A Convergence Time Predictive Model using Machine Learning for LLN

Sakshi Garg
Department of CSE
Amity University, Uttar Pradesh
Noida, India
sakshijyotigarg@gmail.com

Deepti Mehrotra
Department of IT
Amity University, Uttar Pradesh
Noida, India
mehdeepti@gmail.com

Sujata Pandey
Department of ECE
Amity University, Uttar Pradesh
Noida, India
spandey@amity.edu

Hari Mohan Pandey
Department of CS
Edge Hill University
Ormskirk, United kingdom
profharimohanpandey@gmail.com

***Abstract*— The need to interface Low power and Lossy Network (LLN) to the web acquired notoriety with the rise of Internet of Things (IoT). Accordingly, IETF ROLL working group proposed a de-facto IPv6 routing protocol called RPL. RPL provisions 6LoWPAN (IPv6 over Low power and Wireless Personal Area Network) and has been the profound interest among researchers, primarily because of its flexibility to cope with the topology changes and its ability to auto-configure, detect and avoids loops. Since, motes that are deployed in IoT network are battery driven and lossy in nature, network execution is strongly influenced. Consequently, if the network convergence for scalable network can be foreseen, it can be utilized to upgrade the network performance. The idea behind this article is to propose a predictive model that gauges Convergence Time (CT) by performing feature selection by utilizing Machine Learning (ML) strategy for RPL and IoTorii. IoTorii is another such convention proposed in recent literature that scales network better. RPL and IoTorii protocols are simulated on Contiki OS/Cooja simulator using Sky motes. Further, RPL execution precision is tried for Storing and Non-Storing modes both. Similarly, IoTorii performance accuracy is tested for both of its proposed variations: nHLMAC1 and nHLMAC3 addresses. Additionally, the network parameters obtained from the simulation are used for feature selection in predictive modelling. The experiment shows that the prediction model gives the best forecast with 93.619%, 96.962%, 93.112% and 92.635% accuracy for both the protocols with different modes and addresses respectively.**

Keywords—RPL, LLN, IoT, IoTorii, Convergence Time, Linear Regression.

I. INTRODUCTION

Communication networks [1] are continually advancing. Internet of Things (IoT) [2][3] is essentially one such communication network that permits physical entities and individuals to connect, gather and trade information. This interminable transformation has created a riddle comprising of various sorts of gadgets, topologies, conventions, protocols and administrations. Many relevant solutions like SDN [4], NDN [5], cloud and fog computing [6], content forwarding and caching schemes [7] are proposed in literature to optimize the network, yet this paradigm isn't for the most part appropriate to Low power and Lossy Networks (LLNs), because of energy and memory constraints. Therefore, LLNs are ordinarily utilized with IoT paradigm to accumulate and handle information. This will endow network performance manifold and aid in evolving pioneering applications in plenty of realms, including smart cities [8][9] and Internet of Vehicles (IoV) [10].

RPL is the standard protocol introduced by (Internet Engineering Task Force (IETF) Routing Over Low power and Lossy networks (ROLL) working group in 2012 to address these LLN issues [11]. Although, RPL still needs advancement to discourse scalability, mobility, energy Quality of Service (QoS) and network QoS requirements. Many researchers have proposed RPL optimization to enhance network performance [12][13][14], some academicians [15][16][17] have contemplated a redesign of RPL based on its contemporary and forthcoming concerns, while some authors [18][19] have recommended parallel protocols that claim to outperform RPL in light of these issues.

In this article, a predictive model is proposed to estimate the network Convergence Time (CT) for RPL and IoTorii, where IoTorii [19] is another proposed protocol in literature which is also considered to test our model accuracy apart from RPL. A predictive model like such is indispensable to comprehend the network performance beforehand and optimize the network accordingly for enhanced network execution and QoS. This paper realizes the concept of Linear Regression (LR): a Machine Learning (ML) strategy, to build the predictive model by performing feature selection. The contribution is not just limited to exploring the protocols in their native forms but, RPL is simulated for both Storing and Non-Storing modes and the proposed IoTorii protocol is considered with both nHLMAC1 and nHLMAC3 addresses independently. The results justifies the use of the proposed predictive model with 93.619%, 96.962%, 93.112% and 92.635% accuracy.

The paper is further organized as follows: the related work is outlined in Section II, the details of the dataset are discussed in Section III. Section IV discusses the methodology and Section V shows the findings of the study. Finally, the study is concluded with its future scope in Section VI.

II. RELATED WORKS

Literature recommends the use of ML techniques in RPL assessment. Authors in paper [20] proposed a multi-layer perceptron model to study the effect of transmission power on Energy Consumption (EC) and to estimate the optimal transmission range for LLN using RPL. In [21] authors proposed the use of a reinforcement learning model to estimate the link quality and minimize communication overhead caused by the nodes mobility in RPL for IoT networks. Their results showed an improvement in Packet Loss Rate (PRR). Authors in [22] suggested the use of fuzzified metric to select the best parent using machine learning. Their model gave 89% accuracy using random forest

978-1-5386-5541-2/18/$31.00 ©2018 IEEE

classifier. Paper [23] incorporated objective function based on learning automata with RPL to yield Expected Transmission count (ETX) and showed improvement in PRR, EC and communication overhead.

Most of the papers in literature emphasized on using ML techniques for improvising RPL privacy and security. Like, authors in [24] proposed the use of k-nearest neighbor classifier to detect rank attack on RPL network. Similarly, paper [25][26] detected wormhole and rank attacks on RPL in IoT networks using ML techniques. Additionally, authors in [27][28][29] reviewed the assessment of network Intrusion Detection System (IDS) for RPL and their implementation was examined using ML techniques. Many more studies [30][31][32][33] can be found in state-of-art that recommended the use of ML strategies for the identification, evaluation and authentication of diverse attacks on the network using RPL protocol.

These works aimed at refining the RPL performance and mostly privacy and security of RPL but very limited studies could be uncovered that discussed network performance of RPL. This paper conceals this gap and addresses the network performance estimation using predictive model by performing feature selection of network QoS parameters.

III. DATASET

This research have used two kinds of data. Both the dataset contains network QoS parameters. The network multi-variate parameters include CT, Convergence Rate (CR), Hop Count (HC), communication overhead and network density. Data is accumulated for increasing network size from 5 nodes to 200 nodes. Sky mote is configured within Cooja simulation for use within Wireless Sensor Network (WSN). This dataset [34] can be downloaded from IEEE DataPort. An instance of dataset is shown in Fig. 1.

	Nodes	CT	CR	HC	Overhead
0	2	0.019058	1.000000	1.000000	1.500000
1	5	0.198915	1.000000	1.600000	2.000000
2	10	0.309979	0.974242	2.523556	1.980000
3	15	0.344202	0.890559	2.590857	1.993333
4	20	0.396850	0.867483	2.837263	1.995000
5	25	0.418444	0.856060	2.936333	2.000000
6	50	0.692738	0.681638	5.333739	1.983800
7	100	0.825442	0.414542	5.694750	2.000000
8	200	1.383566	0.294715	10.334767	1.994000

Fig. 1. Instance of the pre-processed and pre-cleaned dataset

A. RPL Dataset

RPL dataset is simulated over Cooja simulator on Contiki OS using Sky mote under two modes: Storing and Non-Storing. RPL in storing mode consumes high energy with the increase in network size, while RPL in Non-Storing mode generates high communication overhead with the increase in network size.

B. IoTorii Dataset

IoTorii dataset is also simulated over Cooja simulator on Contiki OS using Sky mote but using two distinct addresses: nHLMAC1 and nHLMAC3. The use of these addresses to develop IoTorii protocol is explained in [19]. nHLMAC1 means one Hierarchical Local Media Access Control (HLMAC) address can be assigned to a node at most, whereas maximum of 3 HLMAC addresses can be assigned to each use using nHLMAC3.

IV. METHODOLOGY

In order to build the predictive model, the steps followed are shown in Fig. 2. The data is explored using the considered dataset. The data is then pre-processed and cleaned to clear the NULL values and fetch the relevant data. Further, the relationship of rest of the network QoS parameters is studied with respect to CT and the most significant features are selected for each dataset. Then, a predictive model is developed for the obtained data based on the problem statement and the model is evaluated for its accuracy and performance.

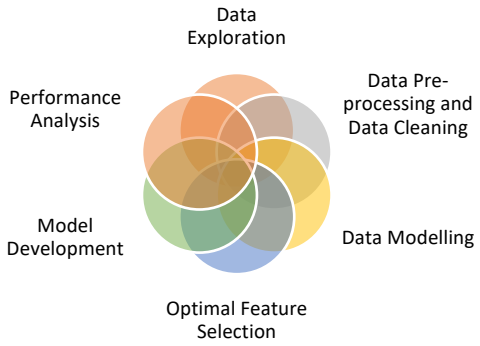


Fig. 2. Predictive Modelling Process

A. Problem Statement

It is indeed that the increase in network lifetime and overall network execution for a protocol with increasing network size, will elevate the performance and QoS for the network, if foreseen and optimized. Limited availability of such predictive model to test the network can be clearly analyzed from the literature. Hence, it breeds the necessity to develop a predictive model that can optimize network performance and offer QoS with high accuracy.

B. Proposed Model

This article commends the use of proposed predictive model as a solution to pre-determine the network performance and accordingly optimize it for superior execution and high QoS. The basic notion behind developing this model is the use of ML strategy, because of its ability to process big-data. Since, the network density is dynamic and capricious, the size of data cannot be pre-accessed. Thus, this model uses multi-variate LR technique by selecting optimal features to build the model with high accuracy.

1) Linear Regression (LR): It is a supervised ML algorithm that is used for predictive analysis. It is a statistical method for continuous data and shows a linear relation between the target/ dependent variable (y) and one or more predictor/ independent variables (x1, x2, … xn). The general equation for multi-variate LR line is:

$$y = m_1x_1 + m_2x_2 + ... + m_nx_n + c \quad (1)$$

where, m1,m2, … mn are coefficients and c is the intercept. The motive is to find the best values of coefficients

an d intercept that fits the LR line best. It can be graphically represented as in Fig. 3.

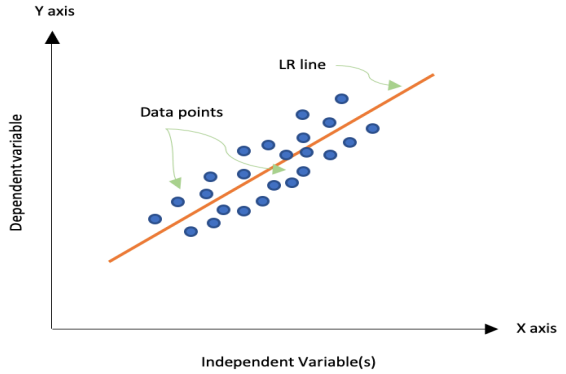


Fig. 3. Line of Linear Regression Represenation

The cost function is calculated using Root Mean Square Error (RMSE) method, which is used to minimize the difference between the actual value of y and predicted value of y. It is expressed as:

$$\text{RMSE} = \sqrt{1/n \sum (y_i - \hat{y}_i)^2} \quad (2)$$

where, yi is the actual value, y'i is the predicted value, for i = 1 to n and n are the total number of observations. The R2 value predicts the model performance and determines the goodness of fit. The range of R2 is 0 to 1 and more the value of R2 is closer to 1, better is the goodness of fit by the model. It is estimated as:

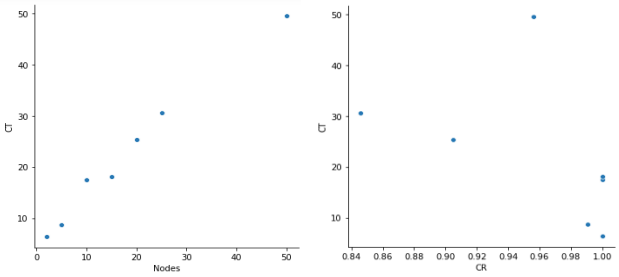
$$R^2 = \textit{Explained Variation} / \textit{Total Variation} \quad (3)$$

V. Results And Discussions

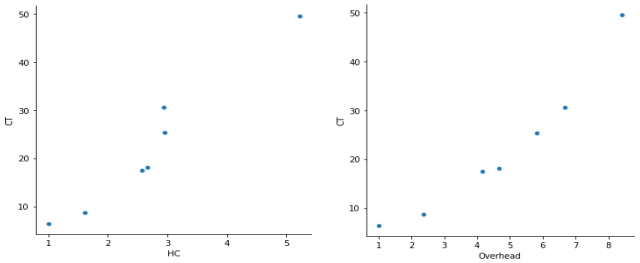
The predictive model is implemented using python script on Anaconda navigator in Jupiter Notebook. The *.xlxs* files of the dataset are pre-processed and converted to readable *.csv* files. The results are separately discussed for both datasets.

A. RPL Dataset

1) RPL in Storing mode: The data modelling graphs of CT with respect to other network QoS parameters are shown in Fig. 4. It is seen that Nodes/ Network density, HC and Overhead has a positive LR line, whereas CR has a negative LR line. Further, feature selection is performed and nodes, CR, HC and overhead are selected as optimal parameters. The coefficient values of each parameters are 1.159, 1.154, -4.023, 0.898 respectively and the intercept value is 8.382.



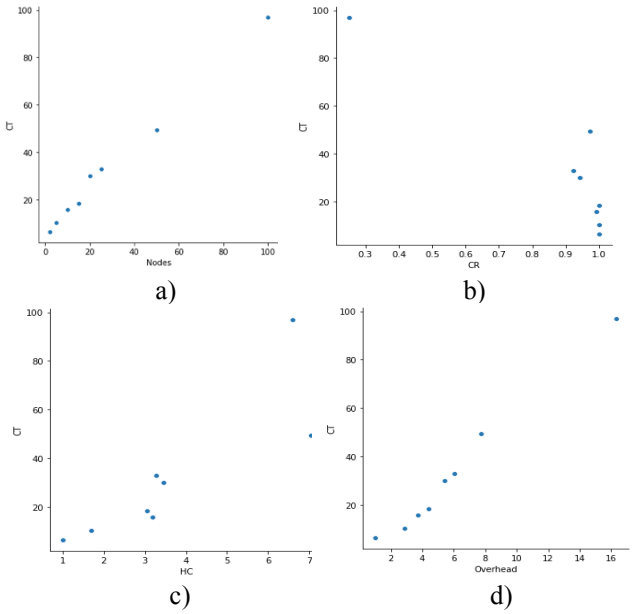
a) b)



c) d)

Fig. 4. Data Modelling of CT with respect to a) Nodes b) CR c) HC and d) overhead for RPL in Storing mode.

2) RPL in Non-Storing mode: The data modelling graphs of CT with respect to other network QoS parameters are shown in Fig. 5. It is seen that Nodes/ Network density, HC and overhead has a positive LR line, whereas CR has a negative LR line. Further, feature selection is performed and nodes, CR, HC and overhead are selected as optimal parameters. The coefficient values of each parameters are 0.571, 0.672, 0.027, 2.319 respectively and the intercept value is 1.733.



a) b)

c) d)

Fig. 5. Data Modelling of CT with respect to a) Nodes b) CR c) HC and d) overhead for RPL in Non-Storing mode.

B. IoTorii Dataset

1) IoTorii with nHLMAC1 address: The data modelling graphs of CT with respect to other network QoS parameters are shown in Fig. 6. It is seen that Nodes/ Network density and HC has a positive LR line, whereas CR has a negative LR line and overhead shows an exponential graph. Further, feature selection is performed and nodes, CR and HC are selected as optimal parameters. The coefficient values of each parameters are -0.001, -0.315, 0.138 respectively and the intercept value is 0.273.

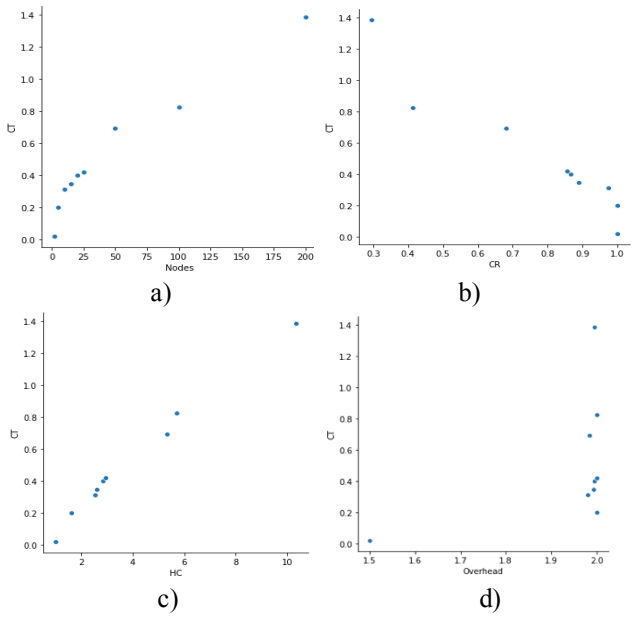


Fig. 6. Data Modelling of CT with respect to a) Nodes b) CR c) HC and d) overhead for IoTorii with nHLMAC1 address

2) IoTorii with nHLMAC3 addresses: The data modelling graphs of CT with respect to other network QoS parameters are shown in Fig. 7. It is seen that Nodes/ Network density and HC has a positive LR line, whereas CR has a negative LR line and overhead shows an exponential graph. Further, feature selection is performed and nodes, CR and HC are selected as optimal parameters. The coefficient values of each parameters are -0.005, -1.477, 0.211 respectively and the intercept value is 1.542.

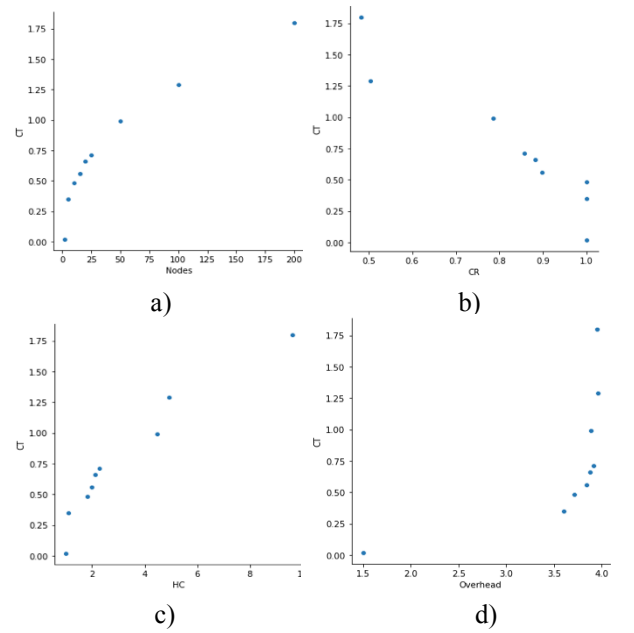


Fig. 7. Data Modelling graphs of CT with respect to a) Nodes b) CR c) HC and d) overhead for IoTorii with nHLMAC3 addresses

The accuracy score for the following study is shown in tabular form in Table 1.

TABLE I. ACCURACY SCORE OF THE PREDICTION MODEL

DATASET		ACCURACY SCORE (%)
RPL Dataset	Storing mode	93.619 %
	Non-Storing mode	96.962 %
IoTorii Dataset	nHLMAC1 address	93.112 %
	nHLMAC3 addresses	92.635 %

This table fairly explains the use of the proposed predictive model to estimate convergence time of the network by optimally selecting the dependent QoS network parameters and improving the overall network lifetime, network performance and offering an advanced network QoS with high precision. The findings also show that the proposed predictive model is equivalently decent with the other RPL like protocol (IoTorii) suggested in literature for LLNs for IoT network.

VI. CONCLUSION AND FUTURE SCOPE

This article evidently underlines the need to develop a predictive model for the IoT network which considers network QoS parameters. Predictive analysis is used with LR technique of ML to build the predictive model. The model is tested for two protocols to effectively justify the performance of the proposed model. The protocols are also simulated for different modes and addresses. The proposed model selects distinct optimal features for both protocols. The accuracy score for the proposed model is equitably great to justify its use for future network CT based predictions and not only for RPL but another protocols like IoTorii too. The average precision score for RPL dataset is 95.291%, while average accuracy score for IoTorii dataset is 92.874%. It can also be inferred that the proposed model performs marginally better for RPL protocol than IoTorii protocol. In future, this system can be trained for RPL and IoTorii data that considers mobile nodes and more network QoS parameters like Packet Delivery Ratio (PDR), EC and latency. Moreover, this work can be extended for IoV network which can prove to be an edge-cutting development in this domain.

REFERENCES

[1] Ramezani, P., & Jamalipour, A. (2017). Toward the evolution of wireless powered communication networks for the future Internet of Things. *IEEE network*, *31*(6), 62-69.

[2] Bagaa, M., Taleb, T., Bernabe, J. B., & Skarmeta, A. (2020). A machine learning security framework for iot systems. *IEEE Access*, *8*, 114066-114077.

[3] Abdelmoneem, R. M., Benslimane, A., & Shaaban, E. (2020). Mobility-aware task scheduling in cloud-Fog IoT-based healthcare architectures. *Computer Networks*, *179*, 107348.

[4] Tariq, A., Rehman, R. A., & Kim, B. S. (2020). EPF—An Efficient Forwarding Mechanism in SDN Controller Enabled Named Data IoTs. *Applied Sciences*, *10*(21), 7675.

[5] Wang, X., & Cai, S. (2020). Secure healthcare monitoring framework integrating NDN-based IoT with edge cloud. *Future Generation Computer Systems*, *112*, 320-329.

[6] Sarangi, A. K., Mohapatra, A. G., Mishra, T. C., & Keswani, B. (2021). Healthcare 4.0: A voyage of fog computing with iot, cloud computing, big data, and machine learning. In *Fog Computing for Healthcare 4.0 Environments* (pp. 177-210). Springer, Cham.

[7] Din, I. U., Hassan, S., Almogren, A., Ayub, F., & Guizani, M. (2020). PUC: Packet update caching for energy efficient IoT-based information-centric networking. *Future Generation Computer Systems*, *111*, 634-643.

[8] Hu, L., & Ni, Q. (2017). IoT-driven automated object detection algorithm for urban surveillance systems in smart cities. *IEEE Internet of Things Journal*, *5*(2), 747-754.

[9] Garg, S., Mehrotra, D., Pandey, S., & Pandey, H. M. (2021). Network efficient topology for low power and lossy networks in smart corridor design using RPL. *International Journal of Pervasive Computing and Communications*.

[10] Garg, S., Mehrotra, D., Pandey, H. M., & Pandey, S. (2021). Accessible review of internet of vehicle models for intelligent transportation and research gaps for potential future directions. *Peer-to-Peer Networking and Applications*, 1-28.

[11] Winter, T., Thubert, P., Brandt, A., Hui, J. W., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, J.P. & Alexander, R. K. (2012). RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks. *rfc*, *6550*, 1-157.

[12] Bendouda, D., & Haffaf, H. (2019, June). QFM-MRPL: Towards a QoS and Fault Management based of Mobile-RPL in IoT for mobile applications. In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)* (pp. 354-359). IEEE.

[13] Thapar, P., & Batra, U. (2018). Implementation of ant colony optimization in routing protocol for internet of things. In *Innovations in Computational Intelligence* (pp. 151-164). Springer, Singapore.

[14] Garg, S., Mehrotra, D., & Pandey, S. (2022). A Study on RPL Protocol with Respect to DODAG Formation Using Objective Function. In *Soft Computing: Theories and Applications* (pp. 633-644). Springer, Singapore.

[15] Gaddour, O., Koubäa, A., Rangarajan, R., Cheikhrouhou, O., Tovar, E., & Abid, M. (2014, June). Co-RPL: RPL routing for mobile low power wireless sensor networks using Corona mechanism. In *Proceedings of the 9th IEEE international symposium on industrial embedded systems (SIES 2014)* (pp. 200-209). IEEE.

[16] Fotouhi, H., Moreira, D., & Alves, M. (2015). mRPL: Boosting mobility in the Internet of Things. *Ad Hoc Networks*, *26*, 17-35.

[17] Bouaziz, M., Rachedi, A., & Belghith, A. (2017, January). EC-MRPL: An energy-efficient and mobility support routing protocol for Internet of Mobile Things. In *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)* (pp. 19-24). IEEE.

[18] Yang, Z., Ping, S., Sun, H., & Aghvami, A. H. (2016). CRB-RPL: A receiver-based routing protocol for communications in cognitive radio enabled smart grid. *IEEE Transactions on Vehicular Technology*, *66*(7), 5985-5994.

[19] Rojas, E., Hosseini, H., Gomez, C., Carrascal, D., & Cotrim, J. R. (2021). Outperforming RPL with scalable routing based on meaningful MAC addressing. *Ad Hoc Networks*, *114*, 102433.

[20] Aboubakar, M., Kellil, M., Bouabdallah, A., & Roux, P. (2020, April). Using machine learning to estimate the optimal transmission range for RPL networks. In *NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium* (pp. 1-5). IEEE.

[21] Ancillotti, E., Vallati, C., Bruno, R., & Mingozzi, E. (2017). A reinforcement learning-based link quality estimation strategy for RPL and its impact on topology management. *Computer Communications*, *112*, 1-13.

[22] Gopika, D., Majumder, P., & Kumar, P. J. (2020, November). FML: Fuzzification with Machine Learning based Parent Node Selection in RPL/6LoWPAN. In *2020 2nd PhD Colloquium on Ethically Driven Innovation and Technology for Society (PhD EDITS)* (pp. 1-2). IEEE.

[23] Saleem, A., Afzal, M. K., Ateeq, M., Kim, S. W., & Zikria, Y. B. (2020). Intelligent learning automata-based objective function in RPL for IoT. *Sustainable Cities and Society*, *59*, 102234.

[24] Neerugatti, V., & Mohan Reddy, A. R. (2019). Machine learning based technique for detection of rank attack in RPL based internet of things networks. *Machine Learning Based Technique for Detection of Rank Attack in RPL based Internet of Things Networks (July 10, 2019). International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN*, 2278-3075.

[25] Jhanjhi, N. Z., Brohi, S. N., & Malik, N. A. (2019, December). Proposing a rank and wormhole attack detection framework using machine learning. In *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)* (pp. 1-9). IEEE.

[26] Yahyaoui, A., Yaakoubi, F., & Abdellatif, T. (2020, June). Machine Learning Based Rank Attack Detection for Smart Hospital Infrastructure. In *International Conference on Smart Homes and Health Telematics* (pp. 28-40). Springer, Cham.

[27] Verma, A., & Ranga, V. (2019, April). ELNIDS: Ensemble learning based network intrusion detection system for RPL based Internet of Things. In *2019 4th International conference on Internet of Things: Smart innovation and usages (IoT-SIU)* (pp. 1-6). IEEE.

[28] Farzaneh, B., Koosha, M., Boochanpour, E., & Alizadeh, E. (2020, April). A new method for intrusion detection on RPL routing protocol using fuzzy logic. In *2020 6th International Conference on Web Research (ICWR)* (pp. 245-250). IEEE.

[29] da Costa, K. A., Papa, J. P., Lisboa, C. O., Munoz, R., & de Albuquerque, V. H. C. (2019). Internet of Things: A survey on machine learning-based intrusion detection approaches. *Computer Networks*, *151*, 147-157.

[30] Osman, M., He, J., Mokbal, F. M. M., Zhu, N., & Qureshi, S. (2021). ML-LGBM: A Machine Learning Model based on Light Gradient Boosting Machine for the Detection of Version Number Attacks in RPL-Based Networks. *IEEE Access*.

[31] Cakir, S., Toklu, S., & Yalcin, N. (2020). RPL attack detection and prevention in the Internet of Things networks using a GRU based deep learning. *IEEE Access*, *8*, 183678-183689.

[32] Foley, J., Moradpoor, N., & Ochenyi, H. (2020). Employing a Machine Learning Approach to Detect Combined Internet of Things Attacks against Two Objective Functions Using a Novel Dataset. *Security and Communication Networks*, *2020*.

[33] Kannimuthu, P., & Thangamuthu, J. (2021). Decision Tree Trust (DTTrust)-Based Authentication Mechanism to Secure RPL Routing Protocol on Internet of Battlefield Thing (IoBT). *International Journal of Business Data Communications and Networking (IJBDCN)*, *17*(1), 1-23.

[34] Elisa Rojas, Hedayat Hosseini, Carles Gomez, David Carrascal, Jeferson Rodrigues Cotrim, April 16, 2020, "IoTorii: Outperforming RPL with scalable routing", IEEE Dataport, doi: https://dx.doi.org/10.21227/cjw4-kr75.