IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 21, 2018, accepted August 27, 2018, date of publication September 18, 2018, date of current version October 17, 2018.
Digital Object Identifier 10.1109/ACCESS.2018.2870889

A Graph-Based Handover Scheduling for Heterogeneous Vehicular Networks

MADE HARTA DWIJAKSARA[1], WHA SOOK JEON[1], (Senior Member, IEEE), AND DONG GEUN JEONG[2], (Senior Member, IEEE)
[1]Department of Computer Science and Engineering, Seoul National University, Seoul 08826, South Korea
[2]Department of Electronics Engineering and Applied Communication Research Center, Hankuk University of Foreign Studies, Yongin-si 17035, South Korea
Corresponding author: Wha Sook Jeon (wsjeon@snu.ac.kr)

This work was supported by the Hankuk University of Foreign Studies Research Fund and by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, and Future Planning under Grant NRF-2017R1A2B2003496.

ABSTRACT The vehicular networks exploiting WiFi have a great advantage in cost and capacity in comparison with the networks using cellular systems. However, to use WiFi in the vehicular environments, the problem of frequent handover (HO) due to the small coverage of access points (APs) should be resolved. In this paper, we propose a new HO technique which is effective in heterogeneous (WiFi/cellular) networks on a real road topology. Specifically, we propose an HO scheduling scheme running on a server, which determines the AP to which a user will be handed over, considering the road topology. Since a user only needs to decide when to initiate the connection to the next AP, a very fast and efficient HO in the vehicular environment can be realized. The design objective of the proposed HO scheduling is to maximize the connection time on WiFi while minimizing the total HO latency and reducing the number of users which contend for the channel within an AP. We show that this objective is well-accomplished with the proposed scheme, by using computer simulation.

INDEX TERMS Handover scheduling, vehicular networks, WiFi, cellular network, graph modeling.

I. INTRODUCTION
The mobile Internet access from moving vehicles has been increasing steadily. Nowadays, many people enjoy various services from a simple news-feed to high rate video streaming, inside moving vehicles [1]. Cellular network, because of its reliability and wide availability, has become a top priority wireless access system for providing such Internet service in vehicular environment. This has led to the drastic growth of data traffic in cellular network, causing a traffic-overload problem. Hence, cellular network operators have been driven to offload the traffic to other alternative technology systems [2]. On the other hand, most of the users usually have the limited access to cellular network due to high-cost issue. Unless the users register for an expensive unlimited data plan, they cannot freely access the Internet through cellular network. Thus, not only the cellular network operators but also the users need alternative technologies being much cheaper.

The IEEE 802.11 standard, also known as WiFi, has become a popular alternative technology of cellular network. Among the available WiFi versions, the IEEE 802.11p is particularly designed for vehicular communication. However, as an alternative of the cellular network, the IEEE 802.11p is still less attractive compared to the IEEE 802.11b/g/n/ac, even in the vehicular environment. First, the IEEE 802.11p network infrastructure is not widely available yet. Hence, to provide the Internet access using the IEEE 802.11p, a huge number of IEEE 802.11p roadside units should be initially installed, which requires a high cost. Second, the commercial end-user devices on the market do not typically support the IEEE 802.11p standard, which causes a compatibility issue. Third, even if it is available, the IEEE 802.11p is not likely able to support high-speed Internet access, due to its low transmission rates.

On the other hand, the IEEE 802.11b/g/n/ac has become a *de facto* global standard for wireless local area network (WLAN), and most of the recent mobile electronic devices are equipped with this WiFi chipset. Furthermore, its access points (APs) have been deployed rapidly, not only indoors but also outdoors so that the WiFi service area also covers the road [3]. Recently, such a deployment condition can be easily found in the campus area, company area, government complex, etc. In such environment, the users have been eager to utilize WiFi whenever possible as an alternative of cellular network. This trend is also supported by the fact that the default connection setting of two major mobile operating systems (i.e., Android and iOS) prefers WiFi to cellular network [4]. Based on these reasons, we focus on the IEEE 802.11b/g/n/ac-based WiFi as an alternative of cellular network in the vehicular environment.

However, when the users access Internet through WiFi in vehicular environment, they may seriously suffer from the frequent occurrence of handover (HO) events with long latency. The frequent occurrence of HO events is caused by the small coverage area of an AP [5]. If a user moves in a vehicular speed, its sojourn time inside the service area of an AP is relatively short and thus, in order to maintain the WiFi connection, the user should often trigger the handover from an AP to another AP. On the other hand, since the WiFi was originally designed for static or semi-static users, it finds a target AP for HO by scanning all channels. Such full channel scanning usually incurs the long HO latency [6].

It is obvious that the frequent HO occurrence and long HO latency may cause severe connection interruption to users. Accordingly, in providing satisfactory services to the users accessing Internet through WiFi inside a moving vehicle, the fast handover with short latency between APs can be essential [7]. In this paper, we design a new HO scheme for effectively realizing fast handover between WiFi APs, by exploiting the characteristic that the movement of a user is predictable with high accuracy on the basis of road topology.

The deployment of APs on a road can be usually modeled as a graph, when the location and service coverage of each AP are known. Since the WiFi network may not cover the whole road area, it is obvious that not only the APs but also the cellular base stations (BSs) should be included in this graph modeling. For convenience of description, when there is no need to explicitly specify AP and BS, we use a point of attachment (PoA) instead of AP/BS. Note that the PoAs, of which service areas are partially overlapped, are referred to as being adjacent to each other. In the graph modeling for PoA deployment on road, each PoA becomes a vertex and the adjacency of two PoAs is represented as an edge between the corresponding vertexes. Since the moving path of a user may be predicted beforehand on the basis of road topology, the PoAs that the user probably passes through while moving on a road also can be easily identified from the PoA adjacency graph for the road. Thus, it is possible to determine a sequence list of target PoAs for HO, in advance before the vehicle of user enters the road. In this paper, we regard a sequence list of target PoAs for HO as a HO schedule. Such a prior HO scheduling can lead to a fast HO because a relatively long time is usually needed to search a next target PoA for HO.

On the other hand, since a real topology of road is actually various and somewhat complex and, furthermore, the moving path of a user changes over time, a single graph modeling of the entire road is ineffective from a HO scheduling perspective. Thus, partitioning a road into several small segments being appropriate to efficient HO scheduling is very important. We, in this paper, suggest a road partitioning scheme and systematically develop a PoA adjacency graph for each road segment. In such a graph, the edge from a PoA a to another PoA b means that the handover from PoA a to PoA b is possible and the weight value of the edge represents the cost of the corresponding HO. Note that the HO cost should be well mapped onto the objectives of HO scheduling, so that the objectives can be achieved by merely selecting a minimal cost path for the given PoA adjacency graph. In this paper, we also transform our HO scheduling objectives to the HO cost, which is assigned to each edge of the graph.

One of our HO scheduling objectives is to minimize the number of HO events for avoiding frequent HO occurrence. We also aim at minimizing the total HO latency that a user will experience while passing through a road segment. To this end, we assign the lower weight to the edge of the HO type with shorter execution process. Note that the HO process between APs with the same IP subnet address (so-called the L2 HO) has fewer execution steps than the HO between APs with respective different subnet addresses (so-called the L3 HO), because the L3 HO requires an additional DHCP process for getting a new IP address. As a result, the L3 HO events need longer execution time and we try to reduce the L3 HO events in HO scheduling. In addition, the AP with fewer users is preferred as a target AP for HO, because the users can get the higher opportunity of accessing the channel from mitigated competition. Also, the APs rather than BS are preferred so that the users can enjoy mobile Internet access with cheaper fee. Thus, we assign the lower weight to the edge for any AP than BS and to the edge for the AP with fewer users.

Based on a directed graph representing both the adjacency relationship of PoAs and the HO cost between them, the proposed HO scheduling problem is formulated as a mixed integer linear programming (MILP) for finding a sequence list of PoAs with a minimum cost and is solved by a central scheduling server.

Before entering a particular road segment, a user requests its HO schedule for the segment to the scheduling server and receives a sequence list of target PoAs together with their working channels. While moving on the road segment, as soon as the connection with its current serving PoA becomes unstable, the user can try a connection to the next PoA, without a long-time HO step for searching a target PoA. This realizes a very fast and efficient handover in high-speed vehicular environment. Moreover, since the cental sever is responsible for HO scheduling of all users, the users are free from the burden calculating their HO schedules, with no need of individually storing the graphs for all road segments.

We summarize the contributions of this paper as follows.

- The HO scheduling scheme for the heterogeneous vehicular network is proposed. The proposed scheme considers the IEEE 802.11b/g/n/ac-based WiFi as an alternative of the cellular network, instead of the IEEE 802.11p.

Thus, it requires a relatively low deployment cost and it is directly applicable.

- A graph model for PoA deployment around a road is systematically developed, considering on a complex structure of real roads, which include straight, curve, intersection, and u-turn.
- A HO schedule of a user is determined in advance before the vehicle of the user enters the road. Moreover, the HO scheduling problem is formulated so as to maximize the connection time on WiFi and minimize the total HO latency while taking account of local load balancing. This makes the fast handover possible and, in turn, enables the users in a moving vehicle to enjoy Internet access with cheap fee.
- A whole HO process such as a prior HO scheduling, a delivery of HO schedule, and a HO event triggering is designed in detail.
- The effectiveness of the proposed HO scheme is assessed under a practical scenario of heterogeneous WiFi/cellular network on a real road topology.

To the best of our knowledge, the proposed scheme is the first HO scheduling scheme for heterogeneous vehicular networks, designed by considering a complex topology of real roads.

The rest of the paper is organized as follows. Section II presents the related works. The system model is elaborated in Section III. Section IV discusses a graph-based modeling as the foundation of the proposed HO scheduling. Section V and Section VI describe the proposed HO scheduling problem and the whole HO process, respectively. In Section VII, the performances of the proposed scheme and some existing HO schemes are assessed by using simulation. The paper is concluded with Section VIII.

II. RELATED WORKS

The HO decision of wireless networks has been extensively studied. The works in [8] and [9] present comprehensive survey regarding HO decision, in which most of HO decision schemes focused on selecting a single target AP for HO on the basis of various selection metrics. Regarding the HO in vehicular networks, the recent works in [10]–[12] deal with the vertical HO, whereas the works in [13]–[17] handle the horizontal HO of IEEE 802.11 networks.

In [10], a cloud-based vertical HO scheme for vehicular networks was proposed, where the problem of selecting a target network for HO was formulated as a coalition formation game having a Nash equilibrium. Each coalition is associated with a particular PoA and, thus, joining a coalition means selecting the corresponding PoA. Marquez-Barja *et al.* [11] formulated a weighted multi-objective optimization problem for a vertical HO in vehicular environments. Four metrics of throughput, packet delay, packet loss, and price (cost) were used for selecting a target PoA. The scheme highly depends on instant information of these metrics for each candidate PoA, sent through the serving PoA. However, in the highly dynamic vehicular environment, this information may not be accurate, which is likely to cause the undesired HO result. Similarly with [11], Wang *et al.* [12] also considered four types of user preference in selecting a target network, such as longer connection duration with network, higher network bandwidth, lower communication cost of network, and service orientation.

Huang and Tzeng [13] proposed two algorithms for the horizontal HO between IEEE 802.11 networks, namely longest distance algorithm (LDF) and least HO frequency algorithm (LHFA). In LDF, a user selects a target AP which can provide the longest service time by utilizing the location information and service radius of neighboring APs. Meanwhile, the LHFA intends to minimize the number of HO events. For this purpose, it firstly constructs a handover-graph according to the current trajectory of the user. Then, using Dijkstra's algorithm, it finds a set of APs to be connected along the trajectory which minimizes the number of HO events. However, it did not discuss the trajectory prediction clearly and assumed that the user always moves on a straight road. When considering various topology of road such as U-turn, bend and intersection, the assumption of straight road is very tight and unpractical, and thus the HO algorithm can be ineffective in real environment.

The scheme in [14] schedules the scanning order of candidate APs on the predicted moving path of the user, based on the expected connection duration. The AP which can provide the longest connection time is scanned first. Since the moving path is predicted with the last three records on general positioning system (GPS) information of the user, if the location information is not accurate, the user may connect to the undesired AP.

Mouton *et al.* [15] utilize the road topology to predict the trajectory of a user. The road is divided into several road sections, which are the parts of road bounded by two intersections or a dead end. A set of candidate APs within a road section is prior decided, based on the received signal strength indicator (RSSI) calculated according to a predefined channel model. However, this scheme did not discuss how to select the target APs from candidate APs. Furthermore, since the selection of candidate APs is merely based on the calculated RSS, it still may cause frequent HO events.

Dwijaksara *et al.* [16] proposed a crowdsourcing-based HO scheme, in which the moving pattern of a user on a particular road is derived from the past moving data of the user reported by itself, and this moving pattern is again used to predict the future moving direction of the user. Then, the AP of which the longest connection time is expected on the predicted future moving path of the user is selected as a HO target. However, this technique requires a complex information configuration of the road, which introduces a significant overhead.

Dwijaksara *et al.* [17] proposed an AP clustering scheme for a fast HO, where some adjacent APs are grouped into a cluster with the same basic service set identifier (BSSID) and then the handover between APs belonging to the same cluster requires only the channel switching. This clustering-based

HO scheme was implemented on mobile devices and APs, and its performance was evaluated through real experiments.

III. SYSTEM DESCRIPTION

In this paper, we consider heterogeneous networks where WiFi systems coexist with cellular networks such as Long Term Evolution (LTE),[1] as depicted in Fig. 1. Since we are interested in efficiently supporting the Internet access of users in vehicles on roads, we concentrate on communication in the areas around roads. The WiFi APs or simply APs are randomly and densely deployed, which causes the service areas of many APs to overlap with each other. On the other hand, the cellular BSs or simply BSs are deployed under a certain cell planning strategy, so that the service area of the cellular network covers the whole road area. Unlike the cellular networks, the service area of the WLANs may not cover some part of road areas. The APs and BSs are connected to the Internet using wired links. In this paper, the AP and BS are referred to as a point of attachment (PoA). In addition, we consider that the whole area is divided into disjoint sub-areas. Therefore, from this point, we can focus on a particular sub-area.

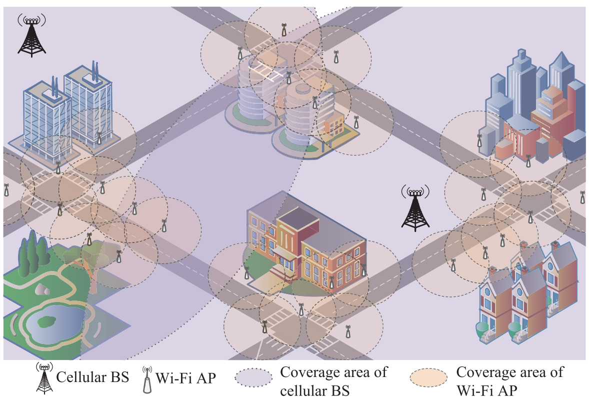

FIGURE 1. Heterogeneous networks in vehicular environments.

Each sub-area is managed by a server, possibly in a similar manner to cloudlet system [18]. The server stores information regarding all PoAs within its sub-area as well as information of the roads belonging to the corresponding sub-area. The set of all PoAs within a sub-area is denoted by $\mathcal{P}$, which consists of the set of APs, $\mathcal{A}$, and the set of BSs, $\mathcal{B}$. Thus, $\mathcal{P} = \mathcal{A} \cup \mathcal{B}$. p_a and R_a respectively denote the location and service radius of PoA-a ($\in \mathcal{P}$). For a given coordinate point p_a, its latitude and longitude components are denoted by $p_{a.x}$ and $p_{a.y}$, respectively.

On the other hand, the roads within a sub-area are divided into several road portions. A road portion is started from an intersection or u-turn area and ended in the next nearest intersection or u-turn area. Thus, a road portion contains no intersection and u-turn areas. The road portion-p is represented by a list of curve points, $\mathcal{L}_p$. When the road portion-p has n curve points and its ith curve points is denoted by l_{p_i}, $\mathcal{L}_p = \left(l_{p_1}, l_{p_2}, ..., l_{p_n}\right)$. Moreover, a straight line between two consecutive curve points is referred to as a moving path. The north azimuth value formed by ith the moving path $(l_{p_i}, l_{p_{i+1}})$ of the road portion-p is denoted by θ_p^i.

A user device or simply a user, is equipped with both WiFi interface and cellular network interface, moves in vehicular speed along a road. The user can track its current position and moving direction using GPS and can connect to both WiFi and cellular network while moving on the road. However, to maximize the connection time on WiFi, the user prefers to transmit data through WiFi whenever possible. To maintain a network connection, the user needs to perform handover to either AP or BS while moving. Thus, horizontal and vertical HO may take place.[2] In the case of horizontal HO between WiFi APs, the HO can be further categorized into two types of L2 and L3 HO, depending on the subnet configuration of the APs, like in [17]. The handover between two APs with the same subnet address is referred to as the L2 HO, whereas the L3 HO means the handover between the APs with respective different subnet addresses. The HO latency from PoA-a to PoA-b is denoted by τ_{ab}. And, $\eta_a(t)$ denotes the number of users connected to PoA-a at time t.

IV. GRAPH-BASED MODELING

Our HO scheduling design is based on a directed graph representing the PoA deployment on a road, where each PoA is represented as a graph node and the adjacency of two PoAs becomes an edge. Let us examine this graph modeling.

A. DIVISION OF ROAD PORTION INTO ROAD SEGMENTS

We first define a road segment as the part of a road portion satisfying the condition that, if the HO from AP-a to AP-b is possible while moving on the road segment, the HO from AP-b to AP-a reversely on the same road segment is not possible. This condition is referred to as a *single direction HO*. With a single direction HO condition, a road segment can be modeled as a graph which has no loop between any two nodes (APs), i.e., the graph cannot have simultaneously the edge from vertex-a to vertex-b and the edge from vertex-b to vertex-a. This no-loop characteristic can lead to more simple but efficient HO scheduling on the road segment.

Let us explain with an example that the *single direction HO* condition may not be held within a road portion. Although a road portion does not contain any intersection or u-turn area (refer to the definition of a road portion in Section III), it may contain a sharp turnout section which excessively changes the moving direction of a user. For instance, a user, which firstly moves toward South direction at the beginning of a road portion, may finally move toward North direction at the end of road portion due to the sharp turnout section.

[1] The cellular networks may also include GSM and UMTS network but, for simplicity we only mention the LTE.

[2] In the case of vertical HO, the user does not actually perform the HO operation. The user can only switch the currently active network interface for data transmission either from LTE to WiFi interface or vice versa.

As the implication, the user, which performs handover from AP-a to AP-b at the beginning of road portion, may possibly perform handover from AP-b to AP-a at the end of road portion.

Now, we partition a road portion into one or more road segments, each of which satisfies the *single direction HO* condition. Let $\mathcal{M}_s$ be the list of curve points which construct the road segment-s. $\mathcal{M}_s = \left(m_{s_1}, m_{s_2}, \ldots, m_{s_k}\right)$, where m_{s_1} and m_{s_k} are the first and last curve points, respectively. In addition, the north azimuth value formed by the moving path $(m_{s_i}, m_{s_{i+1}})$ of the road segment-s is denoted by θ_s^i. We assume that, since a road segment does not contain any sharp turn-out section, the difference between the north azimuth value of any two moving paths within the road segment should not be greater than a certain threshold value, Γ. This threshold is used to detect the sharp turn-out section within a road portion and to divide the road portion into disjoint road segments. The detail is presented in Algorithm 1, where EmptyList() is a function which generates an empty list and InSertList(L, e) is a function which inserts e as the last element of a list L.

Algorithm 1 Construction of Road Segments

1: A list of curve points $\mathcal{L}_p = \left(l_{p_1}, l_{p_2}, \ldots, l_{p_n}\right)$
2: A list of north azimuth values $\left(\theta_p^1, \theta_p^2, \ldots, \theta_p^{n-1}\right)$
3: Threshold of north azimuth value difference, Γ
4: $\mathcal{M}_s \leftarrow$ EmptyList()
5: InsertList$\left(\mathcal{M}_s,\ l_{p_1}\right)$, $\ell \leftarrow 1$, $C \leftarrow \{\}$
6: **for** $i = 2$ to $n-1$ **do**
7:     **if** $|\theta_p^\ell - \theta_p^i| \leq \Gamma$ **then**
8:         InsertList($\mathcal{M}_s$, l_{p_i})
9:     **else**
10:         $C \leftarrow C \cup \left\{ \text{InsertList}(\mathcal{M}_s,\ l_{p_i}) \right\}$
11:         $\mathcal{M}_s \leftarrow$ EmptyList()
12:         InsertList$\left(\mathcal{M}_s,\ l_{p_i}\right)$, $\ell \leftarrow i$
13:     **end if**
14: **end for**
15: $C \leftarrow C \cup \left\{ \text{InsertList}(\mathcal{M}_s,\ l_{p_n}) \right\}$
16: **return** C

We start by setting the first curve point of road portion as the first curve point of the first road segment on the corresponding road portion (line 4, 5). Then, we continuously record all the other curve points which belong to the current road segment, while observing the north azimuth value difference to detect the next road segment. This is done until all curve points of the corresponding road portion are processed (lines 6 – 14). The algorithm returns a set of all road segments for given road portion.

Note that since the HO scheduling problem on each road segment is independent of each other, we initially treat the HO scheduling problem on a single road segment. Then, we will merge all pieces together in Section VI.

B. RELATION BETWEEN PoAs ON A ROAD SEGMENT

Let us consider a road segment-s with a list of curve points, $\mathcal{M}_s = \left(m_{s_1}, m_{s_2}, \ldots, m_{s_k}\right)$. To formulate the HO scheduling problem, we need to discuss the relation between PoAs on road segment-s. We define the *initial PoAs* of a road segment as the PoAs whose service area overlaps with the first curve point. Also, the *final PoAs* of a road segment is defined as the PoAs whose service area overlaps with the last curve point. When $\mathcal{S}_s$ and $\mathcal{E}_s$ respectively denote the set of *initial PoAs* and the set of *final PoAs* for the road segment-s,

$$\mathcal{S}_s = \bigcup_{\forall a \in \mathcal{P}} \left\{ a \;\middle|\; \|p_a - m_{s_1}\| \leq R_a - \Delta \right\}, \tag{1}$$

$$\mathcal{E}_s = \bigcup_{\forall a \in \mathcal{P}} \left\{ a \;\middle|\; \|p_a - m_{s_n}\| \leq R_a - \Delta \right\}, \tag{2}$$

where $\|x-y\|$ is the Euclidean distance between point x and y, and Δ is a distant margin.

The perpendicular distance from the location of PoA-a, p_a, to a straight line formed by the moving path $(m_{s_i}, m_{s_{i+1}})$ is

$$d_a = \frac{\Phi_a}{\|m_{s_{i+1}} - m_{s_i}\|}, \tag{3}$$

where $\Phi_a = |(m_{s_{i+1}}.y - m_{s_i}.y)p_a.x - (m_{s_{i+1}}.x - m_{s_i}.x)p_a.y + m_{a_{i+1}}.x m_{s_i}.y - m_{s_{i+1}}.y m_{s_i}.x|$ and $|\cdot|$ is the absolute value. If $d_a < R_a$, it implies that the service area of PoA-a intersects the straight line formed by the moving path $(m_{s_i},\ m_{s_{i+1}})$. q_a^1 and q_a^2 denote the first and second intersection points, respectively, and are calculated as in [20]. Fig. 2 depicts all possible cases for the intersection between the service area of PoA-a and a straight line formed by the moving path $(m_{s_i}, m_{s_{i+1}})$.

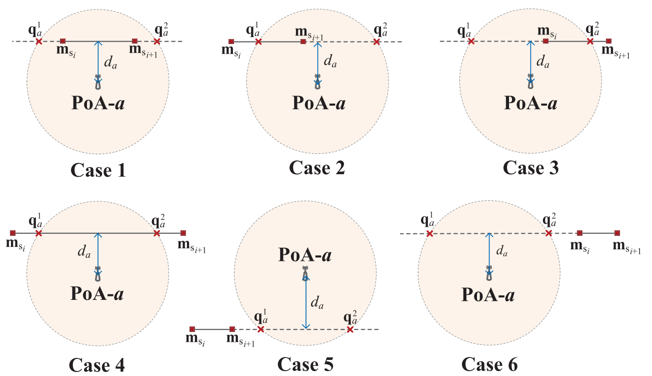

FIGURE 2. All possible cases of intersection between service area of PoA-*a* and the straight line of the moving path (m_{s_i}, $m_{s_{i+1}}$).

Based on the intersection points between the service area of the PoA and the straight line formed by the moving path, we define the moving path being covered by PoA.

Definition 1: A PoA-a covers a moving path $(m_{s_i}, m_{s_{i+1}})$, which is written as COVER$\left(a,\ (m_{s_i}, m_{s_{i+1}})\right)$, if some part of the path $(m_{s_i}, m_{s_{i+1}})$ is located inside the service area of PoA-a. It is satisfied if one of the following conditions is fulfilled.

$$\|m_{s_i} - p_a\| \leq R_a, \tag{4}$$

$$\|m_{s_{i+1}} - p_a\| \leq R_a, \tag{5}$$

$$\Big(d_a < R_a\Big) \wedge \Big(\|m_{s_i} - q_a^1\| + \|q_a^1 - q_a^2\| + \|q_a^2 - m_{s_{i+1}}\| = \|m_{s_i} - m_{s_{i+1}}\|\Big). \tag{6}$$

The equation (4) and (5) indicate whether m_{s_i} or $m_{s_{i+1}}$ is located in the service area of PoA-a (Case 1, 2, and 3 in Fig. 2). The equation (6) holds when both m_{s_i} and $m_{s_{i+1}}$ are located outside the service area of PoA-a, but the straight line formed by those points intersects the service area of PoA-a (Case 4 in Fig. 2). A PoA covering any moving path of the road segment-s (Cases 1 – 4 in Fig. 2) becomes one of the candidate PoAs for HO on the road segment-s.

When the PoA-a covers one or more moving paths of the road segment-s, there are one or two intersection points between the PoA and the covered moving paths, marked by the red-crosses in an example of Fig. 3. The first and second intersection points are denoted by $u_{s,a}$ and $v_{s,a}$, respectively. If the first curve point of road segment-s is within the service area of PoA-a, i.e., if $\|p_a - m_{s_1}\| < R_a$, $u_{s,a} = m_{s_1}$. Also, if the PoA-a contains the last curve point of the road segment-s ($\|p_a - m_{s_k}\| < R_a$), $v_{s,a} = m_{s_k}$. Based on the above discussion, we describe the relationship between PoAs on the road segment-s.

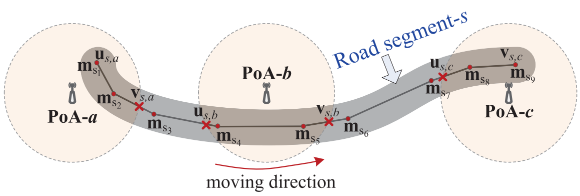

FIGURE 3. **The intersection points between the PoAs and the road segment-*s* where $\mathcal{M}_s = (m_{s_1}, m_{s_2}, \ldots, m_{s_9})$.**

Definition 2: A PoA-b is the next neighbor of PoA-a on the road segment-s, which is written as $\text{NEXT}_s\big(a, b\big)$, if there exists a moving path $(m_{s_i}, m_{s_{i+1}})$, such that PoA-a and PoA-b satisfy the conditions in (7) and (8).

$$\|q_a^1 - q_a^2\| + \|q_b^1 - q_b^2\| > \max_{\phi,\psi \in \{q_a^1, q_a^2, q_b^1, q_b^2\}} \|\phi - \psi\|, \tag{7}$$

$$\Big(h_s(u_{s,a}, m_{s_k}) > h_s(u_{s,b}, m_{s_k})\Big) \vee \Big(h_s(m_{s_1}, v_{s,a}) < h_s(m_{s_1}, v_{s,b})\Big), \tag{8}$$

where $h_s(x, y)$ is the total length of all moving paths from the point x to y.

The condition (7) states that the service areas of two PoAs should continuously cover the moving path (m_{s_i},$m_{s_{i+1}}$). The condition (8) is that the PoA-a should be followed by the PoA-b with regard to the moving direction on the road segment-s. The HO from PoA-a to PoA-b in the road segment-s may take place if the PoA-b is the next neighbor of the PoA-a.

C. DIRECTED GRAPH REPRESENTATION

For given PoA deployment on the road segment-s, we construct a directed graph $G_s = (V_s, E_s)$, where V_s is a set of PoAs whose service area covers at least one moving path of the road segment-s and E_s is a set of all feasible directed edges between PoAs in V_s. Remind that $\mathcal{A}$ is a set of all APs and $\mathcal{B}$ is a set of all BSs. Let $V_{\mathcal{A},s}$ and $V_{\mathcal{B},s}$ respectively denote a set of APs and a set of BSs, covering at least one moving path of the road segment-s.

$$V_{\mathcal{A},s} = \bigcup_{\forall a \in \mathcal{A}} \Big\{a \mid \exists m_{s_i}, m_{s_{i+1}} \in \mathcal{M}_s \text{ such that } \text{COVER}\big(a, (m_{s_i}, m_{s_{i+1}})\big)\Big\}, \tag{9}$$

$$V_{\mathcal{B},s} = \bigcup_{\forall b \in \mathcal{B}} \Big\{b \mid \exists m_{s_i}, m_{s_{i+1}} \in \mathcal{M}_s \text{ such that } \text{COVER}\big(b, (m_{s_i}, m_{s_{i+1}})\big)\Big\}. \tag{10}$$

Then, $V_s = V_{\mathcal{A},s} \bigcup V_{\mathcal{B},s}$. For the sake of convenience, we will also refer a vertex in $V_{\mathcal{A},s}$ as AP and a vertex in $V_{\mathcal{B},s}$ as BS.

Now let us derive all feasible edges in E_s. When the PoA-b is the next neighbor of the PoA-a, there exists a directed edge, e_{ab}, from PoA-a to PoA-b. Let $E_{\mathcal{A},s}$ be a set of all directed edges between two PoAs (i.e., APs) in $V_{\mathcal{A},s}$, and let $E_{\mathcal{B},s}$ be a set of all directed edges between two BSs in $V_{\mathcal{B},s}$.

$$E_{\mathcal{A},s} = \bigcup_{\forall a,b \in V_{\mathcal{A},s}} \Big\{e_{ab} \,\Big|\, \text{NEXT}_s\big(a, b\big)\Big\}, \tag{11}$$

$$E_{\mathcal{B},s} = \bigcup_{\forall a,b \in V_{\mathcal{B},s}} \Big\{e_{ab} \,\Big|\, \text{NEXT}_s\big(a, b\big)\Big\}, \tag{12}$$

If there exists a directed edge from the PoA-a to the PoA-b, it implies that the HO may take place from the PoA-a to the PoA-b. The edges in $E_{\mathcal{A},s}$ or $E_{\mathcal{B},s}$ only specify the horizontal HO either between APs or between BSs. In addition to the horizontal HO, the vertical HO between AP and BS may also take place. Thus, we also should consider the edges which specify the vertical HO. To maximize the connection time on WiFi, the user is allowed to perform the HO from AP to BS only when the current serving AP does not have any other AP as its next neighbor. Thus, the AP can have an outgoing edge to the BS only when the corresponding AP has no outgoing edge to any other AP. Note that an AP belonging to the final PoAs does not have any outgoing edge to the other BSs on the same road segment. Let $V_{\mathcal{A},s}^{-\mathcal{E}_s}$ denote a set of all APs in $V_{\mathcal{A},s}$ excluding the APs in $\mathcal{E}_s$. When E_s^{vo} is a set of all feasible outgoing edges from AP in $V_{\mathcal{A},s}^{-\mathcal{E}_s}$ to BS in $V_{\mathcal{B},s}$,

$$E_s^{\text{vo}} = \bigcup_{\substack{\forall a \in V_{\mathcal{A},s}^{-\mathcal{E}_s}, \forall b \in V_{\mathcal{B},s}, \\ \forall c \in V_{\mathcal{A},s}}} \Big\{e_{ab} \,\Big|\, e_{ac} \notin E_{\mathcal{A},s},\ \text{NEXT}_s\big(a, b\big)\Big\}. \tag{13}$$

Note that a user should always perform handover from BS to AP whenever the AP becomes available, This implies that an AP can have an incoming edge from the BS only when having no incoming edge from any other AP. Let $V_{\mathcal{A},s}^{-\mathcal{S}_s}$ be a set of all APs in $V_{\mathcal{A},s}$ excluding the APs in $\mathcal{S}_s$. Then, a set of all feasible outgoing edges from BS in $V_{\mathcal{B},s}$ to AP in $V_{\mathcal{A},s}^{-\mathcal{S}_s}$ is

$$E_s^{\mathrm{vi}} = \bigcup_{\substack{\forall b \in V_{\mathcal{B},s}, \forall a \in V_{\mathcal{A},s}^{-\mathcal{S}_s}, \\ \forall c \in V_{\mathcal{A},s}}} \left\{ e_{ba} \,\middle|\, e_{ca} \notin E_{\mathcal{A},s},\ \mathrm{NEXT}_s(b, a) \right\}. \tag{14}$$

Finally, a set of all feasible edges, E_s, is as

$$E_s = E_{\mathcal{A},s} \bigcup E_{\mathcal{B},s} \bigcup E_s^{\mathrm{vo}} \bigcup E_s^{\mathrm{vi}}. \tag{15}$$

Fig. 4 depicts an example for the PoA deployment on road segment-s and Fig. 5 shows the corresponding directed graph. Each edge in the graph has a weight representing a cost of the corresponding HO, which will be discussed in the following section.

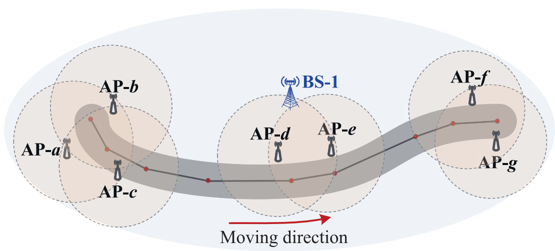

FIGURE 4. Example deployment of PoAs on a road segment.

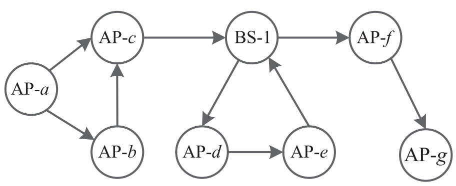

FIGURE 5. A directed graph for the road segment of Fig. 4.

V. HANDOVER SCHEDULING PROBLEM

In this section, we formulate the HO scheduling problem and propose the HO scheduling algorithm.

A. PROBLEM FORMULATION

The primary goal of our scheme is to find the best HO schedule for each road segment, which minimizes the total HO latency while having network connection through WiFi (not cellular) whenever possible and preferring the handover to the AP with fewer users. The HO schedule is a sequence list of target PoAs in the order of the HO actions to be conducted by a user when it moves along a particular road segment. The HO schedule is derived by solving the following HO scheduling problem.

We formulate the HO scheduling problem of road segment-s, using the directed graph $G_s = (V_s, E_s)$. Let z_{ab} be a binary variable indicating whether the edge e_{ab} ($\in E_s$) is selected. Selecting e_{ab} means that the HO from the PoA-a to the PoA-b is scheduled. If e_{ab} is selected, $z_{ab} = 1$; otherwise $z_{ab} = 0$. Let $w(e_{ab}, t)$ be a weight assigned to edge e_{ab} at time t. The HO scheduling problem for the road segment-s at a time t, started from the PoA-υ ($\upsilon \in \mathcal{S}_s$) and ended at any PoA in $\mathcal{E}_s$, is formulated as (16) – (20).

$$\min \sum_{e_{ab} \in E_s} z_{ab}\, w(e_{ab}, t) \tag{16}$$

$$s.t. \sum_{a \in V_s, e_{\upsilon a} \in E_s} z_{\upsilon a} \geq 1, \tag{17}$$

$$\sum_{b \in V_s, e_{ab} \in E_s} z_{ab} - \sum_{c \in V_s, e_{ca} \in E_s} z_{ca} = 0, \quad \forall a \in \Omega_s \tag{18}$$

$$\sum_{a \in V_s} \sum_{b \in \mathcal{E}_s, e_{ab} \in E_s} e_{ab} \geq 1, \quad b \neq \upsilon \tag{19}$$

$$z_{ab} \in \{0, 1\}, \quad \forall a, b \in V_s \tag{20}$$

where Ω_s is the set of all PoAs in V_s excluding the starting PoA and the final PoAs.

The objective (16) seeks for the HO schedule that minimizes the sum of weights. The constraint (17) ensures that at least one of the outgoing edges of PoA-υ should be included since it is a starting PoA. The constraint (18) ensures that if the PoA-a in Ω_s is included in the HO schedule, both its incoming and outgoing edges should be selected together since it is a connecting PoA. The constraint (19) ensures that the HO schedule is ended at one of the PoA in $\mathcal{E}_s$, thus at least one incoming edge to one of these PoAs should be selected. The constraint (20) is self-explanatory.

B. WEIGHT OF EDGE

To solve the above HO scheduling problem, we first should design a weight of each edge, i.e., $w(e_{ab}, t)$ for $e_{ab} \in E_s$. Recall that the objectives of the HO scheduling problem are: (1) maximizing the connection time on WiFi; (2) minimizing the total HO latency taken to perform all HO operations during moving on roads; and (3) preferring the handover to the AP with fewer users. The weight value of each edge should be assigned so that the objectives are fulfilled.

To encourage that a user selects WiFi (not cellular) for communication whenever possible, the highest preference should be given to a directed edge from BS to AP ($\in E_s^{\mathrm{vi}}$). Thus, the edge belonging to E_s^{vi} should have the smaller weight value than the edges with the other types, because of a minimization problem. Next, to achieve the objective of minimizing the total HO latency, the second highest preference is given to a directed edge between APs ($\in E_{\mathcal{A},s}$), since the horizontal HO between APs has much shorter latency than the vertical HO between AP and BS. Lastly, the lowest preference is given to an edge from AP to BS ($\in E_s^{\mathrm{vo}}$) and an edge between BSs ($\in E_{\mathcal{B},s}$), since the connection to BS is unfavorable. Based on the above discussion, the weight of

e_{ab} at a time t, denoted by $w(e_{ab}, t)$, is.

$$w(e_{ab}, t) = \begin{cases} -\text{card}(V_s), & \text{if } e_{ab} \in E_s^{\text{vi}} \\ \alpha \dfrac{\tau_{ab}}{T} + \beta \dfrac{\eta_b(t)}{N_b}, & \text{if } e_{ab} \in E_{\mathcal{A},s} \\ 1, & \text{if } e_{ab} \in E_s^{\text{vi}} \bigcup E_{\mathcal{B},s} \end{cases} \tag{21}$$

In (21), $\text{card}(V_s)$ is the cardinality of the set V_s, τ_{ab} is the HO latency from AP-a to AP-b, and $\eta_b(t)$ is the number of users connected to AP-b at time t. In addition, T is the maximum possible HO latency and N_b is the maximum allowable number of users at PoA-b.

Now, we examine more specifically the weight value according to each edge type, starting from the edges in $E_{\mathcal{A},s}$. The edge from AP-a to AP-b has the weight of $\alpha \frac{\tau_{ab}}{T} + \beta \frac{\eta_b(t)}{N_b}$. The first term accounts for the HO latency and the second term reflects the number of users associated with the target AP, where α and β ($\alpha + \beta = 1$) are the relative importance factors of the first and second terms. Note that any edge in $E_{\mathcal{A},s}$ has a positive value smaller than one, since each term is normalized by its maximum value. The weight is designed with the purpose of reducing the total HO latency (the first term) while preferring the handover to the AP with fewer associated users (the second term). Since the handover to the AP with fewer associated users can help to relieve the channel contention among users within AP, each user may have an opportunity to get the higher throughput. Furthermore, since the number of associated users at each AP is changed over time, this term can give a randomness effect to the HO schedule. Note that the graph representing the PoA deployment on a road segment is a fixed factor. In addition, because the HO latency between PoAs mainly depends on its HO type which is a pre-determined parameter, this term reflecting the number of associated users at each AP is only the time-varying factor in (21). Therefore, without this term, the HO schedule on a particular road segment may always contain the same set of PoAs. This situation is very undesirable since the users are concentrated to some particular APs on a road segment and such load unbalancing may incur the performance degradation.

On the other hand, for maintaining the network connection through WiFi whenever possible, any directed edge from BS to AP should have a weight smaller than zero, since an edge between two APs has a positive value smaller than one. We simply set the weight of any edge in E_s^{vi} to $-\text{card}(V_s)$. As a result, whenever an AP becomes available, the user is forced to connect to the AP.

An edge from BS to other BS is selected only when the BS does not have an AP as its next neighbor. The HO from AP to BS is also scheduled only when the AP has no neighboring AP. Thus, when considering a positive weight value smaller than one for an edge between APs, the weight of an edge belonging to these types can be set to any positive value equal to or larger than one. We simply assign the weight of one to the edges from AP to BS and from BS to BS.

C. HO SCHEDULING ALGORITHM

The problem in (16) – (20) is a mixed integer linear programming (MILP) problem, which can be efficiently and quickly solved using any MILP solver such as CPLEX [21] or LPsolve [22]. Then, the HO schedule is derived from the solution of the problem, which is a set of selected edges. Algorithm 2 returns the HO schedule of a user for the road segment-s, which is a sequence list of target PoAs in order of the HO actions to be conducted by the user while moving along the road segment. Like in Algorithm 1, the function EmptyList() generates an empty list and the function InSertList(L, e) inserts e as the last element of a list L.

Algorithm 2 HO Scheduling on Road Segment-s, Started at PoA-υ

```
1:  Solve the problem (16) – (20) using MILP solver
2:  Let 𝔼 be the set of edges returned by MILP solver
3:  Let 𝕍 be the set of vertices whose edge is in 𝔼
4:  L_{s,υ} ← EmptyList()
5:  InSertList(L_{s,υ}, υ), a ← υ
6:  while 𝔼 ≠ ∅ do
7:      if a ∈ ℬ then
8:          b ← arg min_{b̂∈𝕍, e_{ab̂}∈𝔼} h_s(m_{s1}, v_{s,b̂})
9:      else
10:         find b in 𝕍 such that e_{ab} ∈ 𝔼
11:     end if
12:     InSertList(L_{s,υ}, b), 𝔼 ← 𝔼\{e_{ab}}, a ← b
13: end while
14: return L_{s,υ}
```

Firstly, the starting PoA-υ is inserted to an empty list $L_{s,v}$ as the first element (lines 4 – 5). After that, the HO schedule is derived by repeatedly identifying the next PoA to be connected through the already selected edge until all edges are processed (lines 6 – 13). If the current PoA-a is a BS (line 7), there may be several possible outgoing edges (see Fig. 5). Thus, to avoid skipping any edge in $\mathbb{E}$, we should select the outgoing edge to a PoA which is the closest to the starting curve point of the corresponding road segment (line 8). However, if the current PoA-a is an AP, there is only one outgoing edge from AP-a in $\mathbb{E}$ (line 10). Then, the PoA-b is inserted to the list of target PoAs. An example of the HO schedule for the graph with starting PoA-a in Fig. 5 derived by the Algorithm 2 is $\big($AP-a, AP-c, BS-1, AP-d, AP-e, BS-1, AP-$f\big)$.

VI. HANDOVER SCHEDULING OPERATION

The proposed HO scheduling is composed of an offline phase and an online phase. In the offline phase, the server generates a directed graph for each road segment, by utilizing the pre-inputted information about network configurations and road topologies. It is possible since the graph generation requires only the static information about PoAs and road segments. Meanwhile, since the weight values of edges are periodically updated with the number of associated users reported by each PoA, the HO schedules cannot be computed in advance and should be generated on demand. In the online phase, when receiving the request for HO schedule from a user, the server computes the HO schedule for the user by using Algorithm 2. Note that the HO schedule delivered from the server to the user is merely an ordered list of HO target PoAs. Accordingly, the user should decide the HO triggering points (i.e., when) by itself while moving along a road segment. The block diagram in Fig. 6 summarizes the steps in the proposed HO scheme. Now, we describe the HO schedule delivery and the HO triggering in the online phase.

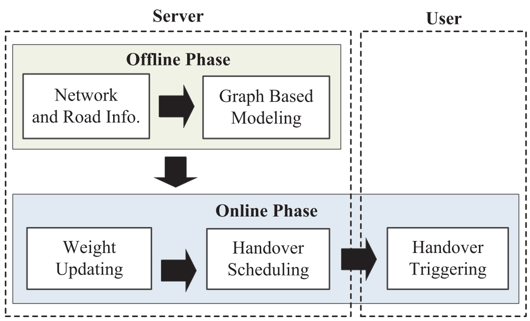

FIGURE 6. Block diagram of the proposed HO scheduling scheme.

A. HO SCHEDULE DELIVERY

When a user connects to the last PoA of the current HO schedule or when it initially joins the network, the user request a new HO schedule to the server through its serving PoA.[3] For this, the user should send a HO request (HO_REQ) packet, which contains the information about its current location and its serving PoA, to the server. Upon receiving this request, the server checks the information in the HO_REQ packet and identifies the possible future road segments to which the user may move, as follows. If a user is currently connected to the final PoA, it implies that the user is close to the end of the road segment. Thus, by observing the current location of the user, the server may figure out the next possible road segments to which the user may move. On the other hand, if a user is not currently connected to the final PoA, it implies that the user has newly joined the network in the middle of a road segment and its next road segment is the current one.

When recognizing the next possible road segments, the server should compute the HO schedules for all possible road segments since it cannot know the road segment of which the user will finally move to. Assume that the PoA-υ is the current serving PoA of the user. For any possible road segment-s, the server gets the HO schedule $L_{s,\upsilon}$, by executing the Algorithm 2 with the starting PoA υ. Then, the server generates the information of (*PoA ID, Channel ID, HO Type*) for each PoA in $L_{s,\upsilon}$, where the *PoA ID* is the unique identifier of the corresponding PoA, the *Channel ID* identifies the working channel of the PoA, the *HO type* indicates whether the type of the HO event from the corresponding PoA to the next PoA is L2, L3, or vertical HO. Let $\mathcal{I}_{s,\upsilon}(i)$ be the information for the ith PoA of $L_{s,\upsilon}$ and let $n_{s,\upsilon}$ be the number of PoAs in $L_{s,\upsilon}$. Then, the HO schedule information for road segment-s is configured as $\mathbb{Z}_{s,\upsilon} = \left(\theta_s^1, \mathcal{I}_{s,\upsilon}(1), \mathcal{I}_{s,\upsilon}(2), \cdots, \mathcal{I}_{s,\upsilon}(n_{s,\upsilon})\right)$, where θ_s^1 is the north azimuth value of the first moving path of the road segment-s. After that, the server replies with a HO reply (HO_REP) packet, which contains the HO schedule information $\mathbb{Z}_{S1,\upsilon}, \mathbb{Z}_{S2,\upsilon}, \cdots, \mathbb{Z}_{S\kappa,\upsilon}$ for all possible road segments $S1, S2, \cdots, S\kappa$.

Upon receiving the HO_REP packet, if there are more than one HO schedule (e.g., at the intersection), the user firstly should select an appropriate HO schedule based on its final moving direction, before updating the HO schedule. It may be better to make the selection as late as possible until the next road segment is manifested, as long as the user does not get out of the service area of the current serving PoA. Based on this reasoning, when the serving PoA of the user is an AP, the user performs the HO schedule selection when it approaches the service area boundary of AP (i.e., when the RSSI value of the serving AP falls below the HO RSSI threshold). Meanwhile, the user being currently served by the BS performs the HO schedule selection when it first detects a new AP, or approaches the cell edge.

When θ_c is the north azimuth value of the current moving direction of the user, the user selects the HO schedule whose north azimuth value is equal to θ_s^*.

$$\theta_s^* = \underset{\theta_i \in \{\theta_{S1}^1, \theta_{S2}^1, \ldots, \theta_{S\kappa}^1\}}{arg\,min} |\theta_i - \theta_c|, \tag{22}$$

B. HO TRIGGERING AND EXECUTION

Since a user gets the HO schedule in advance before entering a road segment, the user knows that it should perform the handover to which PoAs while moving on the road segment. Additionally, the user has to decide when the handover should be tried (HO triggering). We basically consider an RSSI-based HO triggering.

Firstly, when the current PoA is a BS and the next PoA in the current HO schedule is an AP, the user continually checks the accessibility to the next AP while communicating with the serving BS, by intermittently scanning the working channel of the AP. Note that such channel scanning does not affect the data transmission since the user uses different network interfaces for the LTE and WiFi. When detecting that the RSSI value of the next AP is higher than a predefined minimum RSSI threshold, the user should immediately trigger the handover to the AP, regardless of the RSSI status of its serving BS. It allows the user to connect to the WiFi network whenever possible. When the current PoA is a BS and there is no neighboring AP, the horizontal HO mechanism of LTE can be used. One the other hand, if the serving PoA is an AP, the handover operation to the next PoA (either AP or BS) is started when the RSSI of serving AP falls below a predefined threshold.

Once the HO is triggered, the user and the target PoA execute the HO protocol steps (e.g., horizontal HO between

[3] A user joining the network may select its serving PoA based on RSSI.

WiFi APs, horizontal HO between BSs, or vertical HO between AP and BS). The readers can refer to [17] for the execution steps of horizontal HO between WiFi APs and [19] for horizontal HO between LTE BSs. For the vertical HO in our scheme, because the user uses different network interfaces to access WiFi and LTE, it is merely needed to switch the active network interface for data transmission either from LTE to WiFi or vice versa.

Note that if the handover to a particular target PoA fails after a certain number of trials, the user marks the PoA as being unavailable. Then, the user scans the working channels of neighboring PoAs and selects the PoA with the highest RSSI as a HO target PoA. Afterwards, the user executes the handover to the new target PoA by following the proposed HO procedure. In addition, the user may report the unavailable PoA to the server. The server can make a further decision regarding the unavailable PoA, for example by removing it from the graph. Then, it will be excluded in any future HO schedules until it is available.

C. HO SCHEDULE DELIVERY OVERHEAD

In the proposed HO scheme, the burden for computing the HO schedule is moved from user to server, whereas a little communication overhead for delivering the schedule information from server to user is added (the overhead of HO request is negligible). Let us estimate this delivery overhead.

Whenever a user reaches an intersection, the user receives the HO schedules for all possible road segments. For each PoA in a HO schedule, we use the PoA ID of 6 bytes, the channel ID of 1 byte, and the HO type of 1 byte (totally 8 bytes). The HO schedule which contains N_p PoAs uses $(8N_p + 1)$ bytes, including the north azimuth value of 1 byte. If the number of possible road segments at the intersection is N_m, the whole HO schedule information delivered to the user is $(8N_p + 1)N_m$ bytes. Note that the value of N_p depends on the length of road segment. Since the length of a road segment can be designed not to be too long, N_p is likely to be not large. The intersection is also commonly constructed by a few of road segments, implying that the value of N_m is also small. Note that the HO schedule information of a user has the largest size at intersection. As a result, the information amount of any HO schedule is reasonably so small (a few hundred bytes) as to be carried by a single MAC frame. Moreover, since the server sends such the HO schedule information to a user only when it connects to the last PoA of the current HO schedule or initially joins the network, the information delivery overhead of the proposed scheme is expected to be not high.

VII. PERFORMANCE EVALUATION

We evaluate the performance of the proposed scheme through event driven simulation.

A. SIMULATION SETTINGS

We consider a particular region in the Manhattan city (red-marked area) as shown in Fig. 7 whose size is approximately 1000 m × 1000 m. The detail information for the roads on this area is got from OpenStreetMap [23]. The default number of vehicles is 300. We utilize the SUMO tool [24] to generate the traffic model of the vehicles inside the red-marked area. The moving speed of a vehicle is uniformly distributed in the range of 0 ∼ 50 km/h. Furthermore, we set 8 blocks within the area to be clean blocks, in which there is no installed AP. Then, 600 APs are randomly deployed according to a uniform distribution within the area, excluding the clean blocks. It creates several road segments which are not fully covered by the WiFi. On the other hand, we deploy five LTE BSs with the service radius of 500 m (one on each corner and one in the middle of the red-marked area). Thus, the LTE network covers all the road segments.

FIGURE 7. Simulation environment: Manhattan area.

While moving along a road, the vehicles are connected to PoAs for wireless Internet access. During simulation, downlink UDP traffic with the default arrival rate of 4 Mbps is sent to a user. At the same time, the user also sends uplink UDP traffic whose arrival rate is 0.1 Mbps. The APs within the same block are assigned into the same subnet with the probability of 0.8. Thus, the L2 HO takes place only between APs in the same block which belong to the same subnet. The L2 HO execution steps include a single channel scanning, authentication and association process, which take 80 ms to finish. The L3 HO includes L2 HO with an additional DHCP process and is assumed to takes 2 s to finish.[4] Accordingly, $T = 2$ s in (21). In addition, the maximum allowable number of users at each AP is set to 10, i.e., $N_b = 10, \forall b \in \mathcal{A}$ in (21). All packets in the buffer are assumed to be dropped when a user executes the HO.

[4] Actually, these L2 and L3 HO latency values of 80 ms and 2 s respectively were derived through experiment in the real vehicular network environment. The experiment settings were the same with those in [17]. To measure the L2 HO latency, we set the APs to be in the same subnet, thus there was no DHCP process required during the HO. On the other hand, for the L3 HO, we set the APs to be in different subnet and enabled the DHCP server on each AP. The HO was triggered when the RSSI of the current serving AP was equal or less than −80 dBm.

In the simulation, we consider the IEEE 802.11n for WiFi, of which the parameters are presented in Table 1. The APs in proximity are assigned a different operating channel. Thus, co-channel interference between the neighbor APs is assumed to be negligible. For the channel model, we consider pathloss, lognormal shadowing, and multipath fading. The path loss is $33.3 + 36.7 \log_{10} d$ (dB), where d (meters) is the distance between AP and user [25]. The standard deviation of shadowing is 8 dB and the multipath fading follows Jake's spectrum model [26]. The average service radius of WiFi, R_a, is approximately 80 m [27]. For the performance comparison with the proposed scheme, we consider the handover schemes in [11] and [14]. Similarly to the proposed scheme, in the comparison schemes, a user connects to the BS only when there is no available AP. Note that since all schemes access the cellular network in the same manner, the performance differences among three schemes are merely originated from WiFi. Therefore, for the purpose of performance comparison, we measure only the performance of WiFi. The total simulation time is 1100 s.

TABLE 1. Simulation parameters for WiFi.

Parameters	Value
Channel bandwidth	20 MHz
Tx power	20 dBm
Minimum RSSI threshold	−82 dBm
HO RSSI threshold	−80 dBm
Data frame length	1500 bytes
ACK frame length	14 bytes
MAC header	38 bytes
Slot time	9 μs
SIFS time	16 μs
DIFS time	34 μs

B. PERFORMANCE METRICS

The following metrics are used for performance evaluation.

- **Number of HO events** is the total number of L2 and L3 HO events performed by a user moving along the roads during its travel time (i.e., a simulation run time).
- **Ratio of effective connection time** is defined as the ratio of the effective data transmission time excluding the total time taken to finish all of the HO events, for the whole travel time.
- **Fairness index** is calculated using Jain's equation for the total throughput of AP.

$$\mathcal{F} = \frac{\left(\sum_{i=1}^{M} X_i\right)^2}{M \cdot \sum_{i=1}^{M} (X_i)^2}, \tag{23}$$

where M is the total number of APs and X_i is the total throughput of AP-i during the whole travel time. The higher fairness index implies the smaller differences between APs in their throughputs. In other words, the APs are utilized more fairly.

- **Packet drop ratio** is the ratio of the packets dropped by a user during its travel time due to HO events.
- **Average throughput** refers to the average throughput per user, including uplink and downlink one.

C. SIMULATION RESULTS

Table 2 shows the average number of HO events according to the HO type, and the ratio of effective connection time. The default parameter settings are used. We fist discuss the effect of α on the performance of the proposed scheme. Note that α is a weight of HO latency (see (21)). Thus, with the larger α, the proposed scheme tries to reduce the HO latency by preferring the L2 HO with shorter latency compared to the L3 HO, and this allows a user to get the longer effective connection time. Accordingly, in the proposed scheme, as the α gets larger, the average number of the L2 HO events and the ratio of effective connection time increase.

TABLE 2. Comparison of the average number of HO events and the ratio of effective connection time.

Scheme	Average Number of HO Events (L2, L3)	Ratio of Effective Connection Time
Proposed ($\alpha = 0$)	(16.44, 70.15)	86.80 %
Proposed ($\alpha = 0.2$)	(35.30, 51.98)	89.97 %
Proposed ($\alpha = 0.5$)	(35.49, 51.78)	90.00 %
Scheme in [11]	(48.87, 89.66)	83.01 %
Scheme in [14]	(23.18, 72.53)	86.32 %

On the other hand, we can observe in Table 2 that the proposed scheme generates the fewer HO events than the schemes in [11] and [14]. It implies that the proposed scheme more accurately selects the target APs. The scheme in [11] selects a target AP, based on the instant information regarding the current network status sent by the serving AP. This information may not be accurate because of time delay, when it is received by the users. Furthermore, the users in a certain location at a particular time may receive the same information, which causes them to select the same inappropriate AP for HO. The scheme in [14] merely tries to maximize the sojourn time of a user within the AP without considering the HO latency between APs. Thus, it results in the fewer HO events compared to the scheme in [11]. However, since its target PoA selection highly depends on the record of previous moving direction which may not accurately represent the future moving direction, it may still make an inaccurate AP selection.

Figs. 8 and 9 depict the fairness index, the packet drop ratio, and the average throughput, according to the number of users and traffic arrival rate, respectively. First, we discuss the influence of β on the performance of the proposed scheme ($\alpha + \beta = 1$). Note that β is a weight parameter for preferring the AP with fewer associated users, in selecting a HO target AP. Accordingly, as shown in Figs. 8(a) and 9(a), it is obvious that the highest fairness index is achieved when $\beta = 1$, because distributing the users among APs gets the highest priority when $\beta = 1$. Since the contention among users gets

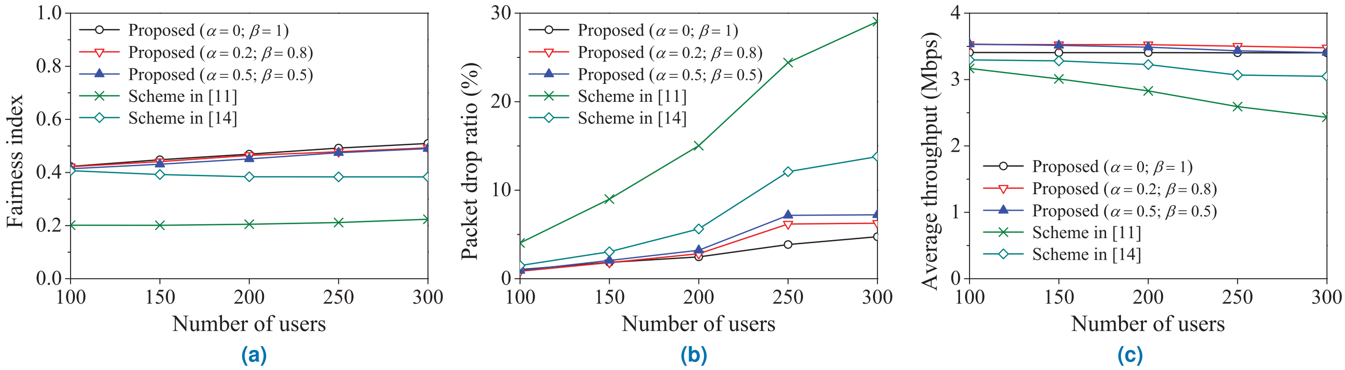

FIGURE 8. **Performance of the system for various number of users. (a) Average fairness index. (b) Average dropped packet of the user. (c) Average throughput of the user.**

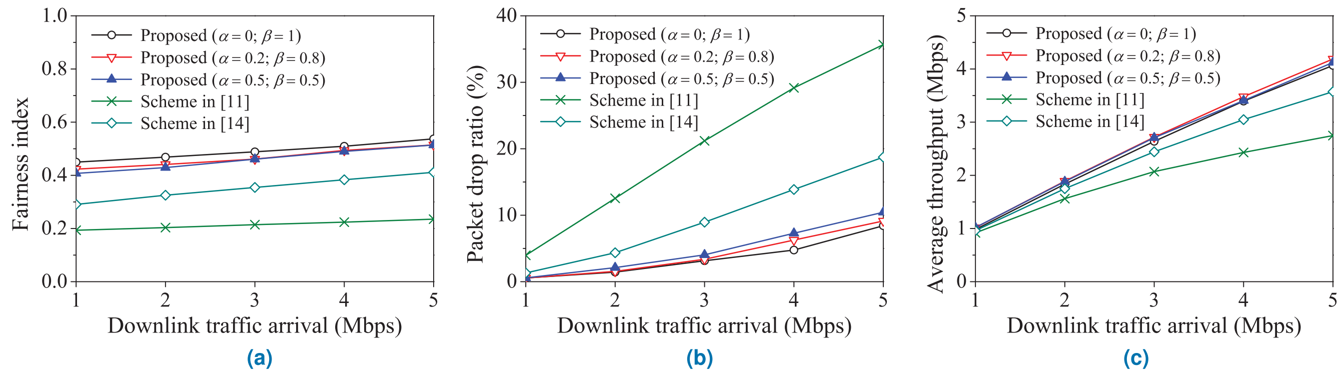

FIGURE 9. **Performance of the system for various downlink traffic arrival rate. (a) Average fairness index. (b) Average dropped packet of the user. (c) Average throughput of the user.**

severer with the smaller β, the number of packets in the buffer increases as β decreases. As a result, with the smaller β, the more packets are dropped during the handover operation and the packet drop ratio increases (see Figs. 8(b) and 9(b)). Meanwhile, in Figs. 8(c) and 9(c), we observe that the throughput of the proposed scheme slightly decreases when α is zero or too high. The reason is that when $\alpha = 0$, the effective connection time ratio decreases (see Table 2) and this causes the decrease of user throughput. A too high α also incurs the high contention and leads to the decrease of throughput. From the above simulation results, it is observed that the proposed scheme achieves better performance when α is not zero while being much smaller than β.

Next, let us compare the performances of the proposed scheme and the existing schemes. We firstly discuss in Fig. 8(a) the throughput fairness performance according to the number of users. In general, the fairness index increases with more users due to averaging effect, because the users are well distributed among the available road segments and more APs have a chance to serve the users. However, regardless of this condition, the fairness index of the scheme in [14] slightly decreases with more users. The reason is that the scheme in [14] merely selects a target AP to maximize the sojourn time of the user within the AP. Thus, since the users moving on the same road segment have a tendency to select the same target APs, as the number of users increases, the users are concentrated to only some particular APs in each road segment, while degrading the fairness performance. On the other hand, as already explained in Table 2, the scheme in [11] may select a target AP based on old information. We observed in simulation that many APs are left unused. As a result, the scheme in [11] has the worst performance among the schemes in throughput fairness.

Fig. 8(b) shows the packet drop ratio. Since the proposed scheme aims at reducing the HO latency while preferring the APs with fewer associated users, the users under the proposed scheme experience the short HO period and less contention among users within AP. As a result, the proposed scheme has the lower packet drop ratio than the schemes in [11] and [14]. Also, as we can observe in Fig. 8(c), the features of the proposed scheme, such as the less contention, longer effective connection time, and lower packet drop ratio, provide the higher throughput to a user, compared to the schemes in [11] and [14].

Next, we examine in Fig. 9 the performances of three schemes according to downlink traffic load. With the higher traffic arrival rate, the fairness in all schemes is improved because the throughput difference between APs decreases (see Fig. 9(a)). In addition, as the downlink arrival rate of a user increases, it is natural that the packet drop ratios of all schemes get higher because each user has more packets in the buffer (see Fig. 9(b)), and the total throughput of user increases because of its increased downlink throughput (see Fig. 9(c)). On the other hand, we can observe in Fig. 9 that the proposed scheme achieves much better performance in all of fairness, packet drop ratio, and the user throughput, than the schemes in [11] and [14]. As already described in Fig. 8, it is caused by the desirable features of the proposed scheme, which are the less contention, longer effective connection time, and shorter HO latency.

In summary, the proposed scheme greatly outperforms the schemes in [11] and [14], for various parameter settings.

VIII. CONCLUSION

In this paper, we have proposed a graph-based HO scheduling scheme for heterogeneous vehicular networks. The proposed scheme exploits the characteristic that the movement of a user is predictable with high accuracy on the basis of road topology. It enables the computation of the HO schedule by a central server in advance considering only the information about the network configurations and road topologies. Using the computed HO schedule, the user only needs to decide when to initiate the connection to the next PoA while moving on the road segment. This technique realizes a very fast and efficient HO in the vehicular environment. Moreover, since the HO scheduling is conducted by the central server, the proposed scheme may take advantage of easily utilizing abundant computing and storage resources in the cloud, and the users are free from the burden of calculating their HO schedules. Regardless of its centralized feature, the proposed scheme introduces very low communication overhead. The simulation results also show that the proposed scheme outperforms the comparison schemes on all performance metrics. Therefore, we believe that the proposed scheme can help strengthening the position of WiFi as an alternative technology of cellular network for providing a low-cost Internet access in the vehicular environment.

REFERENCES

[1] W. Na, N.-N. Dao, and S. Cho, "Mitigating WiFi interference to improve throughput for in-vehicle infotainment networks," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 22–28, Feb. 2016.

[2] N. Lu, N. Cheng, N. Zhang, X. S. Shen, J. W. Mark, and F. Bai, "Wi-Fi hotspot at signalized intersection: Cost-effectiveness for vehicular internet access," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3506–3518, May 2016.

[3] *Wigle Net*. Accessed: May 12, 2018. [Online]. Available: https://wigle.net/

[4] P. Foremski and K. Grochla, "LTE or WiFi? Client-side Internet link selection for smartphones," in *Proc. Int. Conf. Compt. Net.*, Brunów, Poland, May 2015, pp. 43–53.

[5] R. Mahajan, J. Zahorjan, and B. Zill, "Understanding WiFi-based connectivity from moving vehicles," in *Proc. 7th ACM SIGCOMM IMC*, San Diego, CA, USA, Oct. 2007, pp. 312–326.

[6] W. Xu, H. A. Omar, W. Zhuang, and X. S. Shen, "Delay analysis of in-vehicle Internet access via on-road WiFi access points," *IEEE Access*, vol. 5, pp. 2736–2746, 2017.

[7] P. Lv, X. Wang, X. Xue, and M. Xu, "SWIMMING: Seamless and efficient WiFi-based Internet access from moving vehicles," *IEEE Trans. Mobile Comput.*, vol. 14, no. 5, pp. 1085–1097, May 2015.

[8] B. Van Quang, R. V. Prasad, and I. Niemegeers, "A survey on handoffs—Lessons for 60 GHz based wireless systems," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 1, pp. 64–86, 1st Quart., 2012.

[9] A. Ahmed, L. Boulahia, and D. Gaiti, "Enabling vertical handover decisions in heterogeneous wireless networks: A state-of-the-art and a classification," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 776–811, 2nd Quart., 2014.

[10] K. Xu, K.-C. Wang, R. Amin, J. Martin, and R. Izard, "A fast cloud-based network selection scheme using coalition formation games in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5327–5339, Nov. 2015.

[11] J. M. Marquez-Barja *et al.*, "Breaking the vehicular wireless communications barriers: Vertical handover techniques for heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5878–5890, Dec. 2015.

[12] S. Wang, C. Fan, C.-H. Hsu, Q. Sun, and F. Yang, "A vertical handoff method via self-selection decision tree for Internet of vehicles," *IEEE Syst. J.*, vol. 10, no. 3, pp. 1183–1192, Sep. 2016.

[13] S.-C. Huang and Y.-S. Tzeng, "Novel handoff frequency algorithms for vehicular area networks," in *Information Technology Convergence* (Lecture Notes in Electrical Engineering), vol. 253. New York, NY, USA: Springer, Jul. 2013, pp. 3–12.

[14] M. Almulla, Y. Wang, A. Boukerche, and Z. Zhang, "Design of a fast location-based handoff scheme for IEEE 802.11 vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 8, pp. 3853–3866, Oct. 2014.

[15] M. Mouton, G. Castignani, R. Frank, and T. Engel, "Enabling vehicular mobility in city-wide IEEE 802.11 networks through predictive handovers," *Veh. Commun.*, vol. 2, no. 2, pp. 59–69, Apr. 2015.

[16] M. H. Dwijaksara, H. S. Oh, M. Hwang, W. S. Jeon, and D. G. Jeong, "Crowdsourcing-based handover scheme for wireless LAN in vehicular environments," in *Proc. IEEE APWCS*, Incheon, South Korea, Aug. 2017.

[17] M. H. Dwijaksara, M. Hwang, W. S. Jeon, and D. G. Jeong, "Design and implementation of a fast handoff scheme supporting vehicular mobility over IEEE 802.11 WLAN," in *Proc. ACM SAC*, Marrakesh, Morocco, Apr. 2017, pp. 634–641.

[18] Y. Lue, M. J. Lee, and Y. Zheng, "Adaptive multi-resource allocation for cloudlet-based mobile cloud computing system," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2398–2410, Oct. 2016.

[19] V. S. Rao and R. Gajula. (2010). Interoperability in LTE. Continuous Computing, San Diego, CA, USA. Accessed: May 12, 2018. [Online]. Available: https://www.slideshare.net/allabout4g/interoperability-in-lte

[20] Mathworld. *Circle Line Intersection*. Accessed: May 12, 2018. [Online]. Available: http://mathworld.wolfram.com/Circle-LineIntersection.html

[21] *IBM ILOG CPLEX*. Accessed: May 12, 2018. [Online]. Available: http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer

[22] *Lp_Solve 5.5*. Accessed: May 12, 2018. [Online]. Available: http://lpsolve.sourceforge.net/5.5/

[23] S. Coast. *Open Street Map (OSM)*. Accessed: May 12, 2018. [Online]. Available: https://www.open-streetmap.org

[24] German Aerospace Center. *Simulation of Urban Mobility*. Accessed: May 12, 2018. [Online]. Available: http://www.sumo.dlr.de/userdoc/

[25] R. Porat, S. K. Yong, and K. Doppler. *TGah Channel Model*, IEEE Standard 802.11-11/0968r4, Mar. 2015.

[26] W. C. Jakes, *Microwave Mobile Communications*. Piscataway, NJ, USA: IEEE Press, 1994.

[27] H. Nishiyama, M. Ito, and N. Kato, "Relay-by-smartphone: Realizing multihop device-to-device communications," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 56–65, Apr. 2014.

MADE HARTA DWIJAKSARA received the B.S. degree in informatics engineering from the Institut Teknologi Bandung, Indonesia, in 2008, and the M.S. degree in computer science from the Korea Advanced Institute of Science and Technology, South Korea, in 2011.

He is currently pursuing the Ph.D. degree with the Department of Computer Science, Seoul National University, South Korea. From 2011 to 2012, he was a Network Engineer with the Lotte Data Communication Company. His research interests include wireless local area networks and Internet of Things.

WHA SOOK JEON (M'90–SM'01) received the B.S., M.S., and Ph.D. degrees in computer engineering from Seoul National University, Seoul, South Korea, in 1983, 1985, and 1989, respectively.

From 1989 to 1999, she was with the Department of Computer Engineering, Hansung University, Seoul. In 1999, she joined Seoul National University, as a Faculty Member, where she is currently a Professor with the School of Electrical Engineering and Computer Science. Her research interests include resource management for wireless and mobile networks, mobile communications systems, high-speed networks, communication protocols, and network performance evaluation.

Dr. Jeon served on the Editorial Board for the *Journal of Communications and Networks* from 2002 to 2017.

DONG GEUN JEONG (S'90–M'93–SM'99) received the B.S., M.S., and Ph.D. degrees from Seoul National University, Seoul, South Korea, in 1983, 1985 and 1993, respectively.

From 1986 to 1990, he was a Researcher with the Research and Development Center, DACOM, South Korea. From 1994 to 1997, he was with the Research and Development Center, Shinsegi Telecomm Inc., South Korea, where he conducted and led research on advanced cellular mobile networks. In 1997, he joined the Hankuk University of Foreign Studies, Seoul, South Korea, as a Faculty Member, where he is currently a Professor with the Department of Electronics Engineering. His research interests include resource management for wireless and mobile networks, mobile communications systems, communication protocols, and network performance evaluation.

Dr. Jeong served on the Editorial Board for the *Journal of Communications and Networks* from 2002 to 2007.