IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 24, 2019, accepted December 13, 2019, date of publication December 30, 2019, date of current version January 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963139

Latency-Optimal Network Intelligence Services in SDN/NFV-Based Energy Internet Cyberinfrastructure

ARDIANSYAH[1], YONGHOON CHOI[2], (Senior Member, IEEE), MUHAMMAD REZA KAHAR AZIZ[3], (Member, IEEE), KANGWOOK CHO[4], AND DEOKJAI CHOI[1]

[1]Department of Electronics and Computer Engineering, Chonnam National University, Gwangju 61186, South Korea
[2]Department of Electrical Engineering, Chonnam National University, Gwangju 61186, South Korea
[3]Department of Electrical Engineering, Institut Teknologi Sumatera, South Lampung 35365, Indonesia
[4]Department of Market and System Development, Korea Power Exchange (KPX), Naju 58217, South Korea

Corresponding authors: Yonghoon Choi (yh.choi@jnu.ac.kr) and Deokjai Choi (dchoi@jnu.ac.kr)

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1I1A3A01060631).

ABSTRACT Energy internet (EI) is a very complex system with various applications that not only require a high-level of cyber-security but also need low-latency communication. Thus, cyberinfrastructure with latency-optimal network intelligence services (NIS), in which application data flows are deeply examined in real-time, is inevitable. In the future internet system, a set of NIS can flexibly be implemented in network function virtualization (NFV)-based middleboxes that overlay on software-defined networking (SDN) architecture, becoming an SDN/NFV-based cyberinfrastructure. However, how to deploy these middleboxes is a non-deterministic optimization problem, which is complicated and time-consuming. Hence, by focusing on latency minimization, we develop an artificial intelligence (AI)-powered solution consisted of two phases. First, middleboxes placement based on the graph cluster analysis, and second, NIS resource allocation based on the prediction of service usage-ratio in each corresponding cluster. The simulation-based experimental evaluation shows that our proposed strategy using an optimized K-means algorithm outperforms the recent state-of-the-art middleboxes placement approaches. The average end-to-end flow latencies are around 23.81%, 18.44%, and 11.49% lower compared with the simulated annealing method, the basic sequential algorithmic scheme, and the minimum spanning tree procedure, respectively. Besides, the proposed resource allocation scheme optimizes further the latency minimization around 4.24%. We believe that the work presented in this paper will aid the communication service providers (CSP) in providing a secure and low-latency SDN/NFV-based cyberinfrastructure for the EI ecosystem.

INDEX TERMS Energy internet, artificial intelligence, network intelligence, NFV middlebox, SDN architecture.

I. INTRODUCTION

Recently, the penetration of renewable energy generation, such as building-integrated photovoltaics (BIPV), has been increased in many countries [1]–[3]. With renewable energy generation, consumers can evolve into prosumers, a new type of energy stakeholders that can produce and use their own electricity, and also sell their excessed energy to the market. Therefore, various smart grid technologies and applications have been proposed to accommodate the high penetration of prosumers with distributed renewable energy resources (DRERs) and distributed energy storage devices (DESDs) [4], [5]. These smart grid technological advancements bring opportunities to transform the current power system to energy internet (EI), an internet business model of the electricity grid, in which multiple energy and data flows are in dual circulation and coupling among the entire value chain.

The associate editor coordinating the review of this manuscript and approving it for publication was Mubashir Husain Rehmani.

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

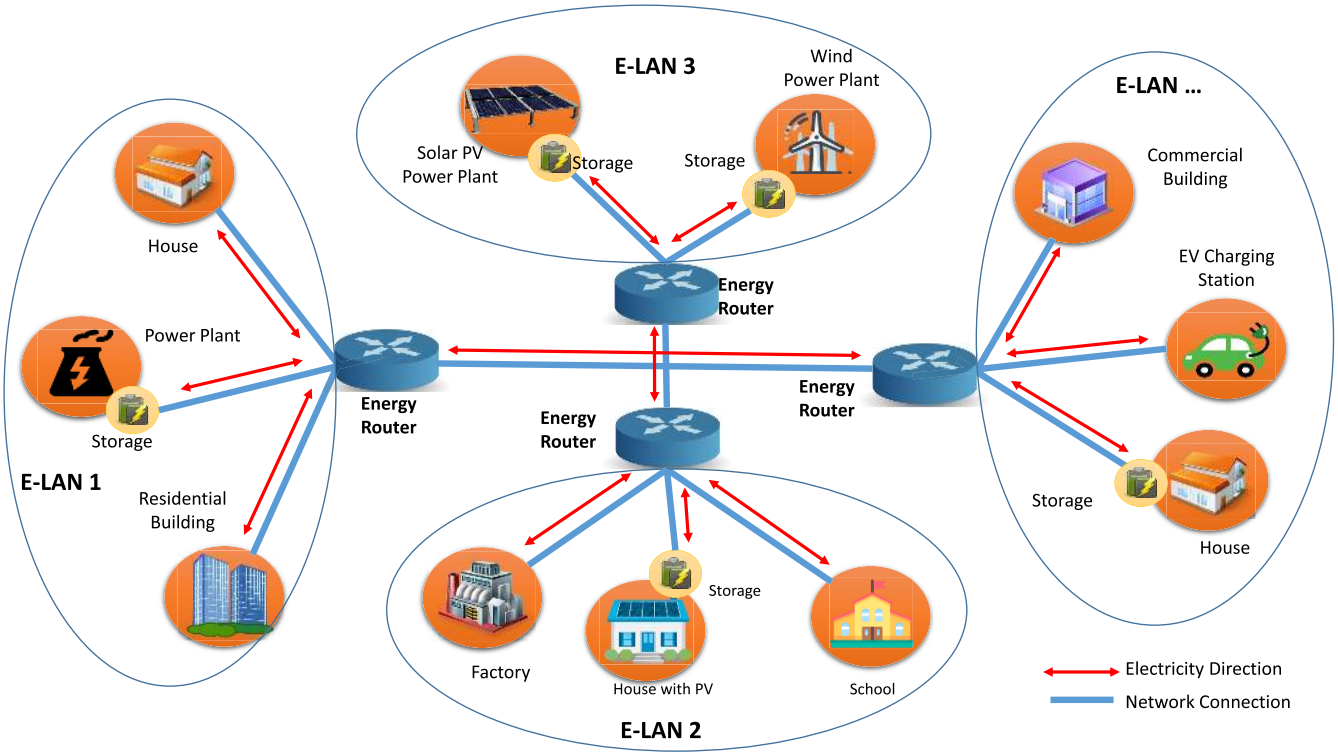

FIGURE 1. Illustration of EI ecosystem [5]. E-LAN is a localized group of energy stakeholders that can operate independently, or connected to the grid to buy/sell energy. Similar to the internet router, the energy router is used as an intermediate device to exchange both energy and data bidirectionally between E-LANs in the network.

In the EI ecosystem, all energy stakeholders can be joined flexibly and seamlessly to the closest energy local area network (E-LAN), as depicted in Fig. 1. Then, using the so-called energy router [6], the energy can be exchanged and transferred between one to another. Hence, the energy sharing economy [7] can be realized, which enables energy consumers to obtain the supplies directly from the nearest producers. Moreover, the cascading failure or blackout could also be resisted, which improves the stability of the whole electricity grid. It is owing to the immense development of the future renewable electric energy delivery and management (FREEDM) system [4], which considerably improves energy utilization containing all novel phases of energy generation, transmission, storage, and distribution. FREEDM system consists of some pivotal technologies such as the intelligent energy management (IEM) software, the distributed grid intelligence (DGI) software, the solid-state transformer (SST), the real-time remote monitoring, and the smart fault isolation device (FID).

It should be noted that the EI has attracted increasing attention of government and institutions in many countries. For rural electrification in Indonesia, a country with more than 17,000 islands, EI is the most promising solution to enable the internet of microgrids [8]. Furthermore, as a response to the Fukushima nuclear crisis, a large group of firms in Japan is starting to explore EI with an expectation to transform the country's electricity system with distributed energy and micro-grid integration [9]. Moreover, the EI was also proposed in Germany following a political decision to shut down all German nuclear reactors by 2022 [10]. Besides, the Chinese government and state grid corporation of China (SGCC) has proposed a proposal so-called "global EI" and then launches an action plan every five years [11]. Also, the EI platform has been launched in Europe as a novel strategy to achieve decarbonizing commitment by 2050 [12]. However, being in its infancy stage, EI business values and social benefits are becoming increasingly apparent with the advances in smart grid technologies. Thus, more research and development need to be performed to support the diverse and rigorous requirements of reliability, flexibility, latency, and security in the EI. To this end, the emerging technical initiative (ETI) on smart grid communications (SGC) has issued a positioning paper in 2018, which included EI as one of the eight research agenda structures [13].

Among novel future internet technologies, software-defined networking (SDN) is expected to be adopted in the building of cyberinfrastructure for end-to-end interactions across the entire value chain in the EI. Utilizing the SDN approach, both energy and data flows can be managed flexibly following the four principles, which are 1) logically centralized management, 2) separation of control, data, and energy planes, 3) programmability, and 4) open interfaces [14]. Therefore, some research works have been conducted recently to develop a framework and evaluate the performance of

SDN-based EI cyberinfrastructure [14]–[16]. However, EI is a very complex system with various applications that have specific and strict functional requirements [17], [18]. Many applications, including distribution automation, load control signaling, and outage alarming, are described to require low-latency communication. Although some other applications, e.g., smart-meter data collection, are more latency tolerant, however, they need to have a high level confidentially, availability, and integrity. Thus, various network intelligence services (NIS) such as network security applications, traffic analysis elements, and deep-packet inspection (DPI) tools are indispensable to be utilized, fulfilling the required cyber-security in EI [18], [19].

In recent years, the ETI for network intelligence has worked together to support and endorse research towards embedding AI, SDN, and network function virtualization (NFV). In the context of cyber-security, taking service policies as inputs, a set of AI-powered NIS can be applied virtually in NFV-based middleboxes that overlay on SDN architecture, and they are becoming SDN/NFV-based cyberinfrastructure [20]. However, the use of this approach to secure and protect application data flows may increase end-to-end flow latency significantly. Considering that the reliability requirement is defined in [21], the successful delivery of the application data flows, but with the latency higher than the defined requirement, can be considered as a failure.

Taking into account the balance between required cyber-security and low-latency communication is essential for many applications in the EI. In this paper, we present our work to provide latency-optimal NIS in SDN/NFV-based EI cyberinfrastructure. In fact, the end-to-end flow latency always depends on the middleboxes' placement in the network. However, some previous research works proved that this problem is a non-deterministic polynomial-time (NP)-hard, which is a complicated and time-consuming decision problem [22]–[25]. Hence, a trade-off optimization method is needed to achieve a heuristic solution. Among existing approaches, graph cluster analysis is the most popular AI-powered solution to solve the problem. However, the recent state-of-the-art graph cluster analysis methods [26]–[28], having at least two drawbacks that can not guarantee the end-to-end flow latency, can be minimized, i.e., 1) randomly choosing the clusters' threshold and 2) arbitrarily selecting the initial center. Hence, the main contributions of this paper are as follows.

1) We introduce the utilization of NIS for fulfilling the main cyber-security requirements in the EI ecosystem. All NIS are virtually implemented in a number of NFV-based middleboxes.
2) We reformulate an objective function for the latency minimization problem. It should be noted that for the time-critical energy control signaling application, the end-to-end latency should be less than 3 ms [6].
3) We consider three main constraints, i.e., the middleboxes' processing power capacity, the forwarding nodes' memory resource, and the communication network configuration. These constraints are the minimum obstacles in such SDN/NFV-based cyberinfrastructure.
4) We develop the AI-powered solution, which consists of two phases. First, an optimized K-means algorithm is utilized to find the latency-optimal middleboxes placement in several clusters. Second, a prediction of NIS usage-ratio is employed to develop a dynamic resource allocation scheme, which optimizes further latency minimization in the corresponding clusters.
5) We evaluate our proposed method along with the recent state-of-the-art approaches, i.e., the simulated annealing [22], the basic sequential algorithmic scheme (BSAS) [26], the minimum spanning tree (MST) [27], and the modified BSAS [28]. The simulation-based experimental comparison is carried out on two network topologies, i.e., FatTree [29] and Abilene [30]. We expect that these two topologies are representing both layered and irregular network structures of SDN/NFV-based EI cyberinfrastructure, respectively.

We believe that the result of this work can be used as a guideline for communication service providers (CSP) to provide a secure but also low-latency cyberinfrastructure for the EI.

The rest of this paper is organized as follows. In the next section, we provide related works to utilize NIS in the EI ecosystem. Section III describes the system model, problem formulation, and recent state-of-the-art methods. Section IV explains the proposed solution, section V presents our evaluation, and finally, section VI concludes this paper.

II. RELATED WORK

A. SERVICE ABSTRACTION MODEL

In recent years, the national institute of standards and technology (NIST) and the open smart grid (OpenSG) network task force have comprehensively analyzed all possible functional requirements of various applications for the future EI. Currently, not less than 1400 application data flows have been specified in detail, including their payload size and type, security, latency, reliability, data transmission frequency, and so forth [31]. On the other hand, several groups work together to specify the quality of services (QoS) requirements for the specific application. For example, the North American synchrophasor initiative network (NASPInet), a working group with the mission to improve power system reliability and visibility through wide-area measurement and control.

The NASPInet has contextualized five classes of data services for synchrophasor applications with specific traffic attributes. As depicted in Table 1 [32], class A is to support the needs of high-performance feedback control applications. Thus, the reliable cyberinfrastructure for this class is critically essential. It should have a fast data rate and very low latency, as well as can guarantee a high level of data availability. Furthermore, classes B and C are for the applications with less strict latency requirements such as the feed-forward estimator enhancement application and the view only appli-

TABLE 1. NASPInet five classes of data services and their specific traffic attributes [32].

Traffic Attribute	*A*	*B*	*C*	*D*	*E*
Low Latency	4	3	2	1	1
Availability	4	2	1	3	1
Accuracy	4	2	1	4	1
Time Alignment	4	4	2	1	1
High Message Rate	4	2	2	4	1
Path Redundancy	4	4	2	1	1

Key: 4-Critically Important, 3-Important
2-Somewhat Important, 1-Not very Important

cation, respectively. Then, class D is to support the need for post-mortem event analysis, and class E is intended for testing, research, and development.

To understand all the requirements above and provide NIS appropriately, the use of a service abstraction model (SAM) is indispensable. Therefore, it is worth to mention a SAM proposed by G.D. Nugraha *et al.*, in [33]. In this case, the service requirements can be represented by three sets of parameters, i.e., content, context, and resources. To be detailed, the content provides the service-related parameters such as payload size and type, maximum delay, minimum bandwidth, and so forth. Furthermore, the context serves the users/applications related parameters concerning interest, such as data transmission frequency, schedule, and location. Last, the resources supply the requirements of network service resources such as networking medium, computing power, memory space, etc. Taking advantage of this model, Fig. 2 depicts the service abstraction template for application data flows in SDN/NFV-based EI cyberinfrastructure.

B. NIS APPLICATIONS

The evolution and growth of internet technologies offer possibilities for CSP to provide better QoS, as well as develop new types of services. Hence, NIS are utilized to capture the detailed information from applications, or users' data flows, to provide the analysis of their demand and to manage the usage once deployed. Some essential applications of NIS, ranging from understanding user behavior analysis to provide intrusion detection, are listed in Table 2.

TABLE 2. Example applications of NIS.

Purpose	Example Applications
Understand Customer Demand	• User behavior analysis • Customer segmentation • Personalized services
Manage Services	• Resources optimization • Quality of experience analysis • Regulatory compliance • Denial of services detection

Recently, many research projects have been conducted to utilize AI techniques for NIS, in terms of traffic classification [34], traffic prediction [35], accelerates service provisioning [36], intrusion detection [37], and so forth. Moreover, for securing cyberinfrastructure against

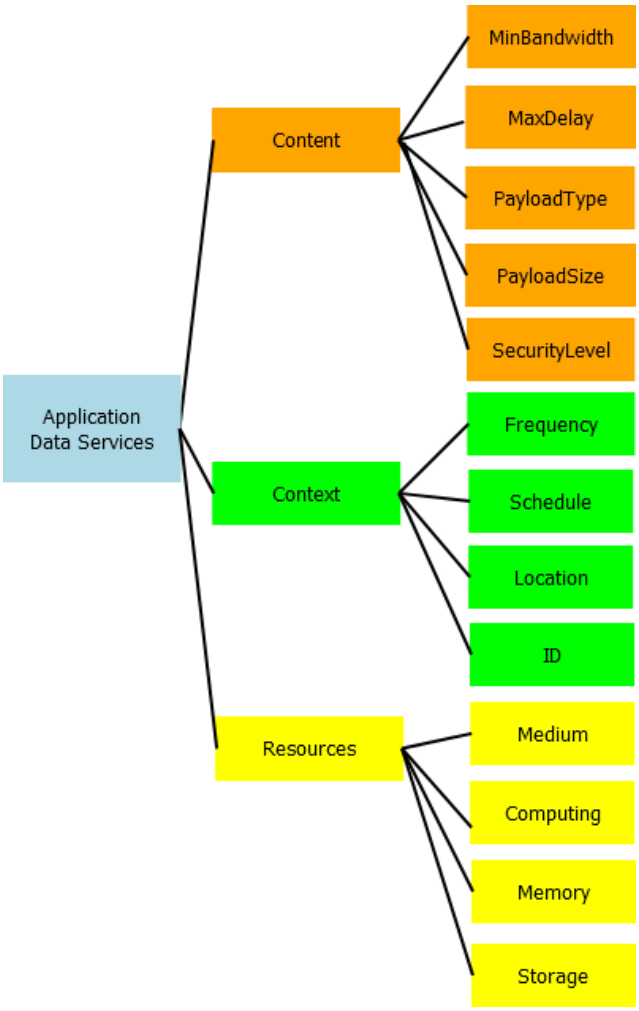

FIGURE 2. Service abstraction template for application data flows in SDN/NFV-based EI cyberinfrastructure. The service abstraction model consists of functional requirements in three sets of parameters, i.e., content, context, and resources.

intruders and other threats, some experiential networked intelligence (ENI) research projects have been started recently combining AI, SDN, and NFV. For example, the SHIELD research project, as described in [38]. They demonstrate an AI-powered framework to detect attacks using a policy-driven control loop intelligently. Adopting this framework to SDN/NFV-based EI cyberinfrastructure, we can develop AI-powered attack detection and mitigation recipes. Through intent-driven and autonomous-driving network, fulfilling the main cyber-security requirements in EI ecosystem are as follows [39]

1) **Attack detection and resilience operation**. It is required to monitor network traffic in real-time, detect abnormal incidents due to various attacks, and continue operations in the presence of attacks using self-healing ability.
2) **Identification, authentication, and access control**. It is essential to ensure that the resources are accessed only by the appropriate entities that are correctly identified.

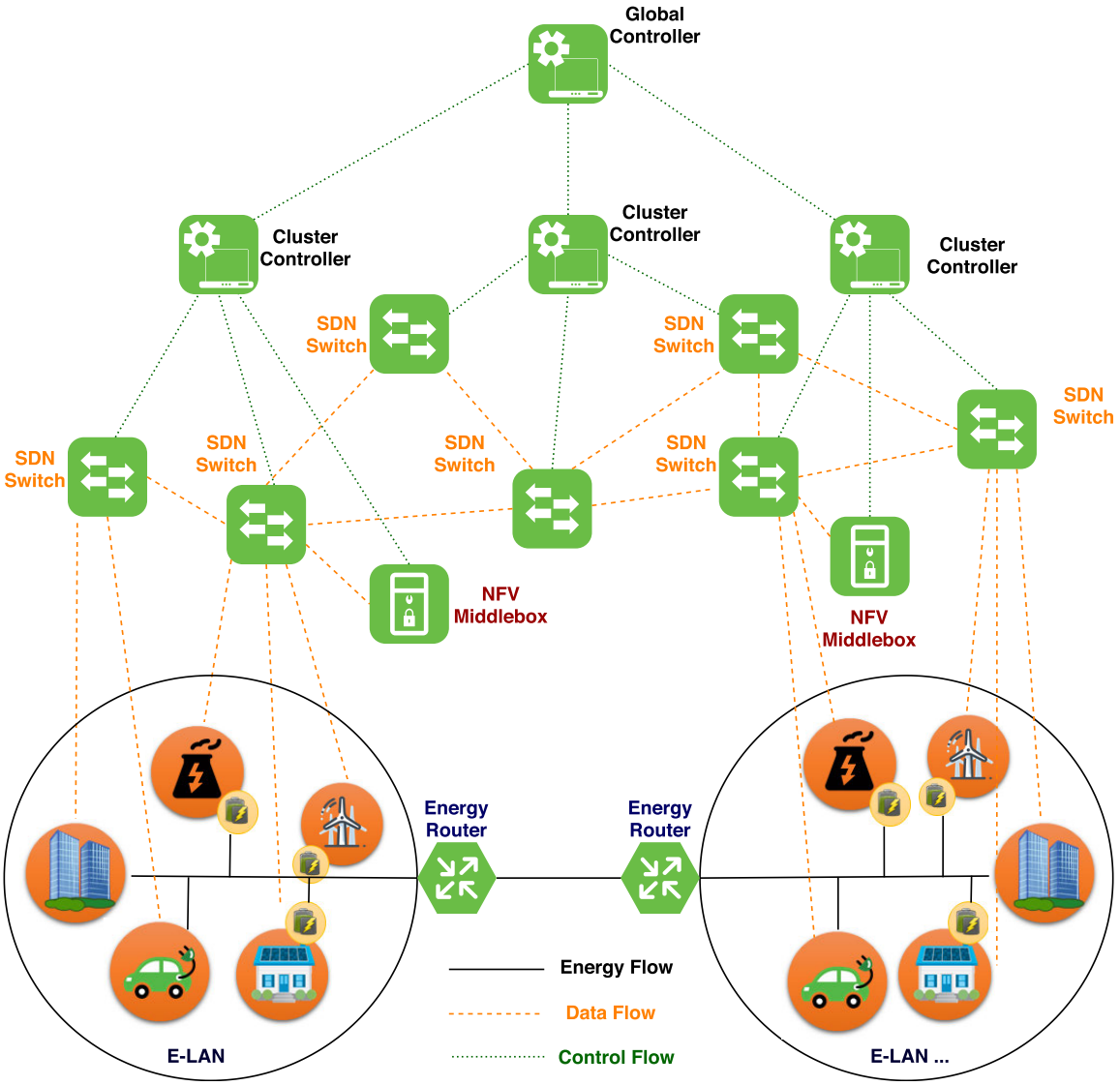

FIGURE 3. Illustration of SDN/NFV-based EI cyberinfrastructure. Both data and energy are flowing inside or inter E-LAN networks. NFV middleboxes are utilized to provide sets of NIS applications, fulfilling the cyber-security requirements in the EI ecosystem. Then, the SDN controllers (cluster and global) managed all flows using AI-driven policy automatically.

As depicted in Fig. 3, SDN/NFV-based EI cyberinfrastructure consists of SDN controllers, SDN switches, and NFV middleboxes that are utilized to securely control and forward application data flows between users/applications in EI ecosystem. However, it should be noted that a latency-optimal NIS is an essential factor for the reliability of SDN/NFV-based EI cyberinfrastructure, as mentioned in the previous section. Hence, secure and low-latency communication are both required for reliable information flowing delivery. However, these objectives usually contradict each other. Therefore, a trade-off solution of the middleboxes deployment strategy to avoid the end-to-end flow latency over than requirement threshold is indispensable.

The next section will be detailed describe the system model and the problem formulation for latency-optimal NIS in SDN/NFV-based EI cyberinfrastructure. Moreover, the detailed comparison of recent state-of-the-art solutions for latency minimization is also be provided.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

Let the SDN/NFV-based EI cyberinfrastructure is represented as a simple directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, v_i, \ldots, v_{N_v}\}$ is the set of nodes and $\mathcal{E} = \{v_i, v_j\}$ is the set of links, where $i = \{1, 2, 3, \ldots, N_v\}$ and $j = \{1, 2, 3, \ldots, N_v\}$ are the subscripts of the node couple and N_v is total number of the node. Let denotes the maximum number of rules can be stored in an SDN switch $\mathcal{S} = \{s_1, s_2, s_l, \ldots, s_{N_s}\} \in \mathcal{V}$ is P_s, thus the number of rules that are currently stored in a switch flow tables is denoted as $p_s \in P_s$. If the set of

all NIS is denoted as $\mathcal{C}_b = \{c_{b_1}, c_{b_2}, c_{b_n}, \ldots, c_{b_{N_{B_c}}}\}$, then an NFV-based middlebox which supplies those services is $\mathcal{B}_c \in \mathcal{V}$. There may be N_q a number middleboxes available in the network, thus let us denote N_{B_c} as the number of NIS middleboxes where $\mathcal{B}_c \in \mathcal{Q}$. Each middlebox has a maximum processing power capacity O_b to perform a set of NIS. This processing power capacity depends on the available central processing unit (CPU) in each middlebox, which is represented in Mbps unit. Table 3 depicts example of resource allocation for NIS.

TABLE 3. Example of resource allocation for NIS [24].

Services	*CPU Required*	*Processing Capacity*
Firewall	4	900 Mbps
Proxy	4	900 Mbps
IDS	4	600 Mbps

Following the SAM as described in the previous section, an application data flow f can be described as $f_n = \{source_n, dest_n, c_{b_n}, o_n, t_n\}$. $source_n$, and $dest_n$ are the source and the destination nodes, respectively. c_{b_n} is the must be visited NIS of a flow's network traffic from source to destination. Furthermore, o_n is the amount of middlebox processing power capacity occupies by a NIS, and t_n is the daily clock periods of requested NIS by a flow. With the knowledge of all information in advance, we can generate $F_{\mathcal{B}_c}$, a set of flows that requires NIS from a middlebox.

B. PROBLEM FORMULATION

Let us define end-to-end flow latency, as

$$D_f^{tot} = \sum_{\forall o_n} \sum_{\forall v_{ij} \in \mathcal{V}} d_{if,bf} + \sum_{\forall o_n} \sum_{\forall v_{ij} \in \mathcal{V}} d_{bf,ef}, \tag{1}$$

where, $d_{if,bf}$ is the aggregate latency from the source ingress-switch to the corresponding NIS middlebox and $d_{bf,ef}$ is the aggregate latency from the corresponding NIS middlebox to the destination egress-switch. The aggregate latency depends on the service processing delay $\alpha = d_{c_{b_n},o_n}$ for each NIS n, and the packet delivery time $\beta = d_{v_i,v_j}$ in each link between two nodes. In a bit more details, the services processing delay and the packet delivery time estimation are described in [40], expressed as

$$d_{c_{b_n},o_n} = \frac{M_{c_{b_n}} * o_n}{O_{b_n}}, \tag{2}$$

$$d_{v_i,v_j} = \frac{Z_{max}}{B_r} + \frac{X_{v_i,v_j}}{L_s}, \tag{3}$$

where M is the number of application data flows which request NIS. For the delay-sensitive application, the satisfaction rate follows the sigmoid utility function, as depicted in Fig. 4. Thus, a precise resource allocation strategy is an avoidable task to increase QoS. Furthermore, Z_{max} is the maximum packet size in bit, B_r is the transmission bit rate in bit/s, X_{v_i,v_j} is the distance or the length of transmission medium in meter, and L_s is the propagation speed in the medium in m/s.

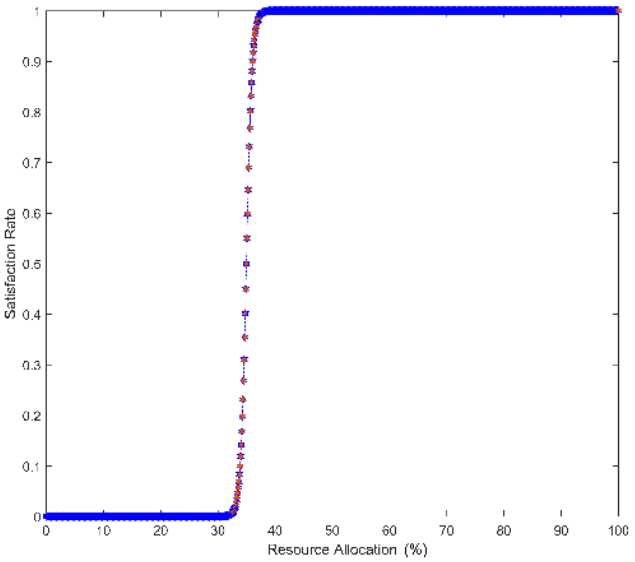

FIGURE 4. The satisfaction rate of each NIS, it should be noted that the percentage of resource allocation affects service processing delay that follows the sigmoid utility function.

To the best of our knowledge, the propagation speed depends on the physical medium of the link, e.g., 2×10^8 m/s for copper wires and 3×10^8 m/s for wireless communication.

It is worth to be mentioned that some existing works have proposed flow routing schemes to manage data flows in the SDN/NFV-based cyberinfrastructure. Hence, a constrained shorted path has been formulated in [41] as

$$r^* = \arg\min_r \{f_C(r) | r \in R_{st}, D(r) \leq D_{\max}\}, \tag{4}$$

that is, finding a forwarding route r from a set of all routes R_{st} that minimizes the objective function $f_C(r)$ such that the delay $D(r)$ to be less than or equal to the threshold value D_{max}. Furthermore, the constraints could be varied, ranging from traffic-chaining ratio, bandwidth consumption, deployment cost, energy consumption, and so on [42]. However, no matter what flow routing scheme is used, the middleboxes deployment provides the most significant effect on network latency. Hence, we need to develop a proper deployment strategy, which minimizes the total latency of each flow $f \in \mathcal{F}_{\mathcal{B}_c}$.

C. EXISTING SOLUTIONS

Some existing approaches have been proposed to deploy NIS middleboxes in SDN/NFV-based cyberinfrastructure with minimum latency. Moreover, Liu *et al.* [22] formulates the latency minimization function as

$$\min \quad D_f^{tot}, \tag{5}$$

$$s.t. \quad x_{n,l} = 1, \quad \forall q_n \in \mathcal{Q}, \ \forall s_l \in \mathcal{S}, \tag{6}$$

$$\sum_{\forall q_n \in \mathcal{Q}} R(q_n) x_{n,l} \leq C(s_l), \quad \forall s_l \in \mathcal{S}, \tag{7}$$

$$x_{n,l} = e, \quad \forall q_n(e) \in \mathcal{Q}, \ \forall s_l(e) \in \mathcal{S}, \tag{8}$$

where $x_{n,l}, n = \{1, 2, 3, \ldots, N_q\}, l = \{1, 2, 3, \ldots, N_s\}$ is the binary variables to represent middlebox placement scheme in

a switch s_l within the set of switch $\mathcal{S}$, with N_q and N_s being the total number of middleboxes and switches, also n and l being their subscripts, respectively. Furthermore, $R(q_n)$ is the required resource to deploy NIS middlebox q_n inside the set of $\mathcal{Q}$ and $C(s_l)$ is resource capacity of each switch inside $\mathcal{S}$.

To provide latency minimization, three constraints are considered, i.e., constraints (6) - (8). The constraint in (6) is to guarantee that each middlebox should be successfully deployed at a location, where $x_{n,l} = 1$ denotes that middlebox q_n is connected to switch s_l, otherwise $x_{n,l} = 0$. Furthermore, the constraint in (7) is to guarantee that the total resource demand for NIS deployment at one location should not exceed the switch resource capacity. Next, the constraint in (8) is to accommodate for middleboxes that can only be deployed in certain places. It is considered that a middlebox may require a power supply and acceleration by some dedicated platforms, which are available only at some locations.

On the other hand, to explain the type m middlebox, Vu and Kim [26] formulates the objective function for latency minimization as

$$\min\ D_f^{tot}, \quad \forall f \in F_m, \tag{9}$$

$$s.t.\ \ r(s_l) \leq R(s_l), \quad \forall s_l \in \mathcal{S}, \tag{10}$$

$$\sum_{f \in F_m} o_f \leq \sum_{n=1}^{N_{q,m}} O_{m_n}. \tag{11}$$

f is a flow from a set of data flows F_m that requires NIS type m from source ingress-switch to destination egress-switch via corresponding middleboxes. Furthermore, $N_{q,m}$ is the number of NIS middleboxes type m in the network, o is the requested processing power, and O is the maximum processing power capacity. In this context, we have two constraints, i.e., constraints (10) and (11). Constraint (10) is the switch memory resource, which is utilized to confirm that a switch has available memory for storing new route table entries. Constraint (11) is the middlebox processing power capacity, to ensure that a corresponding middlebox has enough processing power capacity to process the NIS requested by application data flows.

Reference [26] solves the latency minimization problem with two intuitive properties. The first property is derived from [22], that it is better to deploy the middlebox as close as possible to the most-usage switches. Next, the second property is its own intuitive belief. It may better to divide network such that data flow with a set of ingress-switches are close to each other to share the same middlebox in a cluster. Therefore, they used the BSAS-based clustering algorithm as their proposed solution.

In order to determine a threshold for each cluster, the packet delivery time data between each pair of ingress-switches is utilized. However, this approach has two main drawbacks, i.e., 1) one time and randomly choosing the clusters' threshold, and 2) the arbitrarily initial center selection, not being able to guarantee the end-to-end latency to be shortened. Similarly, the MST-based clustering algorithm, as described in [27], is also very dependent on the proximity threshold, which utilized to remove network edges from the MST cluster, whose lengths are greater than the threshold value. Hence, several successive values are required to be generated [28]. However, this solution needs to run a clustering algorithm many times, which requires high resource and time-consuming to find the best-considered threshold. Hence, graph cluster analysis using a threshold method should be avoided, and a more proper approach is required.

To be more details, Table 4 presents a summary of our investigation on the existing middlebox deployment strategies to support latency-optimal NIS, ranging from probabilistic search-based to graph cluster analysis-based methods. Taking advantage of this comparative analysis, we reformulate the objective function, constraints, and considered topologies for the context of EI.

IV. PROPOSED STRATEGY

Combining both objective functions and constraints in [22], [26], we reformulate the latency minimization problem as follows, as

$$\min\ D_f^{tot} \quad \forall f \in \mathcal{F}_{\mathcal{B}_c} \text{ according to } (1)-(3), \tag{12}$$

$$s.t. \quad x_{N,i} = 1, \quad \forall B_c \in \mathcal{V}, \tag{13}$$

$$\sum_{f_i \in F_{\mathcal{B}_c}} o_n f_i \leq \sum_{j=1}^{N_{B_c}} O_{b_j}, \quad \forall \mathcal{B}_c \in \mathcal{V}, \tag{14}$$

$$p_s \leq P_s, \quad \forall \mathcal{S} \in \mathcal{V}. \tag{15}$$

In this problem, we have three constraints, and those are the constraint (13) - (15). Constraint (13), $x_{N,i} = 1$, otherwise $= 0$, is to guarantee that each middlebox should be successfully connected to any SDN switch in the network. Furthermore, constraint (14) is the middlebox processing power to ensure that a corresponding middlebox having the capacity to process the NIS requested by application data flows. Next, the constraint (15) is the SDN switch memory capacity, which utilized to confirm that a switch has the available resources for storing new rule table entries.

To solve the latency minimization problem described above, we develop the AI-powered strategy, as depicted in Fig. 5. This solution consisted of two phases, i.e., the graph cluster analysis for middleboxes placement and the dynamic resource allocation based on the prediction of NIS usage-ratio in each corresponding cluster.

A. GRAPH CLUSTERING-BASED PLACEMENT

Let a cluster K consists of the SDN ingress-switches of corresponding data flows, that is $K = (s_1, s_2, \ldots, s_f)$, where s_f is the ingress-switch of a flow f. If $\mathcal{S}_{B_c} \in \mathcal{S}$ is the set of ingress-switches of corresponding data flows in F_{B_c}, then to determine the packet delivery time between each pair of ingress-switches, we can calculate the shortest path (SP) delay time between them, as

$$d_{s_i,s_j} = d_{SP(s_i,s_j)}, \quad \text{for } \forall\ s_i,\ s_j \in \mathcal{S}_{B_c}. \tag{16}$$

TABLE 4. Comparative analysis of recent strategies for NIS middleboxes deployment in SDN/NFV-based cyberinfrastructure.

Authors	Environment/Topology	Proposed Solution	Drawbacks
J. Liu *et al.* [22]	Three different topologies, i.e., Abilene, FatTree, Campus Network.	Probabilistic placement framework using the simulated annealing algorithm.	The proposed methods have several drawbacks as follows: 1) high time complexity because the need for working on all switches many times, 2) only focuses on the homogeneous type of middlebox, 3) the random search-based placement did not always guarantee that latency to be minimized, and 4) static resource allocation for each NIS.
T. D. Vu *et al.* [26]	FatTree and Abilene network topologies are utilized to represent both data center and irregular structure in CSP networks.	BSAS-based graph clustering algorithm to minimize latency.	Even though the proposed method has a lower time complexity than [22], however, it has two main drawbacks, i.e., 1) arbitrarily choosing the clusters' threshold and 2) randomly selecting the clusters' initial center, which did not guarantee that the end-to-end flow latency is always be minimized.
Y. T. Woldeyohannes *et al.* [27]	A practical internet service provider (ISP) network topology from Rocketfuel autonomous system number (ASN) 1221.	Graph clustering using the Kruskal's-based MST algorithm.	Similar to the BSAS-based clustering approach [26], arbitrarily choosing the clusters' proximity did not guarantee that the end-to-end flow latency is always be minimized. Moreover, SDN switch memory capacity to store new route table entries is not considered yet.
Ardiansyah *et al.* [28]	Abilene and FatTee network topologies are implemented to represent hierarchical and irregular communication network structure in the EI ecosystem.	Modified BSAS-based clustering algorithm, in which the clusters' threshold is determined using several successive values.	Although it could guarantee a better latency minimization than [26], however, the time-complexity to find the best threshold will increase following the number of available solutions. Moreover, more work on the middlebox resource allocation method is indispensable.
L. Qu *et al.* [43]	A 10-nodes network scheme represents a cloud data center topology.	An incremental approach is proposed to determine the number of required backups for low-latency as well as reliable NIS.	The proposed method did not consider both the capacity of SDN switches memory and the NIS middleboxes processing power. Also, there is a need to implement NIS protection method from failure due to the overload request.
This paper	Following the baseline methods [22], [26], [28], the FatTree and Abilene topologies are used again to represent both layered and irregular communication network structures of EI cyberinfrastructure.	AI-powered solution for latency minimization consists of two steps as follows. First, the middlebox placement strategy using the optimized K-means clustering algorithm. Second, the dynamic resource allocation scheme using the prediction of NIS usage-ratio in each cluster.	Need to consider the energy-saving scenario and more complex network topologies in the future.

Note that the graph cluster analysis using a threshold method should be avoided due to several reasons explained in the previous section, then a more proper solution is unavoidable. In this context, we employ another popular clustering technique, K-means clustering algorithm [44], with some modification for considering several additional conditions as follows

1) Since the objective is to find NIS middleboxes placement with minimum latency, the cluster center initialization method plays a critical role. Therefore, the initialization with careful seeding selection procedure is indispensable.
2) Moreover, the recalculated cluster center should be selected from the SDN ingress-switch. Then, the cluster refinement procedure is performed to re-assign all SDN switches to the appropriate cluster.
3) Also, the distance calculation method needs to accommodate the nodes with an indirect connection or may not physically be connected to the cluster center.

Denote $\mathcal{C} = \{C_1, C_2, C_k, \ldots, C_K\}$ as the set of clusters and let $\mathcal{M} = \{\mu_1, \mu_2, \mu_k, \ldots, \mu_K\}$ is the nearest mean of each cluster, the default K-means algorithm usually uses to partition n_p observations into $K(\leq n_p)$ clusters in which each observations belong to the cluster C_k with the nearest mean μ_k, expressed as the objective function $J(\mathcal{C}, \mathcal{M})$ [44], as

$$J(\mathcal{C}, \mathcal{M}) = \arg\min_{C} \sum_{k=1}^{K} \sum_{n_p \in C_k} ||n_p - \mu_k||^2. \tag{17}$$

Since the results of partitioning in K-means-based clustering is following the Voronoi cells, the Euclidean or Manhattan distance is employed to measure silhouette value for validating similarity and dissimilarity of each point to its own cluster and other clusters [45], as

$$sv(n_d) = \frac{b(n_d) - a(n_d)}{max(a(n_d), b(n_d))}, \tag{18}$$

where, $a(n_d)$ is the average distance from n_d-th point to other points within the same cluster, $b(n_d)$ is the minimum of all average distance from the n_d-th point to the points in each k-th cluster. Let the $sv(n_d)$ range is from -1 to 1. If the $sv(n_d)$ is close to 1, it indicates that the corresponding i-th point lies well with the cluster it belongs.

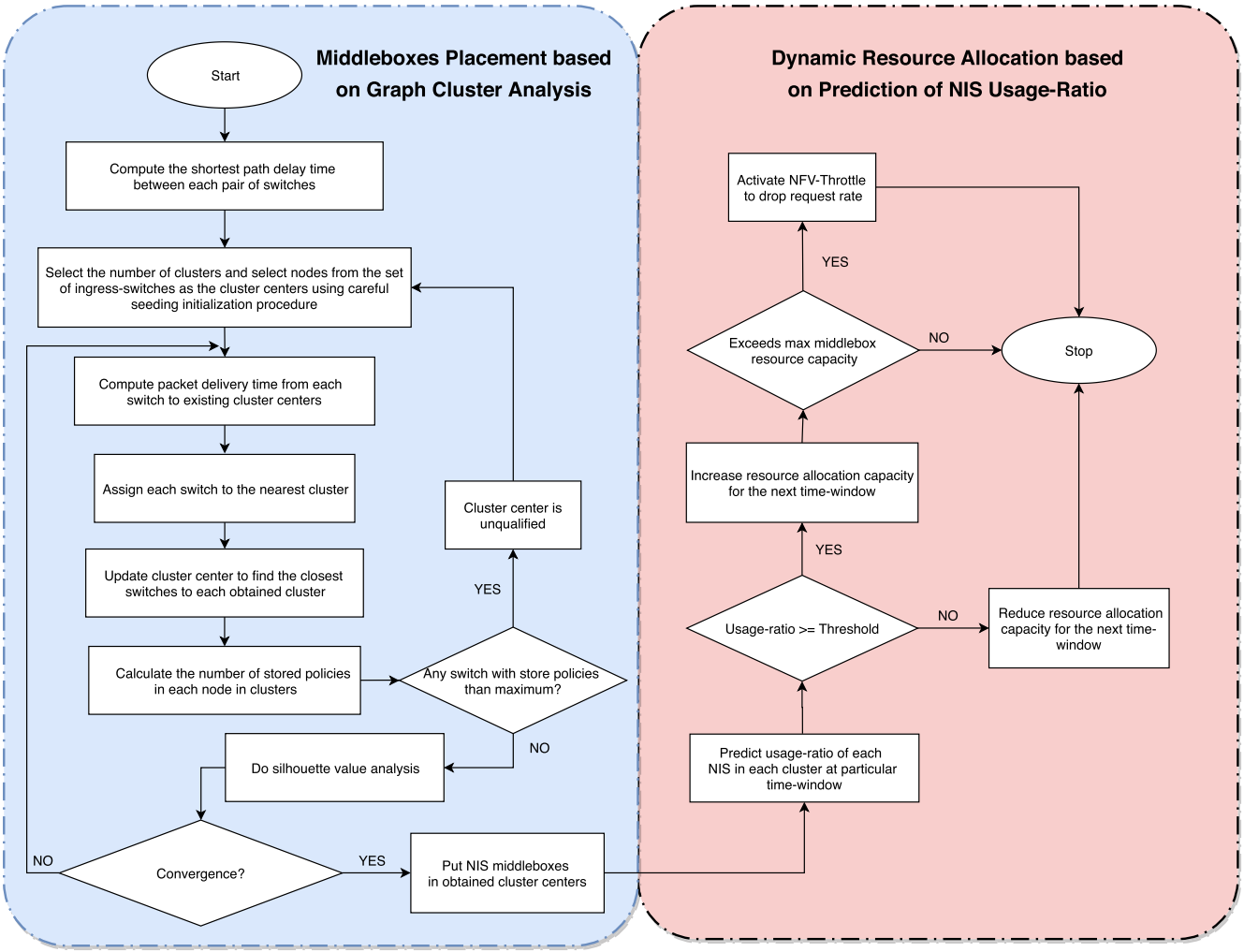

FIGURE 5. Flowchart of proposed strategy for latency-optimal NIS in SDN/NFV-based EI cyberinfrastructure. The graph cluster analysis is performed at the first step to find optimal NIS middleboxes placement in a number of clusters. Then, the second step employs NIS usage-ratio prediction in the corresponding clusters to find optimal resource allocation for each service.

As depicted in Algorithm 1, at first, we collect the SP computation between each pair of flow's ingress-switches. Then, we initialize clusters using a careful seeding initialization procedure, as described in [46]. Furthermore, almost similar to the K-medoids clustering method [47], the proposed graph clustering-based middlebox placement uses the selected node, in which the SDN switch is used as the cluster center instead of using the nearest mean, expressed as the objective function $J(K, Me)$, as

$$J(K, Me) = \arg\min_{K} \sum_{i=0}^{k} \sum_{s_f \in K_i} ||s_f - K_c(i)||^2 \qquad (19)$$

where $Me = \{K_c(0), K_c(1), \ldots, K_c(i)\}$.

The center of each cluster is then updated and validated to minimize the sum of SP delay time to reach all switches in the optimal number of clusters. However, to satisfy the constraint in (15), we need to check and calculate the number of stored policies in each SDN switch. Repeat the steps until it is partitioned into optimal K sub-networks. Then, finally, put NIS middleboxes in each cluster center.

B. DYNAMIC RESOURCE ALLOCATION

After all NIS middleboxes are placed in the optimal position, we then allocate resources for each NIS dynamically. In this context, the resource allocation for each service at a particular time depends on the ratio of those services repeatedly requested by applications/users in a corresponding cluster. Taking advantage of the historical data of application data flows as inputs, the usage-ratio for the next time windows is predicted using the regression trees algorithm as described in [48].

In fact, NIS with the predicted usage-ratio higher than a particular threshold θ_u is subject to be considered as one of important service, similar to

$$UR_{ij} = \frac{n \mid \sum_{t=1}^{T} P_{ijn}(t) \geq \theta_u}{N_t}, \qquad (20)$$

Algorithm 1 Graph Cluster Analysis Using Optimized K-Means Algorithm for NIS Middleboxes Placement.

Input: $G = (S, E)$, F_{B_c}, $S_{B_c} \in S$
Output: NIS middleboxes placement in K clusters

1: **Step 1**: Compute $d_{SP(s_i,s_j)}$ between each pair of switches in G.
2: **Step 2**: Select the number of clusters K and select nodes from $S_{B_c} \in S$ as the initial center of each cluster K_c using careful seeding initialization procedure as in [46].
3: **Step 3**: Compute packet delivery time from each node s_f to existing cluster centers as $d_{s_f,K_c(i)}$. Then, assign each node to the closest cluster.
4: **Step 4**: Update cluster centers $K_{c'}$ to find the closest switch to each obtained cluster, where the sum of shortest path delay time to reach all ingress-switches in a cluster is minimized as expressed in (19).
5: **Step 5**: Calculate the number of stored policies p_s for each node $s_f \in K$
If $\exists s_f \in K$ *such that* $p_s \geq P_s$ **then**
K_c' is unqualified, exclude it then back to Step 2
6: **Step 6**: Validate similarity and dissimilarity of each node to its own cluster and other clusters using silhouette value measurement as depicted in (18) to define the optimal number of clusters.
7: **Step 7**: Repeat Steps 3-6 until it is partitioned into optimal K sub-networks.
8: **Step 8**: Put NIS middleboxes in each cluster centers.

Definitions (re-described for better readability): G, Network graph; S, Set of switches; E, Set of communication links; F_{B_c}, Set of corresponding flows; S_{B_c}, Set of ingress-switches; d_{SP}, The shortest path delay time; K, Number of clusters; K_c, Cluster center; s_f, Ingress-switch of a flow; p_s, Stored rules in a particular switch flow table; P_s, Maximum number of rules capacity in a switch.

where i, j, n, and t are re-used in the remain equations to indicate the targeted cluster, index of NIS application, and the number of time-windows t that services have been operated, respectively. P_{ijn} is the amount of middlebox processing power occupied by each NIS application in previous time-windows from N_t total number of observed time-windows.

Then, the predicted usage-ratio is normalized to the range of UR_{ij} [−1:1]. Using this information, we define the resource allocation capacity of each service for the next time-window is as

$$RA_{ij} = Resv_{ij} + (Resv_{ij} * UR_{ij}), \tag{21}$$

where $Resv_{ij}$ is the percentage of guaranteed CPU allocation for a service at the previous time window. Hence, unlike the existing solution in [43], which the number of NIS is adjusted based on the incremental approach. In our approach, the NIS with a higher predicted usage-ratio obtain higher resource allocation in the next period and vice versa accordingly. Fig. 6 shows an example of the CPU resource allocation for a set of NIS at a particular time-window. Furthermore, to protect NIS from failures due to excess and un-predicted requests, we employ NFV-Throttle procedure [49]. When the volume of the demand exceeds the resource allocation capacity, we evaluate the fraction of the request to drop as

$$\text{drop\_rate} = 100 \cdot \left(1 - \frac{RA_{ij}}{\text{incoming\_request}}\right), \tag{22}$$

if incoming_request $\geq$ max_capacity; otherwise, drop_rate = 0.

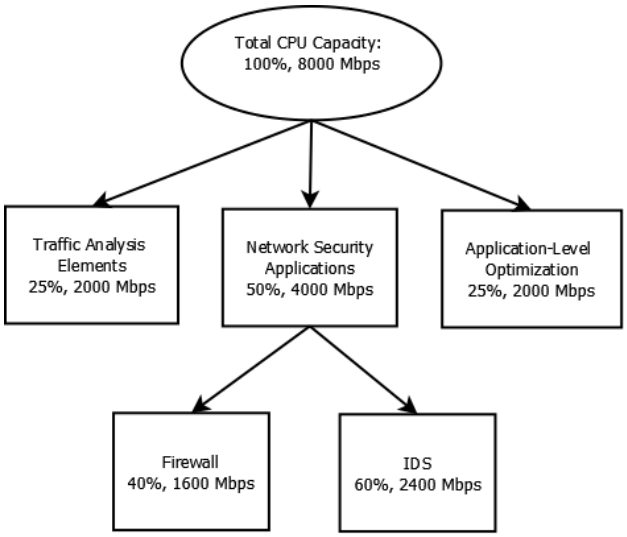

FIGURE 6. An example of dynamic resource allocation based on the prediction of usage-ratio, in which if the NIS application in the corresponding middlebox is considered as an important service, it will get a higher allocated CPU resource at a particular time-window.

TABLE 5. The characteristic of network topologies.

Topology	*FatTree*	*Abilene*
Number of Switches	20	11
Number of Edges	32	14

V. EVALUATION

A. SIMULATION TESTBED

We implement a testbed based on the NFV infrastructure emulation platform (NIEP) [50] in two machines, and each device has 3.40 GHz eight-core CPUs and 8192 MB RAM. In more detail, NIEP utilizes the Mininet [51] and the Click-on-OSv [52] to provide a complete simulation of SDN/NFV-based cyberinfrastructure. Furthermore, we decide to use two network topologies, i.e., Abilene and FatTree, which are explained in the previous section to represent two possible architectures of SDN/NFV-based EI cyberinfrastructure. The characteristics of these network topologies are summarized in Table 5.

For the graph cluster analysis, we set our testbed with several assumptions as follows. First, the communication

TABLE 6. Parameter setting for the performance evaluation.

Parameter	*Value/Range*
Number of network intelligence services	5
Transmission bit rate (Mbps)	100
Propagation speed (m/s)	from 0.6 to 0.9 * 2 x 10^8
Distance between two switches (m)	from 40 to 100
Number of NFV middleboxes	5
Number of flows in each set	30, 35, 40, 45, 50
Processing demand of each flow (Mbps)	From 0.1 to 1
Max number of rules in each switch	25
Max packet size (Bytes)	1500
Usage-ratio threshold (%)	40 - 60

medium between two nodes is a copper-wires with randomly assigned losses following the normal distribution. Second, the transmission bit rate in each switch-port is 100 Mbps, but the distances between the two switches are randomly different. Third, there are five NFV-based middleboxes in each simulation, and those middleboxes could provide five kinds of NIS. Next, there are also five sets of application data flows that randomly request specific services. Last, the maximum number of clusters follows the total number of middleboxes.

Moreover, to evaluate the effects of the resource allocation strategy, we develop the following scenario. The set of application data flows, starting from 30 flows and increase one by one until 70 flows are generated 100 times, respectively. Furthermore, we record the type and the number of requested resources from each generated set. Then, 70% of recorded data are used as training data to develop the usage-ratio prediction model. Using the rest of the recorded data, the usage-ratio is predicted, which can be utilized further to reallocate the CPU resource for each service dynamically. The summary of the parameters setting is described in Table 6.

B. LATENCY MINIMIZATION ANALYSIS

For the first evaluation, we compare the result of our cluster center initialization strategy with the original method of the K-means algorithm. Fig. 7 depicts the sum of SP delay time for various initialization numbers of both approaches. The result shows that the cluster center initialization with careful seeding always guarantees that the packet delivery time between the cluster center and other nodes in the sub-network to be shortened.

Furthermore, we evaluate the end-to-end latency of each application data flow in the network. As depicted in Fig. 8, we simulate experimental comparison between our proposed strategy with recent state-of-the-art solutions for latency minimization problem, those are, 1) the simulated annealing-based method [22], 2) the BSAS-based scheme [26], 3) the MST-based procedure [27], 4) the modified BSAS-based approach [28], and 5) the original K-means algorithm. Moreover, we run all the threshold-based graph clustering methods twice, with and without the careful threshold selection procedure.

The simulation result shows that on both network topologies, our strategy using the optimized K-means clustering

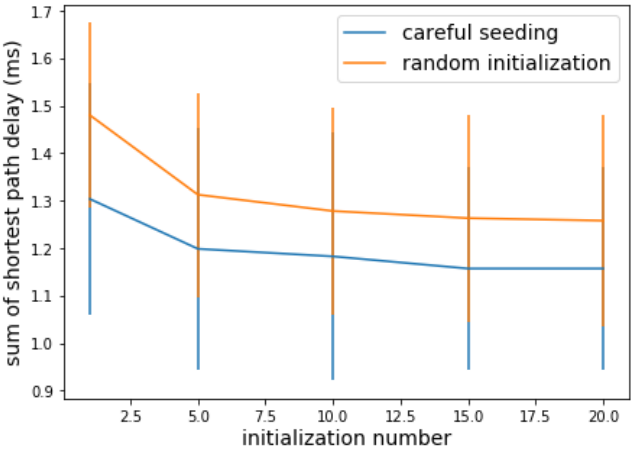

FIGURE 7. Cluster center initialization with careful seeding always provides a lower of the sum of SP delay time result compared to the arbitrarily (random) method of the default K-means algorithm.

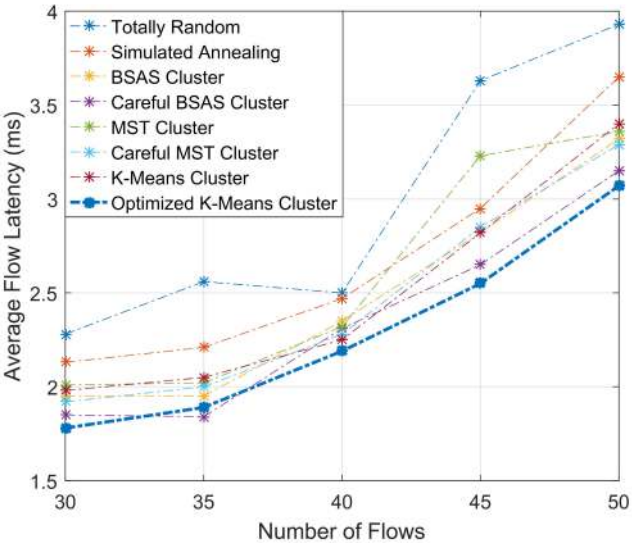

FIGURE 8. The average latency of application data flows in both FatTree and Abilene network topologies. It is shown that our middlebox placement method based on the optimized K-means clustering algorithm is outperform the recent state-of-the-art approaches, ranging from the simulated annealing-based method to the other graph clustering-based procedures, i.e., the BSAS, the MST, and the original K-means algorithms.

algorithm provides an enormous impact on reducing end-to-end flow latency. The average end-to-end flow latency is around 25.22% and 23.81% lower compared with the totally random and the simulated annealing placement methods, respectively. Moreover, the average latency minimization is also improved around 18.44% and 11.49% are compared to other graph clustering-based placement approaches, the BSAS-based scheme, and the MST-based procedure, respectively. Hence, these results prove the intuitive properties and considerations, as described in the previous section. First, with the knowledge of application data services in advance, it is better to put the NIS middleboxes as near as possible to the data flow ingress-switches and divide

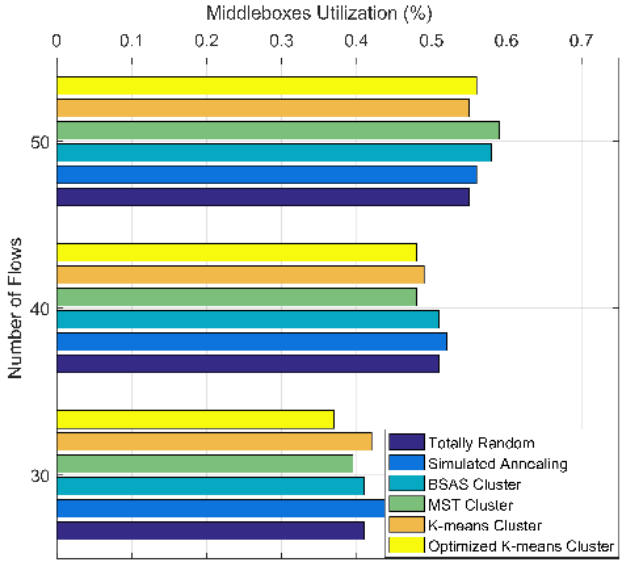

FIGURE 9. The comparison of average NIS middleboxes resource utilization from all placement strategies.

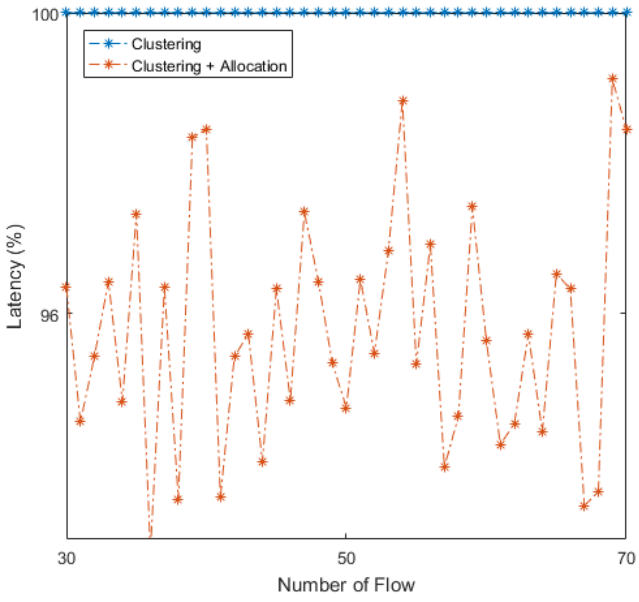

FIGURE 10. The dynamic resource allocation scheme optimize further the latency minimization of graph clustering-based NIS middlebox placement.

them into several clusters. Second, the graph clustering-based placement using an arbitrarily clusters' threshold assignment should be avoided. Third, the cluster center initialization method plays a critical role, in which careful seeding initialization procedure proposed in this paper is essential to find all NIS middleboxes placement with minimum latency in SDN/NFV-based EI cyberinfrastructure.

Next, Fig. 9 shows the average of NIS middleboxes resource utilization from all placement strategies. Based on these results, we conclude that the utilization increases along with the increasing number of application data flow. It should be noted that the middlebox processing power capacity depends on the available CPU in each middlebox. Hence, we apply the proposed dynamic resource allocation mechanism for 1) minimizing the NIS processing delay considering the middleboxes processing capacity constraint in (14), and 2) avoiding the functionality failures due to overload usage of reliability-aware implementation as mentioned in the previous section.

Figure 10 depicts the effects of the proposed dynamic resource allocation scheme in latency minimization. It is shown that the average end-to-end latency decreases compared to just merely the graph cluster analysis approach. Moreover, compared with the incremental resource allocation approach in [43], our dynamic resource allocation scheme optimizes further latency minimization around 4.24%. The best improvement comes from the fifty percent usage-ratio threshold configuration. However, it should be noted that the prediction error of usage-ratio is still around 17.13%. Hence, a better setting of the regression tree algorithm is needed to be implemented to improve the performance. To this end, enhancement using an ensembling method with other AI-driven predictive algorithms or performing a deep learning analysis can be applied in the future.

C. CHALLENGES AND DISCUSSIONS

The latency minimization analysis demonstrates in the previous subsection prove that our proposed strategy can be utilized to provide the latency-optimal NIS in SDN/NFV-based EI cyberinfrastructure. Furthermore, Table 7 depicts the detailed comparison of our contribution to the existing works in the building of cyberinfrastructure for end-to-end interaction across the entire value chain in the EI using the SDN approach. However, some challenges need to be handled in future research, such as:

1) Since there is no unique criterion to define the structure of EI cyberinfrastructure, the energy stakeholders may implement their own network topology or that suggested by the CSP. Therefore, it will be growing both in size and complexity. Hence, a loop-based topology analysis [40] may be needed to be adopted to provide reliable SDN/NFV-based EI cyberinfrastructure with a more dynamic network topology in the future.
2) To analyze application data flows in this work, a detailed SAM for NIS is required. However, to avoid information leaked by non-trusted parties, it would better to also be provided with a privacy-preserving data scheme, e.g., using the so-called differential privacy mechanism [53].
3) Even though the objective of our work is to minimize end-to-end latency. However, more targets, such as energy-saving scenarios, could be implemented in the future. Furthermore, effective resource management based on traffic demand, as depicted in [54], may also be adopted.

TABLE 7. The most recent work in developing SDN-based EI cyberinfrastructure.

Relevant works	Year	Achievements	Limitations
G. Zhang *et al.* [15]	2014	Preliminary research results on SDN-based EI cyber-infrastructure, especially for the IEM application case study.	High-latency in control network, many works need to perform to handle all communication requirements.
W. Zhong *et al.* [14]	2016	SDN-based EI cyberinfrastructure for the electric vehicle case study.	Initial design, thus the latency, security, and scalability requirements are left for the future.
Z. Lu *et al.* [16]	2017	A general framework for the SDN-based EI cyberinfrastructure in China, then basic comparison with the traditional internet protocol (IP) network is also provided.	Security issues for both control and data planes are not considered yet.
This paper	2019	SDN/NFV-based NIS utilization for fulfilling the main cyber-security requirements in the EI ecosystem. Then the AI-powered solution is also developed for latency-optimal implementation.	Energy-saving scenario and a more complex network topology operation are left for future work.

VI. CONCLUSION

In this paper, the utilization of SDN/NFV-based NIS for fulfilling the cyber-security requirements in the EI ecosystem has been introduced. Furthermore, the AI-powered solution has been proposed to deploy NIS with the minimum end-to-end flow latency. This solution consisted of two phases: 1) NIS middlebox placement based on the optimized K-means-based graph clustering analysis, 2) dynamic resource allocation using predicted NIS usage-ratio based on the regression tree analysis. Moreover, the evaluation results have verified that our proposed approach could improve latency minimization significantly in two network topologies, i.e., Abilene and FatTree. The average end-to-end latency is more than 15% lower compared to the state-of-the-art threshold-based clustering algorithm. This result proves our intuitive properties and considerations that the graph clustering-based placement using an arbitrarily clusters' threshold assignment should be avoided, and the cluster center initialization method plays a critical role. However, even though the main objective of this paper is minimizing flow latency, more targets such as energy-saving or more complex topology scenarios can be implemented in the future.

REFERENCES

[1] A. K. Shukla, K. Sudhakar, P. Baredar, and R. Mamat, "BIPV based sustainable building in South Asian countries," *Solar Energy*, vol. 170, pp. 1162–1170, Aug. 2018.

[2] S. Zhang, P. Andrews-Speed, and S. Li, "To what extent will China's ongoing electricity market reforms assist the integration of renewable energy?" *Energy Policy*, vol. 114, pp. 165–172, Mar. 2018.

[3] T. Chen, Q. Alsafasfeh, H. Pourbabak, and W. Su, "The next–generation U.S. Retail electricity market with customers and prosumers—A bibliographical survey," *Energies*, vol. 11, no. 1, p. 8, Dec. 2017.

[4] K. Wang, J. Yu, Y. Yu, Y. Qian, D. Zeng, S. Guo, Y. Xiang, and J. Wu, "A survey on energy internet: Architecture, approach, and emerging technologies," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2403–2416, Sep. 2018.

[5] L. Cheng, T. Yu, H. Jiang, S. Shi, Z. Tan, and Z. Zhang, "Energy internet access equipment integrating cyber–physical systems: Concepts, key technologies, system development, and application prospects," *IEEE Access*, vol. 7, pp. 23127–23148, 2019.

[6] Y. Xu, J. Zhang, W. Wang, A. Juneja, and S. Bhattacharya, "Energy router: Architectures and functionalities toward energy internet," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Oct. 2011, pp. 31–36.

[7] F. Li, R. Li, Z. Zhang, M. Dale, D. Tolley, and P. Ahokangas, "Big data analytics for flexible energy sharing: Accelerating a low–carbon future," *IEEE Power Energy Mag.*, vol. 16, no. 3, pp. 35–42, May 2018.

[8] M. E. Khodayar, "Rural electrification and expansion planning of off-grid microgrids," *Elect. J.*, vol. 30, no. 4, pp. 68–74, May 2017.

[9] J. Boyd, "An internet-inspired electricity grid," *IEEE Spectrum*, vol. 50, no. 1, pp. 12–14, Jan. 2013.

[10] *China National Energy Administration. Germany Builds Smart Energy Network Based On e-Energy Technology Innovation Promotion Program*. Accessed: Sep. 14, 2019. [Online]. Available: http://www.nea.gov.cn/2012-02/14/c_131409715.htm

[11] Z. Liu, *Global Energy Interconnection*. Beijing, China: China Electric Power Press, 2015.

[12] IO. Energy Ecosystem. *Founding Members' Vision: Consumer-at-the-Center*. Accessed: Sep. 20, 2019. [Online]. Available: https://www.ioenergy.eu/about/

[13] H.-P. Schwefel, Y.-J.-A. Zhang, C. Wietfeld, and H. Mohsenian-Rad, "Emerging technologies initiative 'smart grid communications': Information technology for smart utility grids," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Oct. 2018.

[14] W. Zhong, R. Yu, S. Xie, Y. Zhang, and D. H. K. Tsang, "Software defined networking for flexible and green energy internet," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 68–75, Dec. 2016.

[15] G. Zhang, L. Su, Y. Wang, X. Liu, and J. Li, "Research on communication network architecture of energy internet based on SDN," in *Proc. IEEE Workshop Adv. Res. Technol. Ind. Appl. (WARTIA)*, Sep. 2014.

[16] Z. Lu, C. Sun, J. Cheng, Y. Li, Y. Li, and X. Wen, "SDN–enabled communication network framework for energy internet," *J. Comput. Netw. Commun.*, vol. 2017, pp. 1–13, Jun. 2017, Art. no. 8213854.

[17] W. Zhang, J. Li, J. Zhou, and Z. Hu, "The requirement and the key technologies of communication network in internet of energy," in *Human Centered Computing* (Lecture Notes in Computer Science), vol. 9567. Cham, Switzerland: Springer, 2016, pp. 842–848.

[18] K. Wang, X. Hu, H. Li, P. Li, D. Zeng, and S. Guo, "A survey on energy internet communications for sustainability," *IEEE Trans. Sustain. Comput.*, vol. 2, no. 3, pp. 231–254, Jul. 2017.

[19] H. Wang, J. Ruan, Z. Ma, B. Zhou, X. Fu, and G. Cao, "Deep learning aided interval state prediction for improving cyber security in energy internet," *Energy*, vol. 174, pp. 1292–1304, May 2019.

[20] Y. Li and M. Chen, "Software–defined network function virtualization: A survey," *IEEE Access*, vol. 3, pp. 2542–2553, 2015.

[21] D. Griffith, M. Souryal, and N. Golmie, *Smart Grid Communications and Networking: Wireless Network for Smart Grid Application*. Cambridge, U.K.: Cambridge Univ. Press, 2012, pp. 234–262.

[22] J. Liu, Y. Li, Y. Zhang, L. Su, and D. Jin, "Improve service chaining performance with optimized middlebox placement," *IEEE Trans. Serv. Comput.*, vol. 10, no. 4, pp. 560–573, Jul. 2017.

[23] Y. Chen and J. Wu, "NFV middlebox placement with balanced set-up cost and bandwidth consumption," in *Proc. 47th Int. Conf. Parallel Process. (ICPP)*, 2018.

[24] M. A. Raayatpanah and T. Weise, "Virtual network function placement for service function chaining with minimum energy consumption," in *Proc. IEEE Int. Conf. Comput. Commun. Eng. Technol. (CCET)*, Aug. 2018.

[25] W. Ma, J. Beltran, Z. Pan, D. Pan, and N. Pissinou, "SDN–based traffic aware placement of NFV middleboxes," *IEEE Trans. Netw. Serv. Manage.*, vol. 14, no. 3, pp. 528–542, Sep. 2017.

[26] D. T. Vu and K. Kim, "Flow clustering based efficient consolidated middlebox positioning approach for SDN/NFV–enabled network," *IEICE Trans. Inf. Syst.*, vol. E99.D, no. 8, pp. 2177–2181, 2016.
[27] Y. T. Woldeyohannes, A. Mohammadkhan, K. K. Ramakrishnan, and Y. Jiang, "ClusPR: Balancing multiple objectives at scale for NFV resource allocation," *IEEE Trans. Netw. Serv. Manage.*, vol. 15, no. 4, pp. 1307–1321, Dec. 2018.
[28] Ardiansyah, Y. Choi, M. R. K. Aziz, and D. Choi, "Latency minimization for energy internet communications with SDN virtualization infrastructure," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Beijing, China, Oct. 2019. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8909690/authors#authors
[29] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, p. 63, Oct. 2008.
[30] (2015). *Abilene Core Topology*. [Online]. Available: https://itservices.stanford.du/service/network/internet2/abilene
[31] T. Sato, D. M. Kammen, B. Duan, M. Macuha, Z. Zhou, J. Wu, M. Tariq, and S. A. Asfaw, *Smart Grid Standards: Specifications, Requirements, and Technologies*. Singapore: Wiley, 2015.
[32] A. Sundararajan, T. Khan, A. Moghadasi, and A. I. Sarwat, "Survey on synchrophasor data quality and cybersecurity challenges, and evaluation of their interdependencies," *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 3, pp. 449–467, May 2019.
[33] D. N. Gde, Q. Nguyen-Van, T. V. Duc, N. Nguyen-Sinh, P. J. D. Alvin, K. Kim, and D. Choi, "Design of service abstraction model for enhancing network provision in future network," in *Proc. 18th Asia–Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Kanazawa, Japan, Oct. 2016, pp. 1–4.
[34] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescape, "Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 2, pp. 445–458, Jun. 2019.
[35] R. Boutaba, M. A. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar, F. Estrada-Solano, and O. M. Caicedo, "A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities," *J. internet Service Appl.*, vol. 9, no. 1, p. 16, Dec. 2018.
[36] R. I. T. Da Costa Filho, W. Lautenschlager, N. Kagami, M. C. Luizelli, V. Roesler, and L. P. Gaspary, "Scalable QoE-aware path selection in SDN-based mobile networks," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Honolulu, HI, USA, Apr. 2018.
[37] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1153–1176, 2nd Quart., 2016.
[38] G. Gardikis *et al.*, "SHIELD: A novel NFV-based cybersecurity framework," in *Proc. IEEE Conf. Netw. Softwarization (NetSoft)*, Bologna, Italy, Jul. 2017, pp. 3–7.
[39] K. Wang, Y. Zhang, S. Guo, M. Dong, R. Q. Hu, and L. He, "IEEE access special section editorial: The internet of energy: Architectures, cyber security, and applications," *IEEE Access*, vol. 6, pp. 79272–79275, 2018.
[40] A. Leal and J. F. Botero, "Defining a reliable network topology in software–defined power substations," *IEEE Access*, vol. 7, pp. 14323–14339, 2019.
[41] Z. Shu, J. Wan, J. Lin, S. Wang, D. Li, S. Rho, and C. Yang, "Traffic engineering in software-defined networking: Measurement and management," *IEEE Access*, vol. 4, pp. 3246–3256, 2016.
[42] M. Karakus and A. Durresi, "Quality of service (QoS) in software defined networking (SDN): A survey," *J. Netw. Comput. Appl.*, vol. 80, pp. 200–218, Feb. 2017.
[43] L. Qu, M. Khabbaz, and C. Assi, "Reliability–aware service chaining in carrier–grade softwarized networks," *IEEE J. Select. Areas Commun.*, vol. 36, no. 3, pp. 558–573, Mar. 2018.
[44] L. Galluccio, O. Michel, P. Comon, and A. O. Hero, "Graph based k-means clustering," *Signal Process.*, vol. 92, no. 9, pp. 1970–1984, Sep. 2012.
[45] R. C. De Amorim and C. Hennig, "Recovering the number of clusters in data sets with noise features using feature rescaling factors," *Inf. Sci.*, vol. 324, pp. 126–145, Dec. 2015.
[46] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. 18th Annu. ACM-SIAM Symp. Discrete Algorithms (SODA)*, New Orleans, LA, USA, 2007.
[47] E. Schubert and P. J. Rousseeuw, "Faster k-medoids clustering: Improving the PAM, CLARA, and CLARANS algorithms," May 2019, *arXiv:1810.05691*. [Online]. Available: https://arxiv.org/abs/1810.05691
[48] W.-Y. Loh, "Classification and regression trees," *Wires Data Mining Knowl. Discovery*, vol. 1, no. 1, pp. 14–23, Jan. 2011.
[49] D. Cotroneo, R. Natella, and S. Rosiello, "NFV–Throttle: An overload control framework for network function virtualization," *IEEE Trans. Netw. Service Manage.*, vol. 14, no. 4, pp. 949–963, Dec. 2017.
[50] T. N. Tavares, L. Da Cruz Marcuzzo, V. F. Garcia, G. V. De Souza, M. F. Franco, L. Bondan, F. De Turck, L. Z. Granville, E. P. Duarte, Jr., C. R. P. Dos Santos, and A. E. Schaeffer-Filho, "NIEP: NFV infrastructure emulation platform," in *Proc. IEEE 32nd Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, Cracow, Poland, May 2018.
[51] B. Lantz, B. Heller, and N. Mckeown, "A network in a laptop: Rapid prototyping for software-defined networks," in *Proc. 9th ACM SIGCOMM Workshop Hot Topics Netw. (Hotnets)*, Monterey, CA, USA, 2010.
[52] L. Da Cruz Marcuzzo, V. F. Garcia, V. Cunha, D. Corujo, J. P. Barraca, R. L. Aguiar, A. E. Schaeffer-Filho, L. Z. Granville, and C. R. P. Dos Santos, "Click-on-OSv: A platform for running click-based middleboxes," in *Proc. IFIP/IEEE Symp. Integr. Netw. Service Manage. (IM)*, Lisbon, Portugal, May 2017.
[53] Z. Lv, L. Wang, Z. Guan, J. Wu, X. Du, H. Zhao, and M. Guizani, "An optimizing and differentially private clustering algorithm for mixed data in SDN–based smart grid," *IEEE Access*, vol. 7, pp. 45773–45782, 2019.
[54] M. Paliwal and D. Shrimankar, "Effective resource management in SDN enabled data center network based on traffic demand," *IEEE Access*, vol. 7, pp. 69698–69706, 2019.

ARDIANSYAH received the bachelor's degree in computer engineering from Universitas Indonesia (UI), Indonesia, in 2010, and the master's degree in computer science from Chonnam National University (CNU), Gwangju, South Korea, in 2014, where he is currently pursuing the Ph.D. degree with the Graduate Program of Electronics and Computer Engineering.

He was a Graduate Research Assistant with the Advanced Network Laboratory (ANL), CNU, from 2012 to 2014, where he rejoined again, in 2017. He was also a Teaching Staff with the Department of Electrical and Computer Engineering, UI, from 2014 to 2016. His main research interests are related to the analysis, design, and optimization of wireless networked and future internet systems, applied artificial intelligence (AI), and the Internet of Things (IoT) technology for smart energy systems and informatics in built environments (grids, buildings, and cities). He is also a member of the IEEE Communication Society, the ACM Emerging Interest Group on Energy (EIG-Energy), and the IEEE Smart Grid Community.

YONGHOON CHOI (Senior Member, IEEE) received the bachelor's degree in electronics engineering from Sungkyunkwan University, Suwon, South Korea, in 1999, and the master's and Ph.D. degrees in information and communications engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2003 and 2010, respectively.

From 2010 to 2013, he was a Postdoctoral Visiting Scholar with the Department of Electrical Engineering, Stanford University, CA, USA. He was a Research Assistant Professor with the KAIST Institute of IT Convergence, Daejeon, from 2013 to 2014. Since 2014, he has been with the Power ICT Lab, Department of Electrical Engineering, Chonnam National University (CNU), Gwangju, South Korea, where he is currently an Associate Professor. His research interests include the design, analysis, and optimization of wireless/mobile communications networks, cognitive radio systems, network economics, and smart grid communications and networking, which is a convergence of Information and Communications Technology (ICT) and Power and Energy Technology (PET). He is a member of the Korean Institute of Communications and Information Sciences (KICS).

MUHAMMAD REZA KAHAR AZIZ (Member, IEEE) received the bachelor's degree and master's degree *(cum laude)* in electrical engineering (telecommunications) from the Institut Teknologi Bandung (ITB), Bandung, Indonesia, in 2004 and 2012, respectively, and the Ph.D. degree in information science from the Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan, in 2016.

Since 2012, he has been a Teaching Staff with the Institut Teknologi Sumatera (ITERA), Indonesia. He is currently serving as a Reviewer and a TPC Member for IEEE conferences, such as ICC, Globecom, VTC, and others, and several journals, including the IEEE TVT. His research interests include wireless communication systems, signal processing, information theory, radio geolocation, factor graph, antennas, and electromagnetic.

KANGWOOK CHO received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, South Korea, in 1986 and 1988, respectively.

He is currently a Team Leader with the Department of Market and System Development, Korea Power Exchange (KPX), Naju, South Korea. His research interests include power system planning and operations, such as regional load forecasting, optimization of generation and transmission investment, power system probabilistic assessment, and electricity market design.

DEOKJAI CHOI received the bachelor's degree in computer engineering from Seoul National University (SNU), the master's degree in computer science from the Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 1982 and 1984, respectively, and the Ph.D. degree in network management from the University of Missouri–Kansas City, USA, in 1995.

He has been with Chonnam National University (CNU), Gwangju, South Korea, since 1984, where he has also been a Professor with the Advanced Network Laboratory (ANL), Department of Electronics and Computer Engineering (ECE),since 1996. His main research interests include topics on context awareness, pervasive computing, sensor networks, smart grid, and the future Internet. He is currently serving as a Reviewer and a TPC Member for several conferences and journals, including APAN, KNOMS, APNOMS, MDPI Energies, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and so forth.