Mobile Networks and Applications (2020) 25:725–742
https://doi.org/10.1007/s11036-019-01446-w

Active Defense by Mimic Association Transmission in Edge Computing

Shuo Wang[1] · Qianmu Li[1,2] · Jun Hou[3] · Shunmei Meng[1] · Bo Zhang[1] · Cangqi Zhou[1]

Published online: 3 January 2020

Abstract
A large amount of real-time data, including user privacy information, control commands, and other sensitive data, are transmitted in edge computing networks. It requires high-speed and reliable data transmission in dynamic edge computing networks. Traditional methods with passive defense cannot cope with the covert and complicated attacks. Edge computing networks require active defense during data transmission. Existing active defense methods based on dynamic network ignore the connectivity and link quality reduced by attacks and do not adjust defense positively. To maximize the defense revenue in moving adjustment strategy, this paper proposes the model of active defense for edge computing network data interaction. In this model, the network topology mimic association protocol is designed to associate multi-paths and multi-parameters automatically. On one hand, considering the transmission reliability and defensive revenue reduction caused by dynamic network transformation, a real-time multi-feature anomaly detection algorithm based on Non-extensive entropy and Renyi cross entropy is proposed. Based on this, a moving communication path alliance can be constructed pseudo-randomly. On the other hand, this paper proposes a Hidden Markov based state prediction model and a mimic transformation strategy for The Network Topology Mimic Association Graph based on predicted states. Combining these two ways improves the data transmission service quality of the active defense technology in edge computing networks. Experiments are carried in simulated power networks. The results show that our method outperforms the popular methods in terms of transmission efficiency, reliability, and anti-attack performance.

Keywords Network attack · Active defense · Dynamic network · Edge computing

1 Introduction

One of the primary latent risks in a network is the cyber-attack on the network data interaction layer in the form of edge computing. This is due to the large amount of real-time state acquisition data, user privacy information and control command data present in an edg0e computing network. These data play a decisive role in user privacy protection and system decision control [1]. Alternatively, an edge computing network can perform real-time monitoring and control services on the edge of the critical infrastructure, with strict requirements on the performance of real-time data transmission [2, 3]. Considering data security interactions in an edge computing network, it is important to suppress attacks and execute evasive responses before a network attack causes damage [1, 4, 5]. Therefore, edge computing networks urgently require active defense during data transmission.

However, the current network attack methods (CNAMs) such as the advanced persistent threat (APT) are concealed, and the attack principle is complex. Attack monitoring and passive blocking technologies based on traditional misuse detection have been unable to cope with such attacks [6]. For this reason, active

This paper supported by The Fundamental Research Funds for the Central Universities (No.30918012204), Jiangsu province key research and development program(BE2017739), 2018 Jiangsu Province Major Technical Research Project "Information Security Simulation System"(BE2017100), Military Common Information System Equipment Pre-research Special Technical Project (315075701). Industrial Internet Innovation and Development Project in 2019 - Industrial Internet Security On-Site Emergency Detection Tool Project.

✉ Qianmu Li
qianmu@njust.edu.cn

[1] School of Computer Science and Engineering, Nanjing University of Science and Technology, NanjingP.O. Box 210094, People's Republic of China

[2] Intelligent Manufacturing Department, Wuyi University, JiangmenP.O. Box 529020, People's Republic of China

[3] Nanjing Institute of Industry Technology, NanjingP.O. Box 210023, People's Republic of China

defense faces challenges. Fortunately, the self-organizing nature of edge computing networks provides a foundation for active defense of data interaction [7, 8]. By constructing an uncertain and dynamic network environment, the attacker lacks sufficient time to effectively probe the communication path. The dynamic transformation of data transmission network by edge computation can construct this dynamic network environment. This will reduce the effectiveness of the information collected by an attacker prior to the attack. The information collected during the attack will be outdated and invalid. It will increase the cost and complexity during attacker's information collection and detection. The probability of data being attacked can also be reduced.

The active defense technology based on a moving network can solve the current defense problem of data transmission attacks to some extent and increase the costs of cyber-attacks. However, previous technologies do not consider a moving adjustment in the case of reduced network connectivity and link quality caused by an attack [9]. Thus, the defense strategy of a moving adjustment algorithm requires further optimization and improvement.

Therefore, this paper proposes an active defense model for data interaction processes in edge computing based on a network topology mimic correlation. Main contributions are concluded as follows. Figure 1 shows the framework of our research.

(1) The model is achieved by pseudo-randomly constructing a moving communication path alliance under the premise of ensuring service quality. Here, the network topology mimicking association technology is used to simulate the construction of a dynamic multipath communication alliance to prevent network attacks.

(2) This method integrates the network security state and transmission reliability prediction to actively evade network attacks. The model includes the edge-aware node, the edge computing terminal node, and the primary station system. It uses a negotiated moving multipath communication alliance to secure data communication. A network attacker cannot determine the real communication path in the alliance which increases attack costs [9, 10]; thus, it is difficult to implement an effective attack.

(3) In concurrent multipath communication, the network attacker cannot obtain complete transmission data or control instructions [11, 12]. Thus, network security accidents can be avoided, such as data leakage and command tampering in an edge computing network in advance.

(4) In addition, this method also ensures data transfer efficiency.

The rest of this paper is organized as follows: Section 2 discusses relevant studies on moving network technology in mimicry defense. Section 3 gives some relevant definitions and the overall model framework and design for network topology mimic association protocols. In Section 4, this paper describes a mimic transformation method of communication path alliance based on moving threshold anomaly detection. In Section 5, a mimic transformation method utilizing a mimic topology correlation graph based on a network security state prediction is proposed. Section 6 analyzes the security of the model and verify the performance through experiments. Section 7 summarizes the contents of this paper.

2 Related works

In recent years, the moving target defense (MTD) proposed by the US Science and Technology Commission has attracted much attention as a new cybersecurity mimicry defense technology [13]. Moving network technology, as one of the most critical technologies for MTD at the network layer, has a promising application prospect in active defense.

A suitable communication path transformation strategy is crucial for implementation in moving networks. The communication path transformation strategy is used to generate a network management configuration of nodes that are used during the subsequent adjustment period. The randomness of the configuration increases the difficulty for the attacker in predicting the network management configuration.

Recently, the pseudorandom approach has been extended to address the transformation strategy of moving networks.

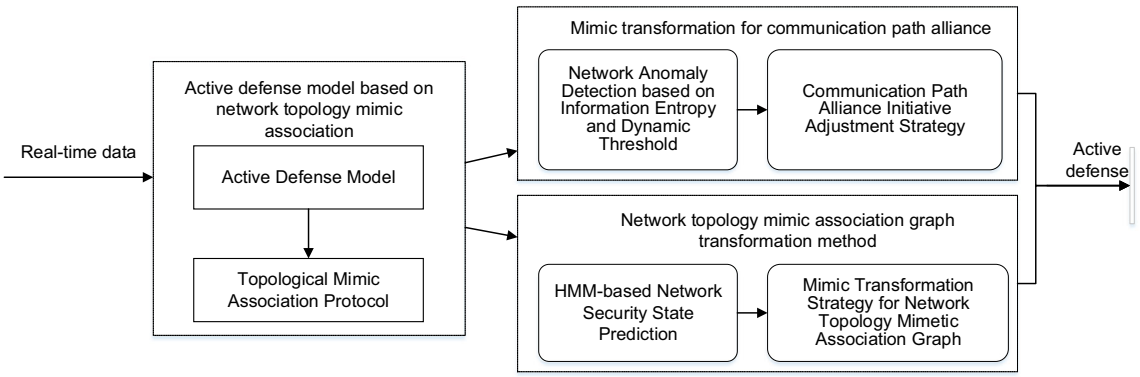

Fig. 1 Framework of active attack defense technology for edge computing network data interactions

Some methods proposed before cope with the randomness of the attack through random port mechanism. Atighetchi et al. [14] proposed a virtual port address association scheme for the client association proxy and a network address translation gateway to fill fake random addresses and ports into the corresponding fields of the data packet. Then, the data stream is redirected to defend against the attack. Once an "expired" node network management configuration is used, the possibility of detection will increase. Badishi et al. [15] developed a random port association mechanism termed random port hopping (RPH). Antonatos et al. [16] established a method for randomizing the network address space based on a transparent address association, which performs a header address translation of data stream packets. This approach maintains the novelty of the address translation table and prevents connection requests outside the service period. Jafarian et al. [17] proposed an OpenFlow random host mutation (35) based on OpenFlow. The authors used OpenFlow to transparently change the IP address of the host to ensure the consistency of the host configuration. These methods perform well but only in static network.

Aimed at the problems of limited hopping space in IPv4 and fixed hopping period, Dunlop et al. [13, 18] proposed moving target defense mechanism based IPv6 (MT6D). In order to enlarge the hopping space, IPv6 address space is adopted. Besides, MT6D uses pseudo-random number to set hopping period so as to improve the randomness. In 2014, Jafarian et al. [19] associated a host IP address with an address block with a short lifetime. The authors proposed a random association method based on the time and space domains to block, spoof and detect attackers.

Based on these, other works also aim to prevent the leakage of MAC address. In 2015, MacFarland et al. [20] hide the link, IP, and port numbers of endpoint by setting up DNS hopping controller so as to prevent the leakage of MAC address. In 2016, Skowyra et al. [21] proposed network identity elimination mechanism called PHARE. It prevents MAC address leakage by randomly transforming header when packets flow out of the endpoint. Moreover, Sun et al. [22] proposed Decoy-Enhanced Seamless IP Randomization (DESIR) to increase the unpredictability. When unauthenticated nodes access the platform, DESIR uses honeypots to observe its behavior. In order to prevent service interruption, DESIR separates the network identifier and transmission identifier of endpoint when it migrates services, thus ensuring the continuity of service provision by reserving the transmission identifier. Pseudorandom functions in moving network are exposed to higher security; however, it is possible that the node network management configuration will collide, in which case, scalability is not desirable.

At present, these researches pay attention to the privacy protection and security transmission of edge computing data. This is generally implemented by password technology and secure transmission protocol. However, most of the researches do not take the real-time requirements of data transmission in edge computing network into account. So, it is difficult to apply them to the real-time security interaction for edge computing data. In addition, the existing works do not consider the impact on data transmission efficiency in the case of network attack. Thus, these methods cannot adjust the data transmission scheme adaptively according to the degree of damage to ensure the transmission efficiency. Therefore, the existing security transmission technology generally belongs to the passive defense technology. They cannot be actively circumvented or actively suppressed by the network attack behavior. Security of the edge computing network cannot be guaranteed.

In general, the implementation of the current moving network technology is simple, but there are several shortcomings: 1) In the existing literature, moving network adjustment strategies primarily focus on static and fixed methods. These approaches cannot be adaptively adjusted in combination with the current network security status. 2) The moving network adjustment strategy needs to compress or amplify the state space of the available node network management configuration. However, current methods with a pseudorandom function have a single control factor, and the generated space of the node network management configuration is difficult to control accurately. Thus, the scalability of the algorithm is weak.

Based on these existing problems, this paper proposed a moving network active defense technology based on network topology mimic association. The proposed method focuses on high security and real-time requirements of data interaction in an edge computing network. For the security threat from unknown vulnerabilities and backdoors, Wu [23] proposed the idea of "cyber mimic defense" based on the principle of uncertainty. The main idea of mimic defense is to reduce the certainty, static and isomorphism of the vulnerability point in the network. Thus, it can increase the attacker's attack difficulty. Then the attacker does not have enough time to probe the target network. Here, the network topology mimic association technology is proposed to improve the active defense efficiency in real-time dynamic data interaction.

3 Secure transmission model based on network topology mimic

3.1 Definition

In this section, the whole edge computing network data interaction model is designed as a three-layer network. The first layer is the sensor node layer consisting of sensing devices. The second layer is the terminal node layer composed of terminals in edge computing. The third layer is the central station. A schematic of the system is shown in Fig. 2. The sensor

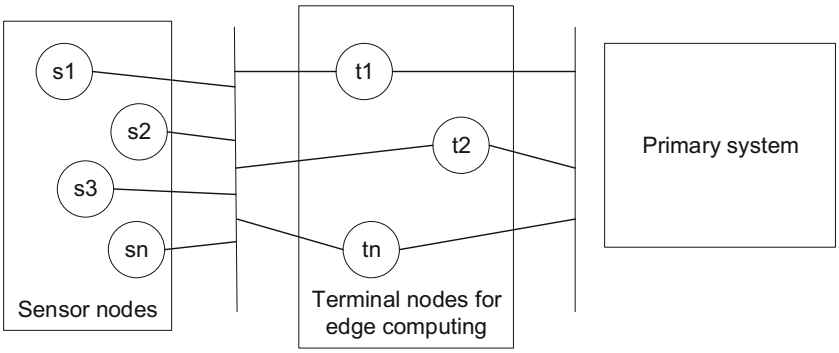

Fig. 2 Abstract model of the edge access network

node directly accesses the edge computing terminal node through a wireless network. The edge computing terminal communicates with the primary station system layer through a wireless or wired network. The edge computing terminal nodes can also communicate with each other. The primary station system layer acts as the control layer for the entire network.

Definition 1: Weighted directed acyclic graph of the edge computing network $G = (V, E, W)$. $V = (v_1, v_2, \ldots, v_n)$ represents the set of nodes $v_i (v_i \in V)$. $E = (e_1, e_2, \ldots, e_n)$ indicates the set of communication paths between node $e_{ij} = < v_i, v_j > (e_{ij} \in E)$. $W = (w_1, w_2, \ldots, w_n)$ indicates the set of weights on the edge e_{ij}. w_{ij} is assigned based on the reliability of the communication path on the edge $e_{ij} = < v_i, v_j > (e_{ij} \in E)$.

Definition 2: Node. $v_i = \{id_i, ip_i, pt_i, \mu_i, ns_i, nd_i, nn_i, plower_i\}$, where Id_i is a unique identifier of the node, which can be assigned based on the node type (perceived node, edge computing terminal node). ip_i is the IP address of node v_i. pt_i is the set of available ports for the node. The port is the address of the layer interaction between the various protocol processes and the transport entity in the application layer. The available range is 0 to 65,535; aside from the first 1024 well-known ports, there remain 64,512 available ports. μ_i is an application layer protocol set supported by node v_i. For example, in a smart power network, the protocols used for communication between sensor node and data interaction primarily include application layer protocols such as the IEC60870–5-101/104 and 62,351 protocols. ns_i is the collection of source nodes for node v_i, and the value of the sensor node is null. nd_i is the set of the next hop node of v_i. nn_i is the neighbor node of v_i, and its value of the sensor node is null. p_{lower_i} is the minimum requirement for node communication reliability.

Each sensor node in the edge computing network assigns its own attribute information when the system is initialized. Adjacent edge computing terminal nodes interacts with each other. Each sensor node records an edge computing terminal node whose signal can be monitored. A sensor node can be subordinate to multiple edge computing terminal nodes. The terminal routing table records ip_i of the edge computing terminal node and the corresponding path node queue of visiting the primary station system. The edge computing terminal node records the addresses of all dependent sensor nodes and simultaneously maintains a network neighbor node table at the same layer. In addition, like the sensor nodes, the edge computing terminal nodes need to maintain each dependent sensor node's path node space queue of the network topology mimic association communication. The primary station system layer constructs a network topology weighted directed graph after obtaining all of the network data.

In this paper, the network topology mimicking association technology is used to simulate the construction of a dynamic multipath communication alliance to prevent network attacks. When the network topology is found to be changed, all nodes exchange information by means of distributed propulsion. Moreover, the sensor node retains the IP addresses of the other edge computing terminal nodes that can be accessed. At the same time, when an abnormality or attack is detected in the network, the edge computing terminal node instructs the sensor node to change the communication path. When a new sensor node joins the subnet, only the members of the sensor node are added while the subnet remains, and the relevant information is directed to the primary station system.

The topology-directed acyclic graph based on the edge computing network consists of multiple subgraphs. Each subgraph consists of a sensor node, all accessible paths between the serving node of the primary station and the associated edge computing terminal nodes. There is only one initial node in each subgraph and one termination node, i.e., the sensor node and the primary station serving node, respectively. In the initialization phase of the network topology mimic association system, the primary station service node generates a space queue of network topology mimic association node for the sensor node according to the subgraph. However, the edge computing network is affected by the network attack. Thus, it needs to dynamically cut off edge computing terminal nodes that are in an abnormal state and add edge computing terminal nodes that are restored to normal. Therefore, not all edge computing terminal nodes or edges in the subgraph can meet the requirements. It is necessarily to adaptively filter the available

paths that meet the reliability requirements. Then, a network topology mimic association graph can be constructed.

Definition 3: Network topology mimic association graph $S_i(t) = \{s_k(t) | 1 \le k \le m\}$. S_i represents a sequence of nonintersecting paths whose reliability meets the requirement from the primary station serving node to the sensor node v_i at time t. m indicates the number of available nonintersecting paths.

An example of the network topology mimic association graph is shown in Fig. 3. In Fig. 3, there are three available nonintersecting paths connecting the primary station serving node and the sensor node. Each path is selected according to the reliable principle. This graph changes with time t.

Definition 4: Nonintersecting communication path $s_k(t) = \{(Sip_k, Spt_k, S\mu_k, Sns_k, Snd_k)^T | 1 \le k \le m\}$. If and only if two acyclic communication paths s_1 and s_2 share sensor node and the primary station serving node but do not share edge computing terminal nodes, they are considered as nonintersecting communication paths. The IP sequence of the path node is represented as $Sip_k = \{sip_1, sip_2, \cdots sip_n\}$. The available communication port sequence is $Spt_k = \{spt_1, spt_2, \cdots spt_n\}$. The available communication protocol sequence is $S\mu_k = \{s\mu_1, s\mu_2 \cdots s\mu_n\}$. The source address sequence is indicated as $Sns_k = \{Sns_1, Sns_2, \cdots Sns_n\}$. The destination address sequence is $Snd_k = \{Snd_1, Snd_2, \cdots Snd_n\}$. Here, n is the number of edge computing terminal nodes in k nonintersecting paths.

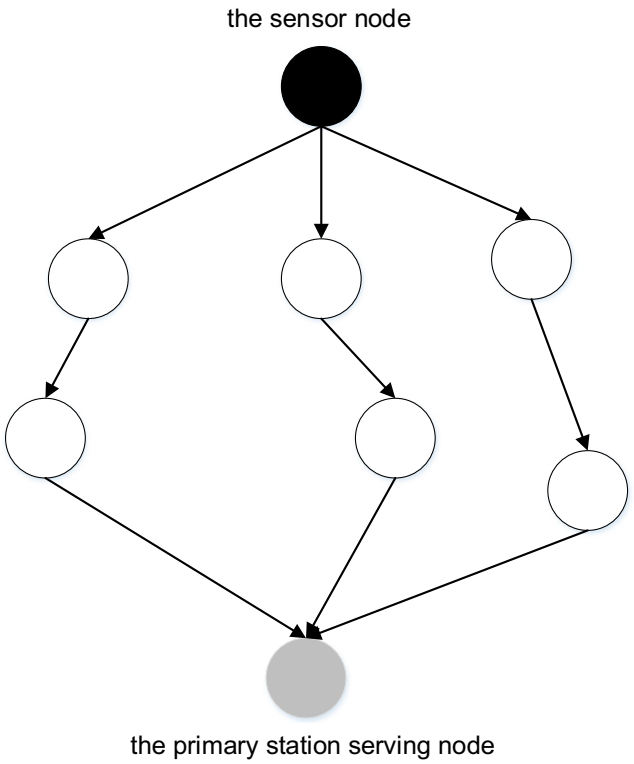

Fig. 3 Network topology mimic association graph

The dynamic communication path alliance is essential to realizing network topology mimic correlation technology. This alliance is determined by the sensor node and the primary station service node according to the network topology mimic association graph and the network security situation. To ensure secure communication, the legitimate communication parties can obtain the necessary information for reorganizing the original data stream at the correct time and on the correct network topology mimic graph.

Definition 5: Moving communication path alliance GS_i. This alliance is a seven-tuple, i.e., $GS_i(t) = \left(S_i, \Phi_{i1}, \Phi_{i2}, C_i, \Re_i, T_S^i, T_{GS}^i\right)$. S_i is the network topology mimic graph. Φ_{i1}indicates a random number for selecting the number of the dynamic communication path. Φ_{i2} represents a random sequence of nonintersecting communication path mapping numbers. C is the association condition. δ is the network topology mimic association transfer relationship. T_S^i is the survival time slot of the network topology mimic association graph. T_{GS}^i is the survival time slot of the moving communication path alliance.

$\Phi_{i1} \le m$ is generated by a pseudorandom function and is used to determine the number of dynamic communication paths.

$\Phi_{i2} = \left(r_i^1, r_i^2, \cdots, r_i^{\Phi_{i1}}\right)$ is a nonintersecting path mapping random number sequence and is also generated by a pseudorandom function. It gives a sequence of dynamic communication path alliance in a nonintersecting path. The association and communication process between two parties depends on Φ_{i2}. The pseudorandom number sequence ensures the randomness of the network topology mimic association negotiation generation and reduces the ability of an attacker to detect the communication path.

C indicates generation conditions for the network topology mimic association negotiation, such as the established negotiation time or a new association negotiation request. $C = (c_1, c_2, \cdots, c_p)$ indicates different trigger conditions.

$\Re \otimes S_i(t_j) \rightarrow^{c_i} S_i(t_{j+1})$ indicates the network topology node configuration process adopted when the communication state changes from $S_i(t_j)$ to $S_i(t_{j+1})$ when the condition for the network topology mimic association is C.

T_S^i is the survival time slot of the network topology mimic association graph. The network topology mimic association graph is replaced at every interval of T_S^i, where the number of handovers is indicated as i.

T_{GS}^i is the dynamic communication path alliance survival time slot. The communication path is updated every interval of T_{GS}^i, where the number of switching steps is indicated as i.

Definition 6: Moving communication path alliance node association configuration $\Omega_i(t)$. This term indicates the configuration of the association factors such as the port and protocol of the edge computing terminal node in the communication path at time t. $\Omega_i(t)$ is defined as $\Omega_i(t) = \Phi_{GS}(t) \times (Spt(t) \times S\mu(t))$. It means the network configuration of k edge computing terminal nodes at communication path i. Spt_krepresents the available port of nodes in the moving communication path alliance. $S\mu_k$is the available protocal. $\Phi_{GS} = \left\{ \left(\Phi_{GS}^{Spt}, \Phi_{GS}^{S\mu} \right)^T \right\}$ indicates sequences of random data. $\Phi_{GS}^{Spt} = \left\{ r_1^1, r_1^2, \cdots, r_1^k \right\}$. $\Phi_{GS}^{S\mu} = \left\{ r_2^1, r_2^2, \cdots, r_2^k \right\}$. Here, r is the function for random data, $rand_i(seed_i)$, $i = 1, 2$.

Ω presents two scalability advantages for the entire system. On the one hand, when it needs to reduce the overhead of the topological linkage of the entire system and dynamic adjustment, the communication path parameters may not have to be updated except for Ω. On the other hand, when the network status of the system is not safe, the dynamic update of Ω will further improve the security of the network topology mimic association. Compared with the pseudorandom function, the network topology mimic association proposed in this paper is more scalable. For example, this association supports both IPv4 and IPv6 and other related factors. At the same time, this approach ensures that the association strategy is controllable and avoids collisions in the node network configuration.

The summary of definitions for main variables is shown in Table 1.

3.2 Framework

The proposed model deploys the network topology mimicking association agent in the primary station system and the sensor node. The structure of the model is shown in Fig. 4.

- The network topology mimicking association agent module is the core. This module controls other modules and available associated communication nodes. It coordinates the communication path between the sensor node and the primary station service node. This module generates a moving communication path alliance.
- After the sensor node and the primary station server node negotiate the network topology mimic association graph, the time synchronization module is used to calibrate the local clock and to enter the network topology mimic association communication mode.
- The traffic distribution module allocates traffic according to the established communication path. Data sent by legal sensor nodes are transmitted to the proxy control module through the currently active communication path. Then, the data are sent to the primary station service node by the traffic reorganization module. The server is also returned to the client by the traffic distribution module and the active path node.
- The delay processing and anomaly detection modules sample the network data stream to evaluate network anomalies and delays. The associated agent control module dynamically changes the mimic association graph configuration of the network topology and the moving communication path alliance according to the evaluation results by using a self-tuning strategy.
- The intrusion detection module detects intrusion based on the redundancy voting mechanism of the mimicry defense model for the edge computing terminal. By comparing the execution results of the heterogeneous redundant execution body, result deviations and network intrusion behavior can be identified.
- The moving communication path alliance and the network topology mimic association graph in the network topology mimic association model change by using an adaptive strategy. This action increases the diversity and randomness of transmission throughout the entire edge computing network and increases the defense strength. In addition, only the available edge computing terminal nodes in the active period can be activated at any time. Each available edge computing terminal node is allocated a node association configuration for the communication path, which will further reduce the possibility that the system communication process will suffer from a network attack.

The primary purpose of network security defense is to pursue higher defense gains under the premise of ensuring network service efficiency. It is to offset higher attack losses with lower protection costs. During data transmission in an edge computing network, the protection cost primarily arises from the network service reliability based on the network topology mimic association graph and the moving communication path alliance adjustment forced by a network attack [24]. In severe cases, an attack will affect regular access and data transmission of the sensor node. Service reliability impairment refers to a reduction in system performance such as data transmission efficiency and computing power, as well as the time cost of system switching. The survival time slot primarily determines the adjustment of the network topology mimic graph and the moving communication path alliance. The service reliability impairment increases as the survival time slot decreases. When the network is not abnormal due to an attack and the survival time slot is infinite, system service reliability impairment due to network topology mimic correlation is minimal. However, if the survival time slot is too long, attackers will have sufficient

Table 1 Definitions of main symbols

G	Weighted directed acyclic graph
V	Node set of edge computing network
E	Communication paths between nodes
W	Weight of the edge
id_i	Identification of node v_i
ip_i	Ip address of node v_i
pt_i	Available port set
μ_i	Application protocol set provided by node v_i
ns_i	Source node set of node v_i
nd_i	Next hop node set of node v_i
nn_i	Neighbor node of node v_i
$plower_i$	The minimum reliability requirements of v_i
S_i	Network topology mimic association graph
Sip_k	Ip address sequences of node in non-intersecting communication paths
Spt_k	Sequences of available port in non-intersecting communication paths
$S\mu_k$	Available protocol of non-intersecting communication paths
Sns_k	Sequences of source address
Snd_k	Sequences of destination address
GS	Dynamic communication path alliance
Φ_{i1}	Random number of paths in dynamic communication path alliance
Φ_{i2}	Random sequence of non-intersecting path map number
C	Negotiation generation conditions of network topology figurative association
$\Re$	Network node configuration process in network topology mimic association transformation
T_s^i	Survival time slot for network topology mimic association graph
T_{GS}^i	Survival time slot for dynamic communication path alliance
$\Omega_i(t)$	Node association configuration for dynamic communication path alliance
I_a	Renyi Cross Entropy
$\delta_{t,f}$	Threshold of network anomaly detection
a_{ij}	Probability of hidden state transition in HMM model
$\psi_t(i,j)$	Transfer probability of hidden states
ζ_t	Forward observation probability
ω_t	Probability of backward observation
γ_t	Probability that the system is in a hidden state at time t
Γ_{TH}^{t+1}	Transfer probability vector of the hidden state
Sec_{TH}^{t+1}	The extent of the cyber threat

time to scan and detect the target system before launching an attack. The attackers could accurately and effectively execute a follow-up attack or a semi-blind attack, and the defense revenue of the system would decrease. In contrast, a high-frequency dynamic adjustment would lead to a greater service reliability impairment, which may also result in reduced defense revenue [11, 25]. Therefore, it is necessary to establish a reasonable active adjustment strategy for the network moving communication path alliance$GS_i(t) = \left(S_i, \phi_{i1}, \phi_{i2}, C_i, \Re_i, T_S^i, T_{GS}^i\right)$, the network topology mimic association graph $S_i(t) = \{s_k(t) | 1 \le k \le m\}$ and the corresponding survival time slots $S_n^-(t+1)$, $S_m^+(t+1)$ as shown in Fig. 5.

3.3 Process of network topology mimic secure transmission

This section designs the network topology mimic association protocol flow. In this step, the server and the client determine the network topology weighted directed graph by negotiation and generate the corresponding network topology mimic association graph. Then, the client pseudo-randomly selects the communication path alliance. The communication parties are allowed to establish independent transport layer connections on multiple dynamic communication paths. In this manner, they can communicate safely according to the established communication path. This process is shown in Fig. 6.

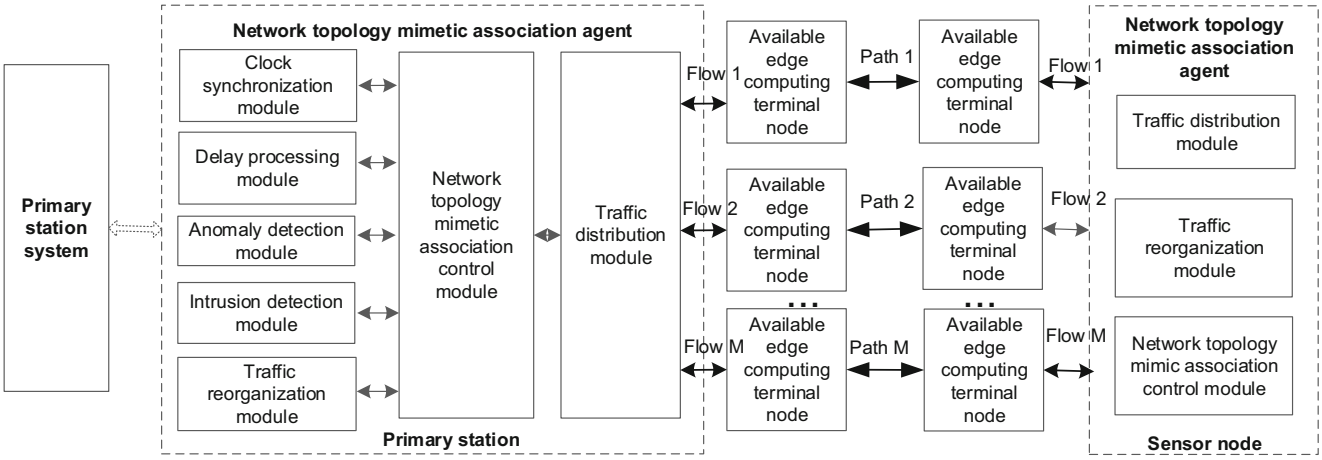

Fig. 4 Model of active defense for edge computing network data interaction

Step 1: When a sensor node which supports the network topology mimic association, accesses the edge access network for the first time and prepares to communicate with the primary station system, the direct access will be denied. Because the edge computing terminal node does not turn on related the access control for data transmission. The sensor node can access only quarantine authentication domain A for identity authentication and trust evaluation. However, once the node authentication and trust evaluation are successful, the edge computing terminal node will open the network access port of the primary station service node.

Step 2: The sensor node sends the regular request message $\mathrm{Req}\{ID_c, Ip_c, ReqID, p_{lower}, mark, T_1\}$ to the primary station node. ID_c is the identity of the sensor node. Ip_c is the IP address of the sensor node. *ReqID* is the corresponding unique ID of each Req message. p_{lower} is the minimum reliability requirement. mark is the support flag of the network topology mimic association. T_1 is the time.

Step 3: The primary station service node records the time T_2 at which the message *Req* is received. If the server does not support the network topology mimic association, the message can be ignored. If the association is supported, the primary station service node switches to the network topology mimic association negotiation mode.

Step 4: The primary station service node initiates a deep search algorithm to find an available path that satisfies p_{lower} between the sensor nodes. Then, a network topology weighted directed graph is generated. $p_{i,j}$ denotes the path reliability between the connecting nodes i and j. $p_{s,t}^k$ denotes the path reliability of the kth path between the primary station serving node s and the sensor node t at time t. In this case, $p_{s,t}^k = \prod_{(i,j)\in k} p_{i,j}$, and $p_{s,t}^k$ should be greater than p_{lower}.

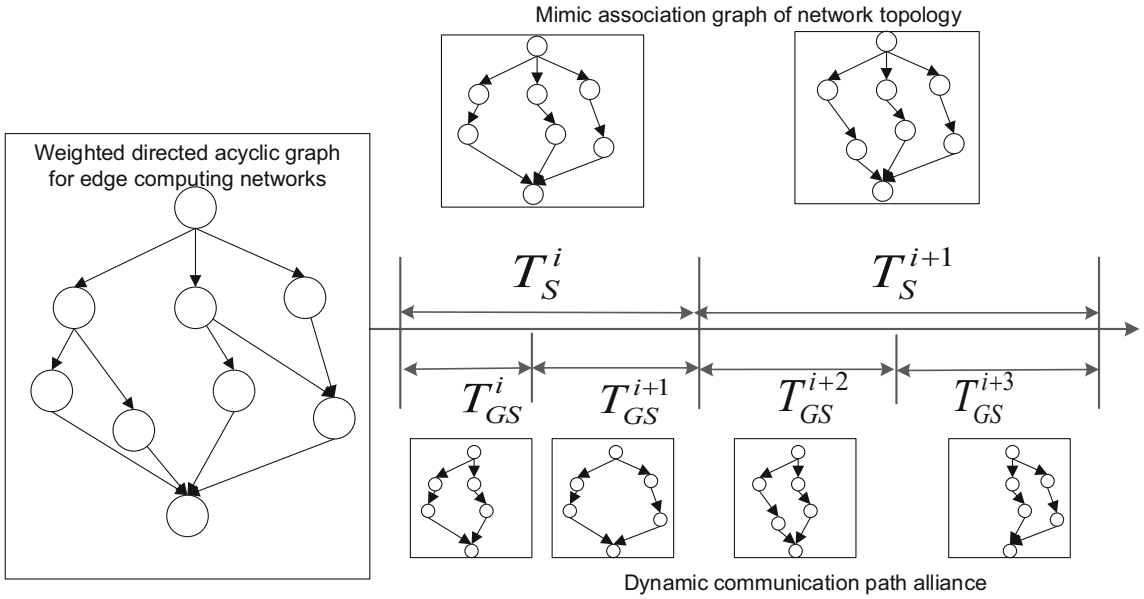

Fig. 5 Survival time slots for the network topology mimetic association graph and moving communication path alliance

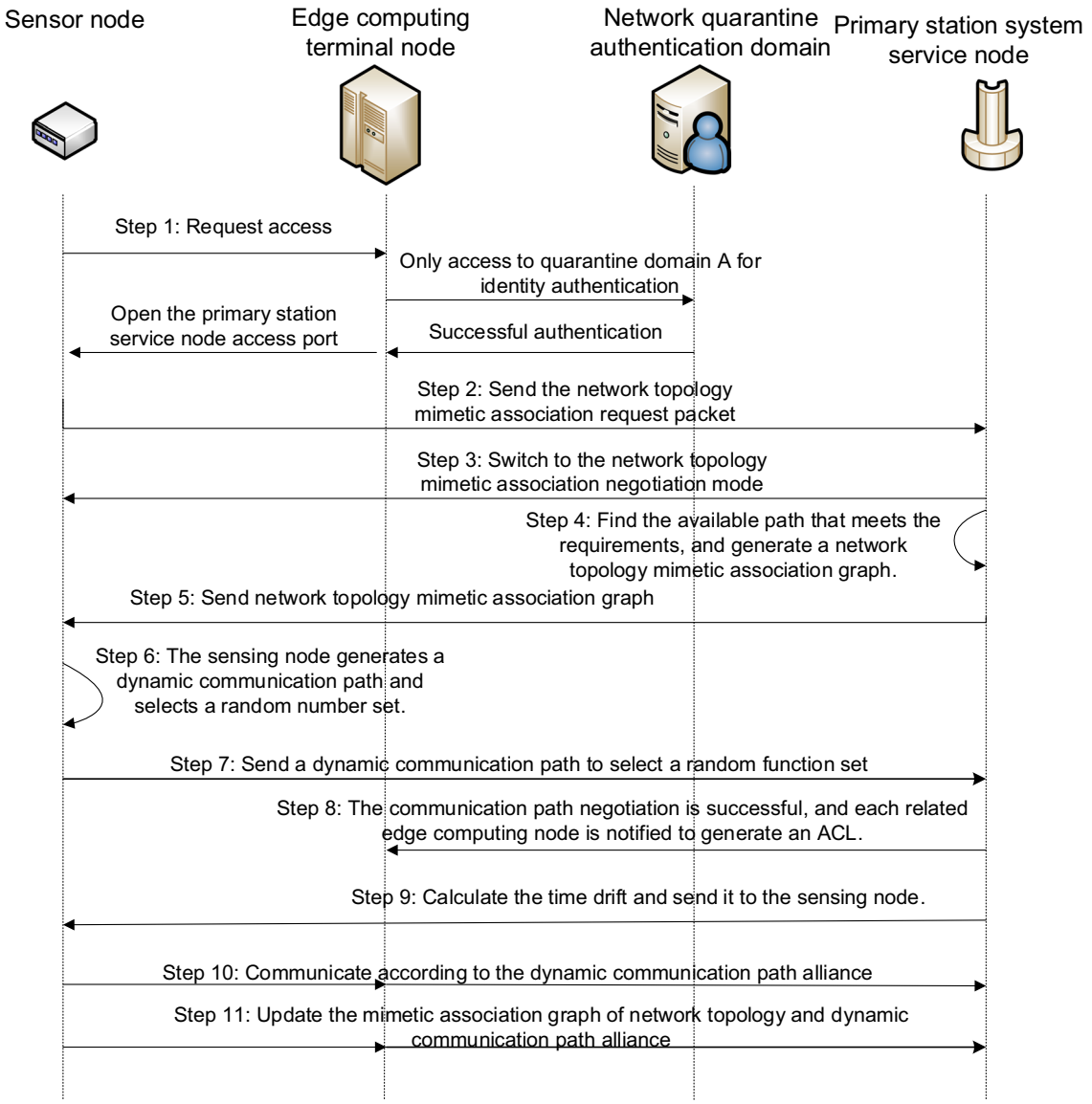

Fig. 6 Network topology mimetic association protocol

Step 5: The primary station service node generates a corresponding network topology mimic association graph $S_i = \{s_k | 1 \le k \le m\}$ based on the network topology weighted directed graph. Next, a response message $\mathrm{Rsp}\{ID_s, S_i, T_3\}$ is sent to the sensor node, including the server identity ID_s, the network topology mimicking association graphS_i, and the response packet sending timestamp T_3.

Step 6: The sensor node records the time T_4 at which the message $\mathrm{Rsp}\{ID_s, S_i, T_3\}$ is received. At the same time, the sensor node generates Φ_{i1}, Φ_{i2}, Φ_{GS} by a random function to determine the network topology mimicking dynamic communication path alliance $GS_i(t)$ and the communication path node association network configuration space $\Omega_i(t)$.

Step 7: The sensor node sends a response message $Rsp\{ID_c, \phi_{i1}, \phi_{i2}, \phi_{GS}, T_5)$ to the primary station serving node.

Step 8: The primary station serving node receives the packet $\mathrm{Rsp}\{ID_c, \Phi_{i1}, \Phi_{i2}, \Phi_{GS}, T_5\}$ and records the time at which the packet is received asT_6. Then, a corresponding ACL is sent to notify all edge computing terminal nodes on the communication path with Ip_c and $\Omega_i(t)$ together.

Step 9: The primary station service node calculates the time drift $\theta = (T_2 - T_1 + T_3 - T_4 + T_6 - T_5)/2$ according to the timestamps T_1, T_2, T_3, T_4, T_5, T_6 and sends θ to the sensor node.

Step 10: The primary station service node adjusts the local time according to the time drift θ by synchronization correction. The sensor node and primary station node implement secure communication according to the established dynamic communication path alliance.

Step 11: When any life cycle of the network topology mimic association, T_S^i or T_{GS}^i, ends normally or abnormally

at the end of the network attack, the network topology mimic association is re-updated.

4 Communication path Alliance mimic transformation method

During data transmission in the edge computing network, if the abnormality detection module changes the moving communication path alliance in the network topology mimic association graph whenever a network attack is detected, the moving communication path alliance switching frequency and the network topology mimic association switching frequency will increase dramatically. This will lead to a decline in network communication efficiency. Cyber-attacks necessitate a process of scanning, lifting, destroying, and so on. Before some of the preliminary steps are completed, the attack does not pose a real threat to the entire system, but it does cause network anomalies to a certain degree [26, 27]. Therefore, in this section, the communication path is adjusted based on a network anomaly metric. When the network anomaly metric exceeds a certain threshold, the moving communication path will be adjusted automatically.

4.1 Network anomaly detection based on information entropy

Many scholars have reported research on measures that can be used for network anomaly metrics. When network traffic is abnormal, changes arise in the distribution of features such as IP addresses and port numbers. However, it is difficult to describe the changes in flow characteristic distributions caused by a network anomaly. In 1948, Shannon first introduced the concept of entropy to information theory and proposed the concept of information entropy. Entropy is an essential concept for measuring the variation of a system parameter distribution and can be used to describe the distribution of network traffic with respect to specific characteristic parameters [26, 27]. When the distribution of the characteristic parameters is more dispersed, the entropy value is larger, and vice versa. Shannon entropy is suitable for describing a system with a normal distribution, while network traffic characteristics present a non-Gaussian distribution. For the network topology mimic association algorithm, the primary station service node T^i_{GS} receives a legal data packet in the association period and should match the network configuration of the communication path GS_i. However, in the case of a network attack, the attack packet prevents the network traffic from matching the network configuration of the communication path GS_i, which will cause some network traffic characteristics to have an abnormal probability distribution. To this end, this paper introduces the Tallies entropy to analyze the unusual characteristics of network traffic. Tallies entropy is Shannon entropy with a full parameter, which is defined as follows:

$$S_q(X) = \frac{1}{q-1}\left(1-\sum_{i=1}^{n}(p_i)^q\right) \tag{1}$$

$p_i = \frac{a_i}{S}$ indicates the probability of occurrence of an event a_i. a_i denotes the number of occurrences of feature elements (such as source IP, destination IP, source port, destination port, etc.) during the observation time, where $\sum_{i=1}^{N}\frac{a_i}{S} = 1$. q is an extensive parameter and plays an important role in the statistical analysis of Tallies entropy. The degree of offset of q from 1 represents the degree of non-extensiveness of the entropy function. The value of q affects the contribution of event a_i toS_q. When q > 1, a large-probability event makes a large contribution. When q < 1, a small-probability event makes a large contribution. If $q \to 1$, the Tallies entropy is consistent with the Shannon entropy.

To accurately measure changes in the security state of the communication path and identify network anomalies, the step size will be set to 0.5 and take 9 values from [−2, 2]. Thus, the distribution state of the characteristic parameter at time t can be expressed as $S_{t,f} = \{S_{q1}, S_{q2}, \cdots, S_{q9}\}$. In this way, the characteristic distribution state of each feature element in the life cycle of the network topology mimic associated communication path is judged by nine different Tallies entropy values. Specifically, no adjustment value is needed for different abnormalities.

In network topology mimic correlation technology, the anomaly detection module collects network traffic over the sampling analysis period $t = T^i_{GS}$ for analysis. The characteristic parameters are extracted from the packet header, such as the source/destination IP address, source/destination port and so on. The Tallies entropy value corresponding to each characteristic parameter at time t is calculated separately, and each Tallies entropy value is normalized as follows:

$$S'_{t,sip} = \left\{s'_{q1}, s'_{q2}, \cdots s'_{q9}\right\}; \tag{2}$$

$$S'_{t,dip} = \left\{s'_{q1}, s'_{q2}, \cdots s'_{q9}\right\}; \tag{3}$$

$$S'_{t,spt} = \left\{s'_{q1}, s'_{q2}, \cdots s'_{q9}\right\}; \tag{4}$$

$$S'_{t,dpt} = \left\{s'_{q1}, s'_{q2}, \cdots s'_{q9}\right\}; \tag{5}$$

Using nine different Tallies entropy values of different feature parameters to determine whether an abnormality has occurred. Whether an abnormality has occurred is judged by comparing the difference between $D_{t,f}$ and $D_{t-1,f}$. In this paper, Renyi cross-entropy is used to measure the difference between two probability distributions [16]. When no anomalies occur, the cross-entropy tends to zero. When an anomaly occurs, the cross-entropy will change abruptly.

For simplicity, taking the characteristic distribution of the source address as an example. Suppose that the anomaly detection module samples and counts data packets for two adjacent time periods to obtain a set of source address sampling data in the network traffic, $sip_{t-1} = \{sip_1, sip_2, \cdots, sip_n\}$ and $sip_t = \left\{sip_1^{'}, sip_2^{'}, \cdots, sip_n^{'}\right\}$. Thus, the Renyi cross-entropy between sip_{t-1} and sip_t is

$$I_a(sip_{t-1}, sip_t) = \frac{1}{1-a} log_2 \sum_{i=1}^{n} \frac{(p(sip_i))^a}{(p(sip_i^{'}))^{a-1}} \tag{6}$$

When $a = 0.5$, the cross-entropy is symmetrical, that is, $I_{0.5}(sip_{t-1}, sip_t) = I_{0.5}(sip_t, sip_{t-1})$. The cross-entropy is then rewritten as

$$I_{0.5}(sip_{t-1}, sip_t) = 2log_2 \sum_{i=1}^{n} \sqrt{p(sip_i)p(sip_i^{'})} \tag{7}$$

Then, the change in the probability distribution for each feature parameter f at time t can be obtained by calculating the Renyi cross-entropy $I_{t,f} = I_{0.5}(S_{t-1,f}, S_{t,f})$.

Therefore, the Renyi cross-entropy of the source/destination IP address and source/destination port at times t-1 and t are calculated respectively:

$$I_{t,sip}\left(S_{t-1,sip}^{'}, S_{t,sip}^{'}\right), \tag{8}$$

$$I_{t,dip}\left(S_{t-1,dip}^{'}, S_{t,dip}^{'}\right), \tag{9}$$

$$I_{t,spt}\left(S_{t-1,spt}^{'}, S_{t,spt}^{'}\right), \tag{10}$$

$$I_{t,dpt}\left(S_{t-1,dpt}^{'}, S_{t,dpt}^{'}\right) \tag{11}$$

An abnormality threshold $\delta_{t,f}$ is introduced. For edge computing networks, the change in network traffic in real time based on a fixed threshold is unreasonable. Therefore, this paper proposes a method to set a dynamic threshold. $I_{t,f}$ is set as the cross-entropy mean value of f at time t and $\sigma_{t,f}$ as the standard deviation of the characteristic $f$cross-entropy at time t. Thus,

$$\delta_{t,f} = I_{t,f} \pm \sigma_{t,f} = \frac{\sum_{t=t-a}^{t-1} I_{t,f}}{a} \pm \sqrt{\frac{1}{a-1}\left[\sum_{t=t-a}^{t-1}\left(I_{t,f} - I_{t,f}\right)^2\right]} \tag{12}$$

Here, a represents a data observation sliding window.

4.2 Strategy for moving communication path alliance mimic transformation

The dynamic adjustment of the life cycle of the moving communication path alliance must meet the principle of "increase slowly and decrease rapidly". That is, when no network abnormality is detected and the probability a network attack is small, the survival time of the moving communication path alliance of the next association cycle slowly increases [28, 29]. Moreover, as the duration of the non-attack state increases, the growth rate of the current moving communication path alliance should also increase to improve the quality of the communication service. When a network abnormality is detected and the probability of a network attack is substantial, the survival time slot of the active communication path alliance in the next period is rapidly reduced. As the abnormal state duration increases, the reduction range of the survival time slot of the active communication path alliance in the next cycle should also increase to ensure communication security.

Assuming that $\sigma_{t,f}^{'}$ is the standard deviation at time t and $\delta^{'}$ is the threshold for a network outlier. Based on expert experience, choosing a function that meets the principle of "increase slowly and decrease rapidly", i.e.

$$g\left(\sigma_{t,f}^{'}\right) = \begin{cases} g_1\left(\sigma_{t,f}^{'}\right), 0 < \sigma_{t,f}^{'} \le \delta' \\ g_2\left(\sigma_{t,f}^{'}\right)\sigma_{t,f}^{'} > \delta' \end{cases} \tag{13}$$

$$\text{With} g_1(\delta^{'}) = g_2(\delta^{'}), g_1^{'}\left(\sigma_{t,f}^{'}\right) < 0, g_2^{'}\left(\sigma_{t,f}^{'}\right) > 0,\ g_1^{'}\left(2\delta^{'} - \sigma_{t,f}^{'}\right) + g_2^{'}\left(\sigma_{t,f}^{'}\right) > 0$$ .The active adjustment strategy is.

$$T_{GS}^{i+1} = \begin{cases} \left(1 + g_1\left(\sigma_{t,f}^{'}\right)\right) * T_{GS}^{i}, 0 < \sigma_{t,f}^{'} \le \delta^{'} \\ \left(1 - g_2\left(\sigma_{t,f}^{'}\right)\right) * T_{GS}^{i}, 0 < \sigma_{t,f}^{'} > \delta' \end{cases} \tag{14}$$

5 Transformation method for the network topology mimic association graph

After the network topology weighted acyclic graph and the network topology mimic association graph is successfully generated, sensor nodes and primary station service nodes select reliable path randomly to communicate safely with multi-path. The network topology mimic association graph and the communication path negotiation confirmation mechanism increase the randomness of the communication path selection and ensure communication efficiency and security by satisfying the reliability requirements [21]. However, the state of the edge computing network changes dynamically. Some accessible paths that do not meet the reliability requirements are improved after the network attack is eliminated and then become available nonintersecting paths. Partially guided paths are subject to network attacks or other uncontrollable factors, which may not satisfy the reliability requirements. Therefore, both parties need to expand or compress the network topology mimic map space according to the established strategy.

When there is a given sequence of observed symbols, the hidden Markov model is suitable to predict the probability of occurrence of a new observed symbol sequence. The hidden Markov model is a stochastic process of the relationship between the observable variable O and the hidden variable S. It is very similar to the abnormal metric (hidden state) and the security state(observable state) of the security situation system [30, 31]. Therefore, using the hidden Markov model can well analyze the network security situation.

Here, this section proposes a hidden Markov based reliability prediction model of network security to realize a network security reliability prediction based on network security anomaly metric data. Based on the security reliability prediction results, the proposed method expands or compresses the network topology mimic association graph and set a reasonable survival time slot T_S^i for the network topology mimic association graph.

5.1 Network security state prediction based on the HMM

The HMM can be described by a quintuple $=(N, M, \pi, A, B)$. In this quintuple, N indicates the number of possible hidden state values in the HMM, which can be recorded as $IS = \{IS_i | 1 \le i \le N\}$. Each hidden state value IS_i corresponds to M observable states O, which is recorded as $O = \{O_i | 1 \le i \le M\}$. Here, π is a $1 \times$ N-order initial probability distribution matrix, indicating the initial probability distribution of the hidden state q_1 for each possible hidden state value for the observable sequence O at time t = 1, $\pi_i = P(q_1 = IS_i)$, $1 \le i \le N$.

$A = (a_{ij})_{N \times N}$ is a hidden state probability transfer matrix for Markov chains. For a first order HMM,

$$a_{ij} = P\left(q_{t+1} = IS_j | q_t = IS_i\right), \sum_{j=1}^{N} a_{ij} = 1, \quad 1 \le i \le N, 1 \le j \le N \tag{15}$$

$B = (b_{im})_{N \times M}$ is a probability matrix of the observed indicators, and the observed probability is $b_{im} = P(O_t = v_m | q_t = IS_i)$, $1 \le i \le N$, $1 \le m \le M$.

To predict the security reliability of all accessible paths in the network topology mimic graph, the method classifies the network security reliability hidden state levels into five categories: safe, mild, general, moderate, and high-risk, expressed as IS_1, IS_2, IS_3, IS_4, IS_5 and assigned to 1, 2, 3, 4, and 5, respectively. Then, the reliability of each accessible path is transferred at a given probability in these five states. At the same time, the network security reliability of each path is defined by two observable indicators, the network transmission efficiency TE and network threat TH. The reliability is expressed as a random variable $x_i (1 \le i \le 2)$. The current security reliability of the entire network is measured from two different dimensions. Then, after time t, the observation sequence $O = \{o_1, o_2, \cdots, o_t\}$ is obtained from observation x_i. To simplify the description, cyber threat prediction is introduced as an example to introduce the prediction algorithm.

First, it is needed to obtain three parameters (π, A, B) from the calculation by means of sample training. Given that O is the sequence of observations for all training samples, to define the probability that the system is in a hidden state IS_i at time t and the system is in state IS_j at time t + 1 as follows.

$$\psi_t(i,j) = P\left(q_t = IS_i, q_{t+1} = IS_j | O, \lambda\right) = \frac{\zeta_t(i) a_{ij} b_{jt+1} \omega_{t+1}(j)}{\sum_k \sum_l \zeta_t(k) a_{k1} b_{1t} \omega_{t+1}(l)} \tag{16}$$

Where $\zeta_t(i) = \left[\sum_{k=1}^{N} \zeta_{t-1}(k) a_{ki}\right] b_{it}$ is the forward observation probability, which indicates the probability of the observation sequence before time t for a hidden state IS_i at time t. The corresponding backward observation probability is $\omega_t(i) = \sum_{k=1}^{N} a_{ki} b_{k,t+1} \omega_{t+1}(k)$.

At the same time, the probability that the system is in a hidden state IS_i at time t is defined as

$$\gamma_t(i) = \sum_{j=1}^{N} \psi_t(i,j) \tag{17}$$

Next, to perform a maximum likelihood estimation of the model parameters by Expectation-Maximization(EM) algorithm. In each iteration, $\psi_t(i,j)$ and $\gamma_t(i)$ are calculated using the E-algorithm for a given λ. Subsequently, the M-algorithm is used to calculate λ in the case of $\psi_t(i,j)$ and $\gamma_t(i)$ until convergence is reached. When there are multiple observation sequences, taking the average to obtain three parameters (π, A, B) as follows:

$$a_{ij} = \frac{\sum_{k=1}^{K} \psi_t^k(i,j)}{\sum_{k=1}^{K} \gamma_t^k(i)} \tag{18}$$

$$b_{im} = \frac{\sum_{k=1}^{K} \psi_t^k(i,m)}{\sum_{k=1}^{K} \gamma_t^k(i)} \tag{19}$$

$$\pi_i = \sum_{k=1}^{K} \gamma_i^k(i) \tag{20}$$

After the model parameters are obtained, using λ to predict the network reliability. When a network threat observation sequence $O_{TH} = \{o_{TH,\ 1}, o_{TH,\ 2}, \cdots, o_{TH,\ t}\}$ is observed in an anomaly detection module of the network topology auto-association system, the Viterbi algorithm is used to calculate the optimal hidden state sequence Q_{TH}. Then, at the next time point (t + 1), the network threat transfers to 5 different hidden states IS_1, IS_2, IS_3, IS_4, IS_5. The transfer probability vectors are

$$\Gamma_{TH}^{t+1}\left(P\left(q_{t+1} = IS_1 | q_t = IS_t\right), \cdots P\left(q_{t+1} = IS_5 | q_t = IS_t\right)\right) \tag{21}$$

On this basis, to multiply the hidden state level transpose vector and Γ_{TH}^{t+1} to calculate the network threat level $Sec_{TH}^{t+1} = \Gamma_{TH}^{t+1} \cdot (1,2,3,4,5)^T$ at the next time point (t + 1) for the

system. Similarly, the transfer probability vectors Γ_{TE} and Sec_{TE}^{t+1} of the network transmission efficiency TE for the next time point (t + 1) can be obtained. Finally, w_{TH} and w_{TE}, the weights of Sec_{TH}^{t+1} and Sec_{TE}^{t+1} respectively can be obtained, based on the expert's experience. Reliability prediction value of the current path at the next time point (t + 1) can be calculated:

$$Sp_{t+1} = w_{TH}Sec_{TH}^{t+1} + w_{TE}Sec_{TE}^{t+1} \tag{22}$$

If Sp_{t+1} is moderate or high-risk, this path should be excluded in the next guided path space. Otherwise, this path will continue to be retained or added to the guided path space.

5.2 Mimic transformation strategy for the network topology mimic association graph

In the network topology mimic correlation graph, it is assumed that there are n available nonintersecting paths at time t whose predicted reliable values being transferred as medium-risk or high-risk paths at time $(t+1)$ in forming the network topology mimic graph $S_n^-(t+1)$. At the same time, there are m non-usable and nonintersecting paths at time t whose predicted reliable values being assessed as safe, mild or general risk at time $(t+1)$ for the network topology mimic association graph $S_m^+(t+1)$. Thus, the next network topology mimic graph is $S_i(t+1) = S_i(t) - S_n^-(t+1) + S_m^+(t+1)$.

At time $(t+1)$, the new path $S_m^+(t+1)$ will be added; if this path is selected as the communication path, only the primary station serving node needs to notify the edge computing terminal node on the path with the relevant ACL and other information. This is according to the network topology mimic association negotiation algorithm. However, for the communication path $S_n^-(t+1)$ at time t, the primary station service node needs to notify the relevant parties to revoke the ACL and other information.

After the network topology mimic graph is adjusted at the completion time (t + 1), a new graph $S_i(t+1) = \{s_k(t+1) | 1 \le k \le m\}$ can be obtained. Then, the overall reliability prediction value corresponding to $S_i(t+1)$ can be obtained as $SA_{S_i(t+1)} = \sum_{i=1}^{m} Sp_{t+1}^i$. The function is then updated, satisfying the principle of "increase slowly and decrease rapidly".

$$h\left(SA_{S_i(t+1)}\right) = \begin{cases} h_1\left(SA_{S_i(t+1)}\right), SA_{S_i(t+1)} = 1 \\ h_2\left(SA_{S_i(t+1)}\right), SA_{S_i(t+1)} \in (2,3) \end{cases} \tag{23}$$

The self-adjusting strategy is as follows:

$$T_S^{i+1} = \begin{cases} \left(1 + h_1\left(SA_{S_i(t+1)}\right)\right) * T_S^i, SA_{S_i(t+1)} = 1 \\ \left(1 - h_2\left(SA_{S_i(t+1)}\right)\right) * T_S^i, SA_{S_i(t+1)} \in (2,3) \end{cases} \tag{24}$$

6 Experiments

6.1 Data preparation

The experiment performs a power system simulation of the network topology mimic association algorithm based on the NS2 network simulation environment. This model uses C++ to write the synchronization module, association module, communication module, attack module, delay processing module, sampling module, anomaly detection module, and deception processing module. It implements the network topology simulation by writing an OTcl script. The number of available IPv4 addresses in the network is 28, and the number of available ports is 1000. The initial correlation period is 120 s. The method supposed that $g_1(x) = -ln(20x + 0.5)$, $g_2(x) = 16x^2 - 0.8x + 0.01$, $h_1(z) = -ln(20x + 0.6)$, $h_2(z) = 16z^2 - 0.64z + 0.064$. To mention that the simulation experiments are conducted in different scenarios with the same resources.

6.2 Experiment metrics

6.2.1 Security analysis

Security is an important indicator for evaluating the advantages and disadvantages of a defense method. This section analyzes the anti-attack capability of the proposed active defense technology for an edge defense network attack based on network topology mimic correlation. The active defense principle for edge computing network attacks based on the network topology mimic association algorithm is shown in Figs. 7 and 8.

- Anti-DoS attack

After the network topology mimic association defense strategy is implemented, the IP address and port of the communication host and the protocol used by the communication parties will be associated after each corresponding time slot. For an attacker who performs a DoS attack, it is necessary to continuously send a large number of service requests to the target host and consume the target host resources. However, the node network configuration of the target host is continuously associated; thus, a DoS attack cannot be initiated [32, 33].

- Anti-following attacks and anti-semi-blind attacks

A following attack is a special type of attack for a network topology mimicking system. When the defender adopts the network topology mimicking association strategy, the attacker will try to detect and locate the current active node network

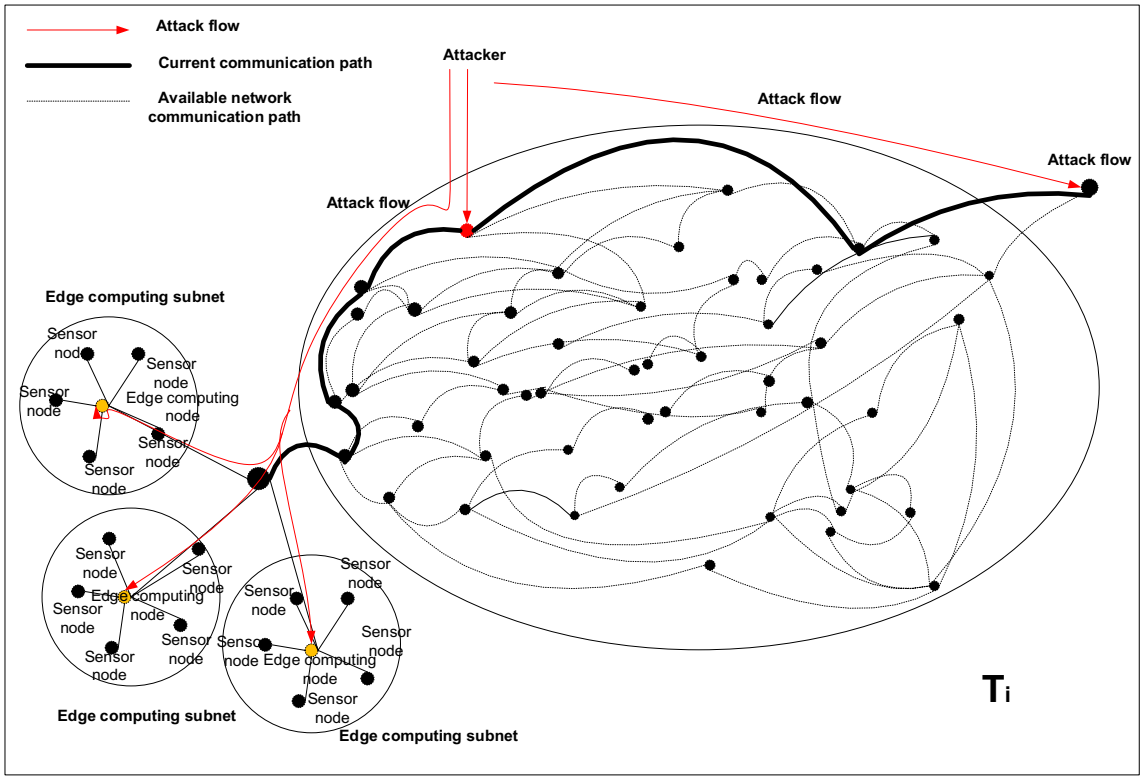

Fig. 7 Defense before network topology mimetic correlation

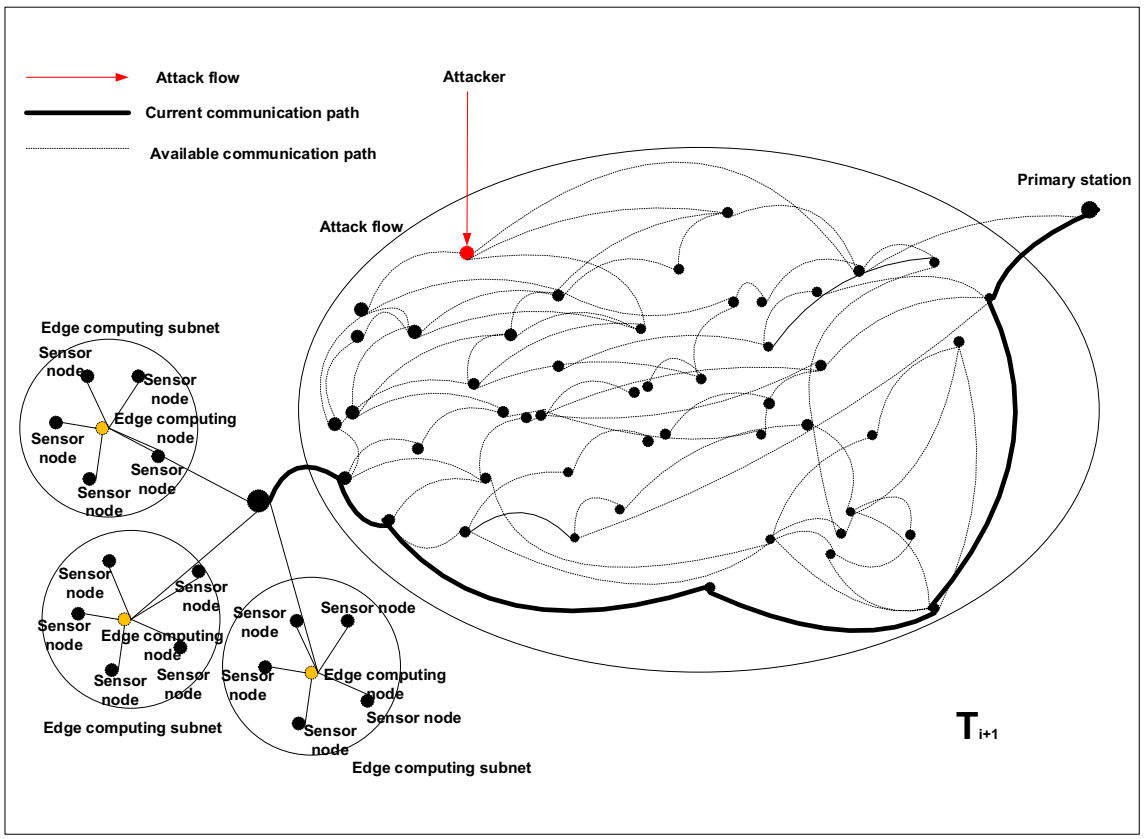

Fig. 8 Defense after network topology mimetic correlation

configuration as the focus of the attack. A blind attack occurs when an attacker cannot locate the current active node network configuration and attacks all available nodes of the node network configuration state space that are detected. The attack strength is evenly distributed across all available nodes. The difference between a semi-blind attack and a blind attack is that the attack intensity of a semi-blind attack is concentrated on a subset of the available nodes while the attack strength of the follow-up attack is concentrated at one point. The network topology mimic association algorithm further increases the difficulty for an attacker to detect and locate the current active node network configuration of the associated system, and thus, the ability to resist follow-up attacks and anti-semi-blind attacks is improved [34].

6.2.2 Network transmission efficiency

Attacks effect the network transmission. Network transmission efficiency can be regarded as the metric to evaluate the state of the network transmission. To evaluate different network transmission efficiency, the most direct and main metric is the network transmission rate. The network topology mimic association (PA NTAA) proposed in this paper is compared with other three algorithms, the non-topology-association algorithm (No NTAA), the simple topology association algorithm (Simple NTAA) and the end-hopping-based topology association algorithm (EH NTAA) proposed in [19, 20]. Experiments compare these algorithms on network transmission efficiency under different attack rates. Average attack rates range from 0 to 100. After that, experiments are also analyzed with and without an attack. Additionally, the transmission efficiencies of the No NTAA and PA NTAA are compared.

6.3 Results

6.3.1 Experiment against DDoS attacks

In this section, the SYN-Flood mode is used to guide a DoS attack. Experiments test the average service response time of the network topology mimic association system under different SYN-Flood attack rates to reflect the service availability performance. Figure 9 shows results for No NTAA, Simple NTAA, EH NTAA, and PA NTAA. The results show that the network topology mimic association strategy proposed in this paper can better resist DoS attacks. This result occurs because the mimic correlation technology of the network topology dynamically measures network anomalies according to the strength of cyber-attacks. Then, the network topology mimic graph and communication paths are automatically adjusted. Adjustments increase the difficulty of hitting a path for DDoS attacks. However, the difference between the results for the association strategy in [19] and PA NTAA is not significant. Moreover, when the mimic map space of the network topology is compressed to almost zero, the DDoS attack enters an unsupervised blind attack state, that is, an average attack on all nodes in the accessible path detected by the attacker.

6.3.2 Experiment against the following attack

In conventional edge computing network communication, the attacker can easily detect the node IP and port number in the communication path. Thus, the following attack at this time is a direct attack. When the attack strength increases, the DoS state will be quickly reached. Under the simple topology association algorithm, the adjustment period is fixed; thus, the communication path is adjusted uniformly according to a fixed term. In the experiments, the attacker's following delay

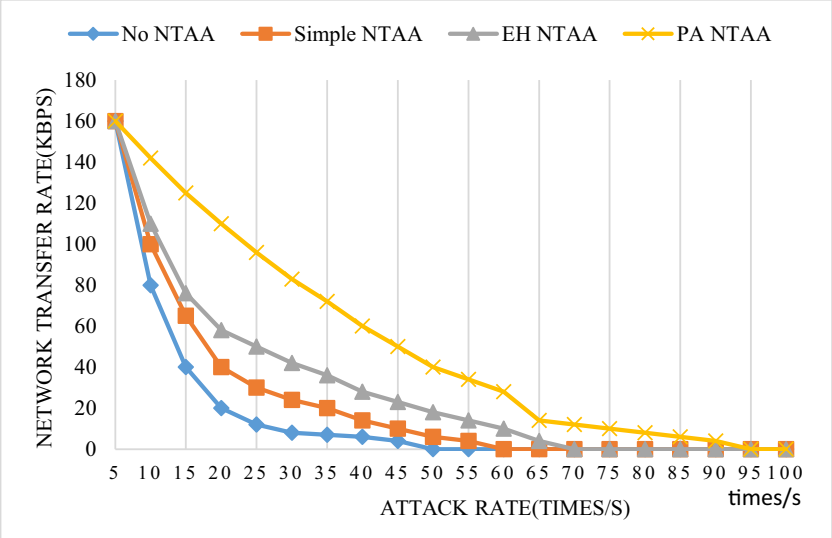

Fig. 9 Results for DDoS attack defense test

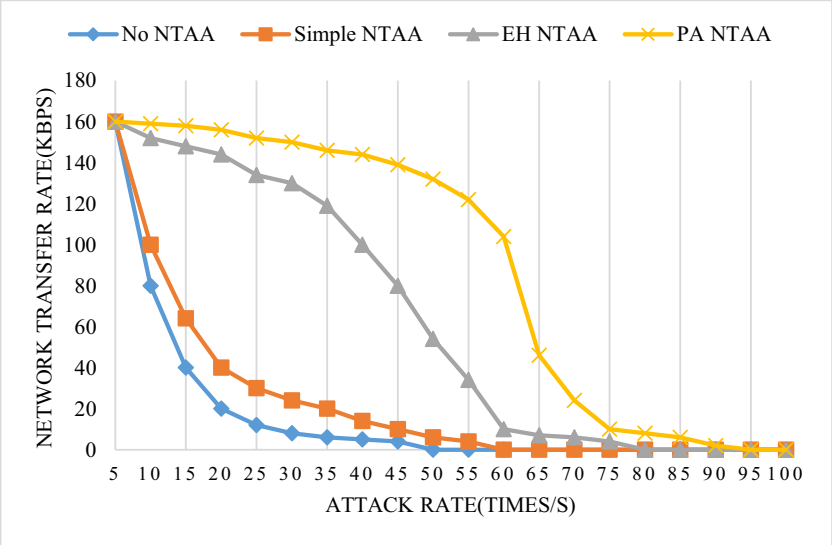

Fig. 10 Results for an accompanying attack defense test

is 2S, which allows the attacker to follow with sufficient time. When the EH NTAA is adopted, in the case of an attack, the adjustment period is reduced by more than half to inhibit the attacker from detecting the communication path. In this paper, when adopting PA NTAA, since both the network anomaly and network security reliability are considered, the adjustment period will be reduced by more than 1/2 in the presence of continuous attacks. In this case, the adjustment period will be reduced more quickly, and the other time delays will increase for the attacker.

The experimental results in Fig. 10 show that the response time for the simple topology association algorithm is better than that for No NTAA. This result arises because the adjustment period of the Simple NTAA is fixed, and the attacker has sufficient time to analyze the current active node and start attacking. The network transmission efficiency of PA NTAA is significantly higher than that of EH NTAA because the adjustment of the network topology mimic graph and communication path is based on network anomaly detection and network security reliability prediction, but not just reduced by more than 1/2. This approach can effectively reduce the transmission efficiency loss caused by the mimic transformation strategy. However, as the attack speed increases, the attack packets occupy a large amount of network bandwidth, causing the network to enter a congested state. Although the attacker cannot identify the attack after the current active node, it will still cause a rapid decline in transmission efficiency.

6.3.3 Experiment against a semi-blind attack

Here, it uses a perceptual node edge access system with 20 communication paths for experiments. It can be seen from Fig. 11 that when the edge of the access node is connected to the network, the network transmission delay increases rapidly as

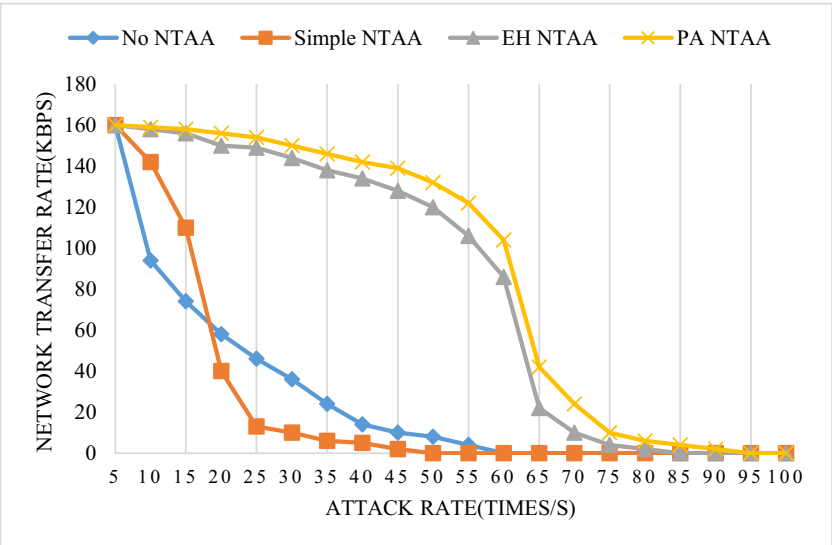

Fig. 11 Results for semi-blind attack defense

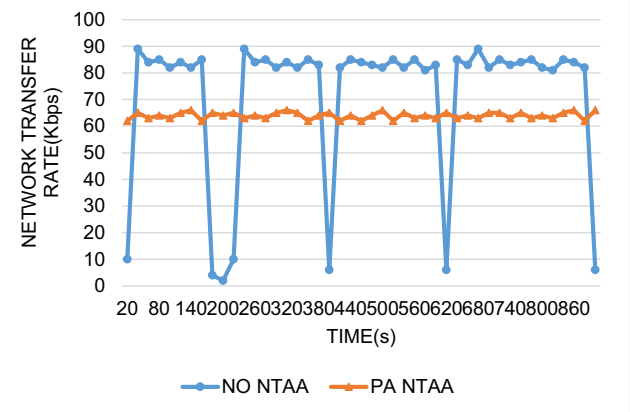

Fig. 12 Experimental analysis results of network transmission efficiency without an attack

the proportion of the received attack path reaches 50%. When the proportion exceeds 60%, the network transmission delay tends to infinity. The average response time of the EH topology association strategy is better than that of the No NTAA but is not as good as that of the Simple NTAA, which is consistent with the analysis presented in [35]. The average response time of the PA NTAA is better than that of the Simple NTAA.

6.3.4 Comparison of with and without attack

The experiment results are also analyzed with and without an attack. Additionally, the transmission efficiencies of the No NTAA and PA NTAA are compared, primarily based on the reporting rate of the primary station node. As shown in Fig. 12, when there is no attack, the transmission efficiency of the PA NTAA is lower than that of the No NTAA during the initial negotiation phase. However, after the negotiation point, the transmission efficiency of the PA NTAA is 43% higher than that of the No NTAA, which is primarily due to the multipath transmission. Moreover, the experimental results show that the adjustment period is 180 s, 198 s, 227.8 s, and 271 s, and the growth rate is approximately 10%, 15%, and 19% of the previous cycle, which satisfies the principle of full growth. In the case of an attack, the transmission efficiency of the No NTAA decreases as the attack strength increases. The transmission efficiency of the PA NTAA can be maintained when the attack strength is not robust. However, as the attack strength increases, the transmission efficiency gradually decreases.

7 Conclusion

Based on a thorough study of the mobile self-organizing characteristics of edge computing networks, the framework combines a moving network transmission with path mimicry adjustment techniques to propose a strict, formal description and definition. An active defense framework for data transmission in an edge computing network based on a link layer and application layer network topology mimic correlation is designed to ensure scalability of the algorithm. To solve the problem of attacks and to improve defense and transmission quality with a moving periodic adjustment of the network, this method proposes a moving communication path alliance and a mimic graph transformation method for network topology. Based on the temporal and spatial dimensions, this work combines moving threshold network anomaly detection and reliability prediction of network security based on the HMM. In this way, a reasonable transformation of the network can be performed. The mimic adjustment overhead can be minimized and active defense problems in the DoS attack, following attack and semi-blind attack can be resolved. Experimental results show that the transmission efficiency of the network topology mimic association algorithm proposed in this paper is higher than that of other popular methods and the reliability and anti-attack performance are significantly improved.

However, there are still some deficiencies in this work. This work is mainly applied in power edge computing network. Experiments are conducted in simulated power network. Due to the particularity of the power network, results are sensitive to parameters and environment in this model. Next step, the method needs to be improved to apply in more common networks.

References

1. Roman R, Lopez J, Mambo M (2018) Mobile edge computing, Fog et al.: a survey and analysis of security threats and challenges. Futur Gener Comput Syst 78:680–698
2. Chen Y, Zhang Y, Maharjan S (2017) Deep learning for secure mobile edge computing. arXiv preprint arXiv:1709.08025
3. Ai Y, Peng M, Zhang K (2018) Edge computing technologies for internet of things: a primer. Digit Commun Netw 4(2):77–86
4. Gershenfeld N, Krikorian R, Cohen D (2004) The internet of things. Sci Am 291(4):76–81
5. Li H, Ota K, Dong M (2018) Learning IoT in edge: deep learning for the internet of things with edge computing. IEEE Netw 32(1): 96–101
6. Yang H (2016) Method for behavior-prediction of APT attack based on dynamic Bayesian game. In: 2016 IEEE international conference on cloud computing and big data analysis (ICCCBDA), pp 177–182, Chengdu, China
7. Wan J et al (2018) Toward dynamic resources management for IoT-based manufacturing. IEEE Commun Mag 56(2):52–59
8. Wang J, Cao J, Ji S, Park JH (2017) Energy-efficient cluster-based dynamic routes adjustment approach for wireless sensor networks with mobile sinks. J Supercomput 73(7):3277–3290
9. Yin Y, Zhang W, Xu Y, Zhang H, Mai Z, Yu L (2019) QoS prediction for Mobile edge service recommendation with auto-encoder. IEEE Access 7:1–1

10. Yin Y, Chen L, Xu Y, Wan J, Zhang H, Mai Z (2019) QoS prediction for service recommendation with deep feature learning in edge computing environment. Mob Networks Appl:1–11
11. Gao H, Zhang K, Yang J, Wu F, Liu H (2018) Applying improved particle swarm optimization for dynamic service composition focusing on quality of service evaluations under hybrid networks. Spec Collect Artic Int J Distrib Sens Netw 14(2):2018
12. Gao H, Huang W, Yang X, Duan Y, Yin Y (2018) Toward service selection for workflow reconfiguration: an interface-based computing solution. Futur Gener Comput Syst 87:298–311
13. Dunlop M, Groat S, Urbanski W, Marchany R, Tront J (2011) MT6D: a moving target IPv6 defense. In: 2011 – MILCOM 2011 military communications conference, pp 1321–1326, Baltimore, MD, USA
14. Atighetchi M, Pal P, Webber F, Jones C (2003) Adaptive use of network-centric mechanisms in cyber-defense. In: Second IEEE international symposium on network computing and applications. NCA 2003, pp 179–188, Cambridge, MA, USA
15. Badishi G, Herzberg A, Keidar I (2007) Keeping denial-of-service attackers in the dark. IEEE Transactions on Dependable and Secure Computing 4(3):191–204
16. Antonatos S, Akritidis P, Markatos EP, Anagnostakis KG (2005) Defending against Hitlist Worms using network address space randomization. Computer Networks 51(12):3471–3490
17. Jafarian JH, Al-Shaer E, Duan Q (2012) OpenFlow random host mutation: transparent moving target defense using software defined networking. Proceedings of the first workshop on Hot topics in software defined networks. ACM, pp 127–132
18. Dunlop M, Groat S, Urbanski W, Marchany R, Tront J (Jul. 2012) The blind man's bluff approach to security using IPv6. IEEE Secur Priv Mag 10(4):35–43
19. Jafarian JH, Al-Shaer E, Duan Q (2014) Spatio-temporal address mutation for proactive cyber agility against sophisticated attackers. Proceedings of the First ACM Workshop on Moving Target Defense. ACM, pp 69–78
20. MacFarland DC, Shue CA (2015) The SDN shuffle. In: Proceedings of the second ACM workshop on moving target defense – MTD '15, pp 37–41, Denver, Colorado, US
21. Skowyra R, Bauer K, Dedhia V, Okhravi H (2016) Have no phear: Networks without identifiers. Proceedings of the 2016 ACM Workshop on Moving Target Defense. ACM, pp 3–14
22. Sun J, Sun K (2016) DESIR: decoy-enhanced seamless IP randomization. In: IEEE INFOCOM 2016 – the 35th annual IEEE international conference on computer communications, pp 1–9
23. Jiangxing Wu. C. M. Defense, "Research on Cyber Mimic Defense, ". J Cyber Secur vol. 1, no. 4, pp. 1–10, 2016
24. Gao H, Miao H, Liu L, Kai J, Zhao K (2018) Automated quantitative verification for service-based system design: a visualization transform tool perspective. Int J Softw Eng Knowl Eng 28(10): 1369–1397
25. Gao H, Mao S, Huang W, Yang X (2018) Applying probabilistic model checking to financial production risk evaluation and control: a case study of Alibaba's Yu'e Bao. IEEE Trans Comput Soc Syst 5(3):785–795
26. Haggerty J, Shi Q, Merabti M (2002) Beyond the perimeter: the need for early detection of denial of service attacks. In: Proceedings of 18th annual computer security applications conference, pp 413–422, Las Vegas, NV, USA
27. Zhang J, Gunter CA (2010) Application-aware secure multicast for power grid communications. In: 2010 first IEEE international conference on smart grid communications, pp 339–344, Gaithersburg, MD, USA
28. Yin Y, Chen L, Xu Y, Wan J (2018) Location-aware service recommendation with enhanced probabilistic matrix factorization. IEEE Access 6:62815–62825
29. Yin Y, Yu F, Xu Y, Yu L, Mu J (2017) Network location-aware service recommendation with random walk in cyber-physical systems. *Sensors* 17(9):2059
30. Liang W et al (2018) A security situation prediction algorithm based on HMM in Mobile network. Wirel Commun Mob Comput 2018: 1–11
31. Wan M, Yao J, Jing Y, Jin X (2018) Event-based anomaly detection for non-public industrial communication protocols in SDN-based control systems. CMC 55(3):447–463
32. Yan Q, Huang W, Luo X, Gong Q, Yu FR (2018) A multi-level DDoS mitigation framework for the industrial internet of things. IEEE Commun Mag 56(2):30–36
33. Bereziński P, Jasiul B, Szpyrka M, Bereziński P, Jasiul B, Szpyrka M (2015) An entropy-based network anomaly detection method. Entropy 17(4):2367–2408
34. Prasanth Vaidya S, Chandra Mouli PVSSR (2017) A robust semi-blind watermarking for color images based on multiple decompositions. Multimed Tools Appl 76(24):25623–25656
35. Zhao C, Jia C (2013) Research on spatial adaptive strategy of end-hopping system. In: 2013 fourth international conference on emerging intelligent data and web technologies, pp 661–666, Xi\"an, Shaanxi, China