IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 28, 2019, accepted September 9, 2019, date of publication September 23, 2019, date of current version October 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2943056

Loose Game Theory Based Anomaly Detection Scheme for SDN-Based mMTC Services

BIZHU WANG[1], YAN SUN[1], AND XIAODONG XU[2], (Senior Member, IEEE)
[1]Department of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K.
[2]National Engineering Laboratory for Mobile Network Technologies, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Yan Sun (yan.sun@qmul.ac.uk)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871045, and in part by the 111 Project of China under Grant B16006.

ABSTRACT In this paper, we exploit the game theory to integrate the strengths of statistics-based detection and the machine learning-based detection for wireless software-defined-networking (SDN) based massive machine-type-communications (mMTC) services. Different from existing game theory approaches that require perfect rationality of the players and the synchronized information shared between the players, the proposed scheme is designed to bear with fewer assumptions. A novel feedback mechanism is proposed to fed machine learning-based detection results into strategy selection for advanced strategy decision making when the maliciousness estimation is not accurate. Besides, a multi-hop architecture is proposed to enable distributed detection instead of common cluster style for enhanced scalability. Various scenarios are set to testify the performance of the proposed scheme. Simulation results show that the proposed scheme relives more pressure of controller, consumes less energy without showing evident influence on the regular packets delivery, under the condition of asymmetric knowledge and multi attack-pattern.

INDEX TERMS Anomaly detection, DDoS, energy-efficient, mMTC, SDN.

I. INTRODUCTION

Under the concept of IoT, the future network is envisaged to accommodate the rapid traffic growth and address new services, including massive machine-type-communications (mMTC) [1]. MMTC covers various applications, such as smart cities, e-health and smart wearables, with tons of often low-complexity, low-power, machine-type devices [2]. As estimated by Ericsson, there will be over 15 billion mMTC devices by 2021, and it is critical to building a management system to control the devices as well as the data generated [1].

Software-defined network (SDN) is expected to provide flexibility for mMTC applications via treating network appliances as compounds of the control plane and data plane [3]. The control plane defines the forwarding strategies while the data plane deals with the actual corresponding packet processing. Through integrating all decision-makings into the SDN controller, the network administrator is able to experiment new network protocol in an cost-effective manner [3].

However, the more open and cloudy architecture designs of SDN introduce security challenges, especially to mMTC network. For example, the DDoS attacker can exploit the centralized architecture to spread viruses rapidly via sending fields-forged packets [4]. Besides, mMTC devices are vulnerable to attacks as they are closely exposed to the public and lack of ability to execute complex security algorithm [5]. Unfortunately, the statistic-related anomaly detection scheme (S-ADS) fails to fulfil the security duties while the periodical traffic uploading in the most of the machine learning based anomaly detection scheme (M-ADS) [6]–[8] imposes a high risk of energy depletion. Thus, a series of game theory based approaches are proposed to work together with S-ADS and M-ADS. It is worth pointing out that the strong assumptions of these game theory based approaches are debatable. Meanwhile, the cluster-based game theory based approaches introduce extra complexity and reduce the flexibility. To deal with the security concerns of SDN architecture in mMTC scenario, we propose a loose game theory based ADS (L-ADS) to enhance the scalability and energy efficiency. The main contributions are summarized as follows:

- L-ADS is designed to bear with less assumptions without degrading the overall performance. For instance, L-ADS keeps the same performance level regardless the pattern adopted by attacker (rational or irrational). The perfect Bayesian Equilibrium (PBE) rule is no longer

The associate editor coordinating the review of this manuscript and approving it for publication was Yi Zhang.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

compulsory. Furthermore, the condition of the same estimation knowledge applying to both attacker and defender is removed as well.
- An novel feedback mechanism is proposed to fed machine learning based detection results into strategy selection for advanced strategy decision making when the maliciousness estimation is not accurate enough.
- The performance of L-ADS is evaluated on SDN enabled mMTC simulation platform. Various scenarios are set to testify the functions of L-ADS under asymmetric knowledge condition and multi-attack pattern condition. Results are compared with other popular ADS for analysis.

II. LITERATURE REVIEW

The statistic-based ADS (S-ADS) and the machine learning based ADS (M-ADS) are both popular solutions in terms of attack detection. However, S-ADS suffers from the performance degradation in large-scale attack or when lacking attacking influence reflected on the statistics [10]. Besides, the M-ADS employs heavy pre-processing and model training, which should be avoided as much as possible when a large number of devices are involved [9]. Therefore, game theory-based approaches [10]–[13] are proposed to work in connection with statistic-based detection as a precursor. The machine learning based detection can be invoked only if necessary. Authors in [11] model the interaction between the attacker and the defender as a non-cooperative, static and complete information gambling to effectively predict the next attack time and the attack target. However, the assumptions of the players' activities in work [11] are rather changeless instead of sequential. Especially the fact that the players do not always with complete information about the system is overlooked. To increase the flexibility of ADS, researchers in [10] propose dynamic strategy selection based on sender type estimation given the statistics of sender's previous behaviour. Even that, the authors in [10] overlooked the impact on performance when the collected statistics are not accurate enough for maliciousness estimation update. Moreover, the assumptions of rational players and synchronized information among the players are noticeably strong in [10], which is rarely the practical case [14]. It is noticeable that most of game theory based approaches are clustered based which requires periodical leader selection for strategy decision making. Such cluster structure further reduce the flexibility against the changing environment if machine learning based ADS is called. To overcome the mentioned issues, in this work, we propose a loose game theory based ADS (L-ADS) for mMTC network with the following features:

- L-ADS works in a distributed manner instead of cluster style. Each node in the framework follows exact the same principle for detection strategy selection. Instant local environmental condition is taken into consideration when local nodes executes the detection algorithm.

Algorithm 1 Operational Process of L-ADS

```
Initialize S-ADS estimation E_co and local M-ADS
estimation E_lo;
while ingress packets do
    Detection Activation Module;
    check fields of packets, check detection_activation ;
    calculates detection-related statistics PRR;
    if detection_activation == 1 then
        Statistics-related ADS Module;
        Obtain pre-detection result X based on PRR;
        Strategy Selection Module;
        Input I = ⟨E_co, E_lo, X⟩;
        Output strategy S from ⟨Pass, Block, LESLA⟩;
        Execute strategy S ;
        if LESLA is activated then
            Machine learning based ADS Module;
            Obtain detection result Y and evolve the
            LESLA model;
            Update E_lo according to Y;
        end
    end
end
if periodical_update == 1 then
    Maliciousness Estimation Module;
    Update E_co based on statistics collected in the last
    interval;
end
```

- L-ADS employs the context-aware maliciousness estimation and the self-evolving machine learning based detection model. Such dynamic game design allows each node to adaptively response to the changing behaviour of attacker.
- L-ADS makes use of previous machine-learning based results in strategy selection. The feedback mechanism enhances the performance under the scenarios of irrational game players, the inaccurate maliciousness estimation and with asynchronous knowledge between attacker and defender, given the assumptions of Bayesian equilibrium are not fully met.

III. SYSTEM DESIGN

As shown in Algorithm 1, the operational process of L-ADS is composed of five modules. The detection activation module judges whether to invoke the detection activity based on the nodes' position in the multi-hop network if no matching flow entry is found. If the decision is positive from detection activation module, the detection-related statistics of the packet in relation with previous hop will be calculated in S-ADS module. The calculation result, as well as the maliciousness estimation will be the input for strategy selection module to decide which action from *Pass*, *Block* and *Activate LESLA* to choose. The *LESLA* module will be evolved if M-ADS is called. Finally, the maliciousness estimation update module

will update both S-ADS based and M-ADS based maliciousness estimations. Eventually the updates will be the feedback to L-ADS as one of the inputs for next round of operation.

A. DETECTION ACTIVATION MODULE

Upon receiving a packet, the detection activation module determines whether or not execute the detection on ingress packets. Considering that there are no matched flow-entries for field-randomly-forged packets, the detection of DDoS attacks is concentrated on the first hop to avoid the numerous triggered 'table-miss' messages. As for other attacks, the first-hop detection is often skipped, yet the detection will be carried out at the last hop to gain balanced energy consumption. It is worth pointed out that the detection activation rule introduces no extra cost since it utilizes the information from the flow-entry mapping and traffic forwarding processes.

B. STATISTICS-RELATED ADS MODULE-PRE-DETECTION

If detection is activated, the mMTC device will conduct pre-detection through the detection-related statistics. Since DDoS attacker can easily explore the SDN architecture in mMTC deployments, L-ADS mainly focuses on the inbound attack. The packet reception rate PRR, which is defined as the ratio of ingress data rate to the total neighbour traffic rate [10], is widely used as an indicator of excessive packet dropping and appearance of malicious users [10]. Therefore, it is adopted in L-ADS as well.

C. MACHINE LEARNING BASED ADS MODULE–LESLA

LESLA is proposed in [15] to enhance the scalability of ADS for mMTC services. The offline training scheme is adopted in LESLA, which only six features are required for detection. Besides, LESLA employs the idea of contrastive pessimistic likelihood estimation (CPLE) to outcome the performance degradation problem in most often-used semi-supervised learning techniques [16]. Through the designed objective function, CPLE enables LESLA to explore the information obtained from online data and to control the improvement over supervised learning. Simultaneously, CPLE drives LESLA to consider the worst case of uncertainty of online-unlabelled data for conservative model update.

Three CPLE-related ADSs are proposed in [15]: CPLE (pessimistic), CPLE (optimal) and CPLE_LESLA. The CPLE (pessimistic) scheme aims to maintain satisfying performance even when all of the online data are out-lier while CPLE (optimal) scheme takes the risk to optimize the new objective function when all estimations are assumed to be correct. LESLA scheme (short for CPLE_LESLA) considers the worst situation in advance but applies the classic log-sum-exp approximation function [17] to avoid a rather conservative model update.

D. MALICIOUSNESS ESTIMATION UPDATE MODULE

Maliciousness estimation update module in L-ADS is to execute adaptive strategy decision-making against the changing behaviour of the attacker. In addition to the most popular statistics-based maliciousness estimation [10], L-ADS builds a local maliciousness estimation based on the results of M-ADS LESLA. Both maliciousness estimations $p(\theta_i)$ are updated as shown in (1) with corresponding detection rate α_k and the false alarm γ_k.

$$p(\theta_i|b_i(t), b_i(t-1)) = \frac{p(\theta_i|b_i(t-1)P(b_i(t)|\theta_i, b_i(t-1))}{\sum_{\tilde{\theta}_i} p(\tilde{\theta}_i|b_i(t-1))P(b_i(t)|\tilde{\theta}_i, b_i(t-1))} \tag{1}$$

where θ is the type of device U_i. $p(\theta_i|b_i(t-1))$ is the latest value of maliciousness estimation. The behaviour of the device U_i at time $t-1$ and t is measured by the detection results $b_i(t-1)$ and $b_i(t)$ respectively. $P(b_i(t)|\theta_i)$ represents the probability of observed detection results given the real type of device U_i. $P(b_i(t)|\theta_i)$ is related with the result and the performance of detection scheme, as well as the real type of packet. For example, $P(a_i(t_k) = Abnormal\,|\theta_i = 1, b_i(t-1)) = p\alpha_k + (1-p)\gamma_k$ when the observed result is *Abnormal* and the last hop is a malicious user attacking network with probability p. The posterior probability in other cases can be calculated in the similar way.

The S-ADS based estimation is updated periodically according to the statistics collected in the latest interval while the M-ADS based maliciousness estimation is updated when M-ADS LESLA module is called by strategy selection module. According to [10], the maliciousness estimation will be reduced to half only when both of the result of S-ADS and M-ADS module is *Normal*. However, in L-ADS, the estimation employs a quantitative update. Therefore, it will contribute to the advanced strategy selection regardless the results from S-ADS and M-ADS.

E. STRATEGY SELECTION MODULE

The main function of strategy selection module is to activate LESLA only when necessary for effective detection and to keep LESLA disabled otherwise to save energy. The transmission from the last hop U_i to the receiver U_j is modelled as a Bayesian game. The type of device is denoted by θ, while the malicious last hop has θ_i equal to 1 and the regular user has θ_i equal to 0. The receiver U_j is always legitimate. The malicious last hop has two strategies [*Attack*, *Not Attack*], yet the regular last hop only has one strategy *Not Attack*. The receiver has three strategies [*Pass*, *Block*, *LESLA*]. The malicious device intends to launch a successful attack without being identified while the receiver aims to detect attacks with limited energy effectively.

Assume the detection rate and the false alarm of S-ADS and M-ADS are represented as α_S, γ_S and α_M, γ_M respectively. Similar to the game theory based ADS in [10], the receiver U_j losses the asset a_j when the receiver chooses to pass the packets and the malicious last hop ($\theta_i = 1$) chooses to launch attacks. In contrast, the device losses the asset a_j if it blocks the regular packets. Besides, the device U_j has α_M possibility of correctly identifying attacks and

TABLE 1. Advanced strategy selection in L-ADS.

S-ADS result	$p_{\theta_{local}}p$	Pass	Block	LESLA
Abnormal	$< min(a_1, a_2)$	$\surd$		
	$min(a_1, a_2) \sim max(a_1, a_3)$			$\surd$
	$> max(a_1, a_3)$		$\surd$	
Normal	$< min(a_4, a_5)$	$\surd$		
	$min(a_4, a_5) \sim max(a_4, a_6)$			$\surd$
	$> max(a_4, a_6)$		$\surd$	

$1-\alpha_M$ probability of missing abnormal packets when *LESLA* is activated. Therefore, the receiver gains $(2\alpha_M - 1)a_j$ payoff with an extra cost of L_{Lj} of activating LESLA. Furthermore, there is γ_M possibility of incorrectly blocking standard packets if *LESLA* is activated. The gain of payoff of the attacker is equal to the loss of receiver plus cost of attacking L_{aj} if it chooses to launch an attack. Otherwise, the attacker gets zero payoffs if not attack. By comparing the payoff function of all possible strategy pair and referring to the Bayesian Equilibrium analysis in [10], the strategy selection in the referenced game theory scheme (R-ADS) [10] is concluded as follows: M-ADS module is activated all the time when the result of S-ADS is *Abnormal* and the receiver chooses to pass the packet when the maliciousness estimation p_θ is below $\frac{\gamma_M a_j + L_{Lj}}{(2\alpha_M + \gamma_M)a_j}$, denoted by $p_{\theta_{th}}$. Otherwise, the device will activate LESLA with the probability of $\frac{a_j - L_{aj}}{2\alpha_M a_j}$.

Different from other ADS schemes, in L-ADS, the result of M-ADS is fed back to build a local maliciousness estimation $p_{\theta_{local}}$ for advanced strategy selection, especially when the Perfect Bayesian Equilibrium (PBE) condition doesn't exist. For instance, the attacker is not rational and acts deviated from BNE rule or the knowledge about the network is not synchronized with the opponent player. In L-ADS, the novel feedback scheme also contributes to correct decision-making when S-ADS based maliciousness estimation becomes less accurate as the performance of S-ADS degrades.

Besides, L-ADS exams the certainty of S-ADS results. For instance, the joint possibility that the incoming packet is malicious and S-ADS result is *Normal* is $p_{\theta_{local}}p(1-\alpha_S)$ while the joint possibility that the incoming packet is regular and S-ADS result is *Normal* is $p_{\theta_{local}}(1-p)(1-\gamma_S) + (1-p_{\theta_{local}})(1-\gamma_S)$. p represents the possibility of launching attack when U_i is malicious. Therefore, the payoff of U_j playing strategy *Pass*, *Block* and *LESLA* when the S-ADS result is *Normal* are

$$\begin{cases} U_j(Pass) = pp_{\theta_{local}}(1-\alpha_S)(-a_j) \\ U_j(Block) = pp_{\theta_{local}}[(1-\alpha_S) + (1-\gamma_S)]a_j - a_j(1-\gamma_S) \\ U_j(LESLA) = pp_{\theta_{local}}[((1-\alpha_S)(2\alpha-1) + (1-\gamma_S)\gamma)a_j \\ +((1-\gamma_S) - (1-\gamma))L_{Lj}] - (1-\gamma_S)(\gamma a_j + L_{Lj}) \end{cases} \tag{2}$$

The case when S-ADS result is *Abnormal* can be calculated in the similar way. By comparing the payoff function, the advanced strategy selection of L-ADS is concluded in Table 1. The alphabets in the table is shown as bellows as shown at the top of next page:

IV. SIMULATION AND ANALYSIS

In this work, a square region with a side length of 600m is built on the OPNET simulator. There are 30 awake devices in the system per second, where the ratio of malicious users r_{mal} ranges from 10%, 30% to 50%. Legitimate devices generate 5 packets per second while the 'DDoS' attackers create 25 packets per second. Packets are forwarded directly if the receiver locates within the transmission range of 150m. Parameters of the network are set according to the customs of mobile ad-hoc network [10]. The experiment implements statistics-related ADS (S-ADS), machine learning based ADS (M-ADS), referenced game theory based ADS (R-ADS) [10] and the proposed L-ADS separately. The performance and limitation of the S-ADS and M-ADS will be briefed first, followed by comprehensive comparisons among various ADSs, in terms of detection rate, false alarm, pre-defined payoff, the load on the controller, the energy consumption and the number of well-delivered packets.

A. ANALYSIS OF S-ADS AND M-ADS

Table 2 lists corresponding detection rate R_{TP} and the false alarm rate R_{FP} if S-ADS is applied to the network with $r\_mal = 10\%$, 30% and 50% respectively. The threshold of *PRR* in regular traffic pattern is first experimented as 0.115, and the threshold of S-ADS, denoted by PRR_{th}, is set at series nearby values to select an optimal threshold for S-ADS. R_{TP} is defined as the ratio of attacks that are correctly recognized while the R_{FP} represents the proportion of regular traffics that are incorrectly blocked. From Table 2, we can find that a lower PRR_{th} is beneficial to improve R_{TP} while a higher PRR_{th} is more appropriate for regular network operation. S-ADS shows decreasing R_{TP} and raising R_{FP} as the increase of r_mal. Besides, S-ADS fails to simultaneously achieve a high R_{TP} and a low R_{FP}.

TABLE 2. Performance of S-ADS under various PRR_{th}.

Metric	R_{TP}			R_{FP}		
PRR_th \ r_mal	10%	30%	50%	10%	30%	50%
0.145	0.94	0.69	0.61	0.38	0.43	0.44
0.155	0.80	0.64	0.55	0.32	0.38	0.35
0.165	0.53	0.57	0.49	0.29	0.34	0.32

TABLE 3. Performance of M-ADS when trained with 100 labelled data and tested on 10000 unlabelled data.

	ACC	R_{TP}	R_{FP}	$B2N$	$B2A$
WQDA(supervised)	0.44	0.86	0.56	0.43	0.57
Self-training	0.38	0.741	0.63	0.37	0.63
CPLE(optimal)	0.80	0.89	0.20	0.80	0.20
CPLE(pessimistic)	0.54	0.89	0.46	0.54	0.46
CPLE_LESLA	0.82	0.86	0.19	0.81	0.189

Table 3 lists the ACC, R_{TP}, R_{FP},$B2N$ and $B2A$ of supervised learning based ADSs, self-training (a popular semi-supervised learning technique) based ADS and three CPLE based ADSs (CPLE pessimistic, CPLE optimal and

$$\begin{cases} a_1 = \dfrac{\gamma_S}{2\alpha_S + \gamma_S} \\ a_2 = \dfrac{\gamma_S(\gamma_M a_j + L_{Lj})}{(\gamma_S\gamma_M + 2\alpha_M\alpha_S)a_j + (\gamma_S - \alpha_S)L_{Lj}} \\ a_3 = \dfrac{\gamma_S((1-\gamma_M)a_j - L_{Lj})}{[\gamma_S(1-\gamma_M) + 2(1-\alpha_M)\alpha_S]a_j + (\alpha_S - \gamma_S)L_{Lj}} \\ a_4 = \dfrac{1-\gamma_S}{2(1-\alpha_S) + (1-\gamma_S)} \\ a_5 = \dfrac{(1-\gamma_S)(\gamma_M a_j + L_{Lj})}{[(1-\gamma_S)\gamma_M + 2\alpha_M(1-\alpha_S)]a_j + [(1-\gamma_S) - (1-\alpha_S)]L_{Lj}} \\ a_6 = \dfrac{(1-\gamma_S)((1-\gamma_M)a_j - L_{Lj})}{[(1-\gamma_S)(1-\gamma_M) + 2(1-\alpha_M)(1-\alpha_S)]a_j + (\gamma_S - \alpha_S)L_{Lj}} \end{cases}$$

$$p = \begin{cases} 1 & when \quad p_\theta < \dfrac{\gamma_M a_j + L_{Lj}}{(2\alpha_M + \gamma_M)a_j} \\ \dfrac{\gamma_M a_j + L_{Lj}}{(2\alpha_M + \gamma_M)a_j p_0} & when \quad p_\theta > \dfrac{\gamma_M a_j + L_{Lj}}{(2\alpha_M + \gamma_M)a_j} \end{cases}$$

TABLE 4. Behaviour of the attacker.

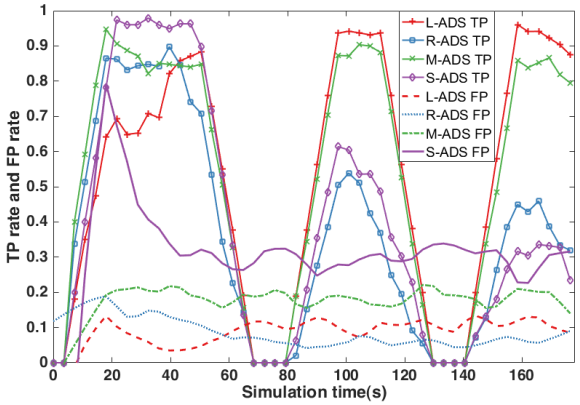
Time(s)	0-30	30-60	60-90
Attack pattern	BNE	continually attack	Not attack
Time(s)	90-120	120-150	150-180
Attack pattern	continually attack	Not attack	BNE

CPLE_LESLA) when evaluated on CTU-13 dataset [19]. *B2N* represents the ratio of background data classified as normal while *B2A* denotes the ratio of background data classified as abnormal. CTU-13 dataset, comprising of background data, normal traffic and abnormal traffic, is widely viewed as the most valuable data set for anomaly detection [18]. The background data is excluded from the training dataset but included as normal traffic in the test dataset for adaptability evaluation. Besides, the detection model is trained with only 100 labelled data but tested on 10000 unlabelled data to show the long-time adaptability.

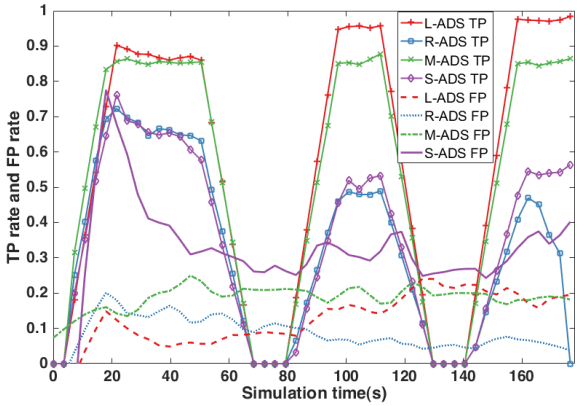
From Table 3, we can find that supervised learning based ADS shows a satisfied R_{TP} and a high R_{FP}. The high R_{FP} is the result of lacking of ability to classify background traffic, as illustrated by nearly equal *B2A* and *B2N*. Self-training based ADS shows the lowest R_{TP} with the highest R_{FP} due to risky evolving. CPLE_LESLA achieves the highest R_{TP} and the lowest R_{FP} when tested on new type of traffic since the self-evolving algorithm of LESLA considers the worst situation in advance as well as applying performance optimization method to avoid rather conservative update.

B. COMPARISONS BETWEEN THE EXISTING ADSS AND THE PROPOSED ADS

In this section, the performance of S-ADS, M-ADS, R-ADS and L-ADS are compared in a system-level. The performance evaluation is carried out when the following two conditions of the game-theory based ADS are not fulfilled simultaneously: a) Both players are rational. b) Knowledge of malicious estimation is synchronized between players.



(a) $r_{mal} = 10\%$

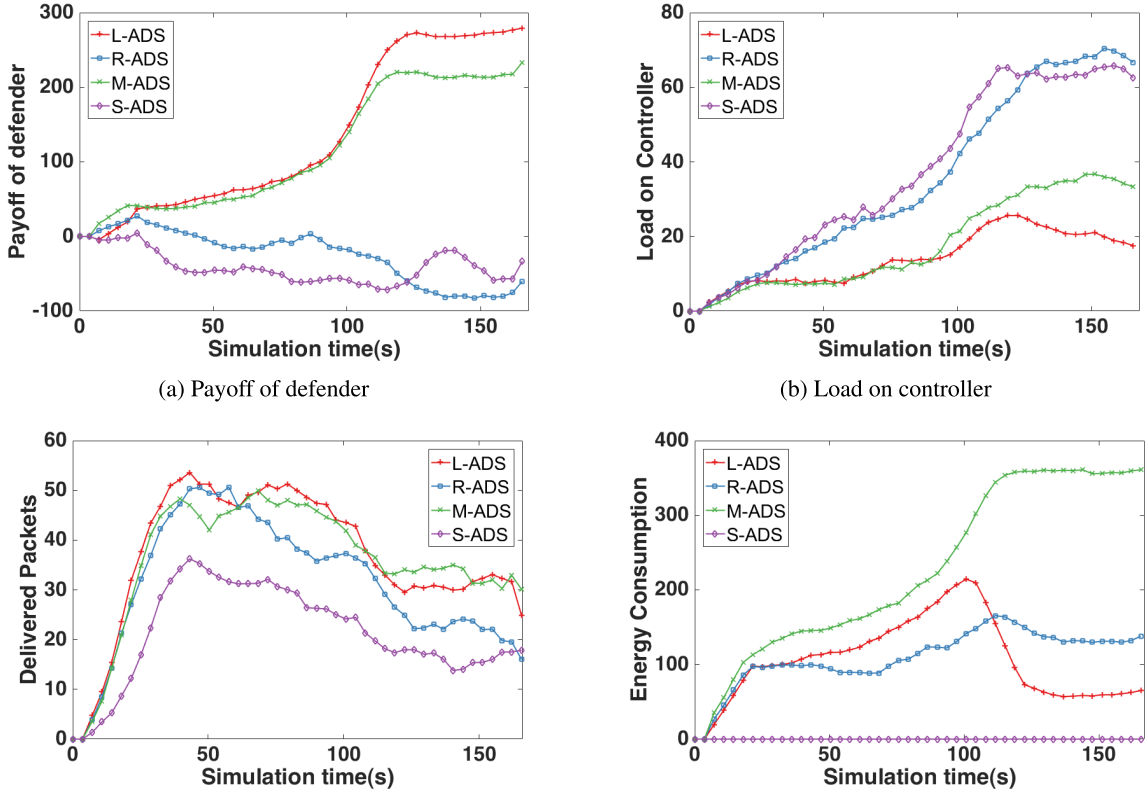


(b) $r_{mal} = 30\%$

FIGURE 1. Dynamic change of sensitivity and false alarm.

1) MULTIPLE ATTACK PATTERN

The real-time behavior of the attacker is defined as shown in Table 4 for analyzing the performance of the L-ADS when the attacker deviates from the BNE strategy. The action of the



(a) Payoff of defender

(b) Load on controller

(c) Number of delivered packets

(d) Energy consumption

FIGURE 2. Dynamic change of performance when $r_{mal} = 50\%$.

attacker is set to cover various patterns as well as achieving the effective attack.

Fig. 1a and Fig. 1b illustrate the result with the dynamic change of detection rate (denoted by *TP*) and the false alarm (denoted by *FP*) of various ADSs when $r\_mal = 10\%$ and $r\_mal = 30\%$ respectively. From Fig. 1, we can get the following observations: a) Detection rate and the false alarm changes as the attack pattern defined in Table 4. When the attacker chooses not to attack, the true positive is 0 since there are no attacks being recognized. b) When $r\_mal = 10\%$, R_{TP} of R-ADS is higher than S-ADS when attacker follows BNE rule (150-180s) while R_{TP} of R-ADS is lower than S-ADS when the action of attacker deviates from BNE (90-120s). When $r\_mal = 30\%$, R-ADS even gets lower R_{TP} than S-ADS since the degrading accurate statistic-based maliciousness estimation makes R-ADS more likely to pass ingress packets instead of activating heavyweight module. c) L-ADS shows a similar R_{TP} as M-ADS regardless the rationality of attacker (150-180s and 90-120s) except the early stage of gambling. Besides, L-ADS shows a similar R_{FP} with M-ADS except for the beginning when $r\_mal = 10\%$, where the information collected from a few abnormal traffics for a short time is not enough to benefit the detection. Through building a M-ADS based maliciousness estimation and comparing the difference between S-ADS based estimation and the M-ADS based estimation, the strategy decision-making is improved in L-ADS. Thus, we can conclude that the L-ADS not only outcomes the inaccurate maliciousness estimation update but maintains satisfying performance even when attacker's behaviour deviates from BNE.

2) ASYNCHRONOUS INFORMATION BASED BAYESIAN GAME

This subsection demonstrates the performance of various ADSs when $r\_mal = 50\%$ and the maliciousness estimation between players are asynchronous: defender makes estimation based on the number of online received packets and online traffic rate while the attacker makes estimation based on the number of packets sent and the traffic rate in the regular pattern.

Fig. 2a shows the payoff of various ADSs. Payoff function defines the whole picture of effective anomaly detection and the efficient activation of the M-ADS module. From Fig. 2a, we can get the following observations: a) L-ADS and M-ADS shows an increasing trend as time elapses. The rising payoffs are the results of more abnormal packets generated and detected due to degrading accurate S-ADS based maliciousness estimation. b) L-ADS gains a higher payoff over M-ADS at the late stage due to the competitive detection rate and the reduced energy

consumption. c) R-ADS gets decreasing payoff since R-ADS is under the assumption of BNE rule and satisfying accurate S-ADS based maliciousness estimation.

Fig 2b shows the load on the controller, which rises for all ADSs along the time. It is because the maliciousness estimation gets lower, followed by more and more attacks launched. Besides, L-ADS is the most effective detection scheme to identify such situation, which takes actions of blocking attacks and reliving the pressure on the controller.

Fig 2c shows the number of packets successfully accepted by the destination nodes. The number of delivered packets of all scheme increases first as the forwarded packets approach the destination node. Then the number of delivered packets of all schemes drops later and then keeps stable. The dropping delivery ratio is the result of more attacks generated and more packets dropped at the buffer. L-ADS shows a similar delivery ratio as M-ADS due to the similar R_{FP}. Besides, the high R_{TP} of L-ADS reduces the possibility of the packet dropping, which is beneficial for standard packet delivery.

Fig 2d shows the energy consumption of executing various ADSs. Energy consumption increases one each time M-ADS is activated while the cost of launching S-ADS is neglected. Both game theory based ADSs reduce the energy consumption via discontinuous activation of M-ADS module. In addition, L-ADS consumes less energy than R-ADS at the late stage since R-ADS activates M-ADS LESLA all the time when the pre-detection result is *Abnormal* while L-ADS chooses the strategy *Pass* or *Block* sometimes. Besides, L-ADS consumes less energy as the information accumulated from the previous results of M-ADS module.

V. CONCLUSION

In this paper, we emphasized the robust and flexible game-theoretical ADS on the SDN platform for mMTC services. Through applying S-ADS as a precursor and activating M-ADS when necessary, L-ADS aims to pursue the efficiency and effectiveness simultaneously. L-ADS fed the previous M-ADS results to decision-making for enhanced performance when the usual assumptions of game-theoretical approach PBE are not fully met. Besides, L-ADS proposes M-ADS based maliciousness estimation to overcome the performance degradation when the result of S-ADS is not qualified to represent the maliciousness of the potential attacker. In addition, a multi-hop localized architecture is proposed to support L-ADS through re-designing the role of defender and attacker in the game, as well as the distributed and self-adaptive M-ADS module. Simulation results demonstrate that L-ADS effectively blocks the attack locally within limited energy budget without showing evident influence on the standard packet delivery even when PBE are not satisfied.

Finally, several significant limitations need to be considered. The game theory-based approach relies on the assumption that the type of user is fixed. If the kind of users is changed during the simulation, the previous experience-based decision-making will suffer from performance degradation. Besides, applying the Markov decision process related technique might improve the estimation accuracy in L-ADS. It is also recommended that further research not only consider the optimal theoretical response but also considers the realistic feedback from the environment, such as the overwhelming warning from the controller and the congestion status of the network, for enhanced generality and performance. More broadly, research is also needed to evaluate the performance of L-ADS in the scale-changing network against more types of attack.

REFERENCES

[1] G. A. Akpakwu, B. J. Silva, G. P. Hancke, and A. M. Abu-Mahfouz, "A survey on 5G networks for the Internet of Things: Communication technologies and challenges," *IEEE Access*, vol. 6, pp. 3619–3647, 2017.
[2] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sep. 2016.
[3] D. Henneke, L. Wisniewski, and J. Jasperneite, "Analysis of realizing a future industrial network by means of software-defined networking (SDN)," in *Proc. IEEE World Conf. Factory Commun. Syst. (WFCS)*, May 2016, pp. 1–4.
[4] X. Ji, K. Huang, L. Jin, H. Tang, C. Liu, Z. Zhong, W. You, X. Xu, H. Zhao, J. Wu, and M. Yi, "Overview of 5G security technology," *Sci. China Inf. Sci.*, vol. 61, no. 8, 2018, Art. no. 081301.
[5] G. M. Køien, "Aspects of security update handling for IoT-devices," *Int. J. Adv. Secur.*, vol. 10, nos. 1–2, 2017.
[6] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, Oct. 2016, pp. 258–263.
[7] Q. Niyaz, W. Sun, and A. Y. Javaid, "A deep learning based DDoS detection system in software-defined networking (SDN)," 2016, *arXiv:1611.07400*. [Online]. Available: https://arxiv.org/abs/1611.07400
[8] K. Malialis and D. Kudenko, "Distributed response to network intrusions using multiagent reinforcement learning," *Eng. Appl. Artif. Intell.*, vol. 41, pp. 270–284, May 2015.
[9] R. Wang, Z. Jia, and L. Ju, "An entropy-based distributed DDoS detection mechanism in software-defined networking," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, vol. 1, Aug. 2015, pp. 310–317.
[10] B. Subba, S. Biswas, and S. Karmakar, "Intrusion detection in mobile ad-hoc networks: Bayesian game formulation," *Eng. Sci. Technol., Int. J.*, vol. 19, no. 2, pp. 782–799, Jun. 2016.
[11] L. Han, M. Zhou, W. Jia, Z. Dalil, and X. Xu, "Intrusion detection model of wireless sensor networks based on game theory and an autoregressive model," *Inf. Sci.*, vol. 476, pp. 491–504, Feb. 2019.
[12] A. Abusitta, M. Bellaiche, and M. Dagenais, "A trust-based game theoretical model for cooperative intrusion detection in multi-cloud environments," in *Proc. 21st Conf. Innov. Clouds, Internet Netw. Workshops (ICIN)*, Feb. 2018, pp. 1–8.
[13] B. Subba, S. Biswas, and S. Karmakar, "A game theory based multi layered intrusion detection framework for VANET," *Future Gener. Comput. Syst.*, vol. 82, pp. 12–28, May 2018.
[14] C. Kiennert, Z. Ismail, H. Debar, and J. Leneutre, "A survey on game-theoretic approaches for intrusion detection and response optimization," *ACM Comput. Surv.*, vol. 51, no. 5, 2018, Art. no. 90.
[15] B. Wang, Y. Sun, C. Yuan, and X. Xu, "LESLA: A smart solution for SDN-enabled mMTC E-health monitoring system," in *Proc. 8th ACM MobiHoc Workshop Pervasive Wireless Healthcare Workshop*, 2018, Art. no. 2.
[16] M. Loog, "Contrastive pessimistic likelihood estimation for semi-supervised classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 3, pp. 462–475, Mar. 2016.
[17] N. Boumal, "Optimization and estimation on manifolds," Ph.D. dissertation, Catholic Univ. Louvain, Ottignies-Louvain-la-Neuve, Belgium, 2014, pp. 47–61.
[18] S. Garcia, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods," *Comput. Secur.*, vol. 45, pp. 100–123, Sep. 2014.
[19] M. Małowidzki, P. Berezinski, and M. Mazur, "Network intrusion detection: Half a kingdom for a good dataset," in *Proc. NATO STO SAS-139 Workshop*, Portugal, 2015.

BIZHU WANG received the B.Eng. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2015. She is currently pursuing the doctoral degree with the Department of Electronic Engineering, Queen Mary University of London, London, U.K. Her current research interests include Ad hoc networks, the Internet of Things, software define networking, and anomaly detection model.

YAN SUN received the B.Eng. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications, in 2001, and the M.Sc. and Ph.D. degrees in electronic engineering from the Queen Mary University of London, in 2003 and 2009, respectively. She joined Siemens Ltd., China, in 2001, as a Network Optimization Engineer and re-joined in 2003, as a System Engineer at Research and Development after received her master's degree in U.K. She then worked in Siemens (later Nokia Siemens Networks, Ltd.) as a Product Manager for five years, responsible for the 3rd generation cellular network equipment product lines. In 2009, she returned to the Department of Electronic Engineering and Computer Science, Queen Mary University of London, as a Lecturer. Her current research interests include ad hoc networks, energy saving for modern mobile networks, software-define networks, and mobile health-care networks.

XIAODONG XU (S'06–M'07–SM'18) received the B.S. degree in information and communication engineering and the master's degree in communication and information system from Shandong University, in 2001 and 2004, respectively, and the Ph.D. degree in circuit and system from the Beijing University of Posts and Telecommunications (BUPT), in 2007, where he is currently a Professor. He has coauthored nine books and more than 120 journal and conference articles. He is also the Inventor or Co-Inventor of 39 granted patents. His research interests include moving networks, D2D communications, mobile edge computing, and caching. He is an Associate Editor of IEEE Access.