Type of Article *(Original Article)*

Cloud Malware Attack Detection Framework using Water Reinforcement Learning

Bhargavi K[1], Adarsh M G[2]
[1]*Department of Computer Science and Engineering, Siddaganga Institute of Technology, Karnataka, India.*
[2]*Department of Computer Science and Engineering, Siddaganga Institute of Technology, Karnataka, India.*

Abstract -

Cloud computing is one of the potential data storage platform which offers computing services over the internet. Both private and public cloud computing platform are prone to various type of cyber-attacks like Distributed Denial of Service (DDoS), hypercall attack, hypervisor DoS, hyperjacking, polymorphic virus, metamorphic virus, Trojan virus, and many more. However providing security to cloud computing is a critical task because vulnerable Application Program Interface (API), data breaches, phishing, data loss, insider threat, etc. Among all forms of attacks in cloud, malware attacks caused by virus, ransomware, worm, and backdoors are found to increase exponentially. Hence there is a need to identify the malware attacks using promising machine learning methodology. In this paper a water wave optimization based reinforcement learning framework is proposed to identify malware attacks. The framework combines the benefits of shallow water wave theory with high adaptability capability of model-free Q-learning to generate malware detection policies. From the experiment results it is observed that the performance of the proposed framework is good with respect to attack detection time, energy consumption, accuracy, and model utility.

Keywords - *Reinforcement learning, water wave optimization, cloud malware, cloud computing, security.*

1. Introduction

Cloud computing offers computing services over the internet [1]. However providing security to cloud computing is a critical task for both public and private cloud service providers. Cloud malware injection attack is potential form cloud-based attack. In which an attacker tries to insert malicious service in the form of malicious virtual machine into the cloud system. The mail purpose of the malware attack is to access the user information by infecting the Infrastructure as a Service (IaaS), Software as a Service (SaaS), and Platform as a Service (PaaS) layers of cloud system [2, 3].

There are four different types of cloud malware injection attacks which include Distributed Denial of Service (DDoS), hypercall attack, hypervisor DoS, hyperjacking, and exploiting live migration. In DDoS attack, the intruder will flood bornets at large scale with malicious traffic which slow down the system. In hypercall attack, the attacker enters inside the network as a guest and requests domain access using hypercall interface. In hypervisor DoS, a hypervisor launch attack over the hypervisor space and gains control over the virtual host. In Hyperjacking, the attacker establishes control over the hypervisor using the rootkit setup on the virtual machine. During live migration the attacker establishes control over the vulnerable migration process to compromise the cloud management system for fake migration, theft of resources, and make random changes in the migrated system [4, 5].

The reinforcement learning is one of the powerful machine learning algorithm which learns to perform optimal policies through reward and punishment mechanism. But the learning algorithm suffer from weakened results because of too much of reinforcement and excessive computation involving large amount of data [6]. Hence the reinforcement learning is enriched with water wave optimization approach referred as Water Reinforcement Learning (WRL) [7]. Water wave optimization approach is a recently designed optimization algorithm motivated by movement of shallow water waves. It uses four different type of operators that are propagation and refraction which is capable enough of highly uncertain high dimensional space of computing domain.

In this paper a novel water reinforcement algorithm inspired by water wave optimization approach is designed which formulates shallow water wave policies to detect the cloud malware attacks at IaaS, SaaS, and PaaS layers. The policies formulated achieves better trade-off between exploration and exploitation activities. Even the weeding out effect caused by use of refraction operator is eliminated. The simple structure and ease implementation of water wave theory helps in achieving high speed of convergence and computation efficiency.

The rest of the paper is organized as follows: Section 2 describes related work carried out in literature. Section 3 gives architecture and algorithm of proposed work. Section 4 deals with experimental setup and results discussion. Finally section 5 draws the conclusion.

2. Related Work

Aslan et al. summarizes the variety of malware detection system developed for cloud system, also highlights the opportunities and challenges involved attack detection [8]. The malware attacks are identified by precisely analyzing the executable programs. Now malware detection process is advanced by performing the detection process both client and server side. Different types of malware attacks identified in cloud are virus, worm, Trojan House, Backdoor, Rootkits, Ransomware, and Obfuscated malware. The existing approaches for malware detection are: Semiparametric-based classification, multi chunk ensemble technique, split screen step process, correlation signature extraction, Q-learning, artificial neural network, convoluted neural network, and hashing. The algorithms proposed to detect the malware are of various type which include signature-base, malware-based, behavior-based, deep-learning-based, and heuristic-based. Most of the approaches suffer from poor accuracy, increased deployment effort, and steep performance cost.

Tian et al. propose a novel deep learning-based approach for malware detection for cloud [9]. A lightweight agent is designed first which collects resource utilization statistics of every virtual machine. The memory forensics analysis technique is applied memory related information from every virtual machine memory layout. A multi convoluted neural network model is constructed to identify the malware viruses. The deployment effort is very much less compared to the existing models and is able to successfully detect multiple type of malicious processes. However the cost of training the complex model is more as it requires hundreds /thousands of GPU-based machines.

Aslan et al. presents an intelligent malware detection system for cloud environment [10]. An intelligent agent is designed which basically creates a dataset to identify the characteristics of the malware virus over each virtual machine. The feature of different types of malware virus are mined and given as input to the rule-based learning agent which effectively separate benign malware virus from malignant malware virus. The detection performs well in identifying the malware whose features are already trained to the learning system. But some of the advanced malware virus exhibit unknown features in that case the learned system fails miserably in identifying the virus. The rule-based engine used to virus classification limits the practical applicability of the approach.

Kumar et al. discuss a clustering approach for malware virus detection in cloud environment [11]. Different type of malwares are identified by performing locality-sensing hashing procedure. The cuckoo sandbox is made use to generate analysis report in a dynamic environment under isolated condition. The prominent features of malware virus is extracted from sandbox. The extracted features are consolidated using the clustering algorithms like random forest, principal component analysis, and chi-square methodologies. The performance is better than non-clustering approaches in terms of increased classification accuracy and decreased false positive rate. However the approach slows down the model and also fails to represent the relationship within the data elements.

Gao et al. propose a malware classification approach based on semi-supervised transfer learning for cloud computing environment [12]. Two important challenges faced in cloud are malware identification and privacy preserving. A novel semi-supervised transfer learning framework is designed to detect the malware. Three main components involved in the framework are malware detection, prediction, and transfer. A recurrent neural network is designed to preserve the privacy of the cloud tenants. The transfer component invokes the prediction component to properly classify the unlabeled dataset. The prediction component accuracy is enhanced by appending Attribute

Selection Measure (ASM) classifier. The attributes exhibiting highest information gain is selected for classification problem. The combination of predicted labels and byte features of the unlabeled dataset form a new training dataset. However the mismatch between pre-training examples and target examples enhances the rate of malware misclassification.

3. Proposed Work

The high level architecture of the proposed water reinforcement learning-based malware attacks detection framework is given in Figure 1.

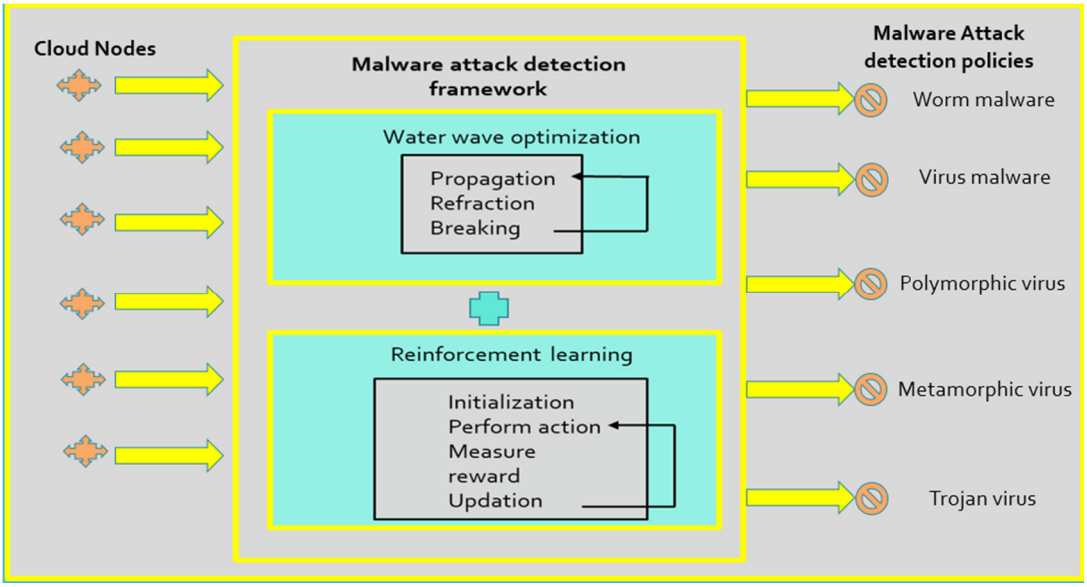

Figure 1: Proposed Water reinforcement learning-based malware attacks detection framework

The cloud nodes are passed as the input for malware attack detection framework. The framework is composed of two sub components that are water wave optimization and reinforcement learning. During water wave optimization the cloud nodes are subjected to propagation, refraction, and breaking process in iterations. During reinforcement learning Q learning is applied over the water wave optimize cloud nodes. That is water wave optimized cloud nodes go through Q state initialization, action performing, reward measurement, and updating of Q states in iterations. The Q learning policies for malware attacks detection is capable enough of producing high quality solutions as it is computationally efficient. The cloud nodes first follows shallow water wave motion principles before entering into the reinforcement learning phase. Further model free learning capability of the Q learning is applied which makes it suitable to operate in complex dynamic cloud scenario.

Algorithm 1: Working of the proposed Water Reinforcement Malware Attack Detection Framework

1. **Start**
2: **Input:** Cloud Nodes $CN = \{cn_1, cn_2, \ldots, cn_n\}$, Population of water waves, Q-function parameters
Output: Malware Attack Detection Policies
MADP= $\{madp_1, madp_2, .., madp_p\}$.
3: Randomly initialize a population P of n waves (solutions):
4: While stop criterion is not satisfied do
5: For each $x \epsilon P$ do
6: Propagation phase
7: Propagate x to a new x'

8: $x'(d) = x(d) + rand(-1,1) * \lambda$L(d)
9: if f(x') > $f(x)$ then
10: if f(x') > $f(x^*)$ then
11: Break x' using below equation
12: $x'(d) = x(d) + N(0,1) * \beta L(d)$
13: Update x^* with x'
14: Replace x with x'
15: Else
16: Decrease $x * h$ by one
17: If $x * h = 0$ then
18: Refract x to a new x'
19:

$$x'(d) = \mathrm{N}\left(\frac{x^*(d) + x(d)}{2}, \frac{|\ x^*(d) - x(d)|}{2}\right)$$

$$\lambda' = \lambda\left(\frac{f(x)}{f(x')}\right)$$

20: Update the wavelengths
21: $\lambda = \lambda * \alpha^{-(f(x)-f_{min}+\varepsilon)/(f_{max}-f_{min}+\varepsilon)}$
22: Return x^*
23: Initialize x^*(Q(S, a)) randomly
24: **for** each episode **do**
25: Initialize S
26: Repeat (for each step of episode):
27: Choose a from S using epsilon greedy policy derived from Q
$q_\pi(S, \pi'(s)) = \sum_a \pi'(a|S) * q_\pi(S, a)$
28: **if** (cn_i ←slow down && crashes) then
$madp_i$ =Worm malware attack
29: **else if** (cn_i ←data loss && More pop-ups) then
$madp_i$ =Virus malware attack
30: **else if** (cn_i ←High client request rate && File infection) then
$madp_i$ =Polymorphic virus attack
31: **else if** (cn_i ←High unusual requests && Infectious node) then
$madp_i$ =Metamorphic virus attack
32: **else if** (cn_i ←Disk errors &&spam interruptions) then
$madp_i$ =Trojan virus attack
28: Take action a, observe r, S'
$x^*(Q(S,a)) = x^*(Q(S,a)) + \alpha[r + \gamma max_a, x^*(Q(S',a'))$- x^*(Q(S, a))]
29: Update $S \leftarrow S'$
30: Until S is terminal
31: **End for**
32: Generate Malware Attack Detection policies
MADP= $\{madp_1, madp_2, .., madp_p\}$.
33: **Stop**

The stepwise working of the proposed framework is given in Algorithm 1. The algorithm is divided into two sub stages that are water wave optimization and Q-learning. During water wave optimization sub stage the cloud nodes

undergo propagation and refraction stages. This aims to yield optimal malware attack detection policies. Further the policies are further fine-tuned using epsilon greedy learning policy of Q-learning agent. Five kinds of malware attacks are detected that are worm malware, virus malware, polymorphic virus, metamorphic virus, and Trojan virus. The worm malware is identified based on the symptoms like cloud nodes slow down and node crashes. The virus malware is identified based on the symptoms like data loss and heavy pop-ups. Polymorphic virus is identifies based on the symptoms like high client requests and frequent file infections. Metamorphic virus are identified based on the symptoms like high unusual client requests and infectious node. Trojan virus are identifies based on the symptoms like disk errors and spam interruptions.

4. Results and Discussion

The proposed framework malware detection framework is 0implemented using JAVA JDK 1.6 with CloudSim 3.3 simulator framework [13]. An array of experiments were conducted on system with Windows 10 operating system. The system is composed of 2GHz dual core with 4 GB main memory running as 64 bit model. The simulation parameters are as follows: Number of data centers are 4, Host (Number of hosts=8, PES=4, MIPS=6000, RAM=20GB, bandwidth=10GB, storage=1 TB), virtual machine (Number of virtual machines=40 to 60, MIPS=1000 to 5000, RAM=1GB to 5GB, Bandwidth=100MB to 500MB, storage=10GB), Cloudlets (Number of cloudlets=500-1000-1500), length=3000 to10000, Type=Heterogeneous, Submission time=Poisson distribution of parameters). The performance of the proposed WRL is compared with the three of the existing works that are Deep Learning (DL) [9], Rule-based Learning Agent (RLA) [10], and Clustering Approach (CA) [11].

Attack detection time

Figure 2: Number of training iterations Versus Attack detection time (ms)

A graph of number of training iterations versus attack detection time is shown in Figure 2. It observed from the graph that the attack detection time of WRL is consistently less over the iterations of training. As it is computationally efficient and capable enough of finding high quality solutions by applying the principles of shallow water including propagation, breaking, and refraction. Whereas the attack detection time of DL is average. As DL is often subjected to biased decisions against certain group of cloud computing scenarios under uncertainty. The attack detection time of both RLA and CA are very high. Because attack detection policies formulated fails to handle ambiguity in the cloud nodes and lack adaptability.

Energy consumption

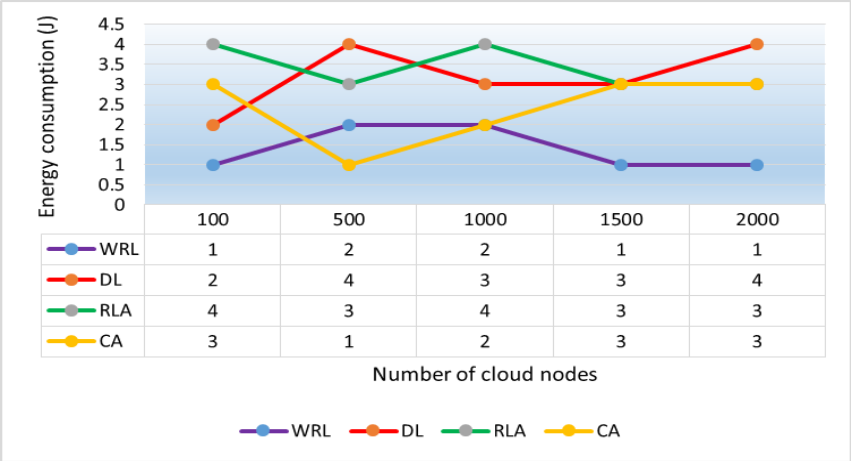

Figure 3: Number of cloud nodes Versus Energy consumption

A graph of number of cloud nodes versus energy consumption is shown in Figure 3. It is inferred from the graph that the energy consumption is less for WRL. Because the water wave models enriched with reinforcement. The energy consumption of DL is very high. The energy consumption of RLA and CA are above moderate. Because of the inability to handle cloud node failures and struggling to adapt to unseen cloud scenarios.

Accuracy

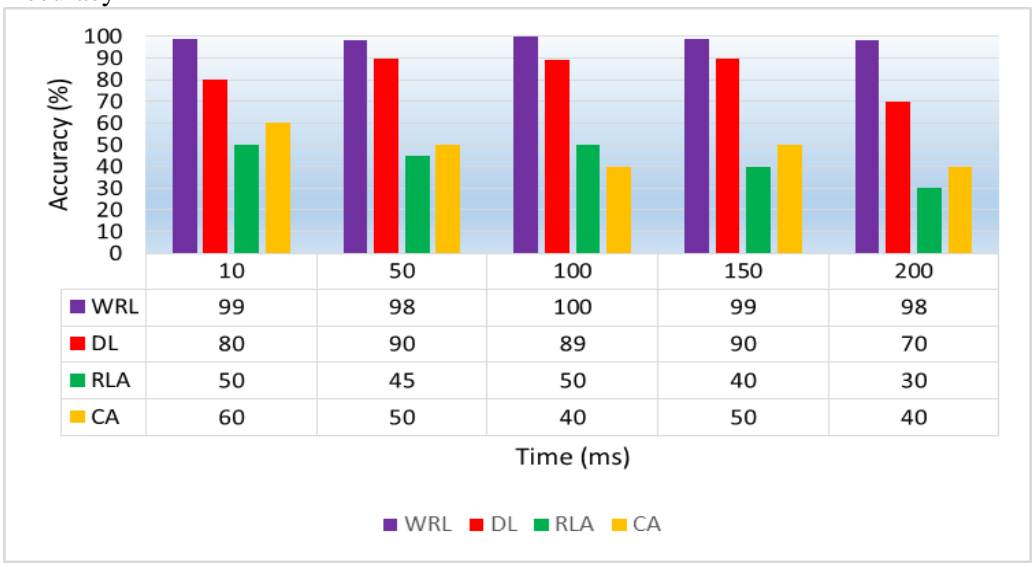

Figure 4: Time Versus Accuracy

A graph of time versus accuracy is shown in Figure 4. The accuracy of WRL is very high over time. As it always move towards global optimization solution and not get trapped into suboptimal solutions. The accuracy of DL is moderate due to poor transparency. Whereas the accuracy of RLA and CA are less due to inflexibility and are sensitive to outliers.

Model utility

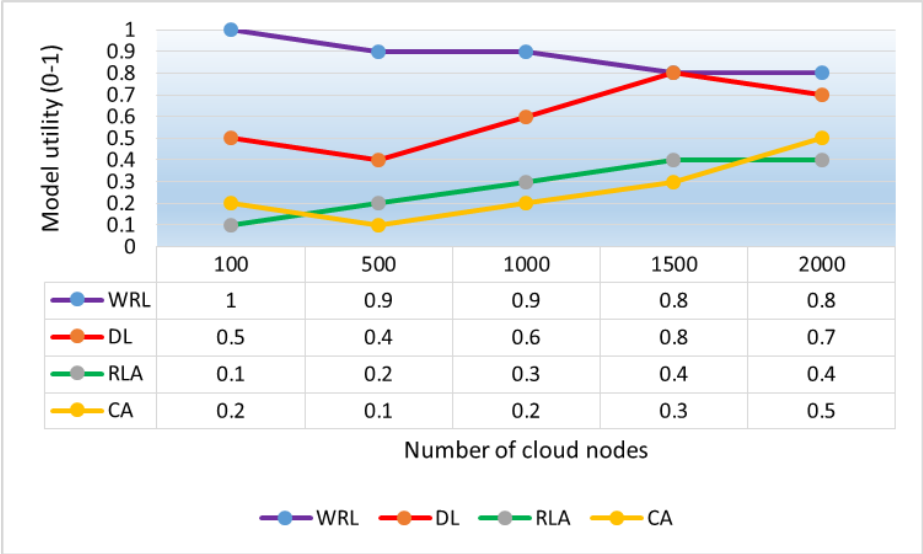

Figure 5: Number of cloud nodes Versus Model utility

The number of cloud nodes versus model utility is shown in Figure 5. The model utility of WRL is very high as it produces high quality solutions by gathering maximum amount of rewards. The model utility of DL is moderate due to high computational requirement and enormous energy consumption. The model utility of RLA and CA are very less due to lesser scalability because of difficulty in handling non-spherical shapes.

5. Conclusion

The paper focus on one of the commonly occurring security attacks in cloud computing which is malware attacks. A novel water wave optimization based reinforcement learning framework to identify different variety of malwares. The performance of the proposed work is compared with the three of the recent existing work incorporating strategy like deep learning, rule-based learning, and clustering approach. From the experimental results it is observed that the performance of the proposed work is good with respect to attack detection time, model utility, accuracy, and energy consumption. As future work the proposed framework will be extended to identify other kind of attacks like data breaches, side channel attack, phishing attack, Man-in-the-Middle attack, and many more.

Conflicts of Interest There is no possible conflict of interest with respect to the submitted manuscript.

Funding Statement

No financial support is received for the submitted manuscript by any funding bodies.

Acknowledgments

I acknowledge the support extended by Siddaganga Institute of Technology during the research and preparation of the manuscript. Both the authors Bhargavi K and Adarsh M G contributed equally to this work.

References

[1] Akbar, H., Zubair, M., & Malik, M. S. (2023). The security issues and challenges in cloud computing. *International Journal for Electronic Crime Investigation*, *7*(1), 13-32.

[2] Gan, C., Feng, Q., Zhang, X., Zhang, Z., & Zhu, Q. (2020). Dynamical propagation model of malware for cloud computing security. *IEEE Access*, *8*, 20325-20333.

[3] Patel, V., Choe, S., & Halabi, T. (2020, May). Predicting future malware attacks on cloud systems using machine learning. In *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing,(HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)* (pp. 151-156). IEEE.
[4] Aslan, Ö., Ozkan-Okay, M., & Gupta, D. (2021). A review of cloud-based malware detection system: Opportunities, advances and challenges. *European Journal of Engineering and Technology Research*, *6*(3), 1-8.
[5] Fui, N. L. Y., Asmawi, A., & Hussin, M. (2020). A dynamic malware detection in cloud platform. *International Journal of Difference Equations (IJDE)*, *15*(2), 243-258.
[6] Clifton, J., & Laber, E. (2020). Q-learning: Theory and applications. *Annual Review of Statistics and Its Application*, *7*(1), 279-301.
[7] Kaur, A., & Kumar, Y. (2022). A new metaheuristic algorithm based on water wave optimization for data clustering. *Evolutionary Intelligence*, *15*(1), 759-783.
[8] Aslan, O. A., & Samet, R. (2020). A comprehensive review on malware detection approaches. *IEEE access*, *8*, 6249-6271.
[9] Tian, D., Ying, Q., Jia, X., Ma, R., Hu, C., & Liu, W. (2021). MDCHD: A novel malware detection method in cloud using hardware trace and deep learning. *Computer Networks*, *198*, 108394.
[10] Aslan, O., Ozkan-Okay, M., & Gupta, D. (2021). Intelligent behavior-based malware detection system on cloud computing environment. *IEEE Access*, *9*, 83252-83271.
[11] Kumar, R., Sethi, K., Prajapati, N., Rout, R. R., & Bera, P. (2020, July). Machine learning based malware detection in cloud environment using clustering approach. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.
[12] Gao, X., Hu, C., Shan, C., Liu, B., Niu, Z., & Xie, H. (2020). Malware classification for the cloud via semi-supervised transfer learning. *Journal of Information Security and Applications*, *55*, 102661.
[13] Sundas, A., & Panda, S. N. (2020, March). An introduction of CloudSim simulation tool for modelling and scheduling. In *2020 international conference on emerging smart computing and informatics (ESCI)* (pp. 263-268). IEEE.
[14] Bhargavi K, Adarsh M G. (2024). “Efficient resource allocation in cloud computing using federated 2Q learning”, Computer Research and Development journal, Voume 24. Issue 9, 2024.