Ad Hoc Networks 11 (2013) 1683–1698

Contents lists available at SciVerse ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Dynamic agent-based hierarchical multicast for wireless mesh networks

Yinan Li *, Ing-Ray Chen

Department of Computer Science, Virginia Tech, Northern Virginia Center, 7054 Haycock Road, Falls Church, VA 22043, United States

ARTICLE INFO

Article history:
Received 16 August 2012
Received in revised form 27 February 2013
Accepted 4 March 2013
Available online 27 March 2013

Keywords:
Hierarchical multicast algorithm
Multicast agent
Wireless mesh networks
Performance analysis

ABSTRACT

We propose and analyze a multicast algorithm named Dynamic Agent-based Hierarchical Multicast (DAHM) for wireless mesh networks that supports user mobility and dynamic group membership. The objective of DAHM is to minimize the overall network cost incurred. DAHM dynamically selects multicast routers serving as multicast agents for integrated mobility and multicast service management, effectively combining backbone multicast routing and local unicast routing into an integrated algorithm. As the name suggests, DAHM employs a two-level hierarchical multicast structure. At the upper level is a backbone multicast tree consisting of mesh routers with multicast agents being the leaves. At the lower level, each multicast agent services those multicast group members within its service region. A multicast group member changes its multicast agent when it moves out of the service region of the current multicast agent. The optimal service region size of a multicast agent is a critical system parameter. We propose a model-based approach to dynamically determine the optimal service region size that achieves network cost minimization. Through a comparative performance study, we show that DAHM significantly outperforms two existing baseline multicast algorithms based on multicast tree structures with dynamic updates upon member movement and group membership changes.

1. Introduction

Wireless mesh networks (WMNs) are emerging in recent years as a promising cost-effective solution for providing last-mile community-based broadband Internet access services. A WMN consists of two types of components: wireless mesh routers (MRs) and mesh clients (MCs) [1]. MRs form a static mesh networking infrastructure called a *wireless mesh backbone* for MCs. MCs are end-user devices with wireless access capability, and unlike MRs, they are usually mobile and may change their locations frequently. A WMN is seamlessly interconnected to the Internet through the gateway functionality of MRs, which can also be used to integrate a WMN with existing wireless networks, for example, mobile ad hoc networks or wireless sensor networks. Generally, one or more MRs in a WMN serve as the Internet gateways and route network traffic originated from or destined to the Internet.

Due to the broadcasting nature of wireless communications and the community oriented nature of WMNs, group communications [2] based on multicasting [3,4] are expected to be a common communication paradigm in WMNs. For example, many popular network applications today are based on a single-source group communication paradigm, and require efficient delivery of various types of contents, e.g., weather forecasts, stock prices, news, and real-time audio/video streams, from a single source to a group of mobile users in WMNs. These applications are multicasting in nature, and therefore can be efficiently

* Corresponding author.
E-mail addresses: yinan926@vt.edu (Y. Li), irchen@vt.edu (I.-R. Chen).

1570-8705/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.adhoc.2013.03.004

implemented by a multicast algorithm. The mobility of multicast group members, however, poses a challenge to the design and development of efficient multicast algorithms in WMNs. More specifically, the multicast algorithm must efficiently support user mobility such that group members can continue to receive subscribed multicast contents and data when they move and change their serving MRs frequently.

In this paper, we propose and analyze a multicast algorithm for WMNs named Dynamic Agent-based Hierarchical Multicast (DAHM) that supports potentially highly mobile users and dynamic multicast group membership. The contribution of the paper is as follows. First, we bring out the design notion of *dynamic agent-based hierarchical multicast* to minimize the overall network communication cost (in terms of traffic incurred) for packet delivery, mobility management, and multicast tree maintenance. This directly contributes to end-to-end delay minimization and throughput maximization for multicast services in WMNs. Second, we bring out the design notion of *integrated mobility and multicast service management* which empowers MRs to serve as multicast agents (MAs) for MCs dynamically. A MC's MA is not only the multicast packet relay station but also the location database of the MC. Under our DAHM design, multicast packet routing is a two-step process. A multicast packet is first routed from the source to the MC's MA through a dynamic multicast tree connecting the source to all MAs and then is routed from the MC's MA to the MC. Packet routing is done efficiently because an MC's MA knows the location of the MC all the time. We achieve integrated mobility and multicast service management by having each MC dynamically determine whether it should select the MR it just enters as its new MA, based on our analysis result. The optimal MA service region for each MC is a critical system parameter for performance maximization in our DAHM design. The third contribution of the paper is that we develop a model-based approach based on stochastic Petri net (SPN) techniques [28] to determine the optimal MA service region size on a per MC basis based on the MC's runtime mobility and service characteristics, so as to maximize multicast performance. The last but not the least contribution is that we demonstrate that DAHM outperforms two existing multicast algorithms in the literature, namely, Regional-Registration based Multicast (RRM) [5] and Dynamic Tree-based Multicast (DTM) through a comparative analysis with simulation validation. RRM is based on a hierarchical tree structure consisting of pure unicast paths, whereas DTM is based on a shortest-path multicast tree structure [6] extended with dynamic updates upon member movement and group membership changes.

The remainder of this paper is organized as follows. Section 2 surveys existing work on multicast routing and algorithms in mobile network environments and particularly in WMNs and contrasts our work with existing work. Section 3 gives a detailed introduction to DAHM. In Section 4 we develop an analytical model for analyzing the performance of DAHM. Detailed performance evaluation and a comparative performance study are given in Section 5, with both analytical results and simulation validation presented. Section 6 discusses issues related to the implementation of DAHM on real mobile devices. The paper concludes with Section 7.

2. Related work

Multicast algorithms and multicast routing protocols within a mobile network environment have been intensively studied for Mobile IP networks [7–10] and mobile ad hoc networks [11]. Due to significant differences in network architectural characteristics and design objectives, however, these algorithms and protocols cannot be applied to WMNs directly without major modification and performance penalty. For example, WMNs lack centralized management facilities such as home agents and foreign agents as in Mobile IP networks. Similarly, although WMNs can be considered as a special type of mobile ad hoc networks, multicast algorithms and routing protocols proposed for such networks are generally not appropriate for WMNs. The reason is that these algorithms and protocols are designed with consideration given specific to characteristics unique to mobile ad hoc networks, e.g., infrastructurelessness, dynamic network topology, energy constraints and weak computing capability of mobile nodes, etc. Therefore, designing new multicast algorithms and routing protocols that take into consideration of the characteristics of WMNs is an important research topic.

The research of multicast in WMNs is still in its infancy. Very recently a few multicast algorithms and routing protocols have been proposed for WMNs [6,12–18]. Zeng et al. [12] proposed two multicast algorithms, namely, the Level Channel Assignment (LCA) algorithm and the Multichannel Multicast (MCM) algorithm, with the objective to improve the multicast throughput in multichannel and multi-interface WMNs. The algorithms focus on the construction of efficient multicast trees that minimize the number of relaying nodes and the total hop count distance of the trees. By using a dedicated channel assignment strategy and partially overlapping channels, interference among channels is reduced and the throughput is improved. Pacifier [13] is a new multicast protocol that pursues high throughput and reliability. Pacifier builds an efficient multicast tree for tree-based opportunistic multicast routing to achieve high throughput, and utilizes intra-flow network coding to achieve high reliability, without the overhead of classic techniques such as Automatic Repeat reQuest (ARQ) and Forward Error Correction (FEC). Pacifier also solves the "crying baby" problem such that the throughput of well-connected nodes is improved without sacrificing the throughput of poorly-connected nodes.

In [6], two primary methods for multicast routing, namely, shortest-path trees (SPTs) and minimum cost trees (MCTs) were investigated and evaluated via extensive simulation using a variety of performance metrics such as packet delivery ratio, throughput, end-to-end delay, delay jitter, and multicast traffic overheads. Based on the comparative simulation results, the author recommended the SPT approach because SPTs performed considerably better than MCTs in terms of these performance metrics. DTM, one of the baseline algorithms for the comparative performance study in this paper, is essentially based on an SPT augmented with the capability to perform dynamic tree updates for supporting member mobility and dynamic group membership. In [14], a cross-layer optimization framework was proposed for maximizing the multicast throughput in WMNs. By realizing that the overall throughput tightly depends on per-link data flow rates (which further depend on link capacities controlled by radio power levels on the physical layer), the paper presented a cross-layer framework spanning the network layer, the link layer, and the physical layer. Within the framework, the multicast routing problem and the wireless medium contention problem are iteratively solved and jointly optimized to generate optimal solutions for the throughput maximization problem.

Ruiz et al. [15] proposed an integrated solution for efficient multicast routing in WMNs connected to the Internet. The solution consists two components: a tree construction algorithm that builds an approximate minimum Steiner tree for efficient multicast routing, and an auto-configuration protocol that configures MRs with topologically correct IP addresses to achieve full compatibility with standard multicast routing protocols used in the Internet. Chakeres et al. [16] examined a wide range of multicast algorithms for WMNs based on the IEEE 802.11 standard. These algorithms provide different degrees of support to fast, efficient, and robust multicast in IEEE 802.11s WMNs. Two of these algorithms are based on broadcast, namely, Default Broadcast (DB) which is the existing multicast algorithm in IEEE 802.11s and Fast Broadcast (FB) which is an enhancement to DB. Another two algorithms are based on unicast, namely, Selective Unicast (SU) and Multiple Unicast (MU), both of which provide robustness by L2 acknowledgments and retransmissions. The last algorithm examined is the so-called Ack-oriented in-mesh Multicast (AM) that also provides robustness by packet acknowledgment and retransmission.

While these algorithms and protocols contributed to various aspects that are key to implementing multicast in WMNs, the critical issue of supporting member mobility and dynamic group membership during the lifetime of a multicast group was not addressed. Specifically, existing algorithms and protocols assume static multicast trees and focus on tree construction algorithms for throughput maximization. This assumption generally is not feasible in real mobile network environments, considering that multicast group members may be highly mobile and they may join or leave the group at arbitrary time. Further, frequent group changes due to member mobility and dynamic group membership can cause the quality and efficiency of a static multicast tree to degrade quickly. In contrast to these algorithms and protocols, DAHM explicitly takes member mobility and dynamic group membership into consideration and dynamically handles mobility management and multicast service management (multicast tree maintenance, group membership management, and multicast packet delivery) in an integrated manner following the design idea of micro-mobility management [23–27].

3. Dynamic agent-based hierarchical multicast

3.1. System model and assumptions

We assume that a WMN has a single Internet gateway, or concisely, a single gateway. We also assume that current and future wireless MRs are powerful enough to host multicast agent software for integrated mobility and multicast service management. Currently available wireless MRs already have good processing capability and expandable memory capacity (via USB-based flash or hard drives) to be used for cooperative data caching in WMNs [19]. Therefore, we assume that they are also capable of performing integrated mobility and multicast service management.

We consider a multicast group that has a single source and dynamic group topology and membership in a WMN. The multicast source can be in the Internet or within a WMN. If the multicast source is an MC within a WMN, the backbone multicast tree is rooted at the source. On the other hand, if the source is a host in the Internet, the backbone multicast tree is rooted at the gateway, as multicast packets will first be routed to the gateway which is responsible for delivering them to the group members.

The multicast group is dynamic with respect to both group member locations because of user mobility and group membership because of member join and leave events. Thus, a multicast group may be characterized by high group dynamics in terms of member locations and group membership. On the other hand, we assume that the source is static.

Within the lifetime of a multicast group, a member may join or leave the group at arbitrary time. We assume that group member join and leave events can be modeled by Poisson processes with rates of λ and μ, respectively. That is, the inter-arrival and inter-departure times are exponentially distributed with averages $1/\lambda$ and $1/\mu$, respectively. We further assume that λ and μ have about the same value such that the multicast group size remains stable over time.

Fig. 1 illustrates the two-level hierarchical multicast structure employed by DAHM. We refer to a multicast group member simply as a member. In Fig. 1, members are labeled as MCs carrying mobile devices. For notational convenience, we refer to an MR serving as the multicast agent for one or more members simply as a multicast agent (MA). Any MR can be an MA when it is selected by a member to serve as the member's MA.

3.2. Overview

DAHM is a dynamic two-level hierarchical multicast algorithm featuring an integrated design that combines backbone multicast routing and local unicast routing. At the upper level of the hierarchy is the multicast tree backbone based on a shortest-path tree (SPT) rooted at the source. The multicast tree consists of MRs with the MRs at the leaves also serving as MAs. In Fig. 1, the multicast tree is connected by thick solid lines. At the lower level of the hierarchy is an MA service area (a subtree rooted at an MA). In Fig. 1, each MA's service area is connected

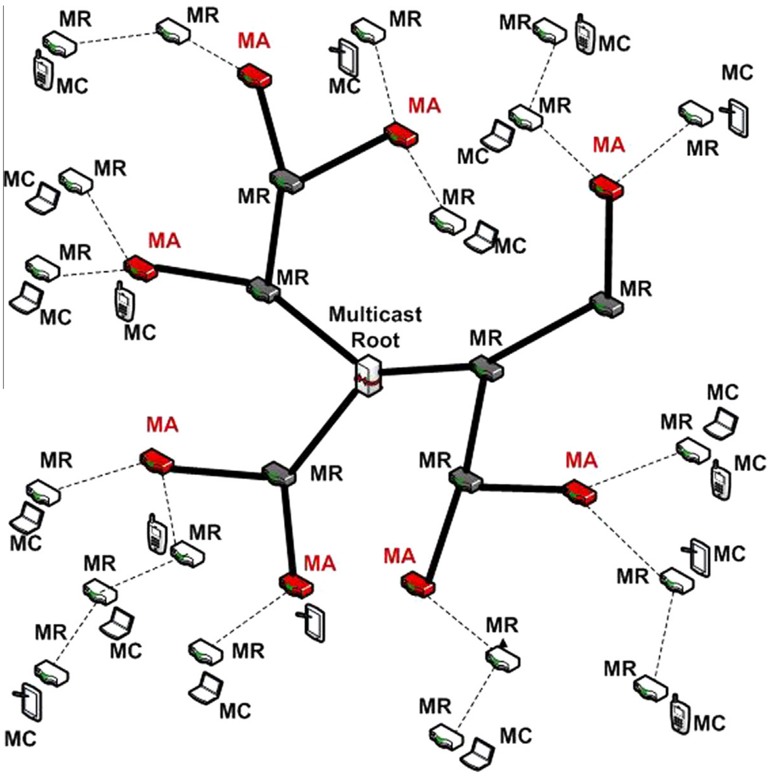


Fig. 1. The two-level hierarchical multicast structure employed by DAHM.

by dotted lines. To a member, the MA's service area is defined in terms of the number of hops (H) the member can be away from its MA. In Fig. 1, the MA subtree at the bottom right has 4 members with their hop distances away from their MA being 1, 1, 1 and 2, respectively. Suppose $H = 5$ for the right bottom most member who is currently 2 hops away from its MA. When it moves from one MR to another MR, it knows that it is still within its MA service area, so it will only inform its address change to the MA instead of making the new MR it just moves into a new MA to avoid the multicast tree maintenance cost. On the other hand if $H = 2$, it will make the new MR it just moves into as its new MA. This will trigger an update to the multicast tree. In general, the multicast tree is updated whenever an MA joins or leaves due to user mobility and group membership changes. Multicast packets are first disseminated from the source to all the MAs via multicast routing through the SPT, and then delivered from the MAs to multicast group members individually via local unicast routing. The reason why unicast routing is used at the lower level rather than multicast routing as in [20,21] is twofold:

- The optimal service region size of an MA that minimizes the overall communication cost, i.e., the optimal threshold $H_{optimal}$ for the number of hops a multicast group member can be away from its MA, can be quite diverse for different group members depending on their mobility and service characteristics, as supported by the analytical and simulation results presented in Section 5. Therefore, group members associated with the same MA can have very diverse hop distances to the MA, making the wireless broadcast advantage no longer valid. Thus, using broadcast routing at the lower level can adversely affect the communication cost, because the overhead of multicast routing can be considerably high especially when a small number of receivers (group members associated with the same MA) are dispersed in a large service area around the sender (the MA).
- Using unicast routing eliminates the need for multicast tree maintenance at the lower level and simplifies mobility management. Suppose that multicast routing is used at the lower level, the need for mobility management as well as multicast tree maintenance would be frequent because multicast group members may have high mobility. Specifically, when a multicast group member moves to a new serving MR, the new serving MR needs to be subscribed to and the old serving MR needs to be unsubscribed from the multicast tree rooted at the MA, thus incurring two tree maintenance operations. When the member moves out of the service region of its current MA and switches to a new MA, not only changes to the multicast trees of both the old and new MAs need to be handled by the corresponding tree maintenance operations, group membership changes also need to be processed. If unicast routing is employed at the lower level, the overhead of multicast tree maintenance and multicast group membership

management at the lower level would be completely eliminated. The saving can be significant, considering that group members can have high mobility and that the number of multicast groups at the lower level can be potentially large.

We use an SPT as the multicast backbone at the upper level as it is shown in [6] that an SPT is superior to a minimum cost tree (MCT) such as an *approximate* minimum Steiner tree (MST) in terms of packet delivery ratio, throughput, average end-to-end delay, and delay average jitter. Another advantage of an SPT over a MST is that the problem of constructing a MST is NP-complete. Additionally, considering SPT instead of sophisticated tree algorithms that strive for high throughput (e.g., [12,13,17]) allows us to focus on the design and analysis aspect of integrated mobility and multicast service management. Indeed, we could replace SPT with a more sophisticated algorithm and the design idea still applies. Here we note that our idea is generic as can be applied to other network services such as mobile data access [22]. Also note that we use hop count as in [12,6] rather than link quality as the metric for multicast routing in the SPT multicast backbone. This is because our focus is on network cost minimization via integrated mobility and multicast service management and the total network cost, defined as the total number of hops of wireless transmissions incurred by DAHM in Section 4.2, is a function of the hop count.

A MA serves as a regional registration point for integrated mobility and multicast service management. Each multicast group member is registered with and serviced by an MA, from which it receives multicast packets via local unicast routing. The multicast group member also sends its updated location information, i.e., the address of its current serving MR, to the MA, whenever it moves and switches to a new serving MR. Each MA maintains a location database that stores the up-to-date location information of each multicast group member it currently services.

A MA and those members it currently services essentially form a local multicast group at the lower level of the hierarchy. Like the multicast backbone, a local multicast group is also dynamic due to user mobility and membership changes. Each MA covers a service region servicing all the members located within the region. The service region size of an MA is a key parameter controlling the trade-off between the communication cost incurred at the upper level and that incurred at the lower level. There exists an optimal service region size that minimizes the overall communication cost. We model the optimal service region size as the optimal threshold for the number of hops a member can be away from its MA, denoted by $H_{optimal}$. This optimal threshold can be determined using the analytical model developed in Section 4. Below we let H and $H_{optimal}$ denote the threshold and the optimal threshold, respectively.

3.3. Member join and leave

3.3.1. Member join

A MC who intends to join a multicast group first selects a serving MR among all MRs within the wireless transmission range based on the wireless link quality, and sends a join request |JOIN_REQ| to the selected serving MR. If the new serving MR is not yet a leaf node of the multicast backbone (that is, if it is not an MA), it needs to join the backbone multicast tree as a leaf node and becomes a new MA for the MC. The MR joins the backbone multicast tree by sending |JOIN_REQ| to the source. Upon receiving |JOIN_REQ| from the MR, the source computes a shortest path to the MR and sends a join acknowledgment |JOIN_ACK| along the path back to it. The MR further forwards |JOIN_ACK| to the MC, confirming that it becomes a new member of the multicast group. By having the MA process member join requests locally, the signaling overhead of member join is significantly reduced. Fig. 2 illustrates the procedure for a member join event. |JOIN_REQ| and |JOIN_ACK| also serve as an association request and an association acknowledgment, respectively.

3.3.2. Member leave

When a member leaves a multicast group, it notifies its MA such that the MA can deregister it. After the member leaves, the MA may no longer service any member, therefore it needs to be removed from the backbone multicast tree. The procedure for a member leave event is illustrated in Fig. 3. More specifically, the leaving member sends a leave request |LEAVE_REQ| to its MA, which responds with a leave acknowledgment |LEAVE_ACK| as a confirmation. If the MA needs to remove itself from the backbone multicast tree because it no longer services any member, it forwards |LEAVE_REQ| to the source. Upon receiving |LEAVE_REQ|

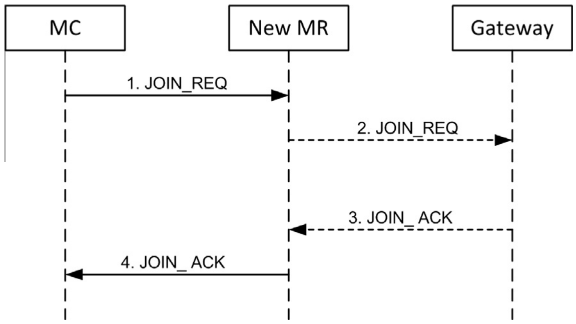


Fig. 2. Message exchange sequence for a member join event.

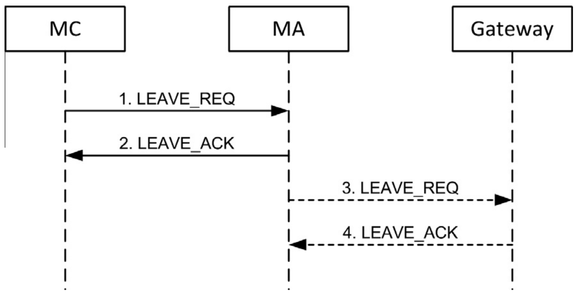


Fig. 3. Message exchange sequence for a member leave event (dashed lines mean conditional message exchanges).

from the MA, the source updates the backbone multicast tree and sends the MA a leave acknowledgment |LEAVE_ACK| in reply to the request.

In some cases, a member disconnects (either voluntarily or involuntarily) and therefore is not able to notify its MA. In DAHM, a member that disconnects is treated as a leaving member. The disconnection of a member can be detected by its MA when the MA tries to deliver multicast packets to the member. Once a member is detected to be disconnected, its MA deregister it and the MA needs to remove itself from the backbone multicast tree if it no longer services any member.

3.4. Mobility management and tree maintenance

In DAHM, when a member moves and changes its serving MR, the following procedure is executed to handle the mobility management and multicast tree maintenance:

- When the member moves and switches to a new MR, it sends an association request |ASSO_REQ| to the new MR. The MR responds with an association acknowledgment |ASSO_ACK| in reply to the request, confirming that the association is completed and the MR becomes the new serving MR of the member.
- If the new serving MR is not an MA and is within the service region of the member's MA, the member sends to its MA a location registration request |LOC_REG_REQ| containing the address of the new MR. The MA updates the member's location information and sends a location registration acknowledgment |LOC_REG_ACK| back to the member. In this way, the MA always knows the up-to-date location information of members within its service region and is therefore able to deliver multicast packets to them individually through unicast routing.
- If the new serving MR is H hops away from the member's current MA, the threshold is reached and the new MR needs to join the backbone multicast tree as a leaf node and becomes the new MA of the member. In this case, a join request |JOIN_REQ| is sent to the source. Upon receiving |JOIN_REQ| from the MR, the source computes a shortest path to the MR and sends a join acknowledgment |JOIN_ACK| along the path back to it. Fig. 4 shows the message exchange sequence in the case that the new MR is H hops away from the member's current MA.
- If the new MR is already an MA, the member switches to the new MA and starts receiving multicast packets from the new MA. Fig. 5 shows the message exchange sequence in the case that the new MR is already an MA.
- After being associated with the new MA, the member sends a deassociation request |DEASSO_REQ| to its old MA, which responds with a deassociation acknowledgment |DEASSO_ACK|.
- If the member's old MA no longer services any member, it removes itself from the multicast tree by sending a leave request |LEAVE_REQ| to the source. Upon receiving |LEAVE_REQ| from the MA, the source updates the backbone multicast tree and sends the MA a leave acknowledgment |LEAVE_ACK| as a confirmation.

3.5. Multicast packet delivery

In DAHM, multicast packets are delivered in a hierarchical manner from the multicast source to the multicast group members within a WMN. More specifically, multicast packet delivery in DAHM follows the following procedure:

- If the source is a host in the Internet, it will first send multicast packets to the gateway, which is then responsible for distributing the packet to the MAs. The gateway can be considered as a virtual source in this case.
- For each multicast packet, the (virtual) source creates a new packet that encapsulates the multicast payload using a multicast address for the destination address field, and disseminates the new packet to the MAs through the backbone multicast tree in multicast routing mode.
- Upon receiving the packet, each MA decapsulates the packet and encapsulates the payload using the address of the serving MR of each member it services for the destination field, and forwards the new packet to the MR via unicast routing. The address of the serving MR of each member can be found in the MA's location database.
- Each MR after receiving the multicast packet decapsulates the packet and delivers the packet to the designated member.

4. Performance model

In this section, we develop a probability model based on SPN techniques [28–31] for evaluating the performance of DAHM. We choose SPN as the tool for performance modeling because: (1) an SPN model is a concise representation of the underlying Markov or semi-Markov chain that may have a large number of states; (2) an SPN model is capable of reasoning the behavior of a member, as it migrates among states in response to system events.

Table 1 lists the parameters and notations used in the following sections. The physical meaning of the mobility rate denoted by σ is the average number of serving MR changes made by a multicast group member per time unit. The time unit used in this paper is second. If a group member moves and changes its serving MR once every 10 min, its mobility rate is $\frac{1}{600}$. The physical meanings of other parameters are clear from the context.

4.1. Stochastic Petri Net (SPN) Model

We assume that a WMN is structured as a two-dimensional $n \times n$ mesh with wraparound on the boundary such that each MR has exactly four neighbors, as illustrated in Fig. 6. Each MR can communicate directly with any of its four neighbors that are within its communication range. A member can change randomly from its current serving MR to any of the MR's four neighbors with equal probabilities of $\frac{1}{4}$. The total number of MRs in the network denoted by N is simply given by $N = n^2$. The average unicast path

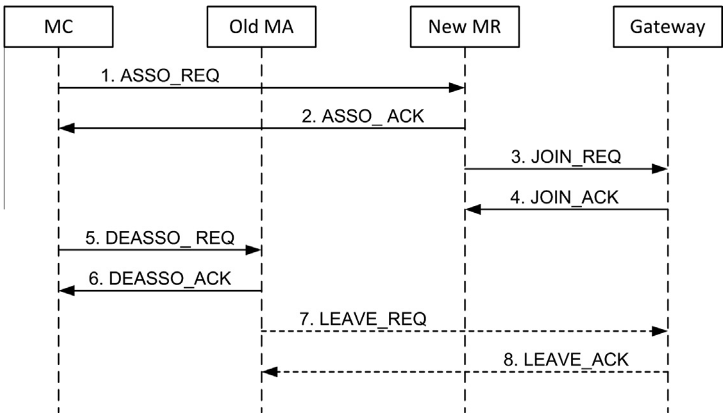


Fig. 4. Message exchange sequence for the case that the new MR is H hops away from the member's current MA.

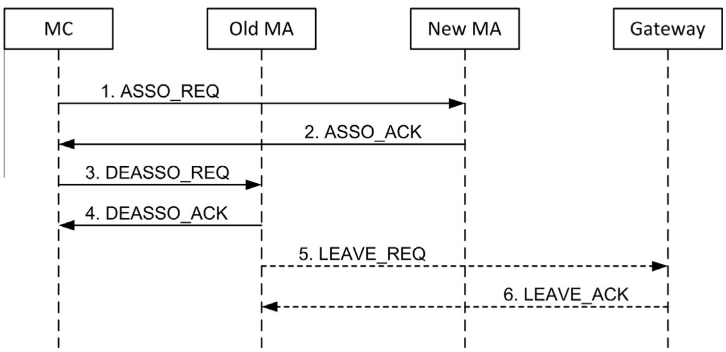


Fig. 5. Message exchange sequence for the case that the new MR is already an MA.

Table 1
Parameters and notations used in performance modeling and analysis.

Parameter	Notation
σ	The average mobility rate of multicast group members
λ_p	The multicast packet rate
SMR	Service to mobility ratio, defined as $\text{SMR} = \frac{\lambda_p}{\sigma}$
λ	The rate of member join events
μ	The rate of member leave events
M	The multicast group size
n	The dimension of the WMN
N	The number of MRs in the WMN
γ	The member density
α	The average unicast path length of the WMN
ω	The arrival rate of a single member to an arbitrary MR
P_{MA}	The probability that an arbitrary MR is also an MA
P_0	The probability that an MR is not covering any member
P_1	The probability that an MR covers exactly one member
P_1^{MA}	The probability that an MA services exactly one member
N_{MA}	The number of MAs
H	The service region size of an MA
T	The multicast tree size in terms of the total number of tree nodes
κ	The multicast scaling factor
L	The expected hop distance from the source to an MA
d	The average degree of inner nodes

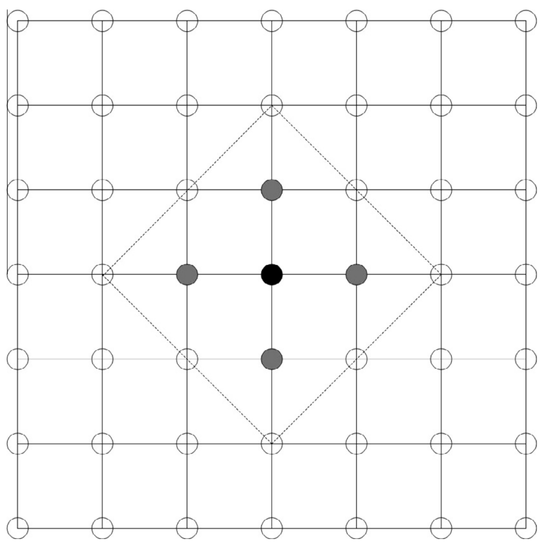


Fig. 6. An $n \times n$ mesh network model.

length (hop count) denoted by α in this $n \times n$ mesh network model is given by [32]:

$$\alpha = \frac{2n}{3} \tag{1}$$

We model the process of arrival and departure of M multicast members to and from an MR using an $M/M/\infty/M$ queue. Fig. 7 depicts the Markov chain for the $M/M/\infty/M$ queueing model, where ω means the arrival rate of a single member to an arbitrary MR, and is given by [7]:

$$\omega = \frac{\sigma}{n^2 - 1} \tag{2}$$

Using the $M/M/\infty/M$ queueing model, the probability P_0 that an MR covers no members and the probability P_1 that an MR covers exactly one member can be derived as:

$$P_0 = \left(1 - \frac{1}{n^2}\right)^M \tag{3}$$

$$P_1 = \frac{M}{n^2}\left(1 - \frac{1}{n^2}\right)^{M-1} \tag{4}$$

The dashed-line square within the mesh structure shown in Fig. 6 illustrates the service region of an MA. Given that the threshold of the number of hops a member can be away from its MA is H, the number of MRs within the service region of an MA on average is $2H^2 - 2H + 1$, in the $n \times n$ mesh network model. The probability denoted by P_{MA} that an arbitrary MR is an MA in DAHM is therefore approximated by:

$$P_{MA} = \frac{1}{2H^2 - 2H + 1} \tag{5}$$

Here we note that P_{MA} given above is only approximate because which MR is chosen as an MA for a multicast member depends on the user's mobility. However, as validated by simulation reported in Section 5.3, this approximation does not affect the result accuracy. A MA services exactly one member if all the MRs within its service region totally service exactly one member. Therefore, the probability denoted by P_1^{MA} that an MA services exactly one member can be calculated as follows:

$$P_1^{MA} = \binom{2H^2 - 2H + 1}{1} \cdot P_0^{2H^2 - 2H} \cdot P_1 \tag{6}$$

At the upper level of the hierarchy, the number of MRs (including MAs) comprising the backbone multicast tree can be derived using the following method. First, the ratio of the total number of multicast links (among MRs) on the tree denoted by L_m over the average unicast path length of the network denoted by α is given by a power-law [33,34] as follows:

$$\frac{L_m}{\alpha} = R^{\kappa} \Rightarrow L_m = \alpha \cdot R^{\kappa} \tag{7}$$

where κ is the *multicast scaling factor*, and is found to be close to 0.7 [33]. R denotes the number of leaves on the multicast tree, i.e., the number of MAs, and is calculated as:

$$R = N_{MA} = P_{MA} \cdot N \tag{8}$$

Given L_m, the total number of MRs (including the MAs) on the backbone multicast tree denoted by T is given as:

$$T = L_m + 1 = \alpha \cdot (N_{MA})^{\kappa} + 1 \tag{9}$$

The expected hop distance L from the source to an MA is the average length of all paths from the source to the MAs, or, equivalently stated, it is equal to the average depth of all MAs (leaves) on the backbone multicast tree rooted at the source. Hence, assuming a perfectly balanced backbone multicast tree, L is calculated as follows:

$$L = \log_d T \tag{10}$$

where d is the degree of an inner node (we use $d = 4$ because each inner node has four neighbors).

The optimal threshold for the number of hops a member can be away from its MA, denoted by $H_{optimal}$, can be determined by using the SPN model. Fig. 8 shows the SPN model for describing the behavior of a single group member. An SPN model consists of places, tokens, and

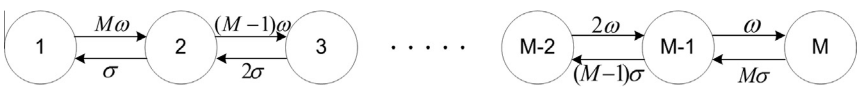


Fig. 7. The Markov chain modeling the process of arrival and departure of M multicast group members to and from an MR.

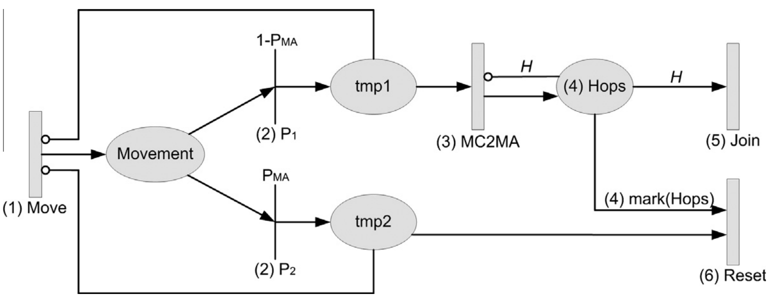


Fig. 8. The SPN model for DAHM.

Table 2
The meanings of places and transitions defined in the SPN model for DAHM.

Symbol	Meaning
Movement	|mark (Movement) = 1| means that the member moves and switches to a new serving MR
Hops	|mark (Hops)| returns the number of hops the member is away from its MA
Move	A timed transition modeling the movement of the member
MC2MA	A timed transition modeling the regional location registration event
Join	A timed transition modeling that the new serving MR joins the multicast tree as a leaf node and becomes a new MA
Reset	A timed transition modeling the event of registering with the new MR that is already an MA

transitions (for modeling events). Table 2 explains the meanings of places and transitions defined in the SPN model.

In Fig. 8 we put in numbers in parenthesis to label the SPN model sequence below. The SPN model is constructed as follows:

1. The event of member movement is modeled by transition *Move*, the rate of which is σ. When a member moves and switches to a new serving MR, a token is put into place *Movement*.
2. The new MR may be either an ordinary MR or an MA. The SPN model distinguishes between these two cases using two immediate transitions *P1* and *P2* that are associated with probabilities $1 - P_{MA}$ and P_{MA}, respectively.
3. In the first case that the new MR is not an MA, the member sends its current MA a |LOC_REG_REQ| message that contains the address of the new serving MR. Upon receiving the message, the MA updates the location information of the member stored in the location database, and acknowledges the location update by a |LOC_REG_ACK| message. The message exchange is modeled by transition *MC2MA*.
4. After transition *MC2MA* is fired, a token is put into place *Hops*. The number of tokens denoted by *mark(Hops)* in place *Hops* represents the number of hops the member is away from its MA.
5. When the number of tokens in place *Hops* reaches the threshold denoted by *H*, i.e., when *mark*(*Hops*) = *H*, transition *Join* is fired, modeling that the new serving MR joins the backbone multicast tree as a leaf node and becomes the new MA of the member. The firing of transition *Join* consumes all the tokens in place *Hops*.
6. In the second case that the new MR is already an MA, the member registers with the new MA, and starts receiving multicast packets from the new MA. This is modeled by transition *Reset*, the firing of which consumes all the tokens in place *Hops*, meaning that the member is now directly serviced by the new MA and the hop counter is reset.

4.2. Performance metrics

We use the *average total communication cost incurred per member per time unit (second)* as the metric for performance evaluation and analysis, and the objective is to minimize this cost. In Section 5.4, we discuss how cost minimization is related to throughput maximization and end-to-end delay minimization. We define the total communication cost as the *total number of hops of wireless transmissions incurred*, corresponding to the total traffic incurred to the network since each hop involves a packet transmission or relay. For example, the service cost incurred per multicast packet delivery per member is given by the average number of hops traveled per packet from the source to any member. More specifically, the average total communication cost incurred per member per time unit by DAHM, denoted by C_{DAHM}, includes the service cost for multicast packet delivery denoted by $\lambda_p \cdot C_s$, the signaling cost for mobility management denoted by $\sigma \cdot C_m$, the signaling cost for processing member join requests denoted by $\lambda \cdot C_j$, and the signaling cost for processing member leave requests denoted by $\mu \cdot C_l$. The equation for calculating C_{DAHM} is therefore given as follows:

$$C_{DAHM} = \lambda_p \cdot C_s + \sigma \cdot C_m + \lambda \cdot C_j + \mu \cdot C_l \quad (11)$$

Here we note that *H* represents the service region size of an MA and hence the optimal service region size is the optimal *H* value ($H_{optimal}$) under which C_{DAHM} in Eq. (11) is maximized. C_s, the service cost incurred per multicast group member per multicast packet delivery in DAHM, consists of two parts. The first part denoted by C_s^1 is the total cost for disseminating the multicast packet from the source to all the MAs through the backbone multicast tree, namely *T*, divided by the multicast group size *M*. The number of wireless transmissions required to deliver a multicast packet from the source to the MAs (i.e., C_s^1) equals *T* because each MR on the tree transmits the packet only once to all its downstream children [35]. The second part denoted by C_s^2 is the average cost for delivering the multicast packet via unicast routing from an MA to a member it currently services. Since a member can be *i* hops away from its MA with probability $P_i (0 \leqslant i \leqslant H-1)$, C_s^2 is given by the probability-weighted average distance between the member and its MA. Therefore, C_s is the sum of the two parts:

$$C_s = C_s^1 + C_s^2 = \frac{T}{M} + \sum_{i=0}^{i=H-1} P_i \cdot i \quad (12)$$

C_m, the mobility management cost incurred per group member, depends on the event triggered by the movement of a member. More specifically, the mobility management cost is incurred when there is an *MA join* (*Join* in the SPN model), *MA reset* (*Reset* in the SPN model), or *MA update* (*MC2MA* in the SPN model) event as follows:

- *MA join*: When the new serving MR of a member is *H* hops away from its MA after a movement, the new serving MR needs to join the backbone multicast tree as a leaf node and becomes the new MA of the member. In this event, the member completes the association with the new MR by sending an |ASSO_REQ| message to it, which responds with an |ASSO_ACK| message as an acknowledgment. The new MR joins the tree by sending a |JOIN_REQ| message to the source, which computes a shortest path to the MR and sends a |JOIN_ACK| message along the path back to it. With probability P_1^{MA}, the member's old MA no longer services any member, and it removes itself from the backbone multicast tree by sending a |LEAVE_REQ| message to the source, which updates the tree and sends the MA a |LEAVE_ACK| message.

- *MA reset*: When the new serving MR of the member is already an MA, the member switches to the new MA. In this event, the member completes the association with the new MA by sending an |ASSO_REQ| message to the MA, which responds with an |ASSO_ACK| message as an acknowledgment. After being associated with the new MA, the member sends a |DEASSO_REQ| message to its old MA, which responds with a |DEASSO_ACK|| message. With probability P_1^{MA}, the member's old MA no longer services any member, and it removes itself from the backbone multicast tree by sending a |LEAVE_REQ|| message to the source, which updates the tree and sends the MA a |LEAVE_ACK| message.
- *MA update*: When a member moves and changes its serving MR, the member sends to its MA a |LOC_REG_REQ| message containing the address of the new serving MR. The MA updates the location information of the member stored in the location database, and acknowledges the location update by a |LOC_REG_ACK| message.

Based on the discussion above, C_m is given by:

$$C_m = \begin{cases} 2 + 2H + (1 + P_1^{MA}) \cdot 2L & \text{ifJoin}'' \\ 2 + 2h + P_1^{MA} \cdot 2L & \text{ifReset}'' \\ 2 + 2h & \text{ifMC2MA}'' \end{cases} \quad (13)$$

where h = *mark*(*Hops*) represents the distance between the member and its MA. We use P_1^{MA} for the probability that the member is the only one currently serviced by its MA. Therefore, once the only member leaves, the MA will no longer service any member, and it should be removed from the backbone multicast tree.

C_j, the signaling cost per member join event, is computed as follows. A MC joins an existing multicast group by sending a |JOIN_REQ| message to its newly selected serving MR. With probability $1 - P_{MA}$, the new serving MR is not yet a leaf node of the multicast backbone, and it needs to join the backbone multicast tree as a leaf node and becomes a new MA for the MC. The MR joins the multicast tree by sending |JOIN_REQ| to the source, which responds with a |JOIN_ACK| message as an acknowledgment. The MR further forwards |JOIN_ACK| to the MC, confirming that it becomes a new member of the multicast group. Therefore, C_j is calculated as:

$$C_j = 2 + (1 - P_{MA}) \cdot 2L \quad (14)$$

C_l, the signaling cost per member leave event, is computed as follows. The leaving member sends a |LEAVE_REQ| message to its MA, which responds with a |LEAVE_ACK| message as a confirmation. With probability P_1^{MA}, the MA needs to remove itself from the backbone multicast tree because it no longer services any member, and it forwards |LEAVE_REQ| to the source, which updates the backbone multicast tree and sends the MA a |LEAVE_ACK| message in reply to the request. Therefore, C_l is calculated as:

$$C_l = 2 + 2h + P_1^{MA} \cdot 2L \quad (15)$$

where h = *mark*(*Hops*) represents the distance between the member and its MA.

5. Performance analysis and numerical results

In this section, we evaluate the performance of the proposed hierarchical multicast algorithms, namely, DAHM, and the effect of various parameters on its performance. We also compare DAHM with two baseline multicast algorithms for WMNs, namely, Regional-Registration based Multicast (RRM) [5] and Dynamic Tree-based Multicast (DTM). Like DAHM, RRM is also a hierarchical multicast algorithm and it also employs MAs for integrated mobility and multicast service management. However, the hierarchical tree structure in RRM is simply a union of pure unicast paths from the source to the group members. Therefore, RRM is a hierarchical unicast-based multicast algorithm. DTM transmits multicast packets through a dynamic shortest-path multicast tree whose leaves are MRs that directly service the members. The multicast tree in DTM is updated to maintain its structural properties every time a member moves and changes its serving MR. Therefore, DTM is essentially based on the existing multicast algorithm that relies on a shortest-path tree [7] augmented with the capability to perform dynamic tree updates for supporting member mobility and dynamic group membership.

To evaluate the effect of user mobility on the performance of the three algorithms, we introduce a parameter called *service to mobility ratio* (SMR) defined as SMR=$\frac{\lambda_p}{\sigma}$. The physical meaning of SMR is the average number of multicast data packets transmitted from the source to a group member during the interval between two serving MR changes of the group member. SMR is an important parameter because it captures the service and mobility characteristics of a group member, both of which can have a significant impact on the operations of DAHM and on the overall network cost.

Table 3 lists the parameters and their values used in performance evaluation. These values are selected to demonstrate diversely sized multicast groups consisting of mobile members characterized by a broad range of SMR. The member join and leave rates are chosen to allow dynamically changing group membership, while maintaining a stable multicast group size. The range of n is selected to model WMNs of reasonably diverse sizes.

Table 3
Parameters and their typical values used in performance evaluation.

Parameter	Meaning	Typical value
M	Multicast group size	[10, 320]
n	Network size	[5, 15]
$\frac{\lambda}{\mu}$	Multicast member join to leave rate ratio	1
SMR	Service to mobility rate	[8, 6000]

5.1. Performance evaluation

In this section we report analytical results obtained from evaluating Eqs. (11)–(15). Specifically, we obtain the analytical results by first assigning a state-dependent cost (specified by Eqs. (12)–(14) or (15)) to each state of the underlying semi-Markov chain of the SPN model, and then computing C_{DAHM} (specified by Eq. (11)) by the state probability-weighted average cost, using the SPNP package [28].

Fig. 9 plots C_{DAHM} as a function of the threshold H, under different multicast group sizes. It can be seen in the figure that there exists an optimal threshold $H_{optimal}$ that minimizes C_{DAHM} for each different M. Fig. 10 further shows C_{DAHM} as a function of the threshold H, under different $n \times n$ network sizes. Again, the optimal threshold $H_{optimal}$ exists for each different n.

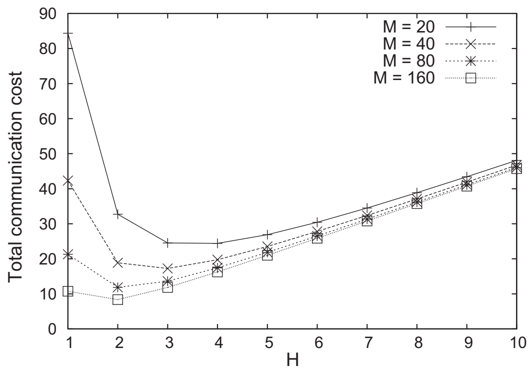


Fig. 9. Cost vs. H, under different multicast group sizes in DAHM (n = 10).

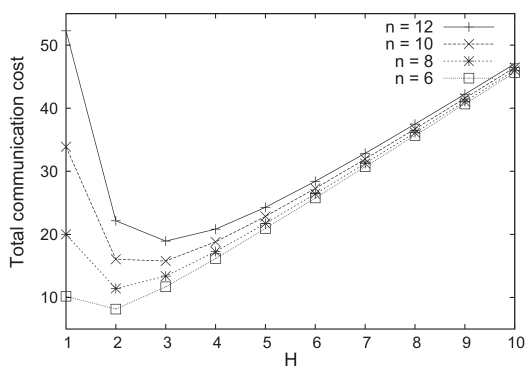


Fig. 10. Cost vs. H, under different network sizes in DAHM (M = 50).

These results demonstrate that the service region size of an MA is key to the performance of DAHM, and there exists an optimal service region size that optimizes the performance of DAHM. The optimal service region size exists because of the tradeoff between the communication cost incurred at the upper level and that incurred at the lower level.

Fig. 11 plots C_{DAHM} as a function of the *member density* denoted by γ, which is defined as $\gamma = \frac{M}{N}$, i.e., the average number of members serviced by one MR. As the figure shows, C_{DAHM} decreases monotonically with increasing γ. This illustrates that multicast efficiency improves as the member density increases because the cost is effectively amortized by the increasing member population. The improvement in multicast efficiency is particularly significant at the upper level because the number of nodes on the backbone multicast tree increases sublinearly with increasing MAs ($\kappa < 1.0$).

Fig. 12 shows the optimal threshold $H_{optimal}$ as a function of γ. It can be seen in the figure that $H_{optimal}$ decreases as γ increases, and it drops to 1 when γ is reasonably large. The service cost C_s for multicast packet delivery in DAHM and accordingly C_{DAHM} decreases with decreasing $H_{optimal}$, because the average distance over which multicast packets are transmitted at the lower level decreases. Therefore, the result conforms to the trend exhibited in Fig. 11.

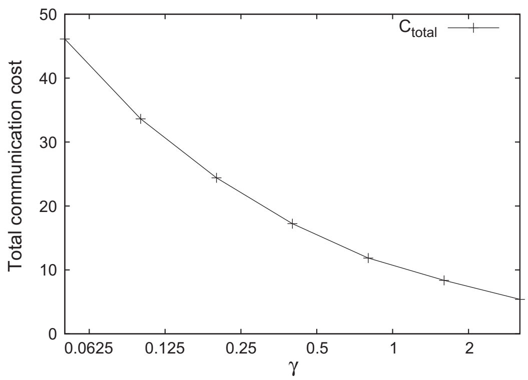


Fig. 11. Cost vs. γ in DAHM.

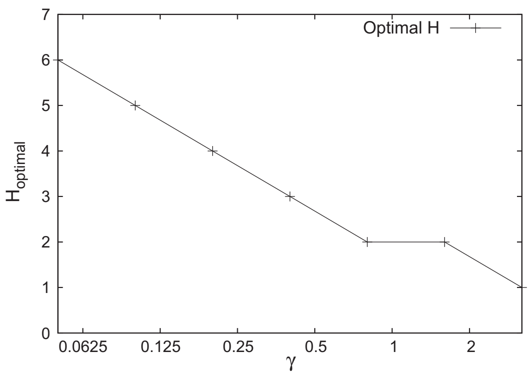


Fig. 12. $H_{optimal}$ vs. γ in DAHM.

5.2. Comparative performance study

In this section, we compare DAHM with RRM and DTM, in terms of the average total communication cost incurred per member per time unit. RRM is a hierarchical multicast algorithm based purely on unicast routing. It is worth emphasizing that because the total communication cost is a per member per time unit metric, even a small cost reduction of 5–10% will be significant over time and over the entire group of members.

For RRM (DTM), the average total communication cost incurred per member per time unit denoted by C_{RRM}(C_{DTM} respectively) consists of the service cost for multicast packet delivery denoted by $\lambda_p \cdot C_s^{RRM}(\lambda_p \cdot C_s^{DTM}$ respectively), the signaling cost for mobility management and multicast tree maintenance denoted by $\sigma \cdot C_m^{RRM}(\sigma \cdot C_m^{DTM}$ respectively), the signaling cost for processing member join requests denoted by $\lambda \cdot C_j^{RRM}(\lambda \cdot C_j^{DTM}$ respectively), and the signaling cost for processing member leave requests denoted by $\mu \cdot C_l^{RRM}(\mu \cdot C_l^{DTM}$ respectively). The following equations calculate C_{RRM} and C_{DTM}:

$$\begin{aligned} C_{RRM} &= \lambda_p \cdot C_s^{RRM} + \sigma \cdot C_m^{RRM} + \lambda \cdot C_j^{RRM} + \mu \cdot C_l^{RRM} \\ C_{DTM} &= \lambda_p \cdot C_s^{DTM} + \sigma \cdot C_m^{DTM} + \lambda \cdot C_j^{DTM} + \mu \cdot C_l^{DTM} \end{aligned} \quad (16)$$

The service cost for multicast packet delivery in RRM denoted by C_s^{RRM} consists of the cost of forwarding the packet from the source to the MAs and the cost of delivering the packet from the MAs to the group members they service, both via unicast routing. Therefore, C_s^{RRM} is calculated as:

$$C_s^{RRM} = \frac{1}{M}(N_{MA} \cdot L + M \cdot \sum_{i=0}^{i=H-1} P_i \cdot i) \quad (17)$$

C_m^{RRM} depends on the event triggered by the movement of a multicast group member. The equation for calculating C_m^{RRM} is the same as that for calculating C_m in DAHM. Additionally, the equations for calculating C_j^{RRM} and C_l^{RRM} are also the same as those for calculating the same cost terms in DAHM, because DAHM and RRM share the same message sequences for multicast structure maintenance and member join and leave events.

The service cost per multicast packet delivery in DTM is equivalent to the number of nodes on the multicast tree because each MR on the tree only transmits the packet once to its downstream children. Therefore, the service cost incurred per member is:

$$C_s^{DTM} = \frac{T_{DTM}}{M} \quad (18)$$

where T_{DTM} denotes the number of tree nodes on the shortest-path multicast tree in DTM. T_{DTM} can be calculated according to the power-law [33,34] as:

$$T_{DTM} = \alpha R^{\kappa} + 1 \quad (19)$$

where R denotes the number of leaf nodes on the multicast tree in DTM, i.e., the number of MRs that service at least one member, which is simply $R = (1 - P_0) \cdot N$.

The tree maintenance cost in DTM consists of the costs of MR association and deassociation, and possibly the costs of multicast tree updates, as calculated by the following equation:

$$C_m^{DTM} = 4 + (P_0 + P_1) \cdot 2L \quad (20)$$

In DTM, when a member joins a multicast group, it establishes the association with a serving MR. With probability P_0, the MR needs to join the multicast tree as a leaf node because it is not already a node on the tree. When a member leaves a multicast group, its association with its current serving MR is canceled. With probability P_1, the member is the only one that the MR services, and the MR needs to remove itself from the multicast tree because it will no longer service any member. Therefore, C_j^{DTM} and C_l^{DTM} are calculated as follows:

$$C_j^{DTM} = 2 + P_0 \cdot 2L \quad (21)$$

$$C_l^{DTM} = 2 + P_1 \cdot 2L \quad (22)$$

Fig. 13 compares the average total communication cost incurred per member per time unit by the three algorithms as a function of the multicast group size M. As can be seen in the figure, the cost decreases as M increases for all three algorithms. The reason is that the member density increases as M increases, given that n is fixed. This observation leads to the generalized conclusion that multicast efficiency improves as the member density increases. It can also be seen in the figure that DAHM is superior to both RRM and DTM.

Fig. 14 compares the total communication cost incurred per member per time unit by the three algorithms as a function of the network size n. As can be seen in the figure, for all the three algorithms, the cost increases with increasing n. This is because the member density decreases as n increases, given that M is fixed. Therefore, this observation also generalizes to the conclusion that multicast efficiency improves as the member density increases. Again, DAHM shows significantly better performance than both RRM and DTM. It is worth emphasizing again that because the total communication cost is a per member per time unit metric, even a small cost reduction of 5–10% will be significant over time and over the entire group of members.

Fig. 15 studies the effect of the mobility rate denoted by σ on the performance of the three algorithms, under

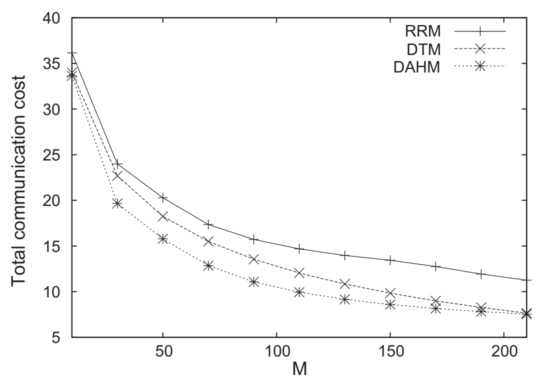


Fig. 13. Performance comparison: cost vs. $M(n = 10)$.

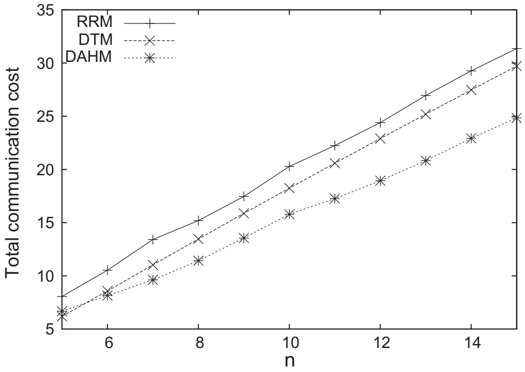


Fig. 14. Performance comparison: cost vs. $n(M = 50)$.

different member densities. As can be seen in the figure, as SMR increases, the costs decrease monotonically because the contribution of the signaling cost for mobility management and multicast tree maintenance to the total communication cost decreases accordingly. As the figures show, DAHM performs consistently better than RRM and DTM over a wide range of SMR and the member density. DAHM copes well with the impact of high user mobility compared with RRM and DTM, due to its capability to dynamically select the optimal service region size of an MA (i.e. $H_{optimal}$) that minimizes the total communication cost. RRM outperforms DTM when the members are highly mobile and the member density is low. However, the advantage diminishes as the member density increases. When the members have high mobility, DTM incurs a substantially larger signaling cost for mobility management and multicast tree maintenance, compared with DAHM and RRM. This is because DTM performs mobility management and tree maintenance every time a member moves and changes its serving MR. Additionally, when the member density is low, i.e., when a small number of members are sparsely distributed within the network, the multicast tree in DTM has a relatively large number of non-leaf MRs, leading to a relatively large cost for multicast packet delivery.

5.3. Simulation validation

Here we conduct simulation experiments to validate the numerical data obtained in the previous sections. We implement the simulation system using a discrete event simulation language called Simulation Model Programming Language (SMPL) [36]. In this simulation system, all operations in DAHM are represented by discrete events associated with costs. For example, location update operations, multicast packet deliveries, member join/leave operations, and multicast tree maintenance operations, are all discrete events. Events are scheduled and executed in event occurrence time order, according to the algorithm description presented in Section 3. The average total communication cost incurred per member per time unit is evaluated and the mean cost is calculated periodically with an interval of 30 min in simulation time. To ensure the statistical significance of simulation results, we use batch mean analysis (BMA) techniques [36]. Each simulation

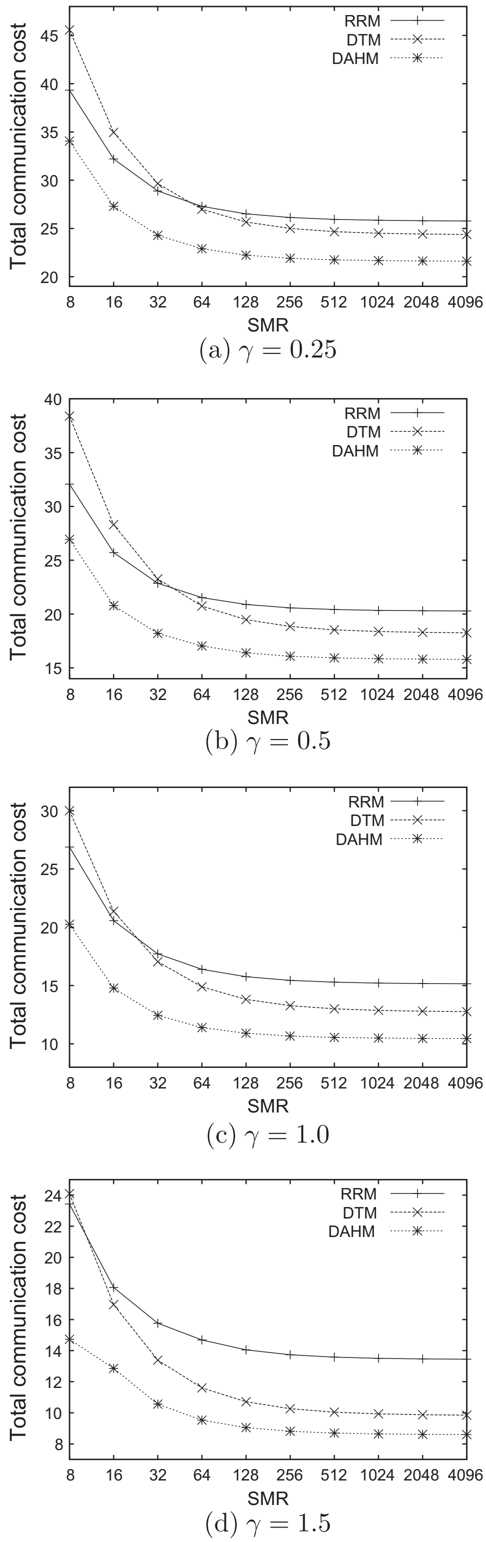


Fig. 15. Performance comparison: cost vs. SMR under different member densities.

batch consists of a large number of runs and therefore a large number of observations for computing an average.

The simulation runs for a minimum of 10 batches, and stops until the calculated mean cost is within 5% from the true mean with a confidence level of 95%.

Fig. 16 shows the analytical results vs. the simulation results for C_{DAHM} as a function of H, under different multicast group sizes. As the figure illustrates, the simulation results show excellent correlations with the analytical results. This justifies that the analytical results are valid and there exists an optimal service region size of an MA, under which DAHM is optimized. Similarly, excellent correlations between the analytical results and simulation results can be seen in Fig. 17, which illustrates the analytical results vs. the simulation results for C_{DAHM} as a function of γ.

Figs. 18 and 19 plot the analytical results vs. the simulation results for the performance comparison among the three algorithms, as a function of M and n, respectively. Again, the analytical results are well correlated with the simulation results in both figures. The perfect correlation between the analytical results and simulation results shown above justifies that the analytical results obtained in the paper are valid.

5.4. Discussion

Based on the analytical and simulation results presented above, we can draw the conclusion that DAHM significantly outperforms both RRM and DTM in a broad spectrum of configurations. This is because DAHM combines backbone multicast routing and local unicast routing into an integrated algorithm, and dynamically determines the optimal service region size of MAs to optimize multicast packet delivery, multicast tree maintenance, and group membership management collectively. Compared with RRM, the packet delivery cost at the upper level of the hierarchy in DAHM is significantly reduced. Compared with DTM, in addition to the reduction of the multicast packet delivery cost, the signaling cost for multicast tree maintenance and membership management in DAHM is significantly reduced.

The performance metric discussed thus far is based on C_{DAHM} (specified by Eq. (11)) which is the number of hops of wireless transmissions incurred per MC per time unit. Below we relate C_{DAHM} minimization with end-to-end delay minimization and throughput maximization. Since every hop incurs a packet transmission by an MR, C_{DAHM} essentially is the amount of traffic incurred to the network per MC. Let $C_{DAHM,i}$ denote the traffic generated by MC i. Then, the total traffic incurred to the network by all MCs is given by $\sum_i C_{DAHM,i}$. Consequently, the average input traffic toward each MR in the system is the total traffic divided by the number of MRs in the network. By utilizing simple arguments of collision theory [38] and queueing theory

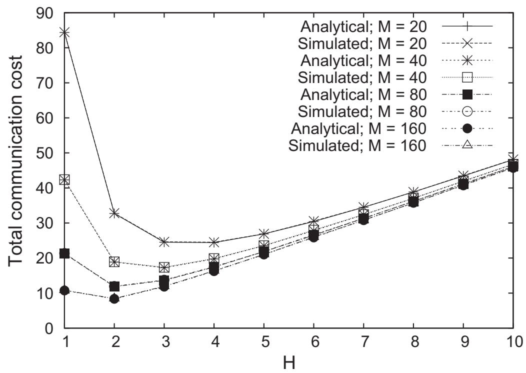


Fig. 16. Analytical modeling vs. simulation: cost vs. H, under different multicast group sizes in DAHM ($n = 10$).

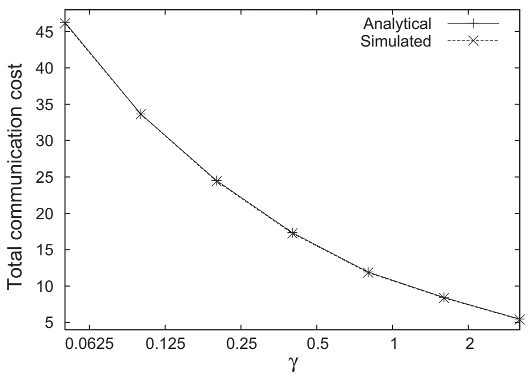


Fig. 17. Analytical modeling vs. simulation: cost vs. γ in DAHM.

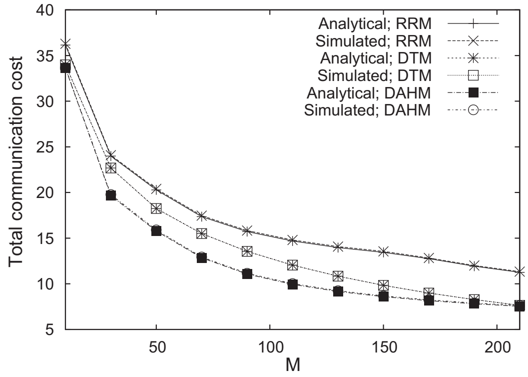


Fig. 18. Analytical modeling vs. simulation: cost vs. $M(n = 10)$.

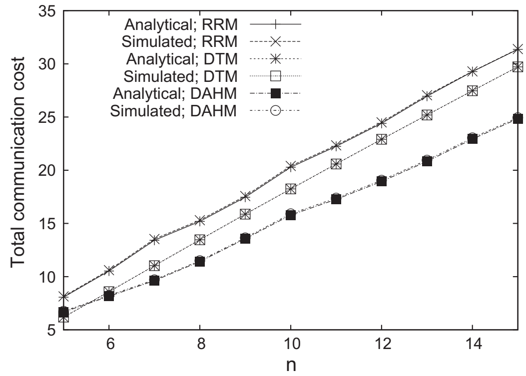


Fig. 19. Analytical modeling vs. simulation: cost vs. $n(M = 50)$.

[39], it can be proven that the per-hop packet delay (including the queueing delay and the retransmission delay because of collision) at any MR is minimized when the input traffic to the MR is minimized. Consequently, the end-to-end delay of a multicast packet to an MC is also minimized. By Little's Law [39] which states that throughput multiplied with response time (end-to-end delay) is equal to the MC population, we can deduce that the network throughput is maximized when the end-to-end delay is minimized, which happens when MC i operates at the optimal $H_{optimal}$ value as identified in our analysis to minimize $C_{DAHM,i}$.

6. Practicability and implementation

We discuss in this section practical issues related to the implementation of DAHM on real mobile devices that can be highly diverse with respect to their computing power and storage capacity. One important issue is how to dynamically determine $H_{optimal}$ at runtime. For powerful mobile devices that are equipped with state-of-the-art processors, the computational procedure developed in this paper can be easily executed to determine $H_{optimal}$ at runtime on a periodic basis. For mobile devices that are less powerful, a simple table-lookup approach can be used to determine $H_{optimal}$ at runtime. The lookup table lists the optimal service region size $H_{optimal}$ for minimizing C_{DAHM} in Eq. (11), given a set of input parameter values as input. The table is built by applying the computational procedure developed in the paper at design time over a perceivable range of input parameter values (as listed in Table 3) specifying the operational and environment conditions. At runtime, $H_{optimal}$ can be determined by looking up in the table using the estimated values of those parameters as keys. Overall, the implementation can be lightweight and very efficient.

To execute the computational procedure presented in the paper, a mobile device needs to first collect data for estimating the values of parameters such as the mobility rate (σ), the multicast packet rate (λ_p), and the rates of member join/leave events (λ and μ). σ can be estimated periodically by an MC by counting the number of serving MR changes during a fixed interval, say, every 30 min. A serving MR change can be detected by a change in the ID number of the current serving MR. Specifically, the MC maintains a counter for the number of serving MR changes, and the counter is incremented whenever the MC changes its serving MR. At the end of each interval, the mobility rate is calculated and the counter is reset. Similarly, the MC can dynamically estimate λ_p by monitoring the sequence numbers of received multicast packets. λ and μ can be monitored dynamically by the source and periodically distributed to the group members.

7. Conclusion

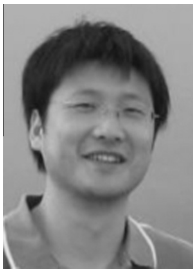
In this paper, we proposed an efficient multicast algorithms for WMNs, namely, Dynamic Agent-based Hierarchical Multicast (DAHM), which supports member mobility and dynamic group membership during the lifetime of a multicast group. DAHM employs a dynamic two-level hierarchical multicast structure, consisting of an upper-level backbone multicast tree rooted at the source with MAs as leaves, and lower-level local multicast groups rooted at the MAs. DAHM leverages and dynamically selects MAs for integrated mobility and multicast service management. MAs are dynamically selected and added to or removed from the backbone multicast tree due to the mobility of multicast group members and dynamic group membership changes. The optimal service region size of an MA that optimizes the performance of DAHM can be dynamically determined using the analytical method presented in the paper. Based on the analytical and simulation results obtained through a comparative performance study, we showed that DAHM significantly outperforms two baseline multicast algorithms for WMNs, namely, Regional-Registration based Multicast (RRM) and Dynamic Tree-based Multicast (DTM).

There are several research directions extending from this work, including (a) investigating how the proposed multicast algorithm can be adapted to support multiple multicast groups simultaneously active in a WMN; (b) utilizing MCs in addition to MRs serving as MAs when a group member cannot find a nearby MR and must rely on other MCs for network traffic relaying; (c) considering the effect of lossy and heterogeneous links of WMNs as in [37] to enhance multicast service performance; (d) conducting more simulation validation based on ns3 in addition to SMPL; and (e) investigating how DAHM can be augmented and optimized to support reliable and secure multicast in WMNs.

References

[1] I.F. Akyildiz, X. Wang, W. Wang, Wireless mesh networks: a survey, Computer Networks 47 (4) (2005) 445–487.
[2] R. Morales, S. Monnet, I. Gupta, G. Antoniu, Move: design and evaluation of a malleable overlay for group-based applications, IEEE Transactions on Network and Service Management 4 (2) (2007) 107–116.
[3] M. Younis, S. Farrag, B. Althouse, TAM: a tiered authentication of multicast protocol for ad-hoc networks, IEEE Transactions on Network and Service Management 9 (1) (2012) 100–113.
[4] T. Pongthawornkamol, I. Gupta, AVCast: new approaches for implementing generic availability-dependent reliability predicates for multicast receivers, IEEE Transactions on Network and Service Management 4 (2) (2007) 117–126.
[5] I.R. Chen, D.C. Wang, Regional registration-based mobile multicast service management in mobile IP networks, Wireless Personal Communications 54 (2009) 635–649.
[6] U.T. Nguyen, On multicast routing in wireless mesh networks, Computer Communications 31 (7) (2008) 1385–1399.
[7] Y. Huh, C. Kim, mMOM: efficient mobile multicast support based on the mobility of mobile hosts, Wireless Networks 12 (2005) 171–178.
[8] C. Lin, K.-M. Wang, Mobile multicast support in IP networks, in: 19th Annual Joint Conference of the IEEE Computer and Communications Societies vol. 3, March 2000, pp. 1664 –1672.
[9] V. Chikarmane, C.L. Williamson, R.B. Bunt, W.L. Mackrell, Multicast support for mobile hosts using mobile IP: design issues and proposed architecture, Mobile Networks and Applications 3 (1998) 365–379.
[10] T.G. Harrison, C.L. Williamson, W.L. Mackrell, R.B. Bunt, Mobile multicast (mom) protocol: multicast support for mobile hosts, in: 3rd Annual ACM/IEEE International Conference on Mobile Computing and Networking, New York, USA, 1997, pp. 151–160.
[11] S.-J. Lee, W. Su, J. Hsu, M. Gerla, R. Bagrodia, A performance comparison study of ad hoc wireless multicast protocols, in: 19th Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 2, 2000, pp. 565–574.

[12] G. Zeng, B. Wang, Y. Ding, L. Xiao, M.W. Mutka, Efficient multicast algorithms for multichannel wireless mesh networks, IEEE Transactions on Parallel and Distributed Systems 21 (1) (2010) 86–99.
[13] D. Koutsonikolas, Y.C. Hu, C. Wang, Pacifier: high-throughput, reliable multicast without "Crying babies" in wireless mesh networks," in: IEEE INFOCOM, Rio de Janeiro, Brazil, April 2009, pp. 2473–2481.
[14] J. Yuan, Z. Li, W. Yu, B. Li, A Cross-Layer optimization framework for multihop multicast in wireless mesh networks, IEEE Journal on Selected Areas in Communications 24 (11) (2006) 2092–2103.
[15] P.M. Ruiz, F.J. Galera, C. Jelger, T. Noel, Efficient multicast routing in wireless mesh networks connected to internet, in: 1st International Conference on Integrated Internet Ad Hoc and Sensor Networks, ACM Press, 2006.
[16] I. Chakeres, C. Koundinya, P. Aggarwal, Fast, efficient, and robust multicast in wireless mesh networks, in: 5th ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, ACM Press, 2008.
[17] S. Roy, D. Koutsonikolas, S. Das, Y. Hu, High-throughput multicast routing metrics in wireless mesh networks, Ad Hoc Networks 6 (2008) 878–899.
[18] J. Dong, R. Curtmola, C. Nita-Rotaru, Secure high-throughput multicast routing in wireless mesh networks, IEEE Transactions on Mobile Computing 10 (5) (2011) 653–668.
[19] S. Das, H. Pucha, Y. Hu, Mitigating the gateway bottleneck via transparent cooperative caching in wireless mesh networks, Ad Hoc Networks 5 (2007) 680–703.
[20] S.M. Das, H. Pucha, Y.C. Hu, Distributed hashing for scalable multicast in wireless ad hoc networks, IEEE Transactions on Parallel and Distributed Systems 19 (3) (2008) 347–362.
[21] D. Koutsonikolas, S.M. Das, Y.C. Hu, I. Stojmenovic, Hierarchical geographic multicast routing for wireless sensor networks, Wireless Networks 16 (2008) 449–466.
[22] Y. Li, I.R. Chen, Adaptive per-user per-object cache consistency management for mobile data access in wireless mesh networks, Journal of Parallel and Distributed Computing 71 (2011) 1034–1046.
[23] Y. Li, I.R. Chen, Design and performance analysis of mobility management schemes based on pointer forwarding for wireless mesh networks, IEEE Transactions on Mobile Computing 10 (3) (2011) 349–361.
[24] Y. Li, I.R. Chen, Mobility management in wireless mesh networks utilizing location routing and pointer forwarding, IEEE Transactions on Network and Service Management 9 (3) (2012) 226–239.
[25] B. Gu, I.R. Chen, Performance analysis of location-aware mobile service proxies for reducing network cost in personal communication systems, Mobile Networks and Applications 10 (4) (2005) 453–463.
[26] I.R. Chen, T.M. Chen, C. Lee, Performance evaluation of forwarding strategies for location management in mobile networks, The Computer Journal 41 (4) (1998) 243–253.
[27] I.R. Chen, T.M. Chen, C. Lee, Agent-based forwarding strategies for reducing location management cost in mobile networks, Mobile Networks and Applications 6 (2) (2001) 105–115.
[28] C. Hirel, B. Tuffin, K.S. Trivedi, SPNP: stochastic Petri Nets version 6.0, in: 11th International Conference on Computer Performance Evaluation: Modelling Techniques and Tools, London, UK, 2000, pp. 354–357.
[29] I.R. Chen, D.C. Wang, Analyzing dynamic voting using Petri nets, in: 15th IEEE Symposium on Reliable Distributed Systems, Niagara Falls, Canada, October 1996, pp. 44–53.
[30] S.T. Cheng, C.M. Chen, I.R. Chen, Dynamic quota-based admission control with sub-rating in multimedia servers, Multimedia Systems 8 (2) (2000) 83–91.
[31] S.T. Cheng, C.M. Chen, I.R. Chen, Performance evaluation of an admission control algorithm: dynamic threshold with negotiation, Performance Evaluation 52 (1) (2003) 1–13.
[32] L. Miller, Connectivity Properties of Mesh and Ring/Mesh Networks, April 2001. <http://w3.antd.nist.gov/wctg/netanal/MeshCon.pdf>.
[33] J.C.I. Chuang, M.A. Sirbu, Pricing multicast communication: a cost-based approach, Telecommuniation Systems 17 (2001) 281–297.
[34] R. Chalmers, K. Almeroth, On the topology of multicast trees, IEEE/ACM Transactions on Networking 11 (2003) 153–165.
[35] P.M. Ruiz, A.F. Gomez-Skarmeta, Approximating optimal multicast trees in wireless multihop networks, in: IEEE Symposium on Computers and Communications, Los Alamtos, CA, USA, 2005, pp. 686–691.
[36] M.H. MacDougall, Simulating Computer Systems: Techniques and Tools, MIT Press, Cambridge, MA, USA, 1987.
[37] G. Zeng, B. Wang, M. Mutka, L. Xizo, E. Torng, Efficient link-heterogeneous multicast for wireless mesh networks, Wireless Networks 18 (2012) 605–620.
[38] I.R. Chen, Y. Wang, Reliability analysis of wireless sensor networks with distributed code attestation, IEEE Communications Letters 16 (10) (2012) 1640–1643.
[39] K.S. Trivedi, Probability and Statistics with Reliability, Queueing, and Computer Science Applications, John Wiley & Sons, Inc., New York, 2002.



Yinan Li received the B.S. degree from Xi'an Jiaotong University in China, the M.S. degree from the University of Tennessee in 2008, and the Ph.D. degree from Virginia Tech in 2012, all in Computer Science. His research interests include wireless networks, mobile ad hoc networks, sensor networks, network security, high performance computing, and dependable computing.

Ing-Ray Chen received the B.S. degree from the National Taiwan University, Taipei, Taiwan, and the MS and Ph.D. degrees in computer science from the University of Houston. He is a professor in the Department of Computer Science at Virginia Tech. His research interests include mobile computing, wireless systems, security, trust management, data management, real-time intelligent systems, and reliability and performance analysis. He is currently serves as an editor for IEEE Communications Letters, IEEE Transactions on Network and Service Management, The Computer Journal, Wireless Personal Communications, Wireless Communications and Mobile Computing, and Security and Network Communications. He is a member of the IEEE and ACM.