Hybrid genetic fuzzy rule based inference engine to detect intrusion in networks. In *Intelligent Distributed Computing*; Springer: Cham, Germany, 2015; pp. 185–198.