Smart Collaborative Balancing for Dependable Network Components in Cyber-Physical Systems

Fei Song, *Member, IEEE*, Zhengyang Ai, Haowei Zhang, Ilsun You, *Senior Member, IEEE*, and Shiyong Li

***Abstract*—The evolution of cyber–physical system (CPS) benefits from substantial supports of many cutting-edge technologies. However, as a significant medium to bridge virtual and reality parts, the dependability of various network components is facing unprecedented challenges and threats. In this article, we propose a smart collaborative balancing (SCB) scheme to dynamically adjust the orchestration of network functions and efficiently optimize the workflow patterns. First, mathematical models of bandwidth allocation for multiuser with appropriate probability distribution are established. Matrix operations are utilized to solve the relevant issues based on individual congestion windows. Invasion defense mechanisms are also provided and discussed. Second, specific procedures of collaboration among different network components are presented. The capabilities of CPS, in terms of bandwidth allocation and invasion defense, are guaranteed via novel queueing policies and access control mechanisms. Third, we build a comprehensive prototype including multiple domains and users for validations. Experimental results in two scenarios illustrate that SCB not only supports service reliability of end hosts with different priorities, but also resists malicious attacks which are targeting the corresponding terminals inside domains. Compared to the benchmarks in software defined networks and traditional Internet, our scheme performs better in both available resource management and abnormal flow recognition aspects.**

***Index Terms*—Cyber–physical systems (CPSs), dependable network components (DNCs), smart collaborative balancing (SCB), software defined networks (SDNs).**

Manuscript received June 20, 2020; revised September 3, 2020; accepted September 17, 2020. Date of publication October 9, 2020; date of current version June 30, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 62072030 and in part by the Soonchunhyang University Research Fund. Paper no. TII-20-3012. *(Corresponding author: Ilsun You.)*

Fei Song and Zhengyang Ai are with the Beijing Jiaotong University, Beijing 100044, China (e-mail: fsong@bjtu.edu.cn; zyangai@bjtu.edu.cn).

Haowei Zhang is with the Beijing Institutein of Technology, Beijing 100081, China (e-mail: 3220195116@bit.edu.cn).

Ilsun You is with the Soonchunhyang University, Asan 31538, Korea (e-mail: ilsunu@gmail.com).

Shiyong Li is with the Yanshan University, Qinhuangdao 066004, China (e-mail: shiyongli@ysu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2020.3029766.

Digital Object Identifier 10.1109/TII.2020.3029766

I. Introduction

CYBER–PHYSICAL systems (CPSs) have been widely investigated by academia, industry, and military in its early stage [1]. On the one hand, the complex connotation of CPS can be understood through several "triple elements" [2]. For instance, computation, communication, and control are regarded as important technologies to realize safety, security, and sustainability [3]. Real-time perception, data service, and dynamic control are significant characteristics [4]. The main target is to establish a closer relationship between analog physics, digital information, and ordinary people [5]. On the other hand, the comprehensive extension of CPS can be observed as the distinctive development routes in different countries [6]. For example, the "industrial Internet" launched by the United States is based on the advantages of cyber field. It enables the physical system to achieve a new leap with the help of powerful network functions, maintains the outstanding position in fictitious economy, and plays more important role within industrial activities. Germany's "Industry 4.0" focuses on taking benefits in physical field, enhancing the capacity of smart manufacturing, maintaining the leading position in real economy, and increasing the portion inside industrial chains. Both complex connotation and comprehensive extension of CPS require high-performance supports from smart networking technologies [7].

Dependable network components (DNCs) are supposed to be execution units for guaranteeing relevant functions of CPS in a reliable network environment [8]. Based on existing deployment patterns, they can be divided into the following two parts: 1) Entity components (computing servers, communication devices, transmission links, etc.) provide stable and scalable infrastructure from the physical perspective [9]. The centralized, distributed, and hybrid connection modes have been utilized in Internet of Things, Internet of Vehicle, and other emerging paradigms. 2) Virtual components (virtual machine, function virtualization, service function chain, etc.) are flexible and resilient from cyber perspective [10]. During the task decomposition process, more service provision and acquisition patterns are convenient when virtual components are successfully established. Through the software defined principle, entity components and virtual components can undertake mutual duties collaboratively with the guidance of controllers [11].

The motivation of this article stems from our industrial applications and practices. When massive users access CPS in large-scale mode, it is challenging to ensure the available bandwidth of each high priority user under limited network resources. Even the allocation policy has been established and accepted by end hosts, simple cyberattacks launched by interdomain hackers may trigger enormous loss in physical infrastructures. Similar incidents have been broadly witnessed in energy, transportation, and other fields, which brings potential security issues for CPS deployments. Therefore, the service reliability and defense capability are challenging in dependable scenarios for current CPS. However, most existing research works only focus on one aspect since it is difficult to achieve the tradeoff between performance and security.

Different with previous work, we insist that the abovementioned two CPS problems can be investigated and handled simultaneously. A smart collaborative balancing (SCB) scheme is proposed to dynamically adjust the orchestration of network functions and guarantee the service reliability. It is suitable for network quality of service (QoS) enhancement and malicious invasion resistance. Novel mathematical models are designed to smartly change the queuing rules inside nodes. Actions among access routers (AR) and switch routers (SR) are created to collaboratively exchange the control and data flows. In packet forwarding procedures, DNCs are established by reasonably balancing various resources of multiple equipment. To the best of authors' knowledge, it is the first SCB scheme for DNCs in CPS scenarios. The main contributions of this article are as follows.

1) Mathematical models for service reliability improvement and defense capability enhancement were proposed based on the probability and differential equation theory. To avoid unnecessary calculation workload, an alternative solution was provided without losing accuracy and generality.
2) Implementation procedures of our scheme, including hierarchical division, encapsulation approaches, interaction details, were given from appropriate perspectives. The characteristics of massive users and network capacity were considered comprehensively.
3) A prototype system was established on top of multiple entity components. Complex virtual components were created to build experiment scenarios. Compared to existing candidate schemes, validation results illustrated that SCB is able to achieve better performance in most cases.

The rest of this article is organized as follows. In Section II, the methodology and consideration focusing on bandwidth allocation and invasion defense are proposed. In Section III, the implementation details of SCB scheme in CPS is presented. In Section IV, validation and discussion are conducted. In Section V, the related work are described and analyzed. Finally, Section VI concludes this article.

II. Methodology and Consideration

There are several function modules in SCB design, which are generated by mathematical theories and security considerations. For bandwidth allocation, throughput formulas for whole group and individual user are proposed, respectively. Due to the difficulty in calculating a specific parameter probability, a subtle method is utilized to separate the model into two parts. The subsequent merging process does not lost the accuracy and generality. For invasion defense, multidimensional improvements are provided via different point of views. Specifically, network space isolation, user identifier uniqueness, bidirectional information control, and multiple protection approaches are discussed to ensure the security requirements of CPS environment.

A. Bandwidth Allocation

Suppose there are n users inside a CPS. All of them are attempting to send packets via network components. Then, the congestion window of whole group is

$$W = (w_1, w_2, \ldots, w_n) \tag{1}$$

where w_i is the window size of the ith user in this group. Each connection has an independent congestion mechanism, and congestion window update will follow

$$\begin{cases} w(t + \text{RTT}) \leftarrow w(t) + w(t) & \text{SS} \\ w(t + \text{RTT}) \leftarrow w(t) + \alpha, \alpha > 0 & \text{CA} \end{cases} \tag{2}$$

$$\begin{cases} w(t + \delta t) \leftarrow (1 - \beta)w(t), 0 < \beta < 1 & \text{TD} \\ w(t + \delta t) \leftarrow 1 & \text{TO} \end{cases}. \tag{3}$$

RTT is the "round trip time" between packets sending and acknowledgements (ACKs) receiving. α and β are additive increase and multiplicative decrease parameters. We utilize abbreviations "SS," "CA," "TD," and "TO" to express the "slow start," "congestion avoidance," "triple duplicated ACKs," and "time out." In each connection, the time intervals of TD and TO events obey the exponential distribution, which means the occurrence of $\{N_{\text{TD}i}(t)\}$ and $\{N_{\text{TO}i}(t)\}$ obeys the Poisson distribution. Then, the arrival rates are expressed as $\lambda_{\text{TD}i}$ and $\lambda_{\text{TO}i}$, respectively. We are able to obtain the group window variation equation

$$\begin{aligned} \mathrm{d}W(t) = &(I - I(W(t))\alpha\text{RTT}^{-1}\mathrm{d}t \\ &+ I(W(t))W(t)\text{RTT}^{-1}\mathrm{d}t \\ &+ (-1)\beta W(t)\mathrm{d}N_{\text{TD}}(t) \\ &+ (I - W(t))\mathrm{d}N_{\text{TO}}(t). \end{aligned} \tag{4}$$

In (4), all parameters are presented according to their matrix forms. For instance, α and β should be understood as

$$\begin{aligned} \alpha &= \text{diag}[\alpha_1, \alpha_2, \ldots, \alpha_n] \\ \beta &= \text{diag}[\beta_1, \beta_2, \ldots, \beta_n] \end{aligned} \tag{5}$$

where α_i and β_i are increasing and decreasing parameters of a congestion window in the ith connection. Different connections provide supports for various services or applications. The values of α_i and β_i should be confirmed before the handshakes. Under standard conditions, $\alpha_i = 1$, $\beta_i = 1/2$. The $I(W(t))$ and its

elements $I_i(w_i(t))$ can be represented as

$$I(W(t)) = \begin{pmatrix} I_1(w_1(t)) & & & \\ & I_2(w_2(t)) & & \\ & & \ddots & \\ & & & I_n(w_n(t)) \end{pmatrix} \tag{6}$$

$$I_i(w_i(t)) = \begin{cases} 1 & w_i(t) \le T_i \\ 0 & w_i(t) > T_i \end{cases}. \tag{7}$$

T_i is the threshold of SS and CA stages. When the congestion window is less than or equal to T_i, the value growth follows an exponential pattern. Otherwise, the value growth will be in linear tendency.

In abovementioned group window variation equation, the dynamic changes of a congestion window $w_i(t)$ will be

$$\begin{aligned} \mathrm{d}w_i(t) = &(1 - I_i(w_i(t)))\alpha_i \mathrm{RTT}_i^{-1}\mathrm{d}t \\ &+ I_i(w_i(t))w_i(t)\mathrm{RTT}_i^{-1}\mathrm{d}t \\ &+ (-1)\beta_i w_i(t)\mathrm{d}N_{\mathrm{TD}i}(t) \\ &+ (1 - w_i(t))\mathrm{d}N_{\mathrm{TO}i}(t). \end{aligned} \tag{8}$$

We can calculate the mathematical expectation of (8)

$$\begin{aligned} E[\mathrm{d}w_i(t)] = &(1 - E[I_i(w_i(t))])\alpha_i \mathrm{RTT}_i^{-1}\mathrm{d}t \\ &+ E[I_i(w_i(t))w_i(t)]\mathrm{RTT}_i^{-1}\mathrm{d}t \\ &+ (-1)\beta_i E[w_i(t)\mathrm{d}N_{\mathrm{TD}i}(t)] \\ &+ E[(1 - w_i(t))\mathrm{d}N_{\mathrm{TO}i}(t)]. \end{aligned} \tag{9}$$

According to classical hypothesis, the distributions of $\{w_i(t)\}$ and $\{N_{\mathrm{TD}i}\}$, $\{w_i(t)\}$ and $\{N_{\mathrm{TO}i}\}$, $\{w_i(t)\}$ and $\{I_i(w_i(t))\}$ are independent to each other. Therefore, (9) can be transformed to

$$\begin{aligned} \mathrm{d}E[w_i(t)] = &\alpha_i \mathrm{RTT}_i^{-1}(1 - P(w_i(t) \le T_i))\mathrm{d}t \\ &+ \mathrm{RTT}_i^{-1}P(w_i(t) \le T_i)E[w_i(t)]\mathrm{d}t \\ &- \beta_i \lambda_{\mathrm{TD}i}E[w_i(t)]\mathrm{d}t \\ &+ (1 - E[w_i(t)])\lambda_{\mathrm{TO}i}\mathrm{d}t. \end{aligned} \tag{10}$$

The expectation of $w_i(t)$ can be expressed as

$$\begin{aligned} E[w_i(t)] = &\frac{\alpha_i \mathrm{RTT}_i^{-1}(1 - P(w_i(t) \le T_i)) + \lambda_{\mathrm{TO}i}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}P(w_i(t) \le T_i)} \\ &+ C_0 e^{-(\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}P(w_i(t) \le T_i))t}. \end{aligned} \tag{11}$$

For proof details, please see Appendix A. In the steady state, we have

$$E[w_i] = \frac{\alpha_i \mathrm{RTT}_i^{-1}(1 - P(w_i(t) \le T_i)) + \lambda_{\mathrm{TO}i}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}P(w_i(t) \le T_i)}. \tag{12}$$

Until now, the available bandwidth of the ith connection is

$$B_{w_i} = \frac{1}{\mathrm{RTT}_i}\frac{\alpha_i \mathrm{RTT}_i^{-1}(1 - P(w_i(t) \le T_i)) + \lambda_{\mathrm{TO}i}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}P(w_i(t) \le T_i)}. \tag{13}$$

The available bandwidth for whole group should be

$$B_w = \sum_{i=1}^{n} \frac{1}{\mathrm{RTT}_i}\frac{\alpha_i \mathrm{RTT}_i^{-1}(1 - P(w_i(t) \le T_i)) + \lambda_{\mathrm{TO}i}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}P(w_i(t) \le T_i)}. \tag{14}$$

Since it is difficult to calculate the value of $P(w_i(t) \le T_i)$, the result of (14) cannot be obtained directly. Therefore, we consider to separate such complicated expressions into two parts according to changing pattern of window size.

For (8), when the growth of window w_i is just in the slow start stage, this formula can be rewritten to

$$\begin{aligned} \mathrm{d}w_i(t) = &w_i(t)\mathrm{RTT}_i^{-1}\mathrm{d}t + (-1)\beta_i w_i(t)\mathrm{d}N_{\mathrm{TD}i}(t) \\ &+ (1 - w_i(t))\mathrm{d}N_{\mathrm{TO}i}(t) \end{aligned} \tag{15}$$

and the corresponding expectation will be

$$\begin{aligned} E[w_i(t)] = &\frac{\lambda_{\mathrm{TO}i}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}} \\ &+ C_1 e^{-(\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1})t}. \end{aligned} \tag{16}$$

In the steady state, we have

$$E[w_i] = \frac{\lambda_{\mathrm{TO}i}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}}. \tag{17}$$

The available bandwidth of the ith connection is

$$B_{w_i} = \frac{\lambda_{\mathrm{TO}i}\mathrm{RTT}_i^{-1}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}}. \tag{18}$$

For (8), if window w_i growth is in the congestion avoidance stage, this formula can be expressed as

$$\begin{aligned} \mathrm{d}w_j(t) = &\alpha_j \mathrm{RTT}_j^{-1}\mathrm{d}t + (-1)\beta_j w_j(t)\mathrm{d}N_{\mathrm{TD}j}(t) \\ &+ (1 - w_j(t))\mathrm{d}N_{\mathrm{TO}j}(t). \end{aligned} \tag{19}$$

Then, the corresponding expectation will be

$$E[w_j(t)] = \frac{\alpha_j \mathrm{RTT}_j^{-1} + \lambda_{\mathrm{TO}j}}{\beta_j \lambda_{\mathrm{TD}j} + \lambda_{\mathrm{TO}j}} + C_2 e^{-(\beta_j \lambda_{\mathrm{TD}j} + \lambda_{\mathrm{TO}j})t}. \tag{20}$$

When the system is in steady state, it will be

$$E[w_j] = \frac{\alpha_j \mathrm{RTT}_j{}^{-1} + \lambda_{\mathrm{TO}j}}{\beta_j \lambda_{\mathrm{TD}j} + \lambda_{\mathrm{TO}j}}. \tag{21}$$

Therefore, the available bandwidth of the jth connection is

$$B_{w_j} = \frac{1}{\mathrm{RTT}_j}\frac{\alpha_j \mathrm{RTT}_j + \lambda_{\mathrm{TO}j}}{\beta_j \lambda_{\mathrm{TD}j} + \lambda_{\mathrm{TO}j}}. \tag{22}$$

If k users are in the slow start stage and $n - k$ users are in the congestion avoidance stage, the obtained bandwidth of whole group should be

$$\begin{aligned} B_w = &\sum_{i=1}^{k} \frac{\lambda_{\mathrm{TO}i}\mathrm{RTT}_i^{-1}}{\beta_i \lambda_{\mathrm{TD}i} + \lambda_{\mathrm{TO}i} - \mathrm{RTT}_i^{-1}} \\ &+ \sum_{j=k+1}^{n} \frac{1}{\mathrm{RTT}_j}\frac{\alpha_j \mathrm{RTT}_j^{-1} + \lambda_{\mathrm{TO}j}}{\beta_j \lambda_{\mathrm{TD}j} + \lambda_{\mathrm{TO}j}}. \end{aligned} \tag{23}$$

B. Invasion Defense

To achieve the reliable transmission of control signaling and service data in CPS applications, a stable system architecture is essential to prevent malicious attacks. We consider the invasion defense required by the CPS from following aspects.

1) Network Space Isolation: Due to the binding limitations of traditional Internet, there is no clear boundary between edge and core networks in CPS, which may trigger potential security issues. In particular, a vulnerable core network is powerless to support the reliability of upstream and downstream information. Therefore, the network space isolation enables the resistance of abnormal access and ensures dependable data transmission.

2) User Identifier Uniqueness: The IP addresses of public networks can be utilized to locate or trace a specific entity. With the increasing of access devices in CPS, more and more drawbacks, such as accurate positioning and differentiated service, have been widely discussed. Therefore, the unique user identifier is requisite to distinguish illegal requests, guarantee the correctness of identifier mapping procedures, and provide flexible state supervision.

3) Bidirectional Information Control: The existing one-way CPS data control technologies mainly focus on device authentications at the source side. It is not easy to detect the legality of equipment at the destination side, simultaneously. Therefore, the bidirectional management of packets exchanging is necessary. Specifically, commands announcement, results feedback, actions monitoring, and information perception can be protected.

4) Multiple Protection Approaches: For the scenario with massive user access, a single prevention method cannot effectively guarantee network security. Therefore, multiple protection approaches are necessary for CPS architecture in complex and changeable network environment. Relevant network challenges, such as distributed denial of service (DDoS) and address resolution protocol flooding, should be addressed from different perspectives as well.

Based on above analysis and considerations, we propose an SCB scheme to handle the bandwidth allocation and invasion defense issues.

III. SCB in CPS

When massive users are involved, it is significant to optimize the resource utilization and enhance the network security for the CPS system. The SCB scheme aims to achieve the efficient and dependable information transmission in CPS by utilizing the novel queueing policies and access control mechanisms. In this section, we will introduce the system structure and main process of SCB in detail.

A. System Structure

The CPS is an extremely complex system that includes many heterogeneous elements and flexible scenarios. As shown in Fig. 1, the structure of our system is divided into the following three layers: physical layer, network layer, and application layer. From down to top, they are able to interact with each other for request collections and reply generations.

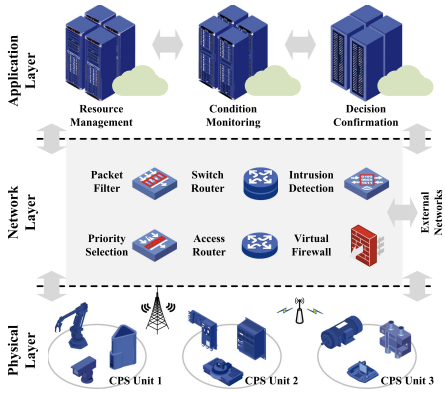

Fig. 1. System structure of SCB.

The physical layer is the foundation of SCB scheme and the bridge between the physical things and the virtual information. It is an important part which directly perceives the circumstance, interacts with the surroundings, and changes the corresponding components. Various CPS units are geographically distributed in a specific area. The member of CPS units may be a single node, such as a smart phone, tablet personal computer, notebook. In addition, The member of CPS units can be a set of nodes, such as robotic arms with supervisory cameras, large electromotor with electric generators. The communications among multiple CPS units are enabled based on SCB gateways. For instance, sensing devices with different hardware structure can exchange data through high-level interpreters.

The network layer is composed of multiple network components, such as SR, AR, packet filter, priority selection, intrusion detection, and virtual firewall. They are the kernel of SCB scheme. In this layer, we focus on the efficient and reliable issues related with node access processes. In the edge network, a terminal has a globally unique access identifier (AID), which contains a specific user level (UL). In the core network, a corresponding routing identifier (RID) is assigned automatically, which is responsible for routing and forwarding. An AR is in charge of providing security protections by storing legitimate users' information and sending UL together with original data to a neighbor SR. The relevant mapping between AID and RID is designed and implemented. The secure isolation for the edge network and core network has been achieved. Meanwhile, SR is also responsible for judging the priority of multiple users and determining the transmission order of data packets buffered inside a queue. Both the QoS guarantee and security improvement can be witnessed in SCB scheme. The interfaces of external network resources are available for further extensions.

The application layer is an integrated platform for users, which encapsulates the detailed information of network layer and physical layer into abstracted service modules. Resource management, condition monitoring, decision confirmation, and other functionalities are investigated and developed. They allow administrators and customers to execute operations directly without considering the details in underlying layers. This layer not only initiates necessary procedures of SCB scheme, but also presents the performance of results.

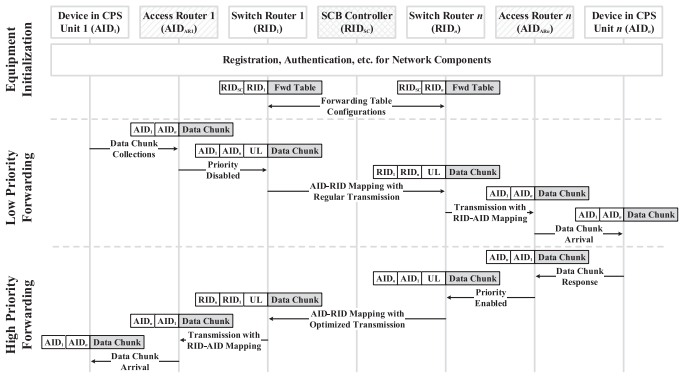

Fig. 2. Main interaction processes of SCB.

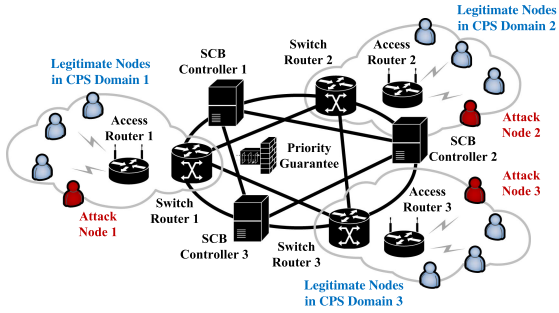

Fig. 3. Network topology of prototype system.

B. Main Procedures

To create an efficient environment, packets buffered inside a transmission queue are smartly scheduled based on dynamic network priorities. To ensure the dependable ability, multidimensional authentication is collaboratively powered by setting identifier mapping mechanism. Both the edge network and core network are covered in SCB scheme. Actions in three phases are illustrated in Fig. 2. Blocks with different stripes represent multiple network components. The corresponding AID or RID are marked inside brackets. In horizontal direction, a symmetrical structure is organized to simplify the descriptions. Some intermediate equipment are omitted due to the space limitations. The packet format, in terms of header and payload, of each packet is illustrated at the receiver side.

1) In equipment initialization phase, all network components have finished the necessary registration, authentication, and other process. Warnings must be generated if failure is encountered. Then, each CPS unit has a unique legitimate identifier, such as AID_1 and AID_2, to access the network. The corresponding AR should store the registration items. The SCB controller distributes a series of initial static forwarding entries for each network component based on topological positions, link status, channel quality, and other information. Forwarding table configurations are delivered from SCB controller (RID_{sc}) to relevant SRs (RID_1, RID_2, . . ., RID_n).

2) In low-priority forwarding phase, the regular transmission is default. The CPS unit 1 uploads collected data chunks inside physical world to the cyber world based on a multidimensional attribute authentication protocol. Identity judgment must be executed at the AR 1 by a quick comparison. Illegal user will be rejected and recorded immediately. Otherwise, $\text{UL} = 0$ should be sent to SR 1 to notify that the priority has been disabled. Then, the mapping table is checked to find the RID information associated with the target AID. After the AID to RID mapping, the packets can be forwarded from SR 1 to SR n via preestablish path. The value of UL should be kept inside the packet during the whole forwarding process. It is important to provision the available bandwidth to all existing users evenly if the values of UL are identical. The RID to AID mapping must be done before the data chunk arrives at the destination.

3) In high-priority forwarding phase, the optimized transmission is enabled. Assuming the data chunk response has been successfully generated by CPS unit n. Since the mapping results of RID_1 and RID_n have been cached in advance, there is no need to trigger the repetitive actions. By sending a path request to the SCB controller, a new transmission workload will be automatically calculated according to the UL information and the state of current network. The corresponding forwarding policies are distributed to intermediate SRs to establish a proper connection. Received routing items are stored inside local list to determine the forwarding directions. Packets with larger UL will be inserted in front of the lower priority queue. The RID pair is encapsulated together with the data chuck by SR n to provide an index for subsequent SRs. The decapsulation procedures must be achieved when the data chunk arrives at the SR 1. Similar with previous operations, the source and destination AIDs are recovered to replace the RIDs. Finally, the packet is correctly received by the CPS unit 1.

IV. Validation and Discussion

To comprehensively evaluate the performance of the SCB scheme, we built a prototype system by combining virtual and physical entities. The network components are divided into multiple physical domains. Through the bandwidth monitoring and attack tools, we validated the superiority of our scheme, in terms of bandwidth allocation and invasion defense. The software defined network (SDN) and traditional network (TN) benchmarks are selected as candidates to highlight the advantages of SCB.

A. Topology and Settings

The network topology of the prototype system is shown in Fig. 3. The edge and core areas are isolated by multiple SRs. The EMC servers and virtual software platform are applied to the system environment. The SCB controller successfully deploys the reliable bandwidth scheduling module. AR maintains access control policies and loads the list of legitimate user databases. SR is in charge of the identification mapping mechanism, control agent functions, and queue update entries. Three legitimate CPS access devices and one illegal attack machine are deployed in each domain. Attack tools are installed in illegal hosts to implement DDoS flows. For bandwidth allocation, the prototype system uses different ULs to concurrently access the network and establish TCP communication with the service platform for individual throughput evaluations. It utilizes the bandwidth

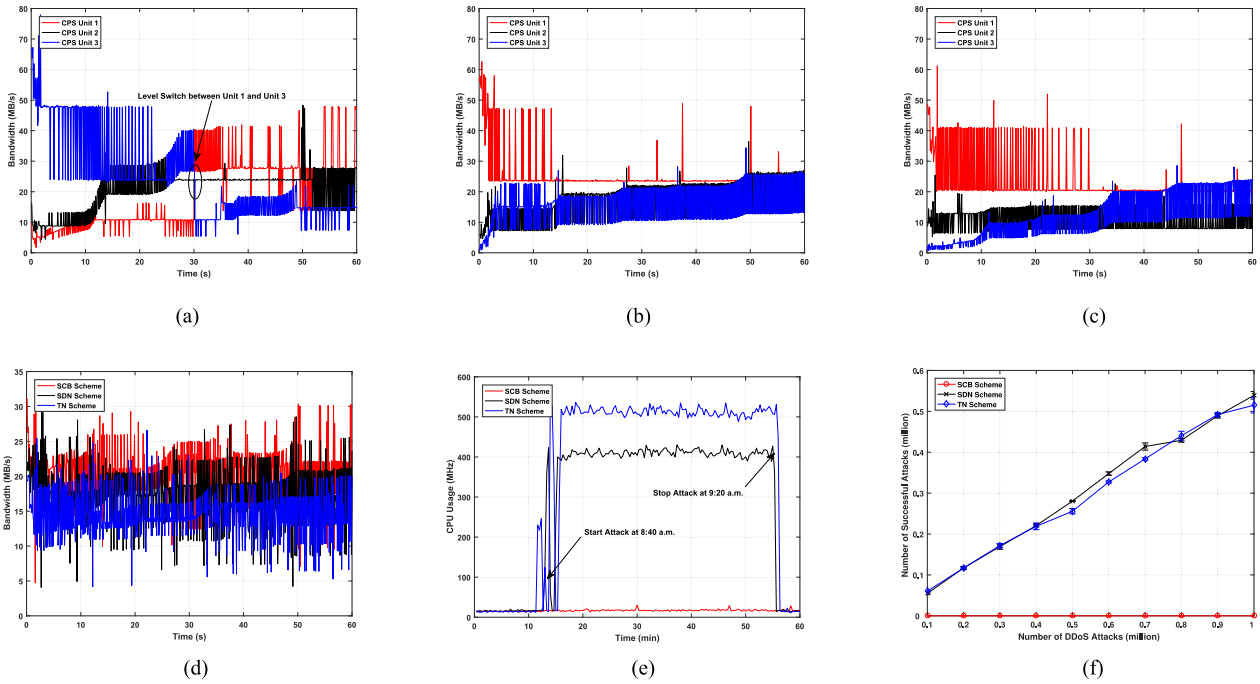

Fig. 4. Validation of bandwidth allocation and invasion defense. (a) CPS with SCB for multiple user priority. (b) CPS with SDN for multiple user priority. (c) CPS with TN for multiple user priority. (d) Average bandwidth for three schemes. (e) Influence of processing capacity. (f) Statistics of successful attacks.

monitoring tool to record the results at the access node. For invasion defense, attackers in different areas are able to initiate a large number of TCP SYN connections to the same target host, simultaneously. By supervising the central processing unit (CPU) usage of the target host and the number of attack packets, the defense capability is verified.

We have completed the registration, authentication, and other necessary procedures for each legitimate host in advance. The range of user level is from 1 to 3. The bandwidth limitation of network is set to 70 MB/s and the maximum transmission unit (MTU) is 1500 Bytes. The service ports of applications inside legitimate terminals are 12 345, 12 346, and 12 347, respectively. In the attack scenario, the packet sending interval of attack hosts is set to 5 us. Each experiment has been repeated multiple times and the total duration is more than 24 hours. The number of collected data items is larger than 220 000.

To emphasize the utilization and protection performance of the SCB scheme, we mainly analyzed and compared the characteristics of the SCB, SDN, and TN schemes in terms of bandwidth allocation and invasion defense.

B. Bandwidth Allocation

In Fig. 4, the bandwidth allocation of three CPS device units in the SCB scheme has been provided. The blue curve shows the situation of the high-level unit (UL = 3). The black and red curves present the fluctuations of the medium-level unit (UL = 2) and the low-level unit (UL = 1), respectively. Each test lasts for more than 60 s and the sampling interval is 0.1 s. Before $t = 30$ s, the available bandwidth of the high-level unit is steadily maintained at the best position. It is clear that the medium-level and low-level units have obtained corresponding bandwidth portions. The controller performs manipulation in queue transmission and schedules forwarding rules dynamically. Therefore, the specific bandwidth is allocated according to different user levels successfully. To emphasize the impact of alternations, we exchange the UL values of Unit 1 and Unit 3 at 30 s without interrupting the packet transmissions. Although the fluctuations are still obvious, mean values of the top and bottom parts are swapped right after user level changes. In the last 30 s, the red curve takes more bandwidth ratio, the black curve is located in the middle, and the blue curve is at the lowest position in most of the time.

In Fig. 4, the distribution results are given when the SDN scheme is applied for these units. Due to the limited forwarding mechanism of the controller, the dynamical demands of services cannot be effectively identified. From the macroscopical perspective, the values of three units show a tendency from dispersion to equalization. Especially, at the beginning of the test, the randomness of entering a buffer lead to serious fluctuation of bandwidth allocation. As the data packets sent via Unit 1, Unit 2, and Unit 3 occupy the queue evenly, the corresponding resources are equally shared. From the microscopic perspective, the data packets generated by Unit 1 first arrive at the queue and take up most of bandwidth. Then, such situation sustains for about 15 s before reaching to the stabilization stage. Compared with the Unit 1, it consumes more time for Unit 2 and Unit 3 to preempt the transmission resources. Until $t = 50$ s, a floating equilibrium point can be obtained by three units, ranging from 10 to 30 MB/s. During the experiment, four pulses within the red curve illustrate the uncertainty of this scheme, which is not beneficial for providing differentiated services.

In Fig. 4, the bandwidth utilization of three units is provided under the traditional network circumstances. Compared to the

previous case, three curves have demonstrated slower convergence rates. Especially for the red one, the whole stable period dominates almost half time. During the rest time of testing, the throughput has been maintained with a reasonable value 20 MB/s. For the Unit 2 and Unit 3, their average bandwidth approximately equals to 12 MB/s and 10 MB/s, respectively. At the beginning of this experiment, data packets from both units still lose the superiority to arrive in the queue earlier than Unit 1, which compels them to obey lower occupation ratio. Moreover, there is no obvious changes for the black curve in the whole process and the value is maintained around 10 MB/s. On the contrary, the trend of the blue curve keeps approaching to the red curve. However, the allocation results show that the red one holds most of bandwidth resources before 35 s. Since the fluctuation of Unit 1 recedes, the curve gradually smooths expect for a rebound at $t = 48$ s. In short, it is powerless for the traditional network to realize deterministic managements.

In Fig. 4, we compare the overall performance of above schemes. According to the statistics, the average bandwidth of SCB scheme (marked with a red curve), is preferable to the other candidates. The changing range is between 5 and 30 MB/s. The bandwidth value shows a significant fluctuation in roughly each 20 s interval. The SDN scheme (marked with a black curve) is inferior to the previous one and slightly outperforms the TN scheme (marked with a blue curve). It's minimum value is around 4 MB/s, while the optimized points approaches to 28 MB/s. Compared with the former two schemes, values of TN scheme illustrate relatively small variations, from 4 to 26 MB/s. In the early stage of this experiment, there is no obvious difference between TN scheme and SDN scheme. After $t = 30$ s, the disadvantages, in terms of flexibility, gradually appear. The average bandwidth values are 26.88, 17.55, and 14.64 MB/s, respectively. Based on the established strategy, advantages of the SCB scheme are quite clear.

In addition, it can be obtained that the part of experimental results fluctuate greatly. This is due to the usage of virtual machines in validation procedures. Based on this condition, the link transmission is unstable with serious jitters. Especially for the virtual network cards, the randomness of data packet forwarding is apparent. However, experimental results are feasible to support the proof of our proposed scheme, i.e., optimizing bandwidth utilization and implementing differentiated services.

C. Invasion Defense

In Fig. 4, the attack tolerance of three schemes in DDoS scenario has been analyzed. We recorded the specific CPU usage of the target hosts. The total testing period lasts for one hour and the attacking duration is set to 40 minutes. To highlight the invasion impact on CPUs, the start and stop points of attack are marked, respectively. The malicious machine launches an attack at 8:40 A.M. and stops it at 9:20 A.M. When the abnormal TCP SYN connection was established, the CPU consumption of the SDN (marked with a black curve) and TN (marked with a blue curve) schemes deteriorated rapidly. The invaded host in TN scheme perceives the attack data packets before the SDN scheme and recovers the CPU status at a slower speed. Compared to other candidates, the TN scheme presents the worst case in terms of the defense performance, with the value about 510 MHz. The CPU usage of SDN scheme lies between the other two schemes, i.e., 400 MHz. For the SCB scheme, the mapping mechanism and access control module has been deployed to achieve the network isolation. Therefore, abnormal data packets sent by illegal machines are blocked outside the core network. The CPU consumption of SCB scheme (marked with a red curve) is not changed significantly.

In Fig. 4, we describe the amount of malicious traffic received by the victim host. Each scheme has been repeated multiple times and the corresponding confidence intervals are provided. As the volume of attack flows increases, the curve values in the SDN scheme and TN scheme show an upward trend. When the number of DDoS data packets is less than 0.4 million, the performance of the SDN scheme is slightly better than that of the TN scheme. With the increase of DDoS flows, the SDN scheme appears periodic deterioration. Due to the lack of defense mechanism inside experiment, SDN and TN schemes fail to effectively resist DDoS data packets. The average attack success rates of them are 55.26% and 56%, respectively. It can be concluded that the received amount is positively correlated with the total number of attack in both schemes. On the contrary, our proposed scheme is able to prevent attacks from the edge network. The red curve illustrates that the malicious traffic has no serious impact on SCB scheme and the average attack success rate is close to 0%.

It can be seen from the abovementioned results that the SCB scheme has certain advantages in terms of bandwidth allocation and invasion defense.

V. Related Work

The existing research work has provided a large number of solutions for not only the service management, but also security enhancement in CPS. We classify them into bandwidth allocation and invasion defense. Typical approaches are analyzed and compared from multiple points of view.

A. Bandwidth Allocation

To integrate various resources, Wang *et al.* [12] proposed a reliable and efficient service composition method. The variation coefficient is adopted to suppress the fluctuations. The best optimal components are chosen based on maximizing the fitness function. Maruf and Azim [13] designed an approach to handle the mixed-criticality tasks by considering the important parameter, such as deadlines and execution times. The authors also proposed a prediction algorithm to offload suitable tasks to remote cloud. Lyu *et al.* [14] presented a redundant communication policy to meet the requirements of state estimation. By incorporating the ISM channels, the proposed scheme provided adequate spectrum capacity and ensured the calculation results. Aksanli and Rosing [15] aim to solve the existing problems in energy management for residential areas. An accurate model for estimating the relationship between family members and power requirements was introduced to efficiently generate several profiles. Wang and Song [16] proposed a novel model to establish resilient controllers for specific units. Try-once-discard protocol was focused and analyzed to reduce the burden of

communication. Mo *et al.* [17] provided a subtle event-driven solution to improve the accuracy of control and reduce the consumption of energy. By utilizing two steps, an optimization problem including actuator schedule and output control was solved.

B. Invasion Defense

To extend the usage scope of adversary models, Sangaiah *et al.* [18] proposed an energy-aware method to achieve confidentiality in smart industrial CPS environment. The authors provided meaningful testimonials to help the users to reduce the computation and communication costs in each query process. Ma *et al.* [19] investigated the dissipativity-based resilient sliding-mode control issue with the considerations of malicious attacks resistance. The operations in physical layer was analyzed to identify the upper bound of the sample-data rate. Farivar *et al.* [20] focused on a class of n-order nonlinear CPS system and assumed that attacks only appear in the forward channel. A hybrid intelligent control mechanism powered by neural networks was developed to compensate abnormal invasions. Wang *et al.* [21] presented a fast CP-ABE scheme for mobile devices not only to lower the storage and computational costs, but also to improve the cyber–physical privacy and security for healthcare information. The correctness of decryption can be verified according to the adoption of signature schemes. Sun *et al.* [22] presented a flexible model based on network state variations to reduce negative influences of DoS attacks. The authors emphasized that parameter setting is critical for condition establishment. Liu *et al.* [23] claimed that the unified network paradigm is not suitable for different requirements of CPS. The advantages of network slicing were identified and illustrated based on a small-scale testbed. Wang and Li [24] studied the synchronization control methodology in dependable scenario. In regular case, an optimal scheme including two control instances was introduced. In attack case, the difficulty of analyzing stability was explained to motivate a new replacement approach.

The existing references have provided some attractive solutions. Unfortunately, they only focus on the service or security issues from a single CPS aspect without comprehensive considerations. Moreover, the user priority and malicious access are not fully analyzed based on industrial demands.

VI. Conclusion

CPS was successfully utilized in many significant industries. Service management and security enhancement were two interactive aspects for massive users in CPS. In this article, we proposed a SCB scheme to strengthen the bandwidth allocation and invasion defense. By establishing mathematical models of congestion window variations, we analyzed the throughput influences of user priority. According to the mechanisms of identity authentication, the malicious attacks can be recognized and restrained. The specific system structure and main procedures of SCB were presented. To validate the feasibility of our scheme, a comprehensive prototype including multiple domains was established based on previous models and mechanisms. Compared to existing candidates, such as SDN and traditional Internet, the SCB scheme performed quite well in different experiments and scenarios.

In the future, we aim to provide on-demand transmission capabilities for CPS units and enable effective connections between the cyber side and physical side. Other performance metrics and hacking patterns will be further investigated.

Appendix
Proof of (11)

To calculate the value of $E[w_i(t)]$, let

$$\begin{aligned}\zeta(t) &= (\mathrm{RTT}_i^{-1}P(w_i(t) \le T_i) - \beta_i\lambda_{\mathrm{TD}i} - \lambda_{\mathrm{TO}i})\\ \iota(t) &= \alpha_i\mathrm{RTT}_i^{-1}(1 - P(w_i(t) \le T_i)) + \lambda_{\mathrm{TO}i}.\end{aligned} \tag{24}$$

Then, we have

$$\frac{\mathrm{d}E[w_i(t)]}{\mathrm{d}t} = \zeta(t)E[w_i(t)] + \iota(t). \tag{25}$$

Leave the $\iota(t)$ at the right-hand side to have

$$\frac{\mathrm{d}E[w_i(t)]}{\mathrm{d}t} - \zeta(t)E[w_i(t)] = \iota(t). \tag{26}$$

The $E[w_i(t)]$ can be expressed by

$$\begin{aligned}E[w_i(t)] &= e^{\int\zeta(t)dt}\left[\int\iota(t)e^{-\int\zeta(t)dt}dt + C_0\right]\\ &= e^{\zeta(t)t}\left[\int\iota(t)e^{-\zeta(t)t}dt + C_0\right]\\ &= e^{\zeta(t)t}\left[-\frac{\iota(t)}{\zeta(t)}e^{-\zeta(t)t} + C_0\right]\\ &= -\frac{\iota(t)}{\zeta(t)} + C_0e^{\zeta(t)t}.\end{aligned} \tag{27}$$

Acknowledgment

The authors would like to thank editors and reviewers for their valuable comments and suggestions.

References

[1] Z. Hong, R. Wang, S. Ji, and R. Beyah, "Attacker location evaluation-based fake source scheduling for source location privacy in cyber-physical systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 5, pp. 1337–1350, May 2019.

[2] K. Guan *et al.*, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part II: Case study for paradigm implementation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9129–9144, Oct. 2018.

[3] K. Guan *et al.*, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part I: Paradigm, significance analysis, and scenario reconstruction," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9112–9128, Oct. 2018.

[4] Y. Zhang, X. Liu, H. Zhang, and C. Jia, "Constructing chaotic systems from a class of switching systems," *Int. J. Bifurcation Chaos*, vol. 28, Feb. 2018, Art. no. 1850032.

[5] H. Hui, C. Zhou, S. Xu, and F. Lin, "A novel secure data transmission scheme in industrial Internet of Things," *China Commun.*, vol. 17, no. 1, pp. 73–88, 2020.

[6] P. Li, R. Li, Y. Cao, D. Li, and G. Xie, "Multiobjective sizing optimization for island microgrids using a triangular aggregation model and the levy-harmony algorithm," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3495–3505, Aug. 2018.

[7] C. Roberts *et al.*, "Learning behavior of distribution system discrete control devices for cyber-physical security," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 749–761, Jan. 2020.

[8] Y. Zhang, P. Shi, C. Lim, H. Zhu, J. Hu, and Y. Zeng, "Chaotification of a class of linear switching systems based on a shilnikov criterion," *J. Franklin Inst.*, vol. 354, no. 13, pp. 5519–5536, 2017.
[9] C. Gong, F. Lin, X. Gong, and Y. Lu, "Intelligent cooperative edge computing in the Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9372–9382, Oct. 2020.
[10] Y. Zhang and T. Jiang, "Finite-time boundedness and chaos-like dynamics of a class of Markovian jump linear systems," *J. Franklin Inst.*, vol. 357, no. 4, pp. 2083–2098, 2020.
[11] P. Li, R. Dargaville, Y. Cao, D. Li, and J. Xia, "Storage aided system property enhancing and hybrid robust smoothing for large-scale PV systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2871–2879, Nov. 2017.
[12] S. Wang, A. Zhou, M. Yang, L. Sun, C. Hsu, and F. Yang, "Service composition in cyber-physical-social systems," *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 1, pp. 82–91, Jan.–Mar. 2020.
[13] M. A. Maruf and A. Azim, "Extending resources for avoiding overloads of mixed-criticality tasks in cyber-physical systems," *IET Cyber-Phys. Syst.: Theory Appl.*, vol. 5, no. 1, pp. 60–70, 2020.
[14] L. Lyu, C. Chen, J. Yan, F. Lin, C. Hua, and X. Guan, "State estimation oriented wireless transmission for ubiquitous monitoring in industrial cyber-physical systems," *IEEE Trans. Emerg. Topics Comput.*, vol. 7, no. 1, pp. 187–201, Jan.–Mar. 2019.
[15] B. Aksanli and T. S. Rosing, "Human behavior aware energy management in residential cyber-physical systems," *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 1, pp. 45–57, Jan.–Mar. 2020.
[16] J. Wang and Y. Song, "Resilient RMPC for cyber-physical systems with polytopic uncertainties and state saturation under TOD scheduling: An ADT approach," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4900–4908, Jul. 2020.
[17] L. Mo, P. You, X. Cao, Y. Song, and A. Kritikakou, "Event-driven joint mobile actuators scheduling and control in cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 11, pp. 5877–5891, Nov. 2019.
[18] A. K. Sangaiah, D. V. Medhane, G. Bian, A. Ghoneim, M. Alrashoud, and M. S. Hossain, "Energy-aware green adversary model for cyberphysical security in industrial system," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3322–3329, May 2020.
[19] R. Ma, P. Shi, and L. Wu, "Dissipativity-based sliding-mode control of cyber-physical systems under denial-of-service attacks," *IEEE Trans. Cybern.*, early access, doi: 10.1109/TCYB.2020.2975089.
[20] F. Farivar, M. S. Haghighi, A. Jolfaei, and M. Alazab, "Artificial intelligence for detection, estimation, and compensation of malicious attacks in nonlinear cyber-physical systems and industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2716–2725, Apr. 2020.
[21] S. Wang *et al.*, "A fast CP-ABE system for cyber-physical security and privacy in mobile healthcare network," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 4467–4477, Jul./Aug. 2020.
[22] Q. Sun, K. Zhang, and Y. Shi, "Resilient model predictive control of cyber-physical systems under dos attacks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4920–4927, Jul. 2020.
[23] Q. Liu, T. Han, and N. Ansari, "Learning-assisted secure end-to-end network slicing for cyber-physical systems," *IEEE Netw.*, vol. 34, no. 3, pp. 37–43, May/Jun. 2020.
[24] N. Wang and X. Li, "Secure synchronization control for a class of cyber-physical systems with unknown dynamics," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 5, pp. 1215–1224, Sep. 2020.

Fei Song (Member, IEEE) is currently a full Professor with the National Engineering Laboratory for Next Generation Internet Technology and the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China. His current research interests include network architecture, system security, protocol optimization, and cloud computing.

Mr. Song is a Technical Reviewer for several journals including the *IEEE Communications Magazine*, IEEE INTERNET OF THINGS JOURNAL, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON SERVICES COMPUTING, and IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING.

Zhengyang Ai received the B.S. degree in information system from the School of Computer and Information Technology, Northeast Petroleum University, Daqing, China, in 2016. He is currently working toward the Ph.D. degree in information system with the National Engineering Laboratory for Next Generation Internet Technology, School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China.

His current research interests include network architecture and security.

Haowei Zhang received the B.S. degree in automatic control from the Beijing Instituted of Technology, Beijing, China, in 2015, and the M.S. degree in entrepreneurship from the University of Liverpool, Liverpool, U.K., in 2017.

He is currently with the China International Engineering Consulting Corporation, Beijing, China. His current research interests include digital economy, semiconductors, robots, and cloud computing.

Ilsun You (Senior Member, IEEE) received the M.S. and Ph.D. degrees in computer science from Dankook University, Seoul, South Korea, in 1997 and 2002, respectively, and the second Ph.D. degree from Kyushu University, Fukuoka, Japan, in 2012.

From 1997 to 2004, he was with the Thin Multimedia, Inc., Internet Security Co., Ltd., and Hanjo Engineering Co., Ltd. as a Research Engineer. He is currently a Full Professor with the Department of Information Security Engineering, Soonchunhyang University, Asan, South Korea. He has served or is currently serving as a main organizer of international conferences and workshops, such as MIST, MobiWorld, and MobiSec. He has focused on 4G/5G security, security for wireless networks and mobile Internet, and IoT security while publishing more than 180 papers in these areas.

Dr. You is a Fellow of the IET. He is the EiC of *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*. He is in the Editorial Board for *Information Sciences, Journal of Network and Computer Applications*, IEEE Access, *Intelligent Automation & Soft Computing*, *International Journal of Ad Hoc and Ubiquitous Computing*, *Computing and Informatics*, and *Journal of High Speed Networks*.

Shiyong Li received the B.Sc. degree from Qingdao University, Qingdao, China, in 2004, the M.Sc. degree from Yanshan University, Qinhuangdao, China, in 2007, and the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2011.

He is currently a Full Professor with the School of Economics and Management, Yanshan University. He is the (Co)-Author of more than 60 papers in mathematics, technique, and management journals. He has been a principal investigator/co-investigator on several research projects supported by the National Natural Science Foundation of China, the National Education Committee Foundation of China, the China Postdoctoral Science Foundation, and other foundations. His research interests include cloud migration for enterprise applications, resource allocation of cloud/edge computing, information systems and electronic commerce, and economics of queues.