Leveraging Q-Learning For Proactive Security Against Adversarial Band Jamming In Wireless Networks

Rahul Choudhary
Assist.Professor, Dept of IT
SVCE, Indore

Abstract- *Wireless Networks are being used in several applications these days such as large scale industries, chemical plants, underground mines, disaster management, military and defense etc. However, due to the large scale of the network and wireless data transfer, the data transmission is often prone to attacks. In this context, Physical Layer Security (PLS) has emerged as an attractive solution for securing wireless transmissions by exploiting the wireless channel characteristics.This work presents a Deep Reinforcement Learning Based Approach for frequency hopping mechanism to ensure security to Wireless Sensor Networks. The evaluation parameters chosen are Average Reward, BER and outage probability. It can be observed from the results that as the spreading factor increases, the BER also increase showing a compromise between security and errors. It has been shown that the proposed system achieves lower BER and outage probability compared to previously existing techniques.*

Keywords- Deep Learning, Deep Reinforcement Learning, Wireless Networks, Spreading Factor, Outage Probability

I. INTRODUCTION

The domain of wireless sensor networks has been vast and huge off late. In the field of wireless communication, WSN has been of great prominence. The utility of the wireless sensors has been enormous in the technologically driven world [1]. The sensors are now being deployed for multiple uses and practical purposes. But with the rampant consumption, the power consumption scenario also raises concern. The parameter of consumption of power by the wireless sensor nodes impacts the over all functioning greatly [2]. So for effective and accurate functioning, the power consumption needs to be checked. Henceforth various research studies are presented on the saving of power pertinent to wireless sensor nodes to bring major improvements in the network lifetime. So this is important for robust and efficient system. [3].

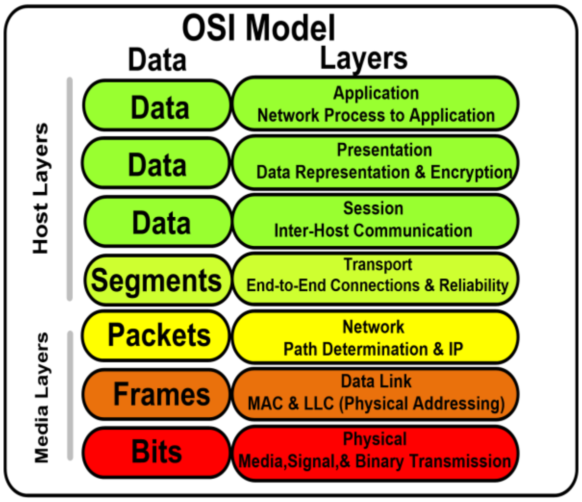

Fig.1 The OSI model

Physical layer security is the security mechanisms adopted in any network where the security patch is implemented directly to act on the bit level or binary transmission level [4].

It is much more secure compared to application level security patches since attackers can bypass upper layers.Physical layer security is also called bit level security [5]. For ad-hoc networks or wireless sensor networks, bit level security is the most secure mechanism to thwart attacks [6].

II. METHODOLOGY

The proposed solution presented in this work can be summarized in the following steps [7]:

1. Design a wireless sensor network.
2. Design a deep reinforcement learning based reward-punishment approach for generation of random frequencies.
3. Implement Frequency Hopping (FH) for bit level transmission.
4. Analyze reinforcement learning parameters.
5. Increase the spreading factor and analyze the effect on the error rate.

6. Compute Bit Error Rate (BER) and Outage Probability for the system.

The fundamental aspect is to first consider machine learning as a tool. Machine Learning is the design of algorithms which can mimic the human thinking process or intelligence. There are several techniques to implement machine learning algorithms such as decision trees, support vector machine (SVM), Fuzzy Logic, Neural Networks etc [8]. Off late, the focus has shifted on deep neural networks due to the fact that they highly resemble the deep layered structure of the human brain and for their extremely fast computation. The use of deep neural networks to train algorithms is termed as deep learning. Typically machine learning algorithms are categorized as [9]:

1. Unsupervised Learning
2. Supervised Learning
3. Semi-Supervised Learning

Figure 2 depicts the difference between artificial neural networks and deep neural networks.

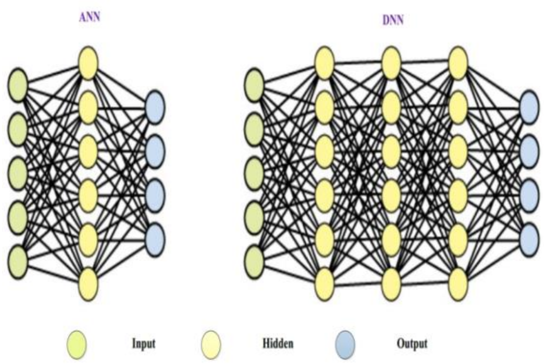

Fig. 2 Artificial Neural Networks and Deep Neural Networks

Reinforcement Learning (RL) can be considered as special category of supervised learning on optimized data. RL is about taking suitable action to maximize reward in a particular situation [10]. It is employed by various software and machines to find the best possible behavior or path it should take in a specific situation. RL is different from contemporary supervised learning in the way that supervised learning the training data has the answer key with it so the model is trained with the correct answer itself [11] Whereas in RL, there is no answer but the reinforcement agent decides what to do to perform the given task. Some salient points regarding RL are [12]:

1. Input: The input is an initial state from which the model will start.
2. Output: There are many possible outputs as there are a variety of solutions to a particular problem [13].
3. Training: The training is based upon the input. The model will return a state and the reward or penalty (punishment) will be decided based on its output [14].
4. The model continuously learns till maximizing the reward.
5. The best solution is decided based on the maximum reward [15].

The agent-environment interaction in RL is depicted in figure 3.

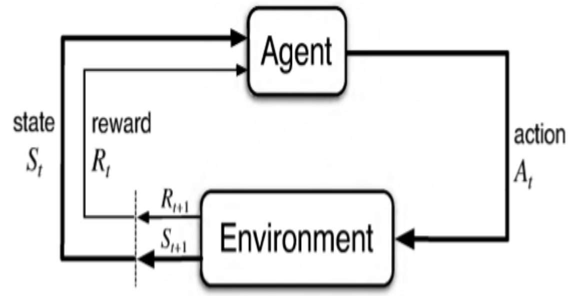

Fig. 3 Agent-environment interaction in RL

Use of RL for Random Frequency Generation:

All binary transmissions in a WSN have some frequency of transmission [16]. Keeping the frequency of transmission static makes the data transmission prone to attacks. To secure data transmission in large WSNs, an apparently changing random frequency synthesizer is needed [17].

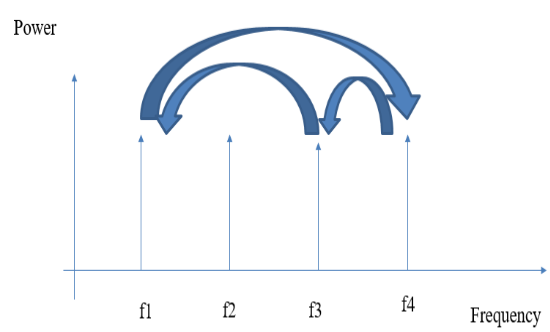

Fig. 4 Concept of Frequency Hopping

Changing the frequency of transmission is also called frequency hopping (FH) [18]. Typically a Pseudo-Random Number Generator (PRNG) is a (usually, deterministic) algorithm which tries to emulate the statistical properties of a sequence of True-Random Numbers (TRNs) [19]. The data sequence generated by a PRNG looks random but is actually deterministic (thus pseudo random). RL can be used as an effective PRNG in the frequency synthesizer [20].

RL is learning what to do in order to accumulate as much reward as possible in a Markov Decision Process (MDP) [21]. The MDP is a random process in which the future states do NOT depend on the previous states but are random in nature [22]. The proposed algorithm to generate a MDP using RL for WSNs is explained subsequently [23]:
For wireless networks, RL can be applied to security tasks such as intrusion detection, jamming resistance, secure routing, and spectrum management. The key strength of RL lies in its ability to adapt to dynamic network conditions and unforeseen attack patterns [24]. Unlike supervised methods that require labeled attack data, RL agents can learn from interactions, making them more robust against novel and stealthy threats. This adaptability is particularly important in mobile ad hoc networks, cognitive radio networks, and IoT environments where attackers constantly change strategies [25].

Q-learning is a model-free RL algorithm that estimates the value (Q-value) of taking a certain action in a given state. Over time, the agent learns an optimal policy that maximizes long-term rewards [26]. In the context of wireless security, Q-learning can be applied for defense strategies against jamming attacks, intrusion detection, authentication, and secure spectrum allocation [27]. For example, a Q-learning-based agent can dynamically switch communication channels in response to jamming signals, thereby maintaining connectivity without predefined rules [28]. The use of RL and Q-learning provides several benefits for wireless network security. First, these approaches are adaptive and scalable, allowing them to respond to new attack strategies without prior knowledge. Second, they reduce computational overhead by learning policies that require less frequent updates compared to traditional monitoring systems. Third, they support distributed learning, which is crucial in large-scale wireless systems such as IoT, where central control is impractical. Moreover, the ability of RL agents to optimize performance under uncertainty makes them highly suitable for environments where attackers exploit unpredictable behaviors.Similarly, Q-learning can help secure routing by enabling nodes to choose paths that minimize the risk of interception or malicious activity [29].

The proposed algorithm is presented next:

Proposed Algorithm:

Start
{

Step.1: *Design a WSN with x=100, y=100 and* $d_{sink} = 150$
Step.2: *Generate random binary data to emulate data transmission.*
Step.3: *Design the PRNG as:*
The agent–environment interaction is made at discrete time-step $t = 0, 1, 2$ ....
At each time-step t, the agent uses the state $S_t \epsilon \boldsymbol{S}$ *given by the environment to select an action,* $A_t \epsilon \boldsymbol{A}$.
The environment answers with a number $R_t \epsilon \boldsymbol{R}$ *called a reward, as well as a next state* S_{t+1}
With increasing iterations, the following sequence is obtained:

$$State\ Space = [\{S_0, A_0, R_0\}, \{S_1, A_1, R_1\}, \dots . \{S_n, A_n, R_n\} \quad (4.7)$$

Step.4: *For all practical cases,* $[S, R, A]$ *are finite sets. For the states to be Markov, the following relation should hold true:*

$$Pr(S_{t+1} = s' : R_{t+1} = r') = Pr\left(S_t = s : R_t = r \overset{t}{\leftarrow} A_t = a\right)$$

Here,
Pr *represents probability*
t *represents present iteration*
$t + 1$ *represents next iteration*
S *represents state*
A *represents action*
R *represents reward*
The condition of equi-probability ensures randomness.
The environment is fed only with the last action, and no other data from the history. This means that, for a fixed policy, the corresponding stochastic process{St}is Markov. This gives the name Markov Decision Process (MDP) to the data (S, A, R, Pr). *Moreover, it is a time-homogeneous Markov process, because p does not depend on t. If the reward function is defined as:*

$$R: (S.f) \underset{A}{\rightarrow} r; f \epsilon F$$

Here,
f *represents the chosen pseudo random frequency*
F *represents the bandwidth*

Step.5: *Through the agent, maximize the reward as:*

$$R_C : max\ E \sum_{t_i}^{t_f} \{R(S_t, A_{t,} f_t)\}$$

Here,

R_C *represents cumulative reward*

E *represents the Expectation or Average operator on Random Variables*

t_i *represents the initial state*

t_f *represents the final state*

Step.6: *Obtain state space and use it for frequency hopping.*

Step.7: *Compute RL Parameters, BER and Outage.*

Stop.

}

The subsequent section discusses the results obtained.

III. EXPERIMENTAL RESULTS

The system has been designed on MATLAB.The results obtained are for the simulations for the designed system which render insight into the performance of the proposed system in terms of the outage probability, the signal to noise ratio and the simulation of the wireless sensor network in terms of the clustering.

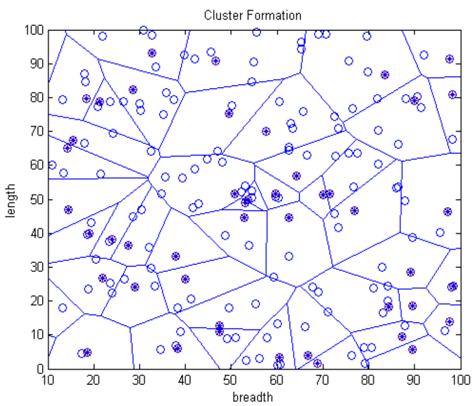

Fig. 5 Formation of Clusters and cluster heads

The above figure depicts the formation of clusters and cluster heads in the network. The dimensions of the network have been chosen as 100mx100m.

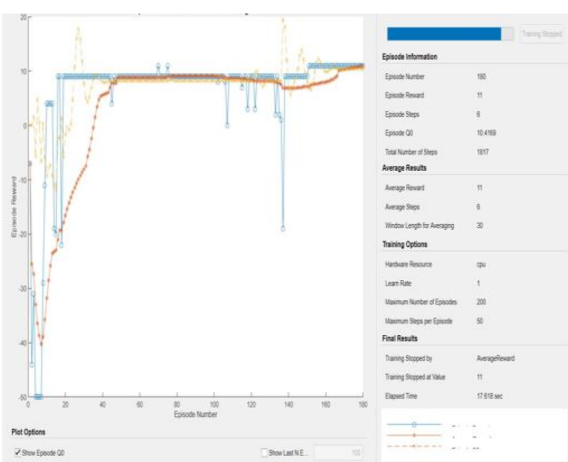

Fig.6 Reinforcement Learning Training Progress

The above graph depicts the reinforcement learning training progress. It can be observed that as the episodes keep increasing, the rewards keep becoming more positive. On the contrary the initial episodes render a negative reward or penalty.

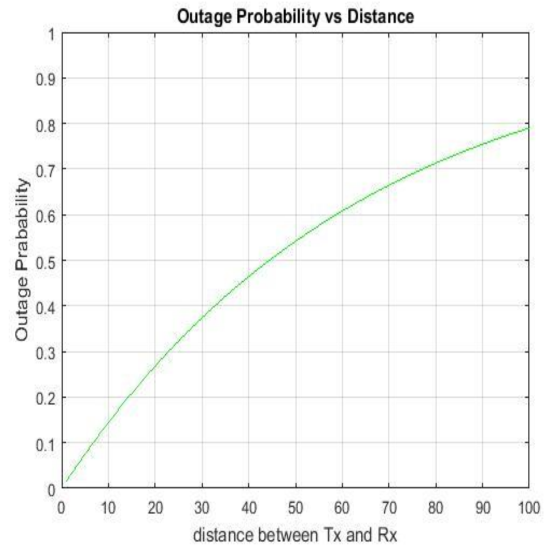

Fig.7 Increase in Outage with Increase in Distance between Tx and Rx

It can be seen from the above graph that the outage probability also depends on the distance between the transmitting end and the receiving end.

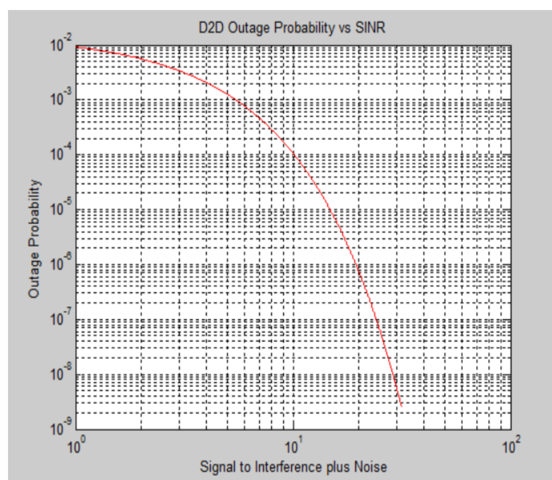

Fig. 8 System Outage Probability

The system outage shows the level of unacceptable quality. In case the system is affected by interference as well as noise effects, then a term called SINR is computed which is the signal to interference plus noise ratio (SINR).

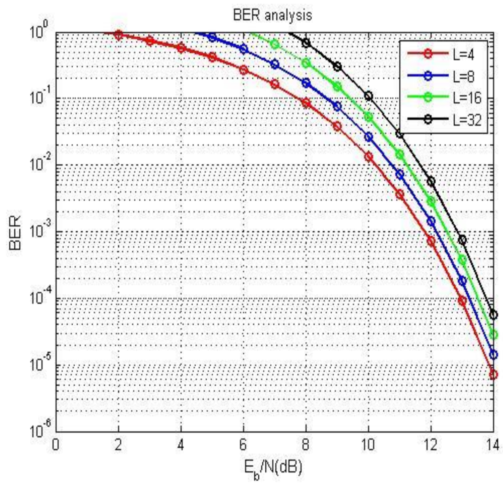

Fig.9BER performance of system

The figure above represents the variation in the BER of the system as a function of (a) SNR designated by Eb/No. It can be seen that as the SNR Increases, the BER decreases. Also as the number of frequencies changed (L) increase, the BER decreases due to the fact that it becomes difficult for the receiver to recover the data.

A summary of results is presented in table 1.

Table.1 Summary of Results

S.No	Parameter	Value
1	Bits in Simulation	10^9
2	Spreading Factor	4-32
3	Channel	Gaussian
4.	ML Model	Reinforcement or Q-Learning
5.	Iterations	100
6	Outage Reached	10^{-7}
7.	BER reached	10^{-6}
8.	Error Rate of Previous Work [25]	10^{-4}
9.	Error Outage of Previous Work [26]	10^{-3}

It can be observed that the proposed work attains lower outage and error rate compared to existing benchmark models in the domain.

IV. CONCLUSION

It can be concluded from the previous discussions that due to the large scale of the network and wireless data transfer, the data transmission is often prone to attacks. As the networks are often ad-hoc in nature with binary transmission, hence it is necessary to employ physical or data link layer security. This work presents a Deep Reinforcement Learning Based Approach for frequency hopping mechanism to ensure security to Wireless Sensor Networks. The reinforcement learning (RL) module is used as the PRNG for frequency hopping. The evaluation parameters chosen are Average Reward, BER and outage probability. The spreading factor has been varied between 4 and 32. It can be observed from the results that as the spreading factor increases, the BER also increase showing a compromise between security and errors. It has been shown that the proposed system achieves lower BER and outage probability compared to previously existing technique.

REFERENCES

[1] A. Albehadiliet al., “Machine Learning-Based PHY-Authentication for Mobile OFDM Transceivers,” in Proc. IEEE VTC 2020-Fall, 2020.

[2] G. Gao, N. Ni, D. Feng, X. Jing and Y. Cao, “Physical Layer Authentication Under Intelligent Spoofing in Wireless Sensor Networks,” Signal Processing, vol. 166, 2020.

[3] L. Liao et al., “Multiuser Physical Layer Authentication in Internet of Things with Data Augmentation,” IEEE Internet of Things J., vol. 7, no. 3, pp. 2077–2088, Mar. 2020.

[4] H. Fang, X. Wang, Z. Xiao and L. Hanzo, “Autonomous Collaborative Authentication with Privacy Preservation in 6G: From Homogeneity to Heterogeneity,” IEEE Network, vol. 36, no. 6, pp. 28–36, Jul. 2022.

[5] R. Xieet al., “A Generalizable Model-and-Data Driven Approach for Open-Set RFF Authentication,” IEEE Trans. Inf. Forensics Security, vol. 16, pp. 4435–4450, Aug. 2021.

[6] A Soni, R Upadhyay, A Jain “Internet of Things and wireless physical layer security: A survey”, Computer Communication, Networking and Internet Security, Springer 2017, pp.115-123.

[7] S. Sullivan, A. Brighente, S. A. P. Kumar and M. Conti, "5G Security Challenges and Solutions: A Review by OSI Layers," in IEEE Access, vol. 9, pp. 116294-116314, 2021

[8] X. Zeng, C. Wang and Z. Li, “CVCA: A Complex-Valued Classifiable Autoencoder for mmWave Massive MIMO Physical Layer Authentication,” presented at IEEE INFOCOM Workshops, 2023

[9] AA Sharifi, M Sharifi, MJM Niya, “Secure cooperative spectrum sensing under primary user emulation attack in cognitive radio networks: Attack-aware threshold selection approach”, vol.70, issue.1, Elsevier 2020.

[10] Syed Hashim Raza Bukhari ,SajidSiraj,Mubashir Husain Rehmani,“ NS-2 based simulation framework for cognitive radio sensor networks”, SPRINGER 2019/.
[11] K. J. PrasannaVenkatesan ,V. Vijayarangan, “Secure and reliable routing in cognitive radio networks”, SPRINGER 2018.
[12] K Gai ,MeikangQiu ,Hui Zhao, “Security-Aware Efficient Mass Distributed Storage Approach for Cloud Systems in Big Data”,IEEE 2017.
[13] JuRen ,Yaoxue Zhang ,Qiang Ye , Kan Yang ; Kuan Zhang ,Xuemin Sherman Shen,” Exploiting Secure and Energy-Efficient Collaborative Spectrum Sensing for Cognitive Radio Sensor Networks”, IEEE 2016.
[14] R.K. Sharma ;,Danda B. Rawat,”Advances on Security Threats and Countermeasures for Cognitive Radio Networks: A Survey”,IEEE
[15] A. Khamaiseh, I. Alsmadi, and A. Al-Alaj, “Deceiving Machine Learning-based Saturation Attack Detection Systems in SDN,” in Proc. IEEE NFV-SDN, 2020.
[16] M. Assis, L. F. Carvalho, J. Lloret, and M. L. Proença Jr., “A GRU Deep Learning System Against Attacks in Software Defined Networks,” J. Network and Computer Applications, vol. 177, p. 102942, 2021.
[17] J. Bhayo et al., “A Time-Efficient Approach TowardDDoS Attack Detection in IoT Network Using SDN,” IEEE Internet of Things J., vol. 9, no. 5, pp. 3612–3630, Mar. 2022.
[18] A. Bahashwan, M. Anbar, S. Manickam, T. Al-Amiedy, M. Aladaileh, and I. H. Hasbullah, “A Systematic Literature Review on Machine Learning and Deep Learning Approaches for Detecting DDoS Attacks in Software-Defined Networking,” Sensors, vol. 23, no. 9, p. 4441, May 2023. ijisae.org.
[19] N. Niknami and J. Wu, “Advanced ML/DL-Based Intrusion Detection Systems for Software-Defined Networks, in Network Security Empowered by Artificial Intelligence, Y. Chen et al., Eds., Adv. in Inf. Security, vol. 107, Springer, Cham, pp. 59-84, Feb.2024.
[20] C. Zhao et al., "Generative AI for Secure Physical Layer Communications: A Survey," in IEEE Transactions on Cognitive Communications and Networking, vol. 11, no. 1, pp. 3-26, Feb. 2024.
[21] B. Ozpoyraz, A. T. Dogukan, Y. Gevez, U. Altun and E. Basar, "Deep Learning-Aided 6G Wireless Networks: A Comprehensive Survey of Revolutionary PHY Architectures," in IEEE Open Journal of the Communications Society, vol. 3, pp. 1749-1809, 2022
[22] M. S. Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, “DDoSNet: A Deep-Learning Model for Detecting Network Attacks,” in Proc. IEEE 21st WoWMoM, Aug. 2020
[23] N. Abuzainab, M. Alrabeiah, A. Alkhateeb and Y. E. Sagduyu, "Deep Learning for THz Drones with Flying Intelligent Surfaces: Beam and Handoff Prediction," 2021 IEEE International Conference on Communications Workshops (ICC Workshops), Montreal, QC, Canada, 2021, pp. 1-6.
[24] S. Alraih, R. Nordin, A. Abu-Samah, I. Shayea and N. F. Abdullah, "A Survey on Handover Optimization in Beyond 5G Mobile Networks: Challenges and Solutions," in IEEE Access, vol. 11, pp. 59317-59345, 2023
[25] Ara and B. Kelley, "Physical Layer Security for 6G: Toward Achieving Intelligent Native Security at Layer-1," in IEEE Access, 2024, vol. 12, pp. 82800-82824.
[26] T. N. Nguyen et al., "Cooperative Satellite-Terrestrial Networks With Imperfect CSI and Multiple Jammers: Performance Analysis and Deep Learning Evaluation," in IEEE Systems Journal, vol. 18, no. 4, pp. 2062-2073, Dec. 2024,
[27] A. Tusha, S. Doğan and H. Arslan, "A Hybrid Downlink NOMA With OFDM and OFDM-IM for Beyond 5G Wireless Networks," in IEEE Signal Processing Letters, 2020, vol. 27, pp. 491-495
[28] C. Miranda, G. Kaddoum, E. Bou-Harb, S. Garg and K. Kaur, "A Collaborative Security Framework for Software-Defined Wireless Sensor Networks," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 2602-2615, 2020
[29] F. Jameel, S.Wyne, I.Krikidi, “Secrecy Outage for Wireless Sensor Network”, IEEE Transactions on Dependable and Secure Computing, vol. 21, no. 7, pp. 1565-1568