Mitigating False Negative Intruder Decisions in WSN-based Smart Grid Monitoring

Safa Otoum, *Student Member, IEEE,* Burak Kantarci, *Senior Member, IEEE* and Hussein T. Mouftah, *Life Fellow, IEEE*

Abstract*—Monitoring the Smart Grid (SG) is highly desired for critical applications such as power quality assessment and transformer monitoring. Due to their low-cost, flexibility and efficiency as well as their widely usage in several critical infrastructure monitoring applications, Wireless Sensor Networks (WSNs) are estimated to be extensively used in SG applications. WSNs-based SG networks are vulnerable to different types of attacks and intruders. In order to operate networks in secured environments, in this paper we analyze our Clustered Hierarchal Hybrid-Intrusion Detection System (CHH-IDS) that is responsible for various attacks injected by known and unknown intruders. As False Positives (FPs) and False Negatives (FNs) are the key performance parameters in IDS, we investigate mitigation of FNs through a two-tier intrusion detection approach, which deals with anomaly and signature detection in parallel. In the presence of such a hybrid mode, utilization proportion between the anomaly detection and signature detection models affect the FN performance. In these two subsystems, Random Forest method is used for signature detection over known attacks and E-DBSCAN (Enhanced Density-Based Spatial Clustering of Applications with Noise) method is used for anomaly detection over unknown attacks. Through simulations that run on real datasets, we validate that the higher the weight of anomaly detection subsystem (i.e. the lower the weight of the signature detection subsystem), the lower the FN rates experienced by the entire H-IDS system. More specifically, we show that FN rates can be significantly reduced by 20.4% when the weight on anomaly detection subsystem is increased from 60% to 70% while the accuracy is expected to be improved through signature detection subsystem by using the Random Forest which has higher detection rate than the E-DBSCAN method.

I. INTRODUCTION

Wireless Sensor Networks (WSNs) have been known as strong candidates to meet the Smart Grid (SG) communication requirements, specifically the continuous monitoring and self-configurability needs. WSNs employ various types of sensors to monitor a wide range of circumstances [1] [2].

SG enables better positioning of energy sources, such as integration of wind and solar with the generation system [3]. It can also help in reducing costs through increased network reliability, and increased processing efficiency, as well as saving energy. Processing efficiency and reliability are important factors attained by accomplishing secure and reliable data aggregation and transmission. SG is vulnerable to various cyberattacks since it relies on the information and communication technologies [4]. Therefore, any intrusion with the control unit may cause irreversible consequences either locally or at larger scale. Moreover, communication lines are also vulnerable to intrusion that can interrupt the communication and manipulate the transmitted signals between the customers and distribution links. Lately, WSNs have been used in SGs because of their flexibility, self-deployment features and low-cost [5] [6].

Some sensors are used to monitor the energy usage or the quality of power lines, which are directly associated to the SG. Integration of WSNs with smart grids has introduced additional security threats to both fields such as flooding and jamming. The security vulnerabilities of SGs can be in the physical domain or in the cyber domain such as intrusions on transmission lines and intrusions on the communication links respectively.

During the past few years and with the extensive growth of communication networks, many new hacking tools and intrusive approaches have appeared, as a result, security is becoming more stimulating. In addition to attacks and intrusions protection techniques such as user authentication and user authorization, intrusion detection is also essential for all-inclusive security of networks, dealing with suspicious activities within networks and automatically detect different intrusion attempts.

The effectiveness of any Intrusion Detection System (IDS) is determined by its possibility of identifying an anomalous condition upon an intrusion attempt [7]. Two different methods used to recognize attacks, namely the anomaly detection method and signature (misuse) detection method. In anomaly detection method, a normal system behaviour profile is formed and any deviation from that defined profile is marked as an anomaly. Signature detection method detects the attacks based on their pre-known patterns [8]. Anomaly detection can detect attacks with a high false positive (FP) rate [9] due to its lower detection rates. On other hand, signature detection results in low FP rate because of its higher detection rates since the attack signatures are pre-defined. In order to combine the advantages of these two methods, we focused on analyzing the impact of each subsystem on the misleading intruder detections (i.e. False Negatives(FNs)) of the Clustered Hierarchal Hybrid-Intrusion Detection System (CHH-IDS) in WSN-based SG monitoring.

The main challenge in intrusion detection is achieving high accuracy of pattern recognition. This eventually affects the accuracy of the system both in terms of FPs and FNs. FP denotes the situation where non-malicious activity is marked as a malicious pattern whereas FN denotes the situation where a non-malicious pattern is marked as malicious. In this paper, we present our Clustered Hierarchal Hybrid IDS (CHH-IDS) model for intrusion detection in WSN-based SG networks. The proposed model consists of two subsystems, namely the anomaly detection subsystem and the signature detection subsystem. Each subsystem runs a different machine

The authors are with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, Canada, Ottawa, ON, K1N 6N5. e-mail: {safa.otoum, burak.kantarci,mouftah}@uottawa.ca

978-1-5090-4372-9/17/$31.00 ©2017 IEEE

learning approach. These techniques are Random Forest and Enhanced DBSCAN (E-DBSCAN). The signature detection subsystem uses the Random Forest algorithm to classify network connections into intrusion and normal data based on a labeled training dataset. The anomaly detection subsystem uses the E-DBSCAN algorithm, which is a density based clustering technique that splits the network nodes into dense regions. We evaluate the CHH-IDS for the WSN-based SG monitoring system under real networks connections which are generated by the Defense Advanced Research Projects Agency (DARPA) and prepared by ACM KDD'99 (special interest group on Knowledge Discovery and Data mining 1999 contest) [10]. Our results show that by proper selection of the weight of each subsystem in the IDS, the FN rate can be reduced down to 1.894%.

The rest of the paper is organized as follows. Section II summarizes the related work. Section III presents CHH-IDS model and its evaluation on KDD'99 datasets. Section IV analyses the simulation results. Finally, Section V concludes the paper and gives future directions.

II. Related Work and Motivation

Addressing security vulnerabilities is the most essential issue in WSN-based SG applications. Although the intrusion detection literature is tremendously rich, there is no "perfect" intrusion detection approach, which can always properly differentiate between malicious and normal activities when known and unknown attacks co-exist. The consequences of FNs are unauthorized activities in the networks whereas FPs may simply block legitimate access. Thus, when real attacks are experienced, True Positives(TPs) are hidden within FPs [11] [12]

The authors in [13] provided an overview of different IDS algorithms, such as artificial neural networks, swarm intelligence, fuzzy sets and soft computing. A collaborative intelligent IDS and a fuzzy inference system were proposed to reduce FPs through fuzzy alert correlation in [14] and [15], respectively, while the authors in [11] reduced both FPs and FNs with their environmental-aware IDS where they integrate the characteristics and the properties of the protected system in traffic analysis such as security policy and network topology.

A system of Attack Session Extraction (ASE) was proposed in [16] to create a pool of traffic traces causing possible FPs and FNs to IDSs. The PCAPLib system is an extended version of the ASE [17], which anonymized users' privacy in the FP and FN traffic traces out of security considerations. However, previous work only focuses on studying how to reduce FPs and FNs in IDSs or how to collect and extract the FP and FN traffic traces from real-world traffic. In [18], the authors introduced an IDS using Bayesian approach for wireless network aiming at minimum FPs and FNs. Their objectives were to recognize signatures of known attacks, match the observed behavior with those signatures and signal an intrusion alarm when match occurs.

In [19], a Behavior-Rule based IDS (BRIDS) for safety critical SG applications has been introduced. BRIDS focuses on detecting the compromised devices that are deployed in Wide Area Networks (WANs), Neighborhood Area Networks (NANs) and Home Area Networks (HANs)in order to support safe and secure applications. The authors analyzed a reference model for the SG by having some nodes monitor the behavior of their neighbors, and derived behavior rules for each device. The study also tackled finding the FN and FP probabilities to be able to define the attack types. Furthermore, the rule-based IDS was applied only in the context of communication networks in which they did not include the physical environment and the closed-loop control structure. They focused only on the unknown attacks where sometimes some rules are incomplete resulting in some attackers not being able to be detected. It is worth mentioning that the power line communications side which is very vulnerable to different types of intruders was also neglected.

To the best of our knowledge, a holistic IDS for WSN-based SG monitoring that works for both known and unknown attacks remains non-addressed. Therefore based on this motivation, we study our CHH-IDS approach by focusing on mitigation of misleading decisions against malicious or legitimate nodes.

III. Clustered Hierarchal Hybrid-Intrusion Detection (CHH-IDS) System for WSN-Based SG Monitoring

We consider a clustered WSN that consists of a central server and N clusters each consisting of M sensor nodes. In each cluster, the Cluster Head (CH) assumes the responsibility of data aggregation. Each sensor node forwards its sensed data to its corresponding CH. The CH groups the data and forwards it to the centralized server.

A. System model

Our topology starts with the selection of CHs, traffic aggregation function and finishes with using the Random Forest and E-DBSCAN methods for the CHH-IDS.

1) Cluster Head Selection: The CH selection is performed by applying the Weighted Cluster Head (WCH) selection algorithm [20], in which the CH is selected based on the comparison of the weight of each node with the other nodes inside its corresponding cluster. In the weighted cluster head selection method, each node is assigned a weight which is a function of its Received Signal Strength (RSS), node degree, and mobility. The WCH selection method passes through the following sequential steps: i) Find the node degree d_n of each node n, ii) Compute the degree of difference Δn, iii) Compute the sum of received signal strength ($SRSS_n$), iv) Compute the node's mobility (M_n), v) Compute the cumulative time, T_n which denotes the time elapsed since node m has been appointed as a CH, and vi) Compute the combined node weight (W_n). The combined node weight equation is represented in (1) below where w_1, w_2, w_3, w_4 are the weighing factors for the corresponding system parameters.

$$W_n = w_1 \Delta n + \frac{w_2}{|1/SRSSn|} + w_3 Mn + w_4 Tn \qquad (1)$$

In the equation, $\Delta n = |dn - \delta|$, d_n is the degree of node n which refers to its neighbors, δ refers to the number of nodes

that a CH can handle while Δn is the degree-difference of node n and $|1/SRSSn|$ is the normalized RSS sum. Each node evaluates its own weight and broadcasts it with its ID. Following upon the broadcast, it compares its weight to those of its neighbors. The node with the minimum weight is selected as the CH [20]. The notations used in the system model are shown in Table I.

Table I
NOTATIONS USED IN THE SYSTEM MODEL

Notation	Description
d_n	Degree of node n
Δn	Degree difference of node n
$SRSS_n$	Sum of received signal strength
M_n	Node mobility
T_m	Cumulative time
W_n	Combined weight
w_1, w_2, w_3, w_4	Weight factors for the corresponding system parameters
δ	Number of nodes that a CH can handle

2) Data Aggregation: In our proposed network, each CH acts as the aggregator for its own cluster. Each CH aggregates the traffic from the other nodes in its cluster and sends these data to the centralized sink. The data aggregation method in [21] has been adopted by our proposed system. The data aggregation method assesses the aggregator's trust score based on the trust score of each node along with the trust evaluation between the aggregator and the nodes. It is worth noting that data aggregation is an essential process in WSNs to eliminate redundancy of sensing data, save energy and minimize communication overhead [21]. We use the function in (2) [21] in our model in order to calculate the trust score of CHs which are represented as the aggregators. In the equation, T_{agg} is trust value of the aggregator, T_n is the trust value of node n, and T^n_{agg} is the trust evaluation between the aggregator and node n.

$$T_{agg} = \frac{(\sum_{n=1}^{k}(Tn+1)\cdot T^n_{agg})}{\sum_{n=1}^{k}(Tn+1)} \quad (2)$$

B. Hybrid intrusion detection model

This section presents a parallel intrusion detection model for WSN-based SG monitoring. In the network under study, the aggregated traffic undergoes two parallel intrusion detection subsystems, namely the anomaly detection subsystem for the unknown attacks and the signature detection subsystem for the known ones, which refers to a hybrid system. The aggregated data (D) is distributed on the two detection subsystems following a time-slotted method in a round-robin fashion as shown in Figure 1. D_{s1} and D_{s2} refers to signature detection data and anomaly detection data respectively.

1) Signature detection: Our proposed signature detection method adopts the Random Forest algorithm as a supervised classification technique [22]. It works in two phases, namely the training and classification phases.

Training phase works offline in order to build the regular patterns and intrusion (i.e. anomalous) patterns by using the training dataset. A labeled training dataset is delivered after preprocessing operations into the intrusion pattern builder, as a result it builds the detection module that needs the intrusive patterns.

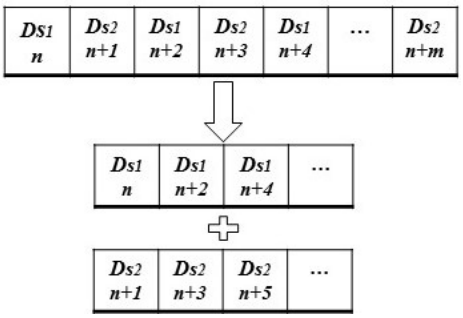

Figure 1. Data distribution between the two intrusion detection subsystems

Classification phase works online to detect intrusions based on the generated patterns from the training phase. In the classification phase, the network traffic is captured by the network aggregators and altered by preprocessing operations, and written into a network features database. At the end, the signature detector decomposes the features database into intrusive features database and normal features database by using the patterns generated in the training phase. An alarm is generated if an intrusion is detected.

Diverse types of intrusions produce various network connections. Majority of the intrusions such as Denial of Service (DoS) produce more connections than non-major intrusions such as User to Root (U2R). This phenomenon leads to higher sensitivity in the detection of major intrusions whereas it does not work in favor of non-major ones. A balanced training dataset can be considered as a solution to this problem. Balanced training dataset is obtained by down-sampling the major intrusions and oversampling the non-major intrusions [22]. The Random Forest algorithm is a classification algorithm that consists of a collection of tree-structured classifiers, in which each the tree sends a unit vote for the most popular class at each input [23]. Each tree is grown up as follows [22] [23]:

- If the size of the training set is Y, a sample population of size Y is taken randomly from the original dataset, and the sample population becomes the training set for growing the tree.
- If there are X input variables, x variables are selected randomly out of the X input variables. Then, the best split on these x is used to split the node. The value of x is held constant during the forest growing. Each tree is grown to the largest extent possible. Random forest-based signature detection is described in Figure 2.

2) Anomaly detection: Anomaly detection subsystem of the CHH-IDS runs the Enhanced-Density Based Spatial Clustering of Applications with Noise (E-DBSCAN) algorithm. DBSCAN is a density clustering algorithm in which it considers clusters as dense regions of objects in the data space that are separated by regions of low density objects [24]. It is one of the most recently used clustering algorithms. It has been updated to derive a new technique in calculating the threshold distance parameter ϵ which is a crucial parameter in the algorithm [25]. DBSCAN can discover clusters of random shapes. However, clusters that

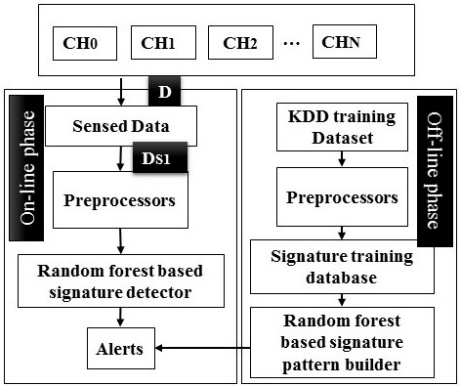

Figure 2. Random forest based signature detection procedure

are close to each other belong to the same class [24]. DBSCAN is dependent on parameters provided by the users, and it is computationally expensive when applied on unbounded datasets. E-DBSCAN is used to perform a performance enhancement on the DBSCAN algorithm [25]. DBSCAN contains two input parameters ϵ and $MinPts$. It also has the following rules:

$\rightarrow \epsilon$-neighborhood $N_\epsilon(x) = y \in X | d(x,y)| \leq \epsilon$ of point x.

$\rightarrow$ A core object has neighborhood of size greater than $MinPts$.

$\rightarrow$ Point y is density-reachable from a core object x.

$\rightarrow$ Points x and y are density connected where x and y are density-reachable from a common core object.

The integration of the E-DBSCAN into WSN-based SG monitoring system is described in Figure 3.

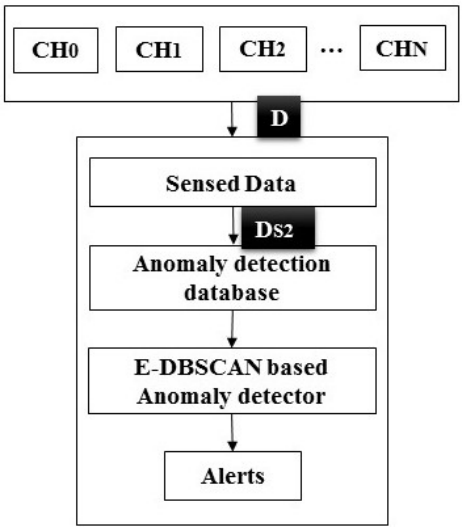

Figure 3. E-DBSCAN based anomaly detection procedure

IV. Performance Evaluation

A. Environment of experiments

The objective of this paper is to mitigate the FNs in an Intrusion Detection System (IDS) for WSN-aided SG monitoring. We have evaluated the performance of CHH-IDS system on NS-3 simulator [26]. We have considered a network of 20 temperature sensors are attached to Power Management Units (PMUs) in the SG network, and the sensors adopt the Hierarchical Dynamic Source Routing (H-DSR) protocol which performs well in Home Area Networks (HANs). The sensors attached to the PMUs make 4 clusters that spread out in a 100m x 100m area. We have run each simulation scenario 10 times, and in the figures, we present the average of these runs with 95% confidence level. The simulation settings along with the assumptions are summarized in Table II.

Table II
Simulation settings

Simulation parameter	Value
Number of nodes	20
Number of clusters	4
Routing protocol	H-DSR
Simulation time	300s
Packet size	250 bytes
Communication range	100m
Trust range	[0,1]
Operational area	100m x 100m
Sensor types	Temperature sensors
Monitored components	PMUs
Attack Types	Defined in KDD CUP'99 Dataset

B. KDD CUP 99 Data set

The KDD (Knowledge Discovery in Data mining) CUP 1999 Dataset is used to validate the efficiency of the balanced CHH-IDS [27]. The KDD CUP 1999 intrusion detection dataset helps designers of IDS with evaluating different methodologies. Attacks in KDD CUP 99 are categorized under four types as follows:

- **Denial of Service (DoS)**: Attacker tries to prevent users from using a service by making some computation or denning legitimate users access to systems.
- **Remote to Local (R2L)**: Attacker does not have an access to the attacked machines, therefore tries to gain access.
- **User to Root (U2R)**: Attacker has local access to the attacked machine but attempts to have extra privileges.
- **Probe**: Attacker tries to gain information about the host by scanning a machine or a networking device in order to determine vulnerabilities that may be exploited later.

The KDD CUP 1999 intrusion detection dataset consists of three components, which are written in detail in Table 3 in the International Knowledge Discovery and Data Mining Tools Competition, "10% of KDD" dataset is hired for the purpose of training. It covers 22 attack types and is a sub set version of the "Whole KDD" dataset. Because of their nature, denial of service attacks account for the majority of the dataset. On the other hand, the "Corrected KDD" dataset provides a dataset with different distributions other than "10% KDD" nor "Whole KDD" and covers 14 additional attacks. Since "10% KDD" is employed as the training set in the original competition, the analysis of hybrid IDS was performed on the "10% KDD" dataset. To carry the experiments effectively, KDD CUP 1999 dataset containing connection records with varying distribution of attack types and normal class has been used in our proposed hybrid IDS. It is worth mentioning that the proportion of data in the testing dataset is not the same as the training dataset

ones, and the test data includes some specific type of attacks which are not in the training set.The KDD data set description is shown in Table III whereas the 22 attacks in training data set are classified in Table IV.

Table III
KDD DATA SET DESCRIPTION

KDD Data set	DoS	R2L	U2R	Probe
Whole KDD	3883370	1126	52	41102
Corrected KDD	229853	16347	70	4166
10% KDD	391458	1126	52	4107

Table IV
ATTACKS IN KDD 99 TRAINING DATASET

Main Attack classes	22 Attacks	Number of attacks
DoS	Smurf,land,pod,Neptune,teardrop,back	6
U2R	Perl,rootkit,buffer_overflow,loadmodule	4
R2L	Imap,guess_passwd,multihop,phf, ftp_write,spy,warezmaster,warezclient	8
Probe	Nmap,portsweep,satan,ipsweep	4

C. Hybrid intrusion detection results analysis

According to our experiments of anomaly and signature detection methods, the results show that signature detection method using Random Forest suffers from high FN rate, on the other hand anomaly detection method using E-DBSCAN achieves a lower FN rate when compared to the signature detection method. Based on this observation, in order to combine the advantages of both anomaly and signature detections methods, we use a hybrid intrusion detection system.

False Negative Ratio (FNR) refers to the percentage of positive (i.e. a node in malicious activity) cases, which inaccurately classified as negative (i.e. a legitimate node) among all estimations, as shown in the following equation.

$$FNR = FN/(TP + FN + TN + FP) \tag{3}$$

Where FN, FP, TN and TP are the False Negative,False Positive,True Negative and True Positive cases, respectively.

In our simulations, we initially set 50% of data to be directed to the anomaly detection subsystem while the other 50% is directed to the signature detection subsystem. The FN performance based on this setting is shown in Figure 4. Figure 4 shows that Random forest-based signature detection subsystem has the highest FN parameters while the overall combined IDS makes a compromise between anomaly and signature detection.

In order to improve the performance of the hybrid model performance, in terms of FNR, we increase the amount of data directed to the anomaly detection subsystem by directing 30% of the data to the signature detection subsystem and 70% of the data to the anomaly detection subsystem. As unknown attacks form the major population of the security attacks, the FNR performance of the hybrid model performance is enhanced as shown in Figure 5.

We have also reversed the setting in the previous tests by directing 70% of the data to the signature detection subsystem and 30% of the data to the anomaly detection subsystem. As unknown attacks form the non-major population of the security attacks, the FNR performance of the overall combined IDS performance is reduced as shown in Figure 6.

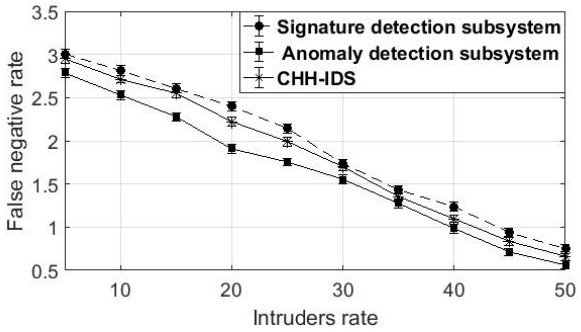

Figure 4. False Negative rate of signature detection and anomaly detection subsystems, and the overall combined IDS system when 50% of the data is handled by the signature detection subsystem and the other 50% of the data is handled by the anomaly detection subsystem

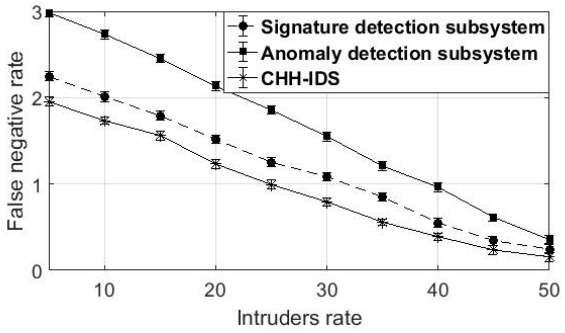

Figure 5. False Negative rate of signature detection and anomaly detection subsystems, and the overall combined IDS system when 30% of the data is handled by the signature detection subsystem and the 70% of the data is handled by the anomaly detection subsystem

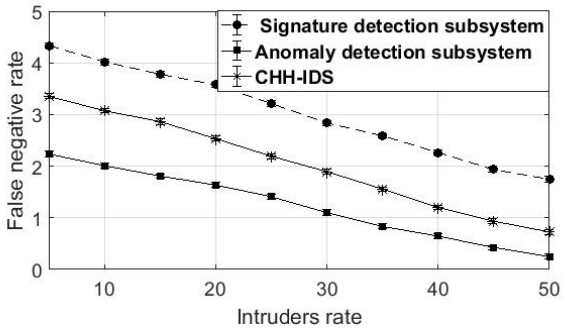

Figure 6. False Negative rate of signature detection and anomaly detection subsystems, and the overall combined IDS system when 70% of the data is handled by the signature detection subsystem and the 30% of the data is handled by the anomaly detection subsystem

In addition to the test case above, other tests have been carried out by setting a ratio of 60%:40% and 40%:60% for anomaly detection subsystem : signature detection subsystem as shown in Figure 7 and Figure 8, respectively.It is clear that by directing higher volumes of data to the anomaly detection subsystem, the performance of the overall combined intrusion detection system can be improved in terms of

FNR. The E-DBSCAN method helps in achieving better performances for our proposed model in terms of FNR while Random Forest method is expected to help in achieving better detection rates. By combining these advantages of the anomaly and signature detection subsystems, robustness of the hybrid intrusion detection model can be guaranteed.

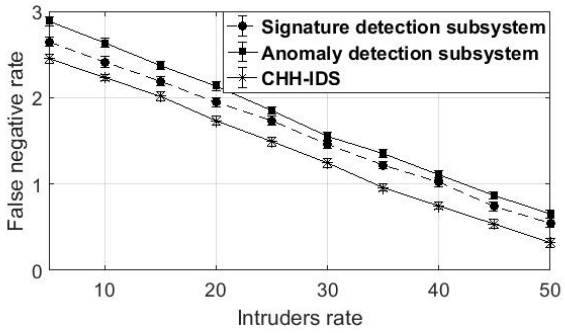

Figure 7. False Negative rate of signature detection and anomaly detection subsystems, and the overall combined IDS system when 40% of the data is handled by the signature detection subsystem and the 60% of the data is handled by the anomaly detection subsystem

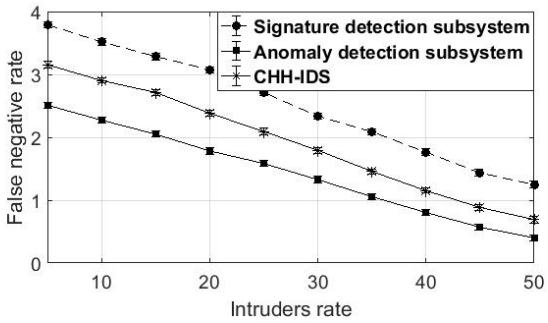

Figure 8. False Negative rate of signature detection and anomaly detection subsystems, and the overall combined IDS system when 60% of the data is handled by the signature detection subsystem and the 40% of the data is handled by the anomaly detection subsystem

V. Conclusion

We have studied the False Negative (FN) performance of a hybrid intrusion detection system which employs an anomaly detection subsystem for unknown attacks and signature detection subsystem for known attacks to SG infrastructure that is monitored through a WSN. The former runs the E-DBSCAN algorithm whereas the latter runs the Random Forest algorithm. By using the KDD CUP99 data set, we have studied the impact of the utilization ratio of each subsystem. Through simulations, we have shown that increasing the volume of the data directed to the anomaly detection subsystem, the False Negative Ratios (FNR) can be significantly reduced as unknown attacks form the majority of the attack population. We have tested various combinations, and have shown that a 10% increase in the volume of the data directed to the anomaly detection subsystem (e.g., 60% to 70%) results in a 20.4% improvement in the FNR. Since Random Forest is expected to have higher detection ratios, it is not possible to direct 100% of the data to the anomaly detection subsystem. We are currently investigating the optimum utilization ratios for each subsystem.

References

[1] R. Kaur and P. Singh, "Review of black hole and grey hole attack," *The International journal of Multimedia & ITS Applications*, vol. 6, no. 6, p. 35–45, 2014.

[2] S. Otoum, M. Ahmed, and H. T. Mouftah, "Sensor medium access control (smac)-based epilepsy patients monitoring system," *IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2015.

[3] "Communications requirements of smart grid technologies." [Online]. Available: https://www.smartgrid.gov/document/communications.requirements_smart_grid_technologies

[4] S. Otoum, B. Kantraci, and H. T. Mouftah, "Hierarchical trust-based black-hole detection in wsn-based smart grid monitoring," *IEEE International Conference on Communications(ICC'17)*, 2017.

[5] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart grid technologies: Communication technologies and standards," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, p. 529–539, 2011.

[6] E. Al-Shaer and M. A. Rahman, "Security and resiliency analytics for smart grids," *Advances in Information Security*, 2016.

[7] H. Sedjelmaci, S. M. Senouci, and M. A. Abu-Rgheff, "An efficient and lightweight intrusion detection mechanism for service-oriented vehicular networks," *IEEE Internet of Things Journal*, vol. 1/6, p. 570–577, 2014.

[8] Y. K., I. Ruiz-Agndez, and P. G., "Integral misuse and anomaly detection and prevention system," *Intrusion Detection Systems*, 2011.

[9] E. Cole, *Network Security Bible*. John Wiley & Sons, Inc., 2011.

[10] "Darpa intrusion detection evaluation." [Online]. Available: http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/index.html

[11] M. Sourour, B. Adel, and A. Tarek, "Environmental awareness intrusion detection and prevention system toward reducing false positives and false negatives," *IEEE Symp. on Comput. Intelligence in Cyber Security*, 2009.

[12] A. Derhab, A. Bouras, M. R. Senouci, and M. Imran, "Fortifying intrusion detection systems in dynamic ad hoc and wireless sensor networks," *International Journal of Distributed Sensor Networks (IJDSN)*, 2014.

[13] S. X. Wu and W. Banzhaf, "The use of computational intelligence in intrusion detection systems: A review," *Applied Soft Computing*, vol. 10, no. 1, p. 1–35, 2010.

[14] H. T. Elshoush and I. M. Osman, "Reducing false positives through fuzzy alert correlation in collaborative intelligent intrusion detection systems — a review," *International Conference on Fuzzy Systems*, 2010.

[15] G. P. Spathoulas and S. K. Katsikas, "Using a fuzzy inference system to reduce false positives in intrusion detection," *Intl. Conf. on Systems, Signals and Image Processing*, 2009.

[16] I.-W. Chen, P.-C. Lin, C.-C. Luo, T.-H. Cheng, Y.-D. Lin, Y.-C. Lai, and F. C. Lin, "Extracting attack sessions from real traffic with intrusion prevention systems," *IEEE Intl. Conf. on Communications*, 2009.

[17] Y.-D. Lin, P.-C. Lin, S.-H. Wang, I.-W. Chen, and Y.-C. Lai, "Pcaplib: A system of extracting, classifying, and anonymizing real packet traces," *IEEE Systems Journal*, vol. 10, no. 2, p. 520–531, 2016.

[18] M. Sharma, K. Jindal, and A. Kumar, "Intrusion detection system using bayesian approach for wireless network," *International Journal of Computer Applications*, vol. 48, no. 5, p. 29–33, 2012.

[19] R. Mitchell and I.-R. Chen, "Behavior-rule based intrusion detection systems for safety critical smart grid applications," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, p. 1254–1263, 2013.

[20] F. Belabed and R. Bouallegue, "An optimized weight-based clustering algorithm in wireless sensor networks," *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2016.

[21] W. Zhang, S. Das, and Y. Liu, "A trust based framework for secure data aggregation in wireless sensor networks," *IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, 2006.

[22] J. Zhang, M. Zulkernine, and A. Haque, "Random-forests-based network intrusion detection systems," *IEEE Trans. on Systems, Man, and Cybernetics, Part C*, vol. 38/5, p. 649–659, 2008.

[23] "Random forests, leo breiman and adele cutler." [Online]. Available: http://www.stat.berkeley.edu/~breiman/RandomForests/

[24] D. Ma and A. Zhang, "An adaptive density-based clustering algorithm for spatial database with noise," *IEEE Intl Conf on Data Mining (ICDM'04)*.

[25] M. Jiang, S. Tseng, and C. Su, "Two-phase clustering process for outliers detection," *Pattern Recognition Letters*, vol. 22/6-7, p. 691–700, 2001.

[26] "ns-3 tutorial." [Online]. Available: https://www.nsnam.org/docs/tutorial/html/

[27] "Kdd cup 1999 data." [Online]. Available: https://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html