Learning Based Routing Path Reconstruction and Network Intrusion Detection

Jieling Li
Xiamen University

Abstract—**A network intrusion detection system is essential for safeguarding the integrity and availability of sensitive assets. Network traffic data contains extensive temporal, spatial, and statistical information. However, existing research has not sufficiently incorporated spatial-temporal multi-granularity data features or explored the mutual reinforcement among different types of data features. To address this, we propose a framework based on the Broad Learning System (BLS) that can learn and integrate features across three granular levels. To comprehensively capture the nuances of traffic data, we construct three granular feature datasets reflecting time, space, and data content characteristics. In this work, we employ fundamental broad learning units to extract abstract features for each granularity, expressing these features in distinct feature spaces to enhance them individually. Additionally, we apply He initialization for feature and enhancement node weights, optimizing the ReLU activation function and improving detection accuracy over random weight initialization. Furthermore, under equivalent configuration conditions, the training and detection times for our Broad Learning System are comparable to or shorter than those of standard BLS.**

Index Terms—**Unmanned aerial vehicle, intrusion detection, broad learning, reinforcement learning.**

I. Introduction

The influence of the Internet continues to increase, and network services are widespread. The user scale and economic benefits of online entertainment, online travel, and online education have increased significantly compared with the past [1]. At the same time, countless network devices and applications, and explosive network data, make the network environment increasingly complex and bring huge hidden dangers to network security [2], [3]. Cybercriminals are becoming more proficient in robbing the benefits of the openness of the Internet, advancing attacks at an alarming rate [4], [5]. Therefore, network security has become an important issue that needs to be addressed in the informatization construction of all sectors [6], [7].

High-speed mobility of nodes in As creates a highly dynamic network topology [8]. Topology changes with neighbor relationships that depend on communication distance or the quality of neighbor links [9]. Most traditional detection protocols. Instead of preselecting a designated relay node on each transmission, an opportunistic detection in the context of ExOR broadcasts a packet [10]. Traditional detection relies on a route discovery process, but with rapid fluctuations in channel conditions, detection information estimated based on average channel quality information may become stale [11]. Nodes do not have to worry about path or link interruptions, as only candidate relays that actually successfully receive packets participate in the forwarding process. The main advantage of opportunistic detection over traditional detection is its flexibility and ease of adapting to network dynamics, which can improve packet delivery rates.

In a complex As environment, optimizing the model needs to be learned and maintained, so that nodes can more effectively make their detection decisions adapt to their dynamic environment. Reinforcement learning (RL) acquires knowledge by exploring interactions with the local environment without external supervision, making it suitable for solving detection challenges in distributed networks. However, it is difficult for single-agent RL to provide a distributed view to identify the resource requirements of each agent. Therefore, the distributed perspective of multi-agent reinforcement learning makes it more suitable for As [12].

The traditional BLS model, which relies on using full data features, is unsuitable for dynamically changing network environments. In this paper, we propose a novel tri-Broad Learning System that incorporates temporal and spatial granularity. First, we divide cyberspace traffic data into spatial-temporal granularity feature datasets according to their characteristics, providing a foundational input dataset. Then, our BLS performs parallel learning across three granularities: time, space, and data content. He initialization is introduced to the feature and enhancement node generation process, further improving predictive accuracy. Finally, classification output is achieved based on the features from the joint mapping. The results indicate that our BLS achieves promising performance and enhances detection accuracy [13]. Our BLS effectively extracts rich information from traffic data at different granularities, enabling efficient fusion learning and accurate classification.

II. Related Works

They evaluated the performance of BLS models employing radial basis function (RBF) and incremental learning for network anomaly and intrusion classification [14]. In this study, the MATLAB implementation of BLS was converted to Python to ensure consistency in comparisons. The authors noted that the incremental BLS algorithm reduces training time by updating weights based solely on new data. However, this approach also demands additional memory.

Following this, they evaluated the performance of BLS models utilizing radial basis function (RBF) and incremental learning for network anomaly and intrusion classification [14]. In this study, the BLS model initially implemented in MATLAB was converted to Python to ensure consistent comparisons. The authors noted that the incremental BLS algorithm reduces training time by updating weights solely based on new data; however, this approach also demands additional memory.

In multi-hop broadcasting protocols, each node forwards the received data packets by broadcasting. For example, the intelligent forwarding protocol in [15] exploits a highly efficient forwarder selection mechanism and eliminates dependence on handshaking mechanisms to improve the transportation safety. An enhanced flooding-based detection protocol in [16] utilizes random network coding and clustering for swarm node networks to reduce the hop count and detection latency. The opportunistic detection protocol in [17] applies ACK-based and timer-based coordination schemes according to the requirements of the service flowing and exploits a dynamic adjustment of the control-message sending-interval to reduce the energy consumption. The probability prediction-based reliable opportunistic detection protocol in [18] predicts the packet queue length in the receiver and uses the prediction result to determine the utility of each relaying node, which can improve the packet delivery ratio and network throughput.

Position-based detection protocols greatly reduce the requirements of topology storage and provide flexibility in the accommodation of the dynamic behavior of As [19], [20]. For example, the geographic load share detection protocol in [21] extends the protocol using a set of the candidates instead of a node to spread traffic and adopts a congestion-aware handover strategy to perform load balancing among neighboring candidates. The predictive optimized link-state detection protocol in [22] exploits GPS information to predict the quality development of wireless links between node nodes and modify the expected transmission count metric to take into account the position and the direction of the node. The adaptive position update strategy for geographic detection in [23] adjusts the frequency of position updates based on the mobility dynamics of the nodes and the forwarding patterns in the network. The energy-aware dual-path geographic detection protocol in [24] provides two node-disjoint anchor lists to shift detection path for load balance and utilizes the location information and the characteristics of energy consumption to make detection decisions, which reduce the energy consumption and improve the packet delivery ratio. The localization and energy-efficient data detection in [25] uses fuzzy logic inference to calculate the locations of unknown nodes and select the cluster head based on the locations of node.

RL algorithms are utilized to improve the detection performance due to the constrained environment of As. For example, the adaptive hybrid communication protocol as presented in [13] combines the directional antennas and position prediction in the MAC layer to overcomes the directional deafness problem and uses Q-learning to update the local detection policies for less delivery delay. The Q-learning-based topology-aware detection protocol as presented in [26] uses two-hop neighbor information to obtain the node position and link quality and adjusts the Hello message interval dynamically to reduce the communication latency. The decentralized deep reinforcement learning-based control framework as presented in [27] uses the actor and critic networks to address instability issue and performs multi-node navigation distributedly to reduce energy consumption. A distributed adaptive opportunistic detection algorithm in [28] uses a reinforcement learning framework to adapt their detection strategies, which minimizes the expected average per-packet.

III. Broad Learning for network intrusion detection

As shown in Fig. 1, we consider d number of nodes in A, whose index sets are represented as $\mathcal{N} = \{1, 2, \ldots, d\}$. All nodes are randomly deployed in 3D space and each node is equipped with a GPS. Each node node operates in the half-duplex mode and can be in one of two states: transmitting state or the receiving state. Except the source node 1 and the destination node d, all other nodes are forwarding nodes, which could serve as both receiver and transmitter nodes. node 1 sends N data packet to node d with power p. Let $i_n^{(k)}$ denote the index of the node which at time slot k transmits packet n. The detection algorithm can be viewed as selecting a sequence of nodes $\{i_n^{(k)}\}$ for relaying packets $n = 1, 2, \ldots, N$.

Let $N_i^{(k)}$ denote the neighbor set of node i, in which all nodes are cooperatively involved in the local forwarding task. All nodes can obtain their one-hop neighbors by sending Hello messages regularly. Each node maintains a neighbor information table, which contains the ID, position coordinates and other information of the neighboring nodes. At time slot k, node i measures the battery level $b_i^{(k)}$, estimates the channel gain with other nodes in the neightbor set, and calculates the relative distance with neighboring nodes $\boldsymbol{R}_{-i}^{(k)} = \{R_{i,j}^{(k)} | j \in N_i^{(k)}\}$, where $R_{i,j} = D_{j,d}/D_{i,d}$. Then node i selects a forwarding set to broadcast the data packet. node $r_i^{(k)}$ then repeat the above steps until the node d receives the data packet. Once node d receiving the data packet, it broadcasts an ACK packet to all nodes after t-th time slot using a feedback channel.

Each node applies the carrier sense multiple access with collision avoidance (CSMA/CA) MAC protocol to perform spectrum access and prevent collisions between nodes. For repeated packet detection, each node needs to maintain a data packet serial number SEQ. Each time the node 1 sends a data packet, SEQ is increased by 1. After receiving the data packet, other nodes record the SEQ to the detection table and perform repeated packet detection based on SEQ. If it has been received, the node discards the packet, otherwise, the node increments the hop count in the data packet by one.

As shown in Fig. 2, we design the modified MAC frame formats. The data packet frame contains the previous forwarding set, forwarding set, Time to live (TTL) and hop count. After receiving the data packet, node i record the hop count and accumulate transmission latency to the detection table. To avoid occupying too many network resources, we need to set a reasonable TTL value to ensure that data packets pass only limited hops. After receiving the packet, the node d broadcasts an ACK and does not receive packets from other nodes. The

ACK packet frame includes the previous forwarding set, hop count, transmission latency and packet delivery rate.

IV. Reinforcement Learning based detection Algorithm

We propose a reinforcement learning based detection algorithm for nodes to efficiently transmit data packet. Each node is an independent agent and learns its detection policy by observing the local network state and communicating with its neighboring nodes. This algorithm enables each node to share the latest updated Q-values and E-values, and exploits the learning experiences from the neighboring nodes, which accelerates the collaborative detection policy optimization.

State: At time slot k, node i evaluates the size of neighbor set $\beta_i^{(k)}$, calculate the relative distance with neighboring nodes to destination $\boldsymbol{R}_{-i}^{(k)}$ from neighbor table, measures the battery level $b_i^{(k)}$ and the channel gain with other nodes denoted by a, and obtains the hop count $c_i^{(k)}$ from data packet head and the receive signal SNR $\xi_i^{(k)}$.

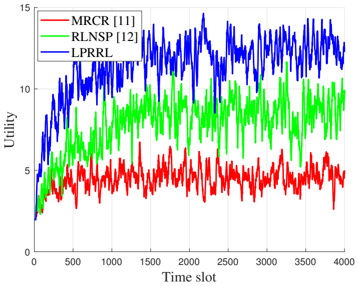
The state $s_i^{(k)}$ consists of the current observed network topology, i.e, the number of neighboring nodes $\beta_i^{(k)}$, the relative distance with neighboring nodes to destination $\boldsymbol{R}_{-i}^{(k)}$, the battery level $b_i^{(k)}$, the channel states with neighboring nodes $\boldsymbol{h}_{-i}^{(k)}$, the receive signal SNR $\xi_i^{(k)}$, and the hop count $c_i^{(k)}$ from data packet head. Thus, the state is given as

$$s_i^{(k)} = \left[b_i^{(k)}, c_i^{(k)}, \beta_i^{(k)}, \xi_i^{(k)}, \boldsymbol{R}_{-i}^{(k)}, \boldsymbol{h}_{-i}^{(k)} \right]. \tag{1}$$

Action: The set of actions is given by the combination of the neighboring nodes, which can be defined as $A_i = \{0,1\}^n$. Among them, 1 means that the neighboring node at the corresponding location in the neighbor set is selected as forwarding node, 0 means the contrary, and $n = \min(|N_i|, \sigma)$ is the number of neighbors that must be considered as possible forwarding nodes. The neighbor set N_i is sorted by the Q-value from neighbor table. Herein, we define the parameter σ that is used when just the σ-th first neighbors in the sorted set N_i must be considered rather than all set of neighbors. This parameter can be used to control the size of A_i and reduce the computation cost to determine the optimal action. Assuming that node i has three neighboring nodes $r_{0,i}$, $r_{1,i}$, and $r_{2,i}$, the set of possible actions is $A_i = \{000, 001, 010, 011, 100, 101, 110, 111\}$, in which $a_i = 110$ means that nodes $r_{0,i}$ and $r_{1,i}$ are selected as forwarders.

Utility: Due to the unstable link caused by the node movement, packet loss may occur. If data packet is received by the destination node d, the packet arrival indicator κ is set as 1, and 0 otherwise. Once node d receiving the data packet, it obtains the end-to-end transmission latency τ and the hop count c, estimates the energy consumption w. Then, node d broadcasts an ACK packet to node i using a feedback channel. node i measures the transmission latency τ_i and the hop count c_i to node d, estimates the energy consumption w_i. If node i does not receive an ACK after expiration time t_n, it obtains a negative value $-\Omega$. The utility $u_i^{(k)}$ is evaluated based on the packet arrival indicator κ, the constant Ω, the transmission latency τ_i, and the energy consumption w_i given by

$$u_i^{(k)} = \begin{cases} \Omega - c_1\tau_i - c_2 w_i, \kappa = 1 \\ -\Omega, \kappa = 0 \end{cases} \tag{2}$$



(a) Utility

Fig. 1. Routing performance.

Let $l(s_i^{(k)}, a_i^{(k)})$ denote the risk level of taking action $a_i^{(k)}$ in the state $s_i^{(k)}$. We assume L risk levels, where risk level L is the most dangerous and zero risk level is safe. The learning agent updates a blacklist denoted by $\mathcal{T} = \{(s_i^{(k)}, a_i^{(k)}) | l(s_i^{(k)}, a_i^{(k)}) = L\}$ to store the most dangerous state-action pairs. Let η_i denote the criterion to measure the risk level i, the learning agent uses a criterion vector $\boldsymbol{\eta} = [\eta_i]_{1 \leq i \leq L}$ according to prior knowledge to identify whether a new explored state-action pair is safe or not. Let denote an indicator that equals to 0 if the argument is true and 1 otherwise. We consider the long-term risk level denoted by $E(s,a)$ for M steps given by

The probability $\pi(s_i^{(k)}, a_i)$ is determined by the risk level and Q-value [29]. The action with higher Q-value and lower risk level will be selected with a higher probability and the action in the blacklist will be forbidden.

At each time slot, the node shares the latest updated Q-value and E-value with the state-policy pair to transfer the learning experiences. The local Q-values and E-values are updated with Q-values from the neighboring nodes via the distributed averaging consensus [30] . More specifically, the node determines the state-policy pairs $\{(\hat{s}_j, \hat{a}_j) | j \in \hat{N}_i\}$ that need to update the Q-value and E-value by discarding the duplicates in the collection of the receiving state-policy pairs $\{(s_j^{(k)}, a_j^{(k)}) | j \in N_i\}$.

V. Simulation Results

As shown in Fig. 1, the proposed scheme improves 85% utility after 4000 time slots.

VI. Conclusion

In this paper, we have proposed a safe reinforcement learning based that treats each unmanned aerial vehicle (node) node as an agent and performs decentralized training and execution. We randomly select a source node 0, which transmits its data to the destination node. All nodes except for the source and destination nodes are relay nodes.

References

[1] S. Gamage and J. Samarabandu, “Deep learning methods in network intrusion detection: A survey and an objective comparison,” *Journal of Network and Computer Applications*, vol. 169, p. 102767, 2020.

[2] Y. Kong, X. Wang, Y. Cheng, and C. Chen, “Hyperspectral imagery classification based on semi-supervised broad learning system,” *Remote Sensing*, vol. 10, no. 5, p. 685, 2018.

[3] J. Fan, X. Wang, X. Wang, J. Zhao, and X. Liu, “Incremental wishart broad learning system for fast polsar image classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 12, pp. 1854–1858, 2019.

[4] M. M. Sakr, M. A. Tawfeeq, and A. B. El-Sisi, “Network intrusion detection system based pso-svm for cloud computing,” *International Journal of Computer Network and Information Security*, vol. 10, no. 3, p. 22, 2019.

[5] J. Gu, L. Wang, H. Wang, and S. Wang, “A novel approach to intrusion detection using svm ensemble with feature augmentation,” *Computers & Security*, vol. 86, pp. 53–62, 2019.

[6] T. S. Yange, O. Onyekware, and Y. M. Abdulmuminu, “A data analytics system for network intrusion detection using decision tree,” *Journal of Computer Sciences and Applications*, vol. 8, no. 1, pp. 21–29, 2020.

[7] C. Khammassi and S. Krichen, “A nsga2-lr wrapper approach for feature selection in network intrusion detection,” *Computer Networks*, vol. 172, p. 107183, 2020.

[8] Q. Wu, Y. Zeng, and R. Zhang, “Joint trajectory and communication design for multi-UAV enabled wireless networks,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2109–2121, Mar. 2018.

[9] L. Hong, H. Guo, J. Liu *et al.*, “Toward swarm coordination: topology-aware inter-UAV routing optimization,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 10 177–10 187, Sep. 2020.

[10] S. Biswas and R. Morris, “ExOR: Opportunistic multi-hop routing for wireless networks,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 4, pp. 133–144, Oct. 2005.

[11] J. Zuo, C. Dong, H. V. Nguyen *et al.*, “Cross-layer aided energy-efficient opportunistic routing in Ad-hoc networks,” *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 522–535, Feb. 2014.

[12] L. Yang, S. Song, and C. P. Chen, “Transductive transfer learning based on broad learning system,” in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2018, pp. 912–917.

[13] Z. Zheng, A. K. Sangaiah, and T. Wang, “Adaptive communication protocols in flying Ad-hoc network,” *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 136–142, Jan. 2018.

[14] Z. Li, A. L. G. Rios, G. Xu, and L. Trajković, “Machine learning techniques for classifying network anomalies and intrusions,” in *2019 IEEE international symposium on circuits and systems (ISCAS)*. IEEE, 2019, pp. 1–5.

[15] H. I. Abbasi, R. C. Voicu, J. A. Copeland *et al.*, “Towards fast and reliable multihop routing in VANETs,” *IEEE Trans. Mobile Comput.*, vol. 19, no. 10, pp. 2461–2474, Oct. 2020.

[16] H. Song, L. Liu, B. Shang *et al.*, “Enhanced flooding-based routing protocol for swarm UAV networks: Random network coding meets clustering,” in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Vancouver, BC, Canada, May 2021, pp. 1–10.

[17] R. Sanchez-Iborra and M.-D. Cano, “Joker: A novel opportunistic routing protocol,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1690–1703, May 2016.

[18] N. Li, J.-F. Martinez-Ortega, V. H. Diaz *et al.*, “Probability prediction-based reliable and efficient opportunistic routing algorithm for VANETs,” *IEEE/ACM Trans. Netw.*, vol. 26, no. 4, pp. 1933–1947, Aug. 2018.

[19] A. Bujari, C. E. Palazzi, and D. Ronzani, “A comparison of stateless position-based packet routing algorithms for FANETs,” *IEEE Trans. Mob. Comput.*, vol. 17, no. 11, pp. 2468–2482, Nov. 2018.

[20] D. Torrieri, S. Talarico, and M. C. Valenti, “Performance comparisons of geographic routing protocols in mobile Ad-hoc networks,” *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4276–4286, Nov. 2015.

[21] D. Medina, F. Hoffmann, F. Rossetto *et al.*, “A geographic routing strategy for north atlantic in-flight internet access via airborne mesh networking,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 4, pp. 1231–1244, Aug. 2012.

[22] S. Rosati, K. Krużelecki, G. Heitz *et al.*, “Dynamic routing for flying Ad-hoc networks,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1690–1700, Mar. 2016.

[23] Q. Chen, S. S. Kanhere, and M. Hassan, “Adaptive position update for geographic routing in mobile Ad-hoc networks,” *IEEE Trans. Mobile Comput.*, vol. 12, no. 3, pp. 489–501, Mar. 2013.

[24] H. Huang, H. Yin, G. Min *et al.*, “Energy-aware dual-path geographic routing to bypass routing holes in wireless sensor networks,” *IEEE Trans. Mobile Comput.*, vol. 17, no. 6, pp. 1339–1352, Jun. 2018.

[25] F. Khelifi, A. Bradai, K. Singh *et al.*, “Localization and energy-efficient data routing for unmanned aerial vehicles: Fuzzy-logic-based approach,” *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 129–133, Apr. 2018.

[26] M. Y. Arafat and S. Moh, “A Q-learning-based topology-aware routing protocol for flying Ad-hoc networks,” *IEEE Internet Things J.*, vol. 9, no. 3, pp. 1985–2000, Feb. 2022.

[27] C. H. Liu, X. Ma, X. Gao *et al.*, “Distributed energy-efficient multi-UAV navigation for long-term communication coverage by deep reinforcement learning,” *IEEE Trans. Mob. Comput.*, vol. 19, no. 6, pp. 1274–1285, Jun. 2020.

[28] A. A. Bhorkar, M. Naghshvar *et al.*, “Adaptive opportunistic routing for wireless Ad-hoc networks,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 243–256, Feb. 2012.

[29] X. Lu, L. Xiao, G. Niu *et al.*, “Safe Exploration in Wireless Security: A Safe Reinforcement Learning Algorithm With Hierarchical Structure,” *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 732–743, Feb. 2022.

[30] D. Lee, N. He, P. Kamalaruban *et al.*, “Optimization for reinforcement learning: From a single agent to cooperative agents,” *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 123–135, May 2020.