AN INTRUSION DETECTION SYSTEM FOR INTERNET OF MEDICAL THINGS

1st Deborah Oladimeji
Faculty of Computer Science
Dalhousie University
Halifax, Canada
db315929@dal.ca

2nd Srinivas Sampalli
Faculty of Computer Science
Dalhousie University
Halifax, Canada
srini@cs.dal.ca

3rd Hiroyuki Ohno
Emerging Media Initiative
Kanazawa University
Kanazawa, Japan
hohno@staff.kanazawa-u.ac.jp

***Abstract*—The terms IoMT (Internet of Medical Things), IoHT (Internet of Health Things) and HIoT (Healthcare Internet of Things) are now all used interchangeably. They describe the connection of medical devices and software applications relating to healthcare information to the Internet using networking technologies. While these technologies bring the promise of improved patient care, improved efficiency, and reduced costs, they also bring new risks as many these connected devices are unmanaged and unprotected. The consequent potential impact is not just on patient data, but on patient care itself. This thesis focuses on providing a highly secure transmission of medical data in IoMT to ensure accuracy and confidentiality of patients' data. We propose a novel intrusion detection system (IDS) based on machine learning (ML) methods which uses both network and biometric parameters as features and can differentiate the normal traffic from attack traffic. Six ML methods were selected for the intrusion detection, namely, Random Forest, K-Nearest Neighbor, Support Vector Machine, Artificial Neural Networks, J48 and Decision Table, and tested against man-in-the-middle and denial of service attacks using a dataset consisting of a combination of about 20,000 normal and attack healthcare data. The dataset was generated on our IoMT test bed that was implemented using four modules, namely, a multi-sensor board, a gateway module, a network module, and a visualization module. The communication between the modules employs a Client Server publish/subscribe messaging transport protocol, MQTT, which is a light weight, simple, easy to implement for constrained devices with limited resources, such as IoMT. Experimental results indicate that our secured healthcare system can detect anomalies in both the network flow and patient's biometric readings. Furthermore, we generated a new healthcare dataset with the combination of biometric data and network traffic available for other researchers for statistical analysis and further research. Finally, we present a comparative summary of the proposed scheme with an existing scheme in terms of accuracy and execution time.**

***Index Terms*—IoMT, Intrusion detection, Machine learning**

I. INTRODUCTION

The advent of Internet of Medical Things (IoMT) has enhanced remote patient monitoring. It reduces unnecessary hospital visits and the burden on health care systems by connecting patients to their physicians and allowing the transfer of health data over a secure network. IoMT has the potential to give more accurate diagnoses, less mistakes and lower costs of care through the assistance of technology, allowing patients to send health information data to doctors. Currently, this is essentially necessary due to the effect of the global pandemic, COVID-19, reducing in-person medical visits which prevents the spread.

The healthcare industry is constantly afflicted by a myriad of cybersecurity-related issues. These issues range from malware that compromises the integrity of systems and privacy of patients to distributed denial of service (DDoS) attacks that disrupt facilities' ability to provide patient care. The security of sensitive data such as protected health information that passes through the IoMT as well as uninterrupted access to the system is a developing concern for healthcare providers. While other critical infrastructure sectors experience these attacks as well, the nature of the healthcare sector's purpose poses unique challenges. It goes beyond financial loss and breach of privacy but has direct impact on human life. For whatever reason attack is being launched on the healthcare system, either large or small, they still pose a danger. As IoMT become an integral part of healthcare, we must find a way to manage them securely and effectively.

In this thesis, we used various machine learning (ML) algorithms to build an efficient Intrusion Detection System (IDS) for healthcare application using a variety of medical sensors. The system comprises of medical sensors from which data are generated, gateway for data gathering, an attacker system to launch attacks, an IDS to analyze flow of traffic for malware detection, and monitoring systems which receives health data for visualization. One of the monitoring systems, in addition to visualization purpose also serves as server to store all medical data. At this point, health data is made available for the medical practitioners. We applied ML methods for DoS and MitM attack detection. The system indicates anomalies in the monitoring graphs to give alert during data manipulation and attempt to disrupt the network flow. This is done by analyzing both the patients' biometric data and network traffic vector. The system reports any threat if any anomalous behaviour is detected in any of the biometric data or network traffic feature.

II. METHODOLOGY

Our testbed has been built using medical sensors attached to the patient's body. We consider a typical architecture of connected medical devices in a hospital networking environment. The data acquisition layer in this system consists healthcare monitoring sensors such as temperature sensor, blood oxygen saturation sensor etc. The sensors in the IoMT network include electrocardiogram (ECG or EKG) sensor, Blood Oxygen Saturation (SpO2) sensor, temperature sensor and blood pressure sensor, placed in and around patients' body. These sensors are attached to a multisensor board via each of their wire connector. These sensor nodes monitor, collect and relay the data to local gateway nodes through the multi-sensor board. The multi-sensor board is connected to a Windows-based computer using a USB port. It also has the capability of connecting via the Bluetooth. MQTT protocol has been used to collect data from the multisensor board. The data is being sent to the cloud via MQTT publish and subscribe feature through which all other monitoring systems connect using the MQTT client. The monitoring systems could be the medical practitioners computers. Only the gateway computer connects wirelessly, the rest of the machines are connected to a switch using Ethernet cables. The switch is then connected to the internet through a router and the gateway computer makes internet connection via Wi-Fi.

Our proposed design is shown in Figure 1, data flows from the medical sensors attached to patient's body through the multisensor board to the gateway, switch and lastly, to the server and other monitoring systems for visualization. While the data travels from the gateway to the server, an attacker may modify the medical data before getting to the server or launch DoS attacks to stop the data from reaching the visualization stage. Meanwhile, network flow and patient biometric data metrics are captured at the IDS computer. The captured data is processed at the IDS for training and testing the machine learning methods for real-time detection of any anormalities. Our system uses Wireshark to capture network traffic and Argus to extract useful flow metrics, while it uses "cu" [1] and "awk" [2] command to collect all biometric data from the multisensor board. All monitoring systems get data by subscribing to the MQTT topic [3].

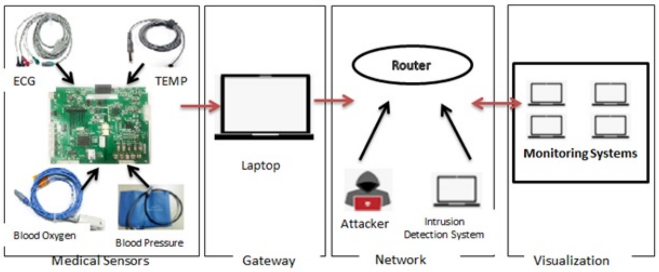


Fig. 1. A Secured healthcare Monitoring System

III. RESULTS AND DISCUSSION

In this section, we present two sets of results to validate that the network parameters and biometric data have an effect for detecting the presence of MitM and DoS attacks in the network. First, we generated dynamic graphs using gnuplot [4], a command-line program that can generate two- and three- dimensional plots. The dynamic graphs gives real-time data view of patient's biometric reading on the monitoring systems. Secondly, we present the results of the metrics used for all ML classifiers.

Figure 2 shows normal heart rate reading without attack while Figure 3 shows heart rate during MitM attack.

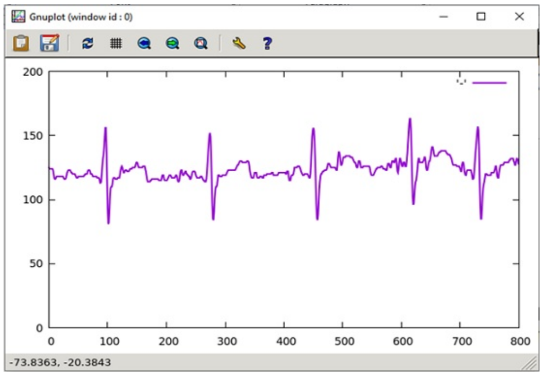


Fig. 2. Result of normal data shown on monitoring systems

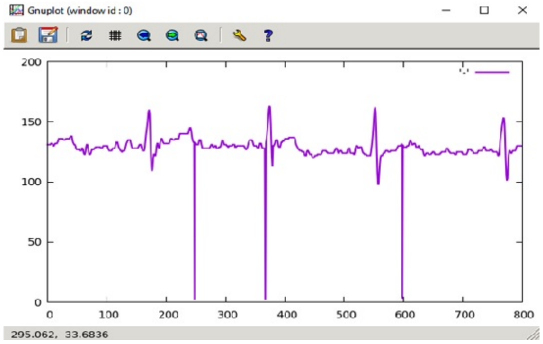
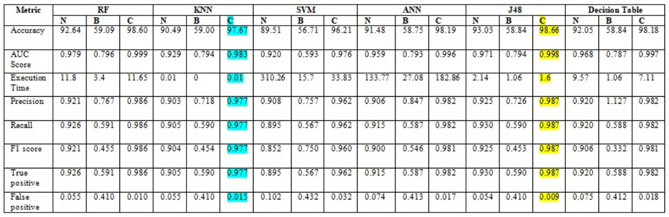


Fig. 3. Data modification/insertion attack shown on monitoring systems

To check the validity of using ML to differentiate between normal and attack biometric data, we compared the six ML methods used based on their performances using Accuracy, Execution time, Area under the ROC (AUC) score, Precision, Recall, F1 score, True Positive, and False Positive. Figure 4 summarizes the performance of machine learning classifiers for the healthcare intrusion detection system across all data where N is the network feature, B is the biometric feature, and C is the combined feature. As shown in Figure 4, all the classifiers perform better with the combined feature as against when only one feature is used. Although J48 gives the highest performance for combined data compared to other methods with an accuracy of 98.66, AUC score of 0.998, precision to be 0.987, recall of 0.987, F1 score of 0.987, true positive to be 0.987, and false positive to be 0.009. However, its execution time is 1.6 which is about 160 times the execution time of KNN. As this is significantly high considering the real-time requirement of the system, KNN which gives the lowest execution time of 0.01 is considered to be the best model for the healthcare monitoring system.



Metric	RF			KNN			SVM			ANN			J48			Decision Table		
	N	B	C	N	B	C	N	B	C	N	B	C	N	B	C	N	B	C
Accuracy	92.64	59.09	98.60	90.49	59.00	97.67	89.51	56.71	96.21	91.48	58.75	98.19	93.03	58.84	98.66	92.05	58.84	98.18
AUC Score	0.979	0.796	0.999	0.929	0.794	0.983	0.920	0.593	0.976	0.959	0.793	0.996	0.971	0.794	0.998	0.968	0.787	0.997
Execution Time	11.8	3.4	11.65	0.01	0	0.01	310.26	15.7	33.83	133.77	27.08	182.86	2.14	1.06	1.6	9.57	1.06	7.11
Precision	0.921	0.767	0.986	0.903	0.718	0.977	0.908	0.757	0.962	0.906	0.847	0.982	0.925	0.726	0.987	0.920	1.127	0.982
Recall	0.926	0.591	0.986	0.905	0.590	0.977	0.895	0.567	0.962	0.915	0.587	0.982	0.930	0.590	0.987	0.920	0.588	0.982
F1 score	0.921	0.455	0.986	0.904	0.454	0.977	0.852	0.750	0.960	0.900	0.546	0.981	0.925	0.453	0.987	0.906	0.332	0.981
True positive	0.926	0.591	0.986	0.905	0.590	0.977	0.895	0.567	0.962	0.915	0.587	0.982	0.930	0.590	0.987	0.920	0.588	0.982
False positive	0.055	0.410	0.010	0.055	0.410	0.015	0.102	0.432	0.032	0.074	0.413	0.017	0.054	0.410	0.009	0.075	0.412	0.018

Fig. 4. Summary of ML models for all data types

IV. Comparison Analysis Anar et al. [5]

Not much work has been done in this area, the only related work done by Anar et al. [5] is thus considered for comparison. Table I shows the summary of the comparison in experimental setup while Table II shows the comparison in results.

TABLE I
COMPARISON ANALYSIS OF EXPERIMENTAL SETUP

Parameters	Existing technique (Anar et al. [5])	Proposed technique
Communication Protocol	Wi-fi using TCP/IP protocol	MQTT protocol to interface IoMT devices
Attack Vector	MitM	MitM DoS – ARP Request ICMP Echo TCP SYN UDP Flood
Output Labels	2 (Normal and attack vector)	6 (Normal and attack data for each attack)
ML Models	RF, KNN, SVM, ANN	RF, KNN, SVM, ANN, J48, DT
Number of Features	34	42
Total Dataset	16000	20000

Anar et al. [5] argued that ANN is the best model for the healthcare system. However, after thorough consideration of the ML models, we concluded that KNN is the best model for IDS in healthcare system. This is majorly due to its low execution time, as time is a very crucial determinant of efficiency of healthcare system. We chose to compare the result of the ML models on combined features only since all methods proves to achieve better results with combined features as against using only either of the two. In addtion to these differences, the major significant differences between the two works is that, our model is able to scale easily, an advantage derived from the use of MQTT.

TABLE II
COMPARISON ANALYSIS OF RESULT

Metrics	(Anar et al. [5])	Proposed IDS
Accuracy	93.42	97.67
AUC Score	0.929	0.983
Optimal Model	ANN	KNN
Execution Time	272.10	0.01

V. Conclusion

In this paper, we have presented a novel Intrusion Detection System for detecting MitM and DoS attacks using ML methods. We have generated a specialized healthcare dataset for IoMT networks which was used to train six machine-learning algorithms: RF, KNN, SVM, ANN, J48, and Decision Table. Our dataset consists of 20 thousand records of normal, MitM and DoS attack packets. We have implemented our proposed solution using MQTT protocol which is highly suitable for resource constrained devices like IoMT. It also provides the additional advantage of remotely monitoring the healthcare network. The attacks considered in this dataset were modification/insertion, data breach, ARP request, ICMP echo, TCP SYN, and UDP flood. Attacks were classified by a model built using cross-validation folds in which the original dataset splits into k equal parts (folds) where k equals 10. The results show that KNN is the best model to use for healthcare monitoring system as it gives the lowest execution time and generated an accuracy rate of 97.67%. Our IDS has a great effect for identifying and detecting the type of attacks and reducing the false alerts. Results show that all ML methods used performed significantly better with combined features compared to using only one of the two features. In our plan for future work, we aim to develop our intrusion detection system using Fuzzy Logic, launching more advanced attacks like DDoS (Distributed Denial of Service) attack, and This and other attacks can be launched by using multiple compromised sources to generate the attacks, and integrate a private MQTT broker on the cloud for more secure and robust solution.

Acknowledgment

Thanks to Dalhousie Faculty of Graduate Studies (FGS) for the consideration for FGS Entrance Scholarship that has provided full financial support for this Masters program.

References

[1] "cu(1): Call up another system - Linux man page," Accessed 2021-05-21. [Online]. Available: https://linux.die.net/man/1/cu

[2] "The GNU Awk User's Guide," Accessed 2021-05-21. [Online]. Available: https://www.gnu.org/software/gawk/manual/gawk.html

[3] "MQTT Client and Broker and MQTT Server and Connection Establishment Explained - MQTT Essentials: Part 3," Accessed 2021-05-05. [Online]. Available: https://www.hivemq.com/blog/mqtt-essentials-part-3-client-broker-connection-establishment/

[4] "gnuplot homepage," Accessed 2021-05-21. [Online]. Available: http://www.gnuplot.info/

[5] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, and R. Jain, "Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study," *IEEE Access*, vol. 8, pp. 106 576–106 584, 2020.