

Network Data Security for the Detection System in the Internet of Things with Deep Learning Approach

Kalubi Kalubi Deiu-merci¹, Mayou²

¹Department of Computer Sciences, Taiyuan University of Technology, Shanxi/Taiyuan
Email: dieumecikalubi@yahoo.com

² Department of Computer Sciences, Taiyuan University of Technology, Shanxi/Taiyuan
Email: yao.ma@hotmail.com

Abstract— we thought to set up a system of interconnection which allows sharing the communication network of data without the intervention of a human being. The Internet of Things system allows many devices to be connected for a long time without human intervention, data storage is low and the level of data processing is reduced, which was not the case with older solutions proposed to secure the data for example: cyber-attack and other systems. But other theories like for example: artificial intelligence, machine learning and deep learning have a lot to show their ability and the real values of heterogeneous data processing of different sizes and many researchers had to work on it. In the case of our work, we have used deep learning theories, to achieve a light data interconnection security solution; we also have TCP/IP protocol for data transmission control, algorithm drillers for classifications. In order to arrive at a good solution; First, we thought of a model for anomalies detection in Internet of Things and we think about the improvement of architectures of the Internet of the existing objects already proposed a system with a light solution and especially multilayer for an IoT network. Second, we analyzed existing applications of machine learning, deep learning to IoT, and cybersecurity. The recent hack of 2014 Jeep Cherokee, iStan pacemaker, and a German steel plant are a few notable security breaches. Finally, from the evaluated metrics, we have proposed the best neural network design suitable for the IoT Intrusion Detection System. With an accuracy of 98.91% and False Alarm Rate of 0.76 %, this research outperformed the performance results of existing methods over the KDD Cup '99 dataset. For this first time in the IoT research, the concepts of Gated Recurrent Neural Networks are applied for the IoT security.

Keywords— data security; Internet of Things; deep learning.

I. INTRODUCTION

By definition, it can be said that networked or interconnected systems often have a long list of interconnected data in parallel; this type of system facilitates the rapid sharing of data as they are distributed and errors are easily reduced according to each connected device. And these kinds of networks are also like a mathematical regression that is linear because the solution between the input variables and the output produced is often captured by nonlinear relations and this final solution which is found from variables results represents the set of many hidden layers of each predefined function. Our studies are very much about the analysis of a new domain that is the Internet of Things with the presentation of architecture and neural networks. We then made the comparison on security issues and privacy between the field of deep learning and that of the Internet of Things. This may be possible with the concepts of machine learning or deep learning because IoT generates a huge amount of heterogeneous data. And we note that with this research we fall into a multilayer architecture and the new technology of the Internet of Things for a unique system. And the algorithms we had to apply during this research are to monitor network data interconnect and classify activities are application attacks for multiple layers of each architecture. And for our research we thought about using the KDD 99 Cup intrusion detection dataset that many researchers who have been working on internet data security consider as a combination of referential data. For all the work related to our work are in part (2), the methodology used for our work is in part (3) and the last part focuses on the implementation and outcome of our work (4).

Motivation

In the world of IoT, the datasets are high-dimensional, temporal and multi-modal. Deep Learning algorithms with robust computation power are more suitable for complex

IoT datasets compared to legacy machine learning techniques. The application of deep learning to the IoT domain, particularly in IoT security is still in the initial stages of research and has a great potential to find insights from the IoT data. With smart use of deep learning algorithms, we believe that IoT solutions can be optimized. For example, recurrent neural networks in deep learning have the capability to learn from previous time-steps of the input data. The data at each time-step is processed and stored and given as input to the next time-step. The algorithm at the next time step utilizes the previous information stored to process the information. Though the neural network structures are complex, the hyper-parameters can be tuned to obtain light-weight functionality for IoT solutions. This hypothesis motivated us to apply deep learning concepts to IoT network security.

Problem statement

The goal of this thesis is to analyze and answer the following research questions:

- What are the security and privacy issues relevant to the IoT environment?
- Does GRU better than the other machine learning approaches for Intrusion Detection on the IoT?
- Does a separate GRU based IDS for each network layer perform better than the all layer GRU?

Contribution

This research can be extended by applying the algorithms on GPU environment on real-time IoT data. Though there are various deep learning algorithms such as deep neural networks, auto encoders, convolutional neural networks and recurrent neural networks, the research problem requires an algorithm that can learn from historical data. Therefore, we have selected the family of recurrent neural networks for the research. Considering the need of building smart and lightweight solutions for the IoT network, we have performed the experiments with only the Gated Recurrent-Unit (GRU) algorithm while the vanilla RNN and LSTM are ignored. We have modified the data by dividing it into various layers such that the same procedure can be applied in an IoT network.

II. RELATED WORKS

There are also several existing works in this area. In this section, we will discuss the most recent work that uses methods and architectures. We were motivated and inspired from this work "Cyber-Physical-Social Based Security Architecture for Future Internet of Things" because after taking a lot of time to study and read this work, we found tremendous benefits from doing our research in this area. Alrawashdeh and Purdy [18] proposed using a RBM with one hidden layer to perform unsupervised feature reduction. The weights are passed to another RBM to produce a DBN. The pre-trained weights

www.ijaers.com

are passed into a fine tuning layer consisting of a Logistic Regression classifier (trained with 10 epochs) with multi-class soft-max. The proposed solution was evaluated using the KDD Cup '99 dataset. The authors claimed a detection rate of 97.90% and a false negative rate of 2.47%. This is an improvement over results claimed by authors of similar papers.

Similarly, Tang et al. [19] also propose a method to monitor network flow data. The paper lacked details about its exact algorithms but does present an evaluation using the NSL-KDD dataset, which the authors claim gave an accuracy of 75.75% using six basic features. Kang and Kang [20] proposed the use of an unsupervised DBN to train parameters to initialise the DNN, which yielded improved classification results (exact details of the approach are not clear). Their evaluation shows improved performance in terms of classification errors. You et al. [16] propose an automatic security auditing tool for short messages (SMS). Their method is based upon the RNN model. The authors claimed that their evaluations resulted in an accuracy rate of 92.7%, thus improving existing classification methods (e.g. SVM and Naive Bayes). In addition, there is other relevant work, including the DDoS detection system proposed by Niyaz et al. [21]. They propose a deep learning-based DDoS detection system for a software defined network (SDN). Evaluation is performed using custom generated traffic traces. The authors claim to have achieved binary classification accuracy of 99.82% and 8-class classification accuracy of 95.65%. However, we feel that drawing comparisons with this paper would be unfair due to the contextual difference of the dataset. Specifically, benchmark KDD datasets cover different distinct categories of attack, whereas the dataset used in this paper focuses on subcategories of the same attack

III. METHODOLOGY

We designed an innovative architecture for an IoT home network that would reduce the size of the datasets for the IDS classifier. We have selected the KDD Cup 1999 Intrusion Detection Dataset for the experiments and proposed an intelligent solution which satisfies the key requirements of the IoT solutions. We have performed the feature engineering using a Random Forest classifier and selected those features with high importance. We performed a rigorous data analysis and prepared the data in the required format before it was used as an input to the model.

Proposed Multi-Layer architecture for IoT network

Out of various security measures, we have selected network security as the use case to prove the defined features are apt for an IoT network. In a regular wireless system, the Intrusion Detection System (IDS) monitors the network data using either a "Signature-based approach" or

an “Anomaly-based approach”. The IDS mounted at a point in the network obtains 27 all the network data and classifies the data into “normal” or “attack”. Other than traditional approaches, Machine Learning (ML) algorithms are applied to a dataset and classification is performed through supervised learning. However, this legacy approach may not be suitable for smart IoT network systems due to their heterogeneity. The security solutions for intrusion detection should be light-weight, multi-layered and have a good amount of longevity. Hence, we developed a multi-layered architecture and applied light-weight machine learning algorithms which can work with better performance for longer periods of the time. An IoT system contains various devices which are placed at different locations with long distances between them. The number of devices involved in IoT systems is higher when compared to a regular wireless or wired system. A single IDS system must have the memory capacity to process the network data among all the devices and must be responsive in a short amount of time. In this case, the performance will be poor in the IoT network due to the high number of devices and the large distance between the devices. Each IDS placed at a TCP/IP layer monitors only the data obtained from the devices that belong to that layer. We chose this architecture as the main architecture of our work because this architecture has many advantages and uses a multilayer system that is a potential system in today's world.

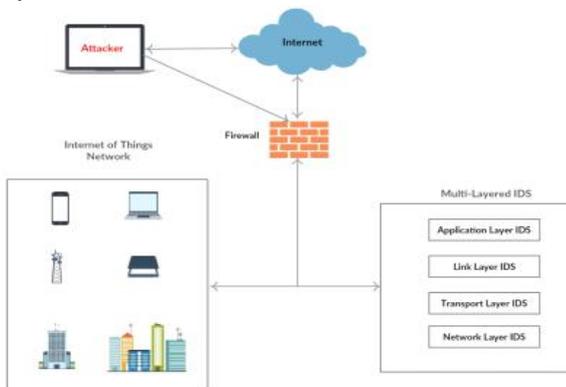


Fig. Multi-Layer architecture for IoT network

IV. EVALUATION AND RESULTS

Feature Selection

We have explained well in the above step-by-step chapters of this research and the random forest classification algorithm used to select the main important features of all the classifiers one by one and Intersecting graphical results for each classifier's characteristics are presented. The "Protocol Type" feature has been selected in all intrusion detection layers.

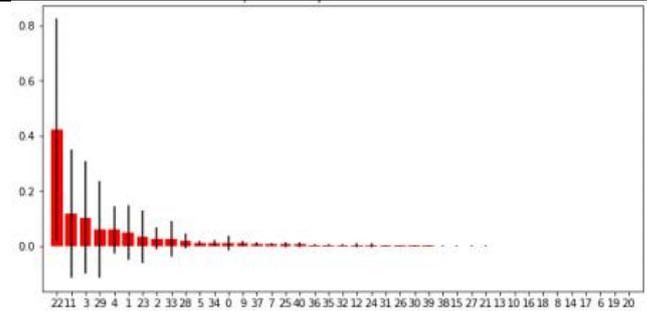


Figure: Feature Importance graph for Application Layer IDS

Data Loading and pre-processing

Functions defined for Data Loading and Data Pre-processing
 DataLoading - Loads the csv data into the system
 DataPreprocessing - Performs below operations on the loaded dataset.

- Dropping the Duplicates.
- Dividing the dataset into features set and the labels set.
- Converting categorical data into numerical data.
- Encoding the normal as '1' and abnormal as '0'.
- Converting input data into float which is required in the future stage of building in the network.
- Adding another column to the labels set - kind of one hot encoding
 i.e normal = '1' is represented as '1 0'
 abnormal = '0' is represented as '0 1'

This is required so that the softmax entropy function can efficiently calculate the accuracy. Loading the data into the system And applying the data preprocessing and feature selection for the dataset. Dividing into train and test datasets, performing the above operations are required before training and testing the model.

- **Normalizing the Input Features and Hyper Parameters:** Here we are not restricting the input size, therefore it batch_size is given as "None", weights and biases are initialized in random using tf.random_normal function. Sizes are defined appropriately as per the logic, the biases output either '1 0' or '0 1'
- **Building the Model:** Before building the model, we have to reshape the inputs in to 3D tensors of size from 2D tensors of size. We can specify a loss function just as easily. Loss indicates how bad the model's prediction was on a single example; we try to minimize that while training across all the examples. Here, our loss function is the cross-entropy between the target and the softmax activation function applied to the model's prediction.

Evaluation Metrics

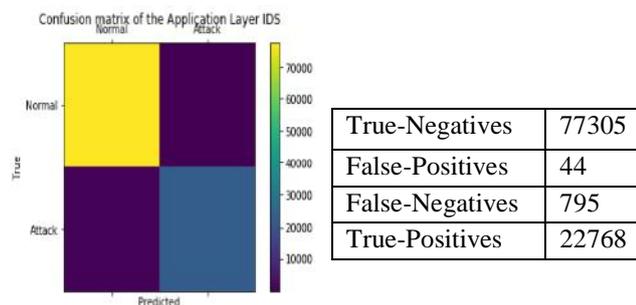
Compared to this part, we make a simple comparison of the following values: the precision of the training, the

rhythms of learning and for a good understanding on the behavior of the model according to the change of these hyper-parameters in relation to the time steps, after testing the performance of each class of IDS classifier by adding the hyper-parameters of the GRU algorithm. We performed a similar type of experiments for all IDS classifiers (all layers, the application layer, the transport layer, and other layer classifiers such as: for the network). And we are much interspersed with the results of the following classifications with their performance.

Classifier performance results with application layer:

With this experience we have achieved the best training accuracy with time steps of "40"; which is our complete result of this layer. The confusion matrix of the optimized (time-steps = 40, learning rate – 0.01) and the corresponding plot for the Application-Layer IDS can be analyzed in Table

Time-Steps	Train Accuracy	Precision	Recall	F-1 Score	Far
10	98.694	0.9994	0.99	0.9977	0.0022
20	98.833	0.9943	0.3296	0.9922	0.0077
30	96.71	0.9996	0.9952	0.9418	0.05812
40	98.706	0.9967	0.9902	0.9983	0.0016



Comparing the results of all classifiers and their layers

The optimized results of all the IDS classifiers are compared and it was found that the performance of the All-Layers IDS classifiers is inferior to the Individual layer IDS classifiers in terms of training accuracy and training time. The light-weight algorithms, when used in a multilayer architecture, perform better which is suitable for an IoT system. The comparison of the results can be found in Table.

Feature Selection Method	IDS Type	Number of features	Training Accuracy	Training Time
Random Forest Classifier	All Layer IDS	12	98.9	50.04 secnds
	Application layer IDS	6	98.7	20.54 secnds
	Transport	6	99.469	18,85

	layer IDS			secnds
	Network layer IDS	6	98.706	30.05 secnds

Comparison of the classifiers performance with existing work

We are at the end of our research. We performed additional analysis and reported that we compared the results with existing research performed by machine learning algorithms on intrusion detection classification, as shown in Table. We can see that our search has exceeded the performance of all existing jobs. We intend to continue in this area of research in the future time.

Algorithm	Precision (%)	Recal l (%)	Accuracy (%)	FAR (%)
DIR, PAYLOAD	74.33	78.55	78.55	75.00
SRC_PORT	95.43	96.32	96.32	95.74
GNNN[22]	87.08	59.12	93.05	12.46
FNN[22]	92.47	86.89	97.35	2.65
RBNN[22]	69.56	69.83	93.05	6.95
K-Mean-KNN[27]	98	98.68	93.55	47.9
GRU RNN[20]	95.72	98.65	97.06	10.01
All layer IDS	98.811	98.42	99.97	0.02
Application layer IDS	99.67	99.02	99.83	0.0016
Transport layer IDS	99.81	99.10 9	99.1	0.1
Network layer IDS	99.67	99.38	99.83	0.16

V. CONCLUSIONS

Our studies are very much about the analysis of a new domain that is the Internet of Things with the presentation of architecture and neural networks. This research focused on the processing of IoT elements where processing power is low with data size that is not as huge. This interdisciplinary research is novel in a way that, it has applied deep learning methods for IoT security. We have proposed light-weight architecture for an Intrusion Detection System (IDS) in an IoT network. Based on TCP/IP layer architecture and the attack types at each layer, we have suggested placing IDS classifiers at each layer. This has reduced the data set size at each classifier and improved the performance in terms of accuracy, recall, training time and false alarm rate. We have applied deep learning algorithms to classify the data at each IDS classifier. This approach has achieved outstanding results with better results than existing work in the literature. Moreover, we have used the full KDD 99'cup 22% data set for the

experiments, unlike previous research work. The training time of Transport Layer IDS, Application Layer IDS and Network Layer IDS is almost half of the All Layer IDS which is important for dynamic IoT networks. The accuracy and false alarm rate of All-Layer IDS is 98.91% and 0.76% respectively which outperformed all other existing IDS classifiers in literature. As the IoT deals with user's personal data and industry's information, it is crucial to implement robust solutions to protect from security threats. This can be possible with the concepts of machine learning and deep learning as IoT generate a humongous amount of heterogeneous data. We have applied Gated-Recurrent-Unit neural networks to the dataset. However, there are many improvised versions of recurrent neural networks such as Dynamic RNN, Bi-Directional RNN which can achieve better performances than basic GRU cells. One can also build a hybrid network using convolutional neural networks and recurrent neural networks to deal with multi-modal data. And here we are at the end and this research that was focused on data security; we say here that our goal was achieved given the end result which was satisfactory. We say that our research has positive results and exceeded the capacity levels of all existing work. We will continue to deepen our knowledge and the suggestions of everyone are welcome.

ACKNOWLEDGEMENT

In terms of gratitude, I first thank my God and my family (my father and mother) for this life that gave me. This research is the result of enormous support from my dear teacher MAYAO. I am thankful for his humble and simple personality, I would like to thank him with all my heart for all the sacrifices, directions, understanding and advice despite the language that does not allow us to communicate well but my teacher was always present for me. I thank all the teachers of my department and those who taught me the Chinese language for their support and encouragement. I am also grateful to all the friends who helped me a lot and motivated me to reach my goal, especially I would like to thank Jean Marie Cimula and Miguel Kakanakou for their love, motivation and encouragement. I think that the man must have the hard spirit to support the realities of life and the determination to accomplish his goals.

REFERENCES

- [1] Xu, K., Wang, X., Wei, W., Song, H., & Mao, B. (2016). Toward software defined smart home. *IEEE Communications Magazine*. Vol. 54(5), pp. 116-122
- [2] Pan, G., Qi, G., Zhang, W., Li, S., Wu, Z., & Yang, L. T. (2013). Trace analysis and mining for smart cities: issues, methods, and applications. *IEEE Communications Magazine*. Vol. 51(6). Pp.120- 126.
- [3] Luo, X., Liu, J., Zhang, D., & Chang, X. (2016). A large-scale web QoS prediction scheme for the Industrial Internet of Things based on a kernel machine learning algorithm. *Computer Networks*. Vol. 101. Pp. 81-89.
- [4] Hasan, M. A. M., Nasser, M., Ahmad, S., & Molla, K. I. (2016). Feature Selection for Intrusion Detection Using Random Forest. *Journal of Information Security*. Vol. 7(03). Pp. 129.
- [5] Li, Y., & Guo, L. (2007). An active learning based TCM-KNN algorithm for supervised network intrusion detection. *Computers & security*. Vol. 26(7). Pp. 459-467.
- [6] Breiman, L., 2001. Random forests. *Machine Learning*. Vol. 45 (1). Pp. 5–32.
- [7] Dua, S., & Du, X. (2016). Data mining and machine learning in cybersecurity. CRC press.
- [8] Alpaydin, E. (2014). Introduction to machine learning. MIT press
- [9] Lichodziejewski, P., Zincir-Heywood, A., & Heywood, M. (2002). Host-based intrusion detection using self-organizing maps. Hawaii, USA. Vol. 2. Pp. 1-29
- [10] S. Harris. (2015). "All in one CISSP Exam Guide," McGraw-Hill. Vol. 7. Pp. 145-467
- [11] A. A. Shah, M. S. H. Kiyhal, M.D. Awan. 2015. "Analysis of Machine Learning Techniques for Intrusion Detection System: A Review," *International Journal of Computer Applications*, Vol. 119. Pp. 19-23. Article (CrossRef Link)
- [12] S. Juma, Z. Muda, M. A. Mohamed, W. Yassin. (2015). "Machine learning techniques for intrusion detection system: a review," *Journal of Theoretical and Applied Information Technology*. Vol. 72. Pp. 422-429
- [13] Hanahan D, Weinberg RA. (2011). Hallmarks of cancer: the next generation. Vol. 144. Pp. 646–74.
- [14] Polley M-YC, Freidlin B, Korn EL, Conley BA, Abrams JS, McShane LM. (2013). Statistical and practical considerations for clinical evaluation of predictive biomarkers. *J Natl Cancer Inst*. Vol. 105. Pp. 1677–83.
- [15] Kim, J., & Kim, H. (2015, August). Applying Recurrent Neural Network to Intrusion Detection with Hessian Free Optimization. In *International Workshop on Information Security Applications*. Vol. 3. Pp. 357-369. Springer, Cham.
- [16] L. You, Y. Li, Y. Wang, J. Zhang, and Y. Yang. (2016). "A deep learningbased RNNs model for automatic security audit of short messages," *International Symposium on Communications and Information Technologies (ISCIT)*. Qingdao, China: IEEE. Vol. 16488389. Pp. 225–229.
- [17] Tesauro, G., Touretzky, D., & Leen, T. (1993) Comparing the Prediction Accuracy of Artificial Neural Networks and Other Statistical Models for Breast

Cancer Survival. *Advances in Neural Information Processing Systems*, Vol. 7. Pp. 1063--1067. The MIT Press.

- [18] K. Alrawashdeh and C. Purdy. (2016). "Toward an Online Anomaly Intrusion Detection System Based on Deep Learning," *IEEE International Conference on Machine Learning and Applications (ICMLA)*. Anaheim, California, USA: IEEE. Pp. 195–200.
- [19] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho. (2016). "Deep learning approach for network intrusion detection in software defined networking," *International Conference on Wireless Networks and Mobile Communications (WINCOM)*. Pp. 258–263.
- [20] M.-J. Kang and J.-w. Kang. (2016). "Intrusion Detection System Using Deep Neural Network for In-Vehicle Network Security". Vol. 11. Pp. e0155781.
- [21] Q. Niyaz, W. Sun, and A. Y. Javaid. 2016. "A deep learning based ddos detection system in software-defined networking (SDN)". Vol. abs/1611.07400. [Online]. Available: <http://arxiv.org/abs/1611.0740>