Optimal Placement of Gateways in Vehicular Networks

Pan Li, Xiaoxia Huang, Yuguang Fang, and Phone Lin

Abstract—In vehicular networks, mobile users in vehicles can access the Internet and enjoy various kinds of services such as Voice over IP (VoIP) and online streaming video. To provide satisfactory quality of service (QoS), we need to appropriately place gateways to link mobile nodes to the Internet or any other information networks. In this paper, we address the problem of optimally placing gateways in vehicular networks to minimize the average number of hops from access points (APs) to gateways, so that the communication delay can be minimized. Moreover, we also investigate how to minimize the total power consumption, and we show that, when the average number of hops is minimized, the average capacity of each AP can be maximized.

Index Terms—Delay, gateways, power consumption, vehicular networks.

I. INTRODUCTION

WIRELESS networks have recently gained more popularity due to low cost and easy and quick deployment. Many users are using wireless networks in their daily lives. For example, in their own homes, users can surf the Internet by accessing an IEEE 802.11 compliant wireless router that is connected to a modem [1]. Recently, new kinds of wireless networks, which are called vehicular networks, have drawn intensive attention. Typically, in vehicular networks, some wireless devices such as routers or sensors are deployed on the roadside, acting as access points (APs). Some APs are connected to gateways with direct links to the Internet such that travelers in vehicles can enjoy various Internet services. We can also mount some cameras on the APs to monitor the traffic situation and to react to accidents quickly. In addition, vehicles can be equipped with some wireless device and, thus, can communicate with each other as well as with roadside APs. The resulting networks are called Vehicular *Ad hoc* NETworks (VANETs). The Federal Communications Commission has already allocated 75 MHz of spectrum at 5.9 GHz for dedicated short-range communications from the roadside to vehicles and in between vehicles [2].

In this paper, we consider the scenario where roadside APs can cover the whole road so that every vehicle is associated with a roadside AP within one hop. In this scenario, no matter how many vehicles there are, the mobile users in vehicles are always able to access the Internet.

Vehicular networks can be 1-D, such as a network along a highway, or 2-D, such as a network covering all the roads in an area. In both cases, for users to access the Internet, we need to place gateways in the network. It is obvious that making every AP connected to a gateway is neither unnecessary nor practical most of the time. Instead, we may want to place just some gateways in the network. Thus, the problem is how to optimally place the gateways in terms of some specific requirements. Similar problems also exist in many other different types of networks, such as wireless sensor networks and wireless mesh networks. Here, in this paper, we study the problem of optimally placing gateways in both 1-D and 2-D vehicular networks. Our objective is to place the gateways in such a way that the average number of hops from APs to gateways can be minimized. The reason we are interested in this formulation is that the communication delay in multihop VANETs is closely related to the number of wireless hops, and the capacity of each AP can be maximized when the average number of hops from APs to gateways is minimized.

More specifically, in 1-D vehicular networks, we find the optimal placement of gateways to minimize the average number of hops from APs to gateways when there is only one gateway and when there are multiple gateways, respectively. We also find the optimal placement of APs to minimize the total power consumption. In 2-D vehicular networks, we investigate two cases. First, for a dense network, we determine the optimal positioning of one gateway. Thus, for those large networks, we can first partition them into small dense clusters and then optimally place one gateway in each cluster. Second, for grid (2-D) networks, we discover the optimal positioning of multiple gateways. For both cases, we have attempted to minimize the power consumption. As we notice, the techniques that are used here can also be applied to the problem in 3-D networks.

P. Li and X. Huang are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA (e-mail: lipanleo@ufl.edu; xiaoxiah@ufl.edu).

Y. Fang is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA (e-mail: fang@ece.ufl.edu).

P. Lin is with the Department of Computer Science and Information Engineering and Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: plin@csie.ntu.edu.tw).

As a final remark, we investigate the power consumption issue at APs in vehicular networks for several reasons. First, along many highways in North America, there are no roadside lights, which means that there are probably no power lines to power the APs. Second, in some scenarios, for example, in areas for some events such as car racing, some temporary vehicular networks are needed. In these cases, the power consumption at APs does matter.

The rest of this paper is organized as follows. We present the related work in the next section. In Section III, we study the problem on how to optimally place gateways in 1-D vehicular networks. The problem for 2-D vehicular networks is investigated in Section IV. We elaborate why we want to minimize the average number of hops in Section V. We finally conclude this paper in Section VI.

II. RELATED WORK

Vehicular networks have emerged as a hot research topic lately. However, the problem of gateway placement in vehicular networks has not been well addressed in the literature.

In [4], Bejerano studies how to use gateways efficiently in *ad hoc* networks. In [8], Hwang *et al.* present an architecture of gateways in sensor networks to provide efficient data querying. Recently, Prasad and Wu [12] address the problem of gateway deployment optimization in cellular wireless-fidelity mesh networks. In their work, they find the optimal positioning of gateways to minimize the network installation costs while maintaining flexibility and an acceptable grade of service.

The problem of how to optimally place the gateways is similar to the facility location problem, which can be traced back to the 17th century. In 1629, Fermat proposed a problem: how to find a point such that the total distance from it to the three vertices of a triangle is minimized. Torricelli and Steiner studied this problem in 1644 and 1837, respectively, and developed some geometric approaches. This problem is known as the Fermat–Torricelli problem, and the optimal point is called Fermat's Point.

Later, this problem was generalized by Weber [15]. The generalized problem was formulated as a model for the optimal location of a facility in the plane that was intended to serve some users, such as the placement of a hospital or a fire station in a town. It was to minimize a function

$$f(x) = \sum_{i=1}^{n} w_i d(x - x_i) \qquad (1)$$

where w_i are positive scalars, x_i are given vectors in R^2, and $d(u)$ is the Euclidean norm of u. This problem is also known as the Steiner–Weber problem, where only one facility is considered.

Francis and White [5] studied the problem when there is only one facility as well as when there are multiple facilities. When there is only one facility, analytical results for the case that $d(\cdot)$ in (1) is the rectilinear distance are given. For the case that $d(\cdot)$ in (1) is the Euclidean distance, they gave an algorithm to find the optimal position. Francis and White
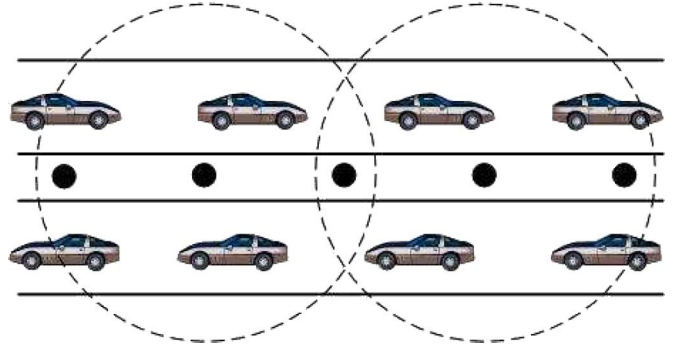

Fig. 1. Example for 1-D vehicular networks.

also presented some analytical results when there are multiple facilities. However, the objective function also accounts for the distance among the multiple facilities, which is not the case in our scenarios regarding the placement of gateways in wireless networks.

Moreover, in this paper, we want to minimize the average number of hops from APs to gateways. In 1-D networks, it is different from the facility problem where distance is taken into consideration, because the number of hops is not necessarily proportional to the distance. We further show in Section III that not all links are of the same length to minimize the total power consumption. We also derive the analytical results for placing multiple gateways in 1-D networks.

In 2-D networks, we propose two new algorithms called Finding the Optimal positioning of Gateway (FOG) and Finding the Optimal positioning of Gateway based on some Approximation (FOGA), respectively, to find the optimal positioning of a single gateway in a dense network. Thus, for large vehicular networks, we can first partition them into small dense clusters [3], [6], [7], [9] and then optimally place one gateway in each cluster. We also derive some analytical results for optimally placing multiple gateways in a grid (2-D) network. The minimization of power consumption is also discussed.

III. OPTIMAL PLACEMENT OF GATEWAYS IN 1-D VEHICULAR NETWORKS

There are many scenarios where APs are placed in 1-D space. For example, the APs can be placed along a road or a highway, as shown in Fig. 1. In this case, every AP is connected via wireless medium to two neighboring APs to ensure network connectivity, and some of them are connected via Ethernet cables to gateways to the Internet. Thus, vehicles on the road can access the Internet through these APs. In this section, we study the optimal placement of gateways such that the average number of hops from APs to gateways is minimized.

A. One Gateway

As shown in Fig. 2, there are n APs along a road numbered from 1 to n, one of which is connected to a gateway. We denote this AP by x. As mentioned before, AP x is connected to the gateway via an Ethernet cable. So, we consider the number of
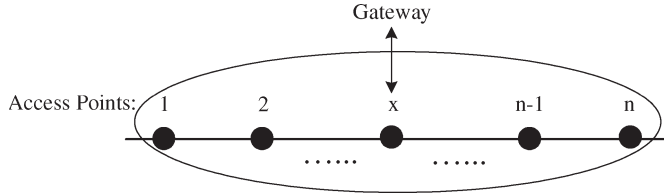
Gateway

Access Points: 1 2 x n-1 n

Fig. 2. Topology of a 1-D vehicular network with one gateway.

Gateway Gateway Gateway

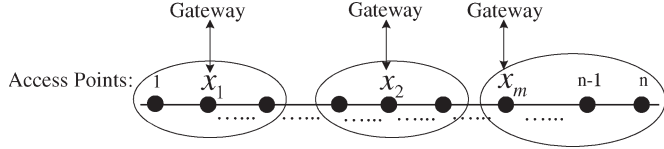Access Points: 1 x_1 x_2 x_m n-1 n

Fig. 3. Topology of a 1-D vehicular network with multiple gateways.

hops from x to the gateway to be 0. Let $H(x)$ be the average number of hops from all the APs to the gateway. Then, we obtain

$$H(x) = \frac{1}{n} \cdot [(x-1) + (x-2) + \cdots + 1 + 0$$
$$+ 1 + 2 + \cdots + (n-x)]$$
$$= \frac{1}{n} \cdot \left[\left(x - \frac{n+1}{2} \right)^2 + \frac{n^2-1}{4} \right].$$

To minimize $H(x)$, we have

$$x = \left\lfloor \frac{n+1}{2} \right\rfloor \quad \text{or} \quad x = \left\lceil \frac{n+1}{2} \right\rceil \quad (2)$$

and

$$H_{\min}(x) = \begin{cases} (n^2-1)/4n, & n \text{ is odd} \\ n/4, & n \text{ is even} \end{cases}$$

where $H_{\min}(x)$ is the minimum value of $H(x)$. $H_{\min}(x)$ can be rewritten as

$$H_{\min}(x) = \frac{1}{n} \left\lceil \frac{n^2-1}{4} \right\rceil = \frac{1}{n} \left\lfloor \frac{n^2}{4} \right\rfloor. \quad (3)$$

From (2) and (3), we discover that the minimum average number of hops from APs to the gateway is on the order of $O(n)$ and is achieved when the gateway is connected to the median AP $\lfloor (n+1)/2 \rfloor$.

B. Multiple Gateways

Next, we turn to the situation when there exist multiple gateways. As shown in Fig. 3, there are n APs, $m(m > 1)$ of which are connected to gateways. We denote these m APs by x_i, where $1 \leq i \leq m$ and $1 \leq x_i \leq n$. In this network, every AP always transmits packets to and receives packets from the nearest gateway. Assuming that there are n_i APs

routing through the gateway that is connected to AP x_i, then $\sum_{i=1}^{m} n_i = n$. We have

$$H(x) = \frac{1}{n} \sum_{i=1}^{m} n_i H_i(x_i)$$

where $H_i(x_i)$ is the average number of hops from the n_i APs to their gateway connected to AP x_i, and $H(x)$ is the average number of hops from all APs to their gateways.

1) When $C = n/m$ is an odd integer, then

$$H(x) \geq \frac{1}{n} \left\{ \left\lceil \frac{n_1^2 - 1}{4} \right\rceil + \left\lceil \frac{n_2^2 - 1}{4} \right\rceil + \cdots + \left\lceil \frac{n_m^2 - 1}{4} \right\rceil \right\}$$
$$\geq \frac{1}{n} \left\{ \frac{n_1^2 + n_2^2 + \cdots + n_m^2}{4} - \frac{m}{4} \right\}$$
$$\geq \frac{m}{4n} \cdot \left(\frac{n_1 + n_2 + \cdots + n_m}{m} \right)^2 - \frac{m}{4n}$$
$$= \frac{n}{4m} - \frac{m}{4n}.$$

The minimum value is achieved when $n_i = n/m = C$.

2) When $C = n/m$ is an even integer, then

$$H(x) \geq \frac{1}{n} \left\{ \left\lfloor \frac{n_1^2}{4} \right\rfloor + \left\lfloor \frac{n_2^2}{4} \right\rfloor + \cdots + \left\lfloor \frac{n_m^2}{4} \right\rfloor \right\}$$
$$\geq \frac{1}{n} \left\{ \frac{n_1^2 + n_2^2 + \cdots + n_m^2}{4} \right\}$$
$$\geq \frac{m}{4n} \cdot \left(\frac{n_1 + n_2 + \cdots + n_m}{m} \right)^2$$
$$= \frac{n}{4m}.$$

The minimum value is achieved when $n_i = n/m = C$.

3) When n/m is not an integer but there exists $k(1 \leq k \leq m)$ such that $C = (n-k)/m$ is an odd integer, then

$$H(x) \geq \frac{1}{n} \left\{ \left\lceil \frac{n_1^2 - 1}{4} \right\rceil + \left\lceil \frac{n_2^2 - 1}{4} \right\rceil + \cdots + \left\lceil \frac{n_{m-k}^2 - 1}{4} \right\rceil \right.$$
$$\left. + \left\lfloor \frac{n_{m-k+1}^2}{4} \right\rfloor + \left\lfloor \frac{n_{k+2}^2}{4} \right\rfloor + \cdots + \left\lfloor \frac{n_m^2}{4} \right\rfloor \right\}$$
$$\geq \frac{n_1^2 + n_2^2 + \cdots + n_m^2}{4n} - \frac{k}{4n}.$$

Since it is impossible to make $n_1 = n_2 = \cdots = n_m$, the minimum value is achieved when $n_i = C$, where $i \in S$ (S is a subset of $\{1, 2, \ldots, m\}$ that contains $m - k$ elements), and $n_j = C + 1$, where $j \in \bar{S}$, which is

$$H(x) \geq \frac{kC^2 + (m-k)(C+1)^2}{4n} - \frac{k}{4n}$$
$$= \frac{(n-k+m)(n-3k+m)}{4mn}.$$

4) When n/m is not an integer but there exists $k(1 \leq k \leq m)$ such that $C = (n - k)/m$ is an even integer, then

$$
\begin{aligned}
H(x) \geq \frac{1}{n} & \left\{ \left\lceil \frac{n_1^2 - 1}{4} \right\rceil + \left\lceil \frac{n_2^2 - 1}{4} \right\rceil + \cdots + \left\lceil \frac{n_k^2 - 1}{4} \right\rceil \right. \\
& \left. + \left\lfloor \frac{n_{k+1}^2}{4} \right\rfloor + \left\lfloor \frac{n_{k+2}^2}{4} \right\rfloor + \cdots + \left\lfloor \frac{n_m^2}{4} \right\rfloor \right\} \\
\geq & \frac{n_1^2 + n_2^2 + \cdots + n_m^2}{4n} - \frac{k}{4n}.
\end{aligned}
$$

The minimum value is achieved when $n_i = C + 1$, where $i \in S$ (S is a subset of $\{1, 2, \ldots, m\}$ that contains k elements), and $n_j = C$, where $j \in \bar{S}$, which is

$$
\begin{aligned}
H(x) & \geq \frac{k(C+1)^2 + (m-k)C^2}{4n} - \frac{k}{4n} \\
& = \frac{n^2 - k^2}{4mn}.
\end{aligned}
$$

5) When $m > n$, obviously, $H_{\min}(x) = 0$, because every AP can be connected to one gateway via a wired line.

From the results that we have obtained so far, we conclude that the minimum average number of hops from APs to gateways, i.e., $H_{\min}(x)$, is on the order of $O(n/m)$. In other words, $H_{\min}(x)$ will approach infinity as $n \to \infty$, unless m increases as fast as n. When m increases faster than n, $H_{\min}(x)$ will be 0.

C. Power Consumption

In this section, we study the problem of power consumption at APs. Our objective is to minimize the total power consumption of APs by carefully adjusting the distances among them after we obtain the optimal positioning of the gateways.

We first present the power propagation model that we will use in our study [13], i.e.,

$$
P_r(d) = P_t h(G_t, G_r, h_t, h_r, L, \lambda) \frac{1}{d^\gamma} \tag{4}
$$

where P_t and P_r are the transmitted power and the received power, respectively, G_t and G_r are the antenna gains for the transmitter and the receiver, respectively, h_t and h_r are the antenna heights of the transmitter and the receiver, respectively, d is the distance between the transmitter and the receiver, $L(L \geq 1)$ is the system loss factor that is not related to propagation, λ is the wavelength, $h(\cdot)$ is a function, and γ is the path loss exponent.

Assume that the APs are in the same wireless environment, and hence, they can successfully receive packets if the received powers are above the same threshold P_r^{thresh} and every AP can use power control to make the received power exactly P_r^{thresh} to save power. Moreover, we assume that the transmitter's antenna gain G_t and the receiver's antenna gain G_r are the same for all transmissions, respectively. The antenna heights of all transmitters and receivers are the same as well for analysis

purposes. Thus, according to (4), for a transmission of distance d, transmission power P_t is

$$
P_t = P_r^{\text{thresh}} d^\gamma \cdot T' = d^\gamma T
$$

where T' is a constant that is determined by G_t, G_r, h_t, h_r, L, and λ.

Denote an arbitrary AP that is not connected to a gateway directly by k. Then, power consumption P_k for relaying one packet from AP k to one gateway through other APs is

$$
P_k = \sum_{i=1}^{j} P_{t,i} = \sum_{i=1}^{j} d_i^\gamma \cdot T
$$

where $P_{t,i}$ is the transmission power for the ith hop from AP k to the gateway, and $1 \leq k \leq n$, $1 \leq j \leq n - 1$. For an AP that is connected to a gateway directly, we can assume that the power consumption is not a major concern and, hence, is negligible.

As a result, the total power consumption P for all the APs to relay one packet to a gateway is

$$
P = \sum_{k=1, \notin \{x_i\}}^{n} P_k \tag{5}
$$

where $\{x_i\}$ is the set of APs that are connected to gateways. In what follows, we further investigate the problem in a few cases.

1) One Gateway: We first consider the situation that there is only one gateway. As shown in Fig. 2, APs are denoted by $1, 2, \ldots, n$, and the gateway is connected to the AP x, which is, at the optimal position, derived in Section III-A. Let d_i be the distance between AP i and AP $i + 1$, where $1 \leq i \leq n - 1$. Then, (5) will lead to

$$
\begin{aligned}
P = & \left(d_1^\gamma + d_2^\gamma + \cdots + d_{x-1}^\gamma \right) \cdot T \\
& + \left(d_2^\gamma + \cdots + d_{x-1}^\gamma \right) \cdot T + \cdots \\
& + d_{x-1}^\gamma \cdot T + d_x^\gamma \cdot T + \cdots \\
& + \left(d_x^\gamma + d_{x+1}^\gamma + \cdots + d_{n-1}^\gamma \right) \cdot T \\
= & \left[d_1^\gamma + 2d_2^\gamma + \cdots + (x-1)d_{x-1}^\gamma \right. \\
& \left. + d_{n-1}^\gamma + 2d_{n-2}^\gamma + \cdots + (n-x)d_x^\gamma \right] \cdot T.
\end{aligned}
$$

As a result, we obtain

$$
P \geq (n-1) \left\{ \left(\prod_{i=1}^{x-1} i d_i^\gamma \right) \left(\prod_{j=1}^{n-x} j d_{n-j}^\gamma \right) \right\}^{\frac{1}{n-1}} T.
$$

The equality holds when $i d_i^\gamma = j d_{n-j}^\gamma$ for all $i \in [1, x-1]$ and $j \in [1, n-x]$. Since

$$
\sum_{i=1}^{x-1} d_i + \sum_{j=1}^{n-x} d_j = L
$$

where L is the length of the road of interest, then the minimum value of P can be achieved when

$$
d_i = \begin{cases}
\left(\sum_{i=1}^{x-1} i^{-\frac{1}{\gamma}} \right. \\
\left. + \sum_{i=x}^{n-1} (n-i)^{-\frac{1}{\gamma}} \right)^{-1} L, & \text{for } i = 1 \\
i^{-\frac{1}{\gamma}} d_1, & \text{for } i \in [2, x-1] \\
(n-i)^{-\frac{1}{\gamma}} d_1, & \text{for } i \in [x, n-1]
\end{cases}.
$$

From this result, we conclude that to minimize the total power consumption P, the links that are closer to the gateway should have shorter distances, which is consistent with our intuition, because these links have more packets to transmit.

2) Multiple Gateways: Next, we consider the situation that there are multiple gateways, as shown in Fig. 3. The APs are denoted by $1, 2, \ldots, n$, and $m(m > 1)$ of them are connected to gateways, which are denoted by $\{x_i\}$, where $i \in [1, m]$, $x_i \in [1, n]$, and which are situated at the optimal positions that were derived in Section III-B.

Assume that every AP relays packets to the nearest gateway in terms of the number of hops. Let S_i denote the set of APs relaying packets to the gateway that is connected to AP x_i. Then

$$
S_i = [t_{i-1} + 1, t_i], \qquad \text{for } 1 \le i \le m
$$

where

$$
t_i = \begin{cases}
0, & \text{for } i = 0 \\
\lfloor (x_i + x_{i+1})/2 \rfloor, & \text{for } i \in [1, m-1] \\
n, & \text{for } i = m
\end{cases}.
$$

Let d_i be the distance between AP i and AP $i+1$, where $i \in [1, n-1]$. Then, (5) results in

$$
\begin{aligned}
P = & \left[d_1^\gamma + 2d_2^\gamma + \cdots + (x_1 - 1)d_{x_1-1}^\gamma \right. \\
& \left. + d_{t_1-1}^\gamma + 2d_{t_1-2}^\gamma + \cdots + (t_1 - x_1)d_{x_1}^\gamma \right] \cdot T \\
& + \left[d_{t_1+1}^\gamma + 2d_{t_1+2}^\gamma + \cdots + (x_2 - 1)d_{x_2-1}^\gamma \right. \\
& \left. + d_{t_2-1}^\gamma + 2d_{t_2-2}^\gamma + \cdots + (t_2 - x_2)d_{x_2}^\gamma \right] \cdot T + \cdots \\
& + \left[d_{t_{m-1}+1}^\gamma + 2d_{t_{m-1}+2}^\gamma + \cdots + (x_m - 1)d_{x_m-1}^\gamma \right. \\
& \left. + d_{n-1}^\gamma + 2d_{n-2}^\gamma + \cdots + (n - x_m)d_{x_m}^\gamma \right] \cdot T.
\end{aligned}
$$

So, we have

$$
P \ge (n-1) \prod_{k=1}^{m} \left(\left(\prod_{i=1}^{x_k-1} i d_{t_{k-1}+i}^\gamma \right) \left(\prod_{j=1}^{t_k-x_k} j d_{t_k-j}^\gamma \right) \right)^{\frac{1}{n-1}} T.
$$

The equality holds when $i d_{t_{k-1}+i}^\gamma = j d_{t_l-j}^\gamma$ for all $i \in [1, x_k - 1]$, $j \in [1, t_k - x_k]$, and $k, l \in [1, m]$. Since

$$
\sum_{k=1}^{m} \left(\sum_{i=1}^{x_k-1} d_{t_{k-1}+i} + \sum_{j=1}^{t_k-x_k} d_{t_k-j} \right) = L
$$

where L is the length of the road, then we have

$$
\sum_{k=1}^{m} \left(\sum_{i=1}^{x_k-1} i^{-\frac{1}{\gamma}} + \sum_{j=x_k}^{t_k-x_k} (t_k - j)^{-\frac{1}{\gamma}} \right) d_1 = L.
$$

Thus, the minimum value of P can be achieved given the expression for d_i, shown at the bottom of the page, where $k \in [0, m]$.

From this result, we conclude that, to minimize total power consumption P, the links that are closer to the gateways should have shorter distances.

IV. OPTIMAL PLACEMENT OF GATEWAYS IN 2-D VEHICULAR NETWORKS

In this section, we study the problem of finding the optimal placement of gateways in 2-D vehicular networks such that the average number of hops from APs to the gateways is minimized.

A. Dense Network

Generally speaking, it is very difficult to find the optimal positioning of multiple gateways in a large 2-D network. However, for some cases, we can first partition a large network into several small dense clusters and then determine the optimal placement of one gateway for each cluster. One example is given in Fig. 4. In this section, we attempt to find the optimal position of one gateway in a dense network.

Denote the location of AP i by (x_i, y_i), where $1 \le i \le n$. Let $g = (x, y)$ denote the location of the gateway and D denote the area where APs are placed. Here, we assume that the gateway is not necessarily connected to an AP via an Ethernet cable. Instead, it can be put it anywhere in the network. Then, the average number of hops from APs to the gateway, which is denoted by $H(g)$, satisfies

$$
H(g) \ge \frac{1}{n} \sum_{i=1}^{n} \left\lceil \frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{r(n)} \right\rceil \tag{6}
$$

where $r(n)$ is the common transmission range for all n APs that can assure connectivity. The equality holds when every

$$
d_i = \begin{cases}
\left(\sum_{k=1}^{m} \left(\sum_{i=1}^{x_k-1} i^{-\frac{1}{\gamma}} + \sum_{j=x_k}^{t_k-x_k} (t_k - j)^{-\frac{1}{\gamma}} \right) \right)^{-1} L, & \text{for } i = t_k + 1 \\
i^{-\frac{1}{\gamma}} d_1, & \text{for } i \in [t_k + 2, x_{k+1} - 1] \\
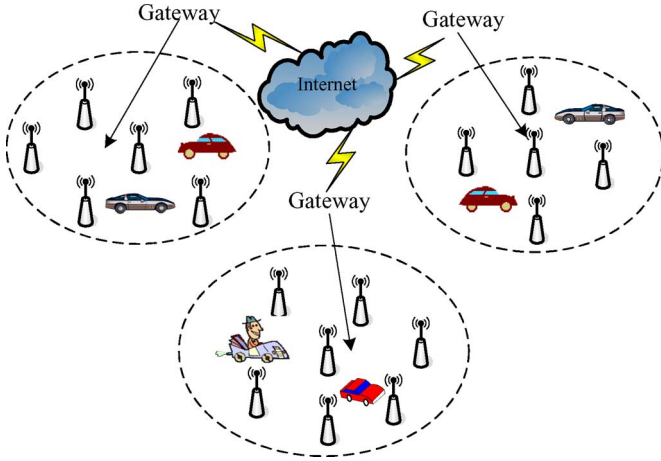(t_{k+1} - i)^{-\frac{1}{\gamma}} d_1, & \text{for } i \in [x_{k+1}, t_{k+1} - 1]
\end{cases}
$$

Fig. 4. One example for partitioning a large 2-D vehicular network into three small dense clusters.

AP can transmit packets to the gateway along a path, which is approximately a straight line.

For a dense network, we can approximately assume that the equality in (6) holds. Thus, to minimize $H(g)$, it is equivalent to

$$\text{Minimize}\left\{ f(x,y) = \sum_{i=1}^{n} \sqrt{(x-x_i)^2 + (y-y_i)^2} \right\}$$

$$\text{s.t.} (x,y) \in D.$$

Taking the partial derivative of $f(x,y)$ with respect to x and y, respectively, we obtain

$$\frac{\partial f(x,y)}{\partial x} = \sum_{i=1}^{n} \frac{2(x-x_i)}{\sqrt{(x-x_i)^2 + (y-y_i)^2}} \tag{7}$$

$$\frac{\partial f(x,y)}{\partial y} = \sum_{i=1}^{n} \frac{2(y-y_i)}{\sqrt{(x-x_i)^2 + (y-y_i)^2}}. \tag{8}$$

To determine the optimal position of the gateway, we need

$$\frac{\partial f(x,y)}{\partial x} = 0 \quad \frac{\partial f(x,y)}{\partial y} = 0. \tag{9}$$

Notice that (7) and (8) are undefined if, for any i, $(x,y) = (x_i,y_i)$. One modification of the partial derivative approach that is shown in (9) is proposed by Kuhn [10], which is defined as follows:

$$R(x,y) = (0,0) \tag{10}$$

where

$$R(x,y) = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y} \right)$$

if $(x,y) \neq (x_i,y_i)$ for $i = 1,2,\ldots,n$, and

$$R(x,y) = \begin{cases} \left(\frac{u_k-1}{u_k} s_k, \frac{u_k-1}{u_k} t_k \right), & \text{if } u_k > 1 \\ (0,0), & \text{if } u_k \leq 1 \end{cases}$$

if $(x,y) = (x_k,y_k)$ for some k, $k \in [1,n]$, and

$$s_k = \sum_{\substack{i=1 \\ \neq k}}^{n} \frac{(x_k - x_i)}{\sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}}$$

$$t_k = \sum_{\substack{i=1 \\ \neq k}}^{n} \frac{(y_k - y_i)}{\sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}}$$

$$u_k = \left(s_k^2 + t_k^2 \right)^{1/2}.$$

Consider a complex plane with imaginary unit j. Let $z_i = x_i + jy_i$, where $1 \leq i \leq n$, denote the positions of the APs and $z = x + jy$ denote the position of the gateway. Then, $\overrightarrow{zz_i} = \overrightarrow{z_i} - \overrightarrow{z}$ is a vector pointing to point z_i from point z. As a result, when $z \neq z_i$ for $i = 1,2,\ldots,n$, (10) can be reformulated as

$$\sum_{i=1}^{n} (\overrightarrow{zz_i}/|\overrightarrow{zz_i}|) = 0$$

i.e.,

$$\sum_{i=1}^{n} \cos \alpha_i + j \sum_{i=1}^{n} \sin \alpha_i = 0 \tag{11}$$

where $\alpha = \arg \overrightarrow{zz_i}$, i.e., the augment of $\overrightarrow{zz_i}$, and

$$\cos \alpha_i = \frac{\text{Re}(\overrightarrow{zz_i})}{|\overrightarrow{zz_i}|} \quad \sin \alpha_i = \frac{\text{Im}(\overrightarrow{zz_i})}{|\overrightarrow{zz_i}|} \tag{12}$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real and imaginary parts of a complex number, respectively.

We can interpret (11) in the following way. Imagine that, on the horizontally placed plane D, there is an object, which is subject to n unit forces, i.e., with the magnitude of 1, coming from positions $z_i (1 \leq i \leq n)$. According to *the principle of minimum total potential energy* [14], the object will displace to a position that minimizes its potential energy, where it achieves equilibrium, i.e., (11) holds. Thus, we arrive at the algorithm for finding the optimal position of gateway (FOG) shown in Table I. One example with six nodes is shown in Fig. 5.

We have evaluated the performance of FOG by simulations. We generate nodes in an area of 10×10 according to the Poisson Point process and then find the optimal positioning using the FOG algorithm. In the simulation, we set the parameter $Threshold$ to be equal to the parameter $step$. The iteration times, i.e., the numbers of steps from the beginning point to the optimal point, are shown in Fig. 6. In Fig. 6(a), we keep the parameter $step$ fixed and increase the Poisson intensity λ from 1 to 10. We find that the number of iterations first decreases as λ increases. This indicates that, when the number of nodes increases, the optimal point approaches the average point $(\sum_{i=1}^{n} x_i/n, \sum_{i=1}^{n} y_i/n)$. We observe that the number of iterations increases after $\lambda \geq 7$. We contend that this is because

<div style="text-align:center">

TABLE I
ALGORITHM FOR FOG

</div>

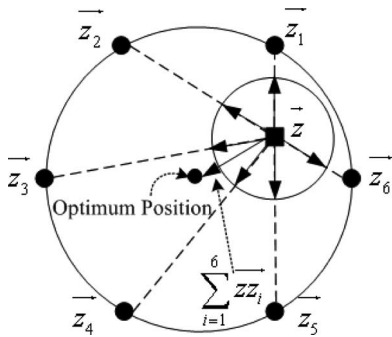|	
01.	Initialization: We begin to search from this position: $x = \sum_{i=1}^{n} \frac{x_i}{n}$, $y = \sum_{i=1}^{n} \frac{y_i}{n}$.
02.	Iteration: compare (x, y) with (x_i, y_i).
03.	IF $(x, y) \neq (x_i, y_i)$, for $i = 1, ..., n$,
04.	calculate $Scos = \sum_{i=1}^{n} \cos \alpha_i$, and $Ssin = \sum_{i=1}^{n} \sin \alpha_i$, where $\cos \alpha_i = \frac{(x_i - x)}{\sqrt{(x-x_i)^2 + (y-y_i)^2}}$, and $\sin \alpha_i = \frac{(y_i - y)}{\sqrt{(x-x_i)^2 + (y-y_i)^2}}$.
05.	ELSE IF $(x, y) = (x_k, y_k)$, for some $k \in [1, n]$,
06.	IF $u_k > 1$, where u_k is defined in (11),
07.	calculate $Scos = \frac{u_k - 1}{u_k} \sum_{i=1, \neq k}^{n} \cos \alpha_i$, and $Ssin = \frac{u_k - 1}{u_k} \sum_{i=1, \neq k}^{n} \sin \alpha_i$.
08.	ELSE
09.	$Scos = 0$, and $Ssin = 0$.
10	END IF
11.	END IF
12.	IF $|Scos| \leq Threshold$ and $|Ssin| \leq Threshold$
13.	Go to 18.
14.	ELSE
15.	Recalculate x and y: $x = x + step \cdot Scos$, and $y = y + step \cdot Ssin$
16.	Go to 2.
17.	END IF
18.	Finish.

Fig. 5. Example to illustrate the operation of FOG.

the parameter *step* is too big for the algorithm to efficiently converge to the optimal point. This is clearly shown in Fig. 6(b), where the number of iterations decreases as *step* decreases from 0.01 to 0.006. The number of iterations increases again when *step* decreases further, because, at that time, *step* is too small for the algorithm to efficiently converge to the optimal point. As mentioned in [5], the value of $f(x, y)$ is relatively flat in the region near the optimal location. Thus, a very good solution is usually obtained if we place the gateway close to the optimal position. In other words, we can obtain a very good solution by choosing appropriate parameters in the algorithm, i.e., *step* and *Threshold*.

From the simulations, we discover that, when the nodes are distributed in an area according to the Poisson point process, the optimal position approaches the average position, i.e., $x \to \sum_{i=1}^{n} x_i/n$, and $y \to \sum_{i=1}^{n} y_i/n$, as the number of nodes n increases. We also find that the parameter *step* has an effect on the FOG algorithm. Moreover, FOG is very efficient since the results are calculated on a time scale of 0.01 s on a computer with one 1.6-GHz central processing unit and one 512-MB random access memory.

So far, we have discussed the alternative approach that was proposed by Kuhn [10]. Next, we present another modification of the partial derivative approach that was proposed in [5], i.e.,

$$T(x, y) = (0, 0)$$

where

$$T(x, y) = \left(\sum_{i=1}^{n} \frac{(x - x_i)}{\sqrt{(x - x_i)^2 + (y - y_i)^2 + \epsilon}} \right.$$
$$\left. \sum_{i=1}^{n} \frac{(y - y_i)}{\sqrt{(x - x_i)^2 + (y - y_i)^2 + \epsilon}} \right) \quad (13)$$

and ϵ is an arbitrarily small positive constant. We can see that (13) is always defined, and when $\epsilon \to 0$, $T(x, y) \to ((\partial f(x, y)/\partial x), (\partial f(x, y)/\partial y))$, i.e., the original definition of partial derivative.

Similar to that previously discussed, we also develop the algorithm for finding the optimal position of gateway based on some approximation (FOGA), as shown in Table II.

We have carried out simulations to evaluate the performance of FOGA. We generate nodes in an area of 10×10 according to the Poisson Point process, with an intensity of 10. The parameter *step* is set to be equal to the parameter *Threshold*, which is 0.01. The iteration times, i.e., the numbers of steps from the beginning point to the optimal point, with different values of ϵ, are shown in Fig. 7. We can see that, when ϵ decreases, the results of FOGA converge to those of FOG. Moreover, when $\epsilon = 0.0001$, FOGA attains the same results as FOG.

B. Grid Network

There are some scenarios where a 2-D vehicular network is a grid network. One example is shown in Fig. 8, where a vehicular network is deployed to cover several intersecting roads. Many services such as Voice over IP, Mobile TV, or e-mail services can be supported. The topology of this network with two gateways is shown in Fig. 9. Assume that there are a total of n APs and that m of them are connected to gateways. For analysis purposes, we assume that all APs use the same transmission range $r(n)$, which is equal to the length of one cell in the grid. We need to find the optimal positioning of these gateways so that the average number of hops from APs to gateways is minimized.

Denote the APs by $z_i = x_i + jy_i$ and the gateways by $g_k = p_k + jq_k$, where $1 \leq i \leq n$ and $1 \leq k \leq m$. Each gateway is connected to one AP via an Ethernet cable, as shown in Fig. 9. Thus, $\{p_k, k = 1, 2, \ldots, m\} \subseteq \{x_i, i = 1, 2, \ldots, n\}$, and $\{q_k, k = 1, 2, \ldots, m\} \subseteq \{y_i, i = 1, 2, \ldots, n\}$. Assuming that the routing protocol chooses the path with the lowest number of hops, then the number of hops between AP z_i and gateway g_k, which is denoted by $H_{i,k}$, is

$$H_{i,k} = \frac{1}{r_n} \left(|x_i - p_k| + |y_i - q_k| \right).$$

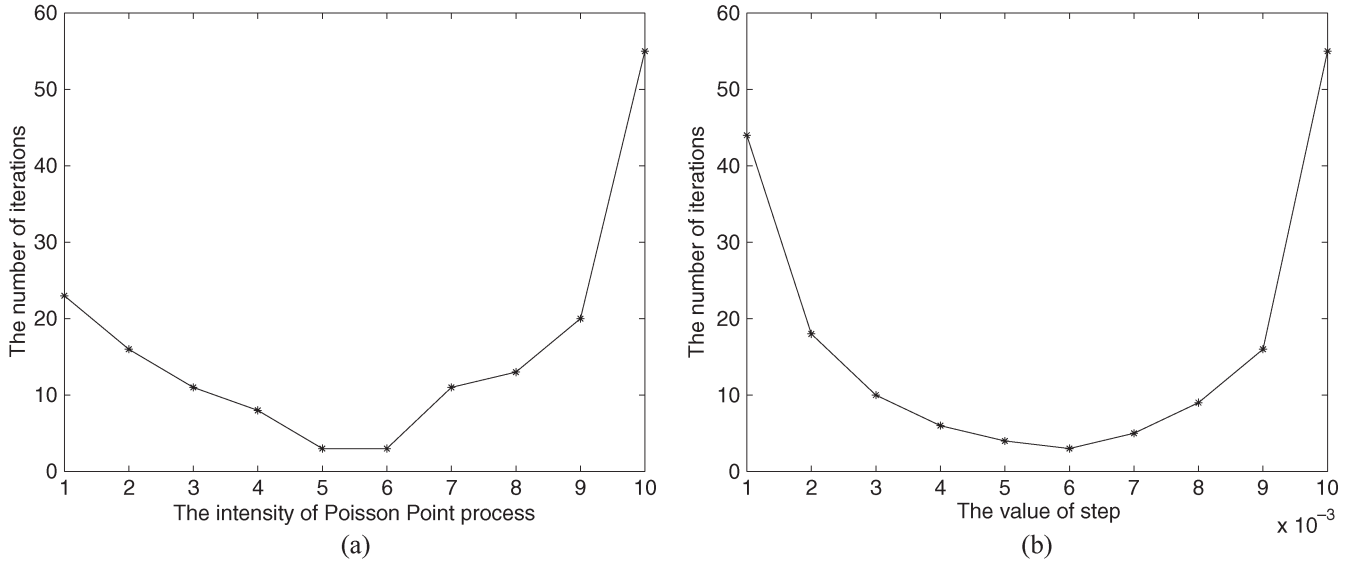(a)   (b)

Fig. 6.   (a) Number of iterations with $step = 0.01$ and intensity of Poisson point process increasing from 1 to 10. (b) Number of iterations with the intensity equal to 10 and $step$ increasing from 0.001 to 0.01.

TABLE II
ALGORITHM FOR FOGA

1.   Initialization: We begin to search from this position:
     $x = \sum_{i=1}^{n} \frac{x_i}{n}, y = \sum_{i=1}^{n} \frac{y_i}{n}$.
2.   Iteration: calculate $Scos = \sum_{i=1}^{n} \cos \alpha_i$, and
     $Ssin = \sum_{i=1}^{n} \sin \alpha_i$, where $\cos \alpha_i = \frac{(x_i - x)}{\sqrt{(x-x_i)^2 + (y-y_i)^2 + \epsilon}}$,
     and $\sin \alpha_i = \frac{(y_i - y)}{\sqrt{(x-x_i)^2 + (y-y_i)^2 + \epsilon}}$.
3.   IF $|Scos| \leq Threshold$ and $|Ssin| \leq Threshold$
4.   Go to 9.
5.   ELSE
6.   Recalculate x and y: $x = x + step \cdot Scos$, and
         $y = y + step \cdot Ssin$
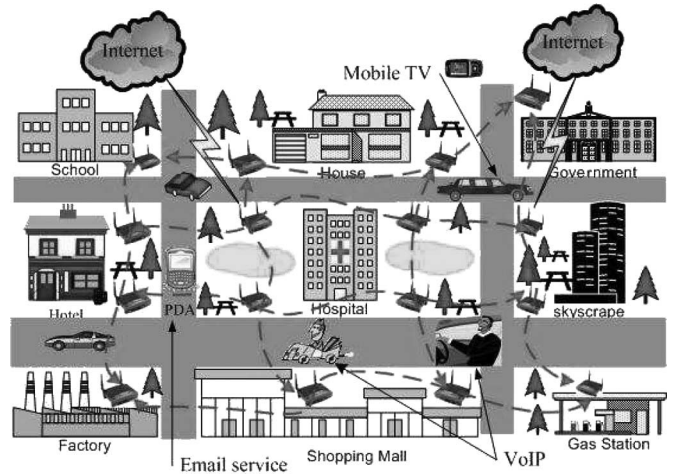7.   Go to 2.
8.   END IF
9.   Finish.

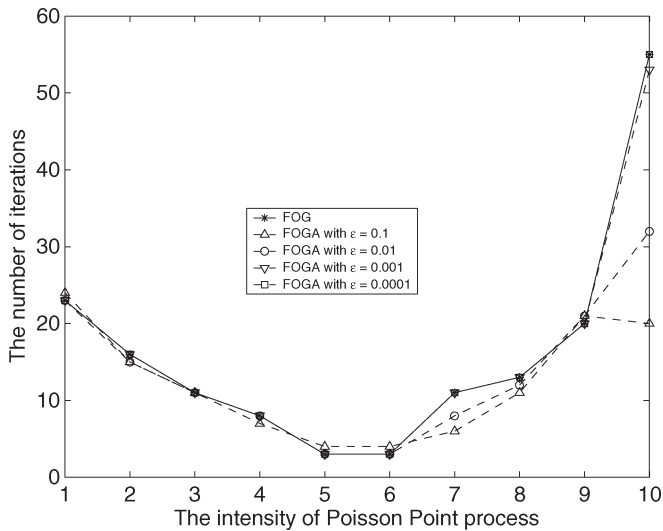Fig. 8.   One example for a 2-D grid vehicular network.

Fig. 7.   Number of iterations with $step = 0.01$ and intensity of Poisson point process increasing from 1 to 10. The results of FOG and FOGA with ϵ increasing from 0.0001 to 0.1 are shown in the figure.
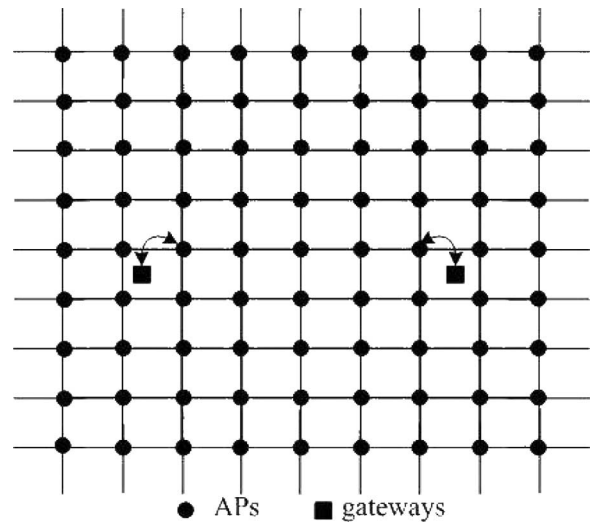

Fig. 9.   Topology of a 2-D grid vehicular network with two gateways.

Thus, the average number of hops from an AP to a gateway, which is denoted by $H(g_1, \ldots, g_m)$, is

$$H(g_1, \ldots, g_m) = \frac{1}{n} \sum_{k=1}^{m} \sum_{i \in S_k} H_{i,k}$$

$$= \frac{1}{nr(n)} \sum_{k=1}^{m} \sum_{i \in S_k} (|x_i - p_k| + |y_i - q_k|)$$

where $S_k (1 \leq k \leq m)$ is the set of APs that transmit packets through gateway g_k.

Since $r(n)$ is a constant, given the number of APs n, then

$$\underset{\{p_k, q_k\}, k=1, \ldots, m}{\text{Minimize}} H(g_1, \ldots, g_m)$$

$$\iff \underset{\{p_k\}, k=1, \ldots, m}{\text{Minimize}} \frac{1}{n} \sum_{k=1}^{m} \sum_{i \in S_k} |x_i - p_k|$$

$$+ \underset{\{q_k\}, k=1, \ldots, m}{\text{Minimize}} \frac{1}{n} \sum_{k=1}^{m} \sum_{i \in S_k} |y_i - q_k|.$$

As a result, finding optimal positioning for m gateways in a 2-D network can be transformed to the same problem in a 1-D network, which can be solved using approaches that are similar to those in Sections III-A and B.

C. Discussions

In both networks shown in Sections IV-A and B, respectively, the average transmission power P for transmitting one packet from an AP to a gateway is

$$P = P_t \cdot H(G) \propto r^\gamma(n) \cdot H(G)$$

where P_t is the transmission power, which is the same for all APs, G is the set of gateways in the network, $H(G)$ is the average number of hops from an AP to a gateway, and γ is the path loss exponent. Thus, when $H(G)$ is minimized, the average transmission power P is also minimized. Moreover, if the cost is proportional to the distance, then it is also proportional to the number of hops, because we assume that every hop has the same length. Thus, when $H(G)$ is minimized, the total cost is also minimized. Moreover, the techniques that are shown in this section can also be applied to the networks in 3-D spaces.

V. MORE DISCUSSIONS

In previous sections, we have found the optimal positioning of gateways in vehicular networks, so that the average number of hops from APs to gateways is minimized. There are a couple of additional benefits from our study.

First, the end-to-end delay can be reduced. Since the end-to-end delay in wireless networks is due to the contention, the processing and queueing at each intermediate node and the number of hops are a good metric for it. Thus, when the average number of hops is reduced, the end-to-end delay can also be reduced.

Second, the average capacity of APs can be maximized. Assume statistically that each AP generates (or receives) traffics at the same date rate, which is denoted by λ. We denote the total one-hop capacity of the network by C, which is the same as that defined in [11]. Then, we have

$$\lambda \cdot H(x) \cdot n \leq C \quad (14)$$

where $H(x)$ is the average number of hops, as we defined before, and n is the total number of APs in the network. From (14), we can obtain

$$\lambda \leq \frac{C}{nH(x)} \leq \frac{C}{nH_{\min}(x)}$$

where $H_{\min}(x)$ is the minimum average number of hops, as we derived before. We can easily observe that, when the average number of hops is minimized, the capacity of APs can be maximized.

VI. CONCLUSION

In this paper, we have addressed the problem of optimally placing one or multiple gateways in both 1-D and 2-D vehicular networks to minimize the average number of hops from APs to gateways. Thus, the communication delay can be decreased. Moreover, the capacity of each AP can be maximized. More specifically, in 1-D vehicular networks, we have obtained the analytical results for optimal placement of gateways. In 2-D vehicular networks, we have developed two algorithms called FOG and FOGA, respectively, to determine the optimal positioning of one gateway in dense networks. Simulation results show that our algorithms are efficient. We have also given some analytical results for finding the optimal placement of multiple gateways in 2-D vehicular grid networks and discussed how to minimize the total power consumption.

REFERENCES

[1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 1999. IEEE Standards Working Group.
[2] *Standard Specifications for Telecommunications and Information Exchange Between Roadside and Vehicle Systems—5 GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer Specifications*, 2003.
[3] A. Amis, R. Prakash, T. Vuong, and D. Huynh, "Max-min d-cluster formation in wireless ad hoc networks," in *Proc. 19th Annu. Joint Conf. IEEE Comput. Commun. Soc.*, Mar. 2000, pp. 32–41.
[4] Y. Bejerano, "Efficient integration of multi-hop wireless and wired networks with QoS constraints," in *Proc. ACM MobiCom*, Atlanta, GA, 2002, pp. 215–226.
[5] R. Francis and J. White, *Facility Layout and Location: An Analytical Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
[6] M. Gerla and T. Tsai, "Multicluster, mobile multimedia radio networks," *ACM Baltzer J. Wirel. Netw.*, vol. 1, no. 3, pp. 255–265, 1995.
[7] G. Gupta and M. Younis, "Load-balanced clustering of wireless sensor networks," in *Proc. IEEE Int. Conf. Commun.*, May 2003, pp. 1848–1852.
[8] K. Hwang, J. In, N. Park, and D. Eom, "A design and implementation of wireless sensor gateway for efficient querying and managing through World Wide Web," *IEEE Trans. Consum. Electron.*, vol. 49, no. 4, pp. 1090–1097, Nov. 2003.
[9] B. Kernigham and S. Lin, "An efficient heuristic procedure for partitioning graphs," *Bell Syst. Tech. J.*, vol. 49, no. 2, pp. 291–307, 1970.

[10] H. Kuhn, "On a pair of dual non-linear programs," in *Non-Linear Programming*. Amsterdam, The Netherlands: North-Holland, 1967, pp. 38–54.

[11] J. Li, C. Blake, D. S. D. Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *Proc. 7th Annu. Int. Conf. Mobile Comput. Netw.*, Rome, Italy, Jul. 2001, pp. 61–69.

[12] R. Prasad and H. Wu, "Gateway deployment optimization in cellular Wi-Fi mesh networks," *J. Netw.*, vol. 1, no. 3, pp. 31–39, 2006.

[13] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall PTR, 2002.

[14] J. N. Reddy, *Energy Principles and Variational Methods in Applied Mechanics*, 2nd ed. Hoboken, NJ: Wiley, 2002.

[15] A. Weber, *Theory of the Location of Industries*. Chicago, IL: Univ. Chicago Press, 1957.
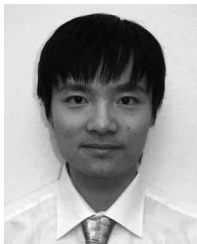
Pan Li received the B.S. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 2005. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville.

His research interests include medium access control, routing algorithms, and performance analysis in wireless networks.
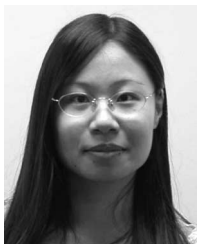
Xiaoxia Huang received the B.S. and M.S. degrees in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 2000 and 2002, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 2007.

She is currently with the Department of Electrical and Computer Engineering, University of Florida. Her research interests include mobile computing, QoS, and routing in wireless *ad hoc* networks and wireless sensor networks.
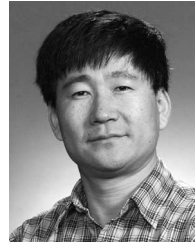
Yuguang Fang received the Ph.D. degree in systems engineering from Case Western Reserve University, Cleveland, OH, in 1994 and the Ph.D. degree in electrical engineering from Boston University, Boston, MA, in May 1997.

He was an Assistant Professor with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, from July 1998 to May 2000. He joined the Department of Electrical and Computer Engineering, University of Florida, Gainesville, as an Assistant Professor, in May 2000 and was promoted to Associate Professor with tenure in August 2003 and to Full Professor in August 2005. He holds a University of Florida Research Foundation Professorship that runs from 2006 to 2009. He has published more than 200 papers in refereed professional journals and conference proceedings. He received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He received the Best Paper Award at the IEEE International Conference on Network Protocols (ICNP) in 2006 and the IEEE TCGN Best Paper Award at the IEEE High-Speed Networks Symposium, IEEE Globecom, in 2002.

Prof. Fang has served on several editorial boards of technical journals including the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON MOBILE COMPUTING, and Association for Computing Machinery *Wireless Networks*. He has also been actively participating in professional conference organizations such as serving as the Steering Committee Co-Chair for QShine 2005, the Technical Program Vice Chair for the IEEE Conference on Computer Communications (INFOCOM) 2005, the Technical Program Symposium Co-Chair for IEEE Global Communications Conference 2004, and a member of the Technical Program Committee for IEEE INFOCOM (1998, 2000, and 2003–2008).

Phone Lin received the B.S. and Ph.D. degrees from National Chiao Tung University, Taiwan, R.O.C., in 1996 and 2001, respectively.

From August 2001 to July 2004, he was an Assistant Professor with the Department of Computer Science and Information Engineering (CSIE), National Taiwan University, Taipei, Taiwan. Since August 2004, he has been an Associate Professor with the Department of CSIE and the Graduate Institute of Networking and Multimedia, National Taiwan University. His research interests include personal communications services, wireless Internet, and performance modeling.

Dr. Lin is an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the *Wiley Journal on Wireless Communications and Mobile Computing*, a Guest Editor for *IEEE Wireless Communications* Special Issue on Mobility and Resource Management, and a Guest Editor for the Association for Computing Machinery (ACM)/Springer MONET Special Issue on Wireless Broad Access. He was the recipient of many research awards, including the Research Award for Young Researchers from the Pan Wen-Yuan Foundation in Taiwan in 2004, the K. T. Li Young Researcher Award honored by ACM Taipei Chapter in 2004, the Wu Ta You Memorial Award of National Science Council in Taiwan in 2005, and the Fu Suu-Nien Award of National Taiwan University in 2005 for his research achievements.