

Bandwidth Efficient Video Multicasting in Multiradio Multicellular Wireless Networks

De-Nian Yang, *Member, IEEE*, and Ming-Syan Chen, *Fellow, IEEE*

Abstract—In this paper, we propose a new mechanism to select the cells and the wireless technologies for layer-encoded video multicasting in the heterogeneous wireless networks. Different from the previous mechanisms, each mobile host in our mechanism can select a different cell with a different wireless technology to subscribe each layer of a video stream, and each cell can deliver only a subset of layers of the video stream to reduce the bandwidth consumption. We formulate the Cell and Technology Selection Problem (CTSP) to multicast each layer of a video stream as an optimization problem. We use Integer Linear Programming to model the problem and show that the problem is NP-hard. To solve the problem, we propose a distributed algorithm based on Lagrangean relaxation and a protocol based on the proposed algorithm. Our mechanism requires no change of the current video multicasting mechanisms and the current wireless network infrastructures. Our algorithm is adaptive not only to the change of the subscribers at each layer, but also the change of the locations of each mobile host.

Index Terms—Multicast, layer-encoded video, heterogeneous wireless networks.

1 INTRODUCTION

THE success of wireless and mobile communications in the 21st century has resulted in a large variety of wireless technologies such as second- and third-generation cellulars, satellite, Wi-Fi, and Bluetooth. The heterogeneous wireless networks combine various wireless networks and provide universal wireless access. Users in the heterogeneous wireless networks are usually covered by more than one cell to avoid connection drop and service disruption. In addition, more mobile terminals in the wireless networks are likely to own multiple wireless technologies. Therefore, the heterogeneous wireless networks provide the mobile hosts with many choices for the cell and the wireless technologies to access the Internet.

Video delivery in wireless networks is becoming an important multimedia application due to the proliferation of the Web-based services and the rapid growth of wireless communication devices. For a video stream delivered to a single mobile host, a video server can adaptively adjust the encoder to accommodate the delay, jitter, and packet loss of the networks [1], [2], [3]. For a video stream delivered to multiple mobile hosts with diverse requirements and network conditions, the video stream can be encoded at the highest resolution and divided into *layers* such that each receiver can decode the stream on the preferred rate and resolution with a set of layers [4], [5]. The most significant layer, that is, the base layer, contains the data representing the most important features of the video, whereas the additional layers, that is, the enhancement layers, contain data that progressively refine the reconstructed video

quality. The layers are distributed to receivers via multicast channels in wireless networks [6].

Previous works on layer-encoded video multicast mainly focus on deciding the set of subscribed layers for the receivers with heterogeneous requirements and network conditions. McCanne et al. [7] propose a protocol to deliver the required layers to each of the receivers. Vickers et al. [8] adaptively adjust the number of layers and the bit rate of each layer according to the network condition. Kim and Ammar [9] develop a quality adaptation scheme that maximizes the perceptual video quality through minimizing quality variation. Liu et al. [10] find the optimal number of layers and the amount of bandwidth for each layer to maximize the total utilities of all mobile hosts. Zhao et al. [11] determine the optimal power allocation and the transmission rate of the base station to deliver each layer of a video stream in CDMA networks. Therefore, previous works assume that the network operators have selected the set of cells to multicast a video stream. Moreover, each cell must multicast the layers of a video stream in the progressive manner. In other words, a cell cannot deliver the enhancement layers without multicasting the base layer of a video stream.

In this paper, different from the previous works, we focus on the selection of the cells and the wireless technologies to multicast each layer of a video stream in the heterogeneous wireless networks, where each selected cell can deliver only a *subset* of layers of the video stream to reduce the total bandwidth consumption. In our approach, each mobile host can select a different cell with a different wireless technology to subscribe each layer of a video stream. Consider Fig. 1 as an example, where A, B, C, and D are the mobile hosts. Here, we assume that the bandwidth allocated to each Universal Mobile Telecommunications System (UMTS) and each Wi-Fi cell can support one and two layers of a video stream. The solid line connecting a mobile host and a base station represents that the mobile

• The authors are with the Department of Electrical Engineering, National Taiwan University, #1 Roosevelt Rd., Sec. 4, Taipei, Taiwan 106, ROC. E-mail: {dnyang, mschen}@cc.ee.ntu.edu.tw.

Manuscript received 11 Sept. 2005; revised 29 July 2006; accepted 17 Dec. 2006; published online 28 June 2007.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-0272-0905. Digital Object Identifier no. 10.1109/TMC.2007.70725.

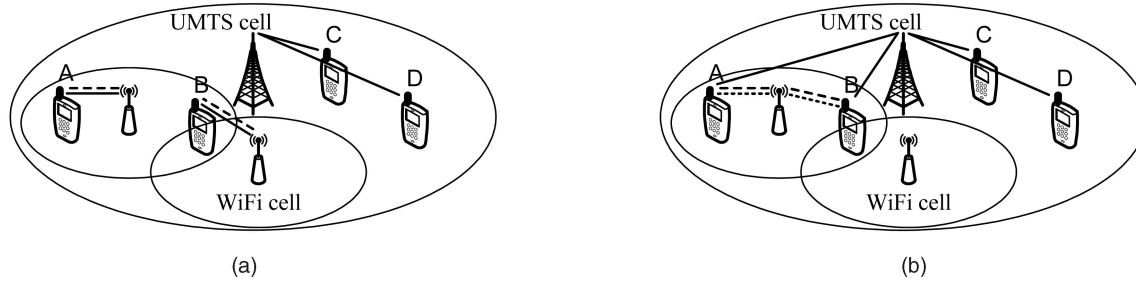


Fig. 1. Comparison of two different selections of cells and technologies to deliver each layer of a video stream. (a) Each mobile host subscribes all layers from a single cell. (b) Each mobile host can subscribe each layer from a different cell.

host subscribes Layer 1, the base layer, of a video stream from the cell. The long dash line and the short dash line correspond to Layer 2 and Layer 3, the enhancement layers, of the video stream. For the network operators, each layer consumes less network bandwidth in our approach. The three cells multicast Layer 1 in Fig. 1a. In Fig. 1b, however, the two Wi-Fi cells do not multicast Layer 1 because mobile hosts A and B subscribe the layer from the UMTS cell. Note that mobile hosts A and B subscribe Layer 2 and 3 from a single Wi-Fi cell to minimize the total bandwidth consumption in Fig. 1b. For the users, each mobile host can subscribe more layers in our approach. Mobile hosts A and B subscribe two layers if they subscribe all layers from only the Wi-Fi cells in Fig. 1a. On the contrary, mobile hosts A and B subscribe three layers in Fig. 1b. They subscribe Layer 1 from the UMTS cell and Layers 2 and 3 from the Wi-Fi cells. Subscribing the layers from the cells with multiple wireless technologies may consume more power. Some mobile hosts are concerned with the video quality but others are concerned with the power consumption. To capture the heterogeneity of users, therefore, the number of wireless technologies allowed to be used by each mobile host can be different in this paper.

Explicitly, we formulate in this paper the selection of the cell and the wireless technology to multicast each layer of a video stream as an optimization problem, which is referred as the Cell and Technology Selection Problem (CTSP) in the heterogeneous wireless networks for layer-encoded video multicasting. The objective of the problem is to minimize the total bandwidth cost of the selected cells and the wireless technologies. We design a mechanism, which includes an Integer Linear Programming (ILP) formulation, a distributed algorithm, and a network protocol to solve the CTSP. The network operators can use our ILP formulation for network planning. We show that the problem is NP-hard and design an algorithm LAGRANGE, which is based on Lagrangean relaxation [12] on our ILP formulation. Lagrangean relaxation has been widely used in various network design, control, and pricing problems [13], [14], [15], [16], [17], and the convergence of Lagrangean relaxation has been proved in the literature [12]. We adopt Lagrangean relaxation in our algorithm instead of other optimization techniques due to the following reasons: First, our algorithm decomposes the original problem into multiple subproblems such that each subproblem can be solved by each mobile host individually. In other words, the algorithm can be implemented in the distributed

manner, and the important merit of algorithm LAGRANGE enables us to design a network protocol accordingly. Second, the algorithm adapts to the change of the subscribers at each layer of a video stream and the change of locations of each mobile host. The algorithm iteratively reduces the total bandwidth consumption according to the current subscribers of each layer at each location. Third, the algorithm provides the lower bound on the total bandwidth cost of the optimal selected cells and the wireless technologies. For a video stream with a large number of subscribers, the lower bound obtained by our algorithm provides a benchmark to compare with any algorithm for the problem since using the ILP formulation here to find the optimal solution is computationally infeasible. In addition, the algorithm requires no change of the current video multicasting mechanisms and the current wireless network infrastructures.

Algorithm LAGRANGE can be regarded as a rerouting mechanism. The network may use different cells to multicast each layer of a video stream when the network condition changes. Note that rerouting mechanisms have been designed for unicast communications in circuit switched networks [18], optical networks [19], and satellite networks [20] to reduce the total bandwidth consumption. However, designing a rerouting mechanism for layer-encoded video multicast in the heterogeneous wireless networks is more difficult. First, the selection of cells for any two layers is correlated because each mobile host can use only a limited number of cells and wireless technologies. Second, we show that the multicast communication makes our problem become NP-hard in Section 2. Since our mechanism can be regarded as a rerouting mechanism, a mobile host in our mechanism needs the sophisticated handover mechanisms such as those in the literature [21], [22], [23], [24] to subscribe the video stream from the proper cells in different situations.

Overall, the contributions of this paper are many-fold:

- We consider the heterogeneity of mobile hosts. Some mobile hosts would like to improve the video quality by subscribing more layers from more cells with proper wireless technologies. Other mobile hosts may concern the power consumption and thereby subscribe a limited number of layers with a single cell.
- For each wireless technology, our mechanism reduces the number of cells to multicast a video

stream by clustering the mobile hosts. Therefore, we can reduce the wireless bandwidth consumption even when the operators own only one wireless network.

- Our mechanism chooses the proper wireless technologies to minimize the total bandwidth cost of all wireless networks. For a set of nearby mobile hosts, our algorithm minimizes the bandwidth cost by using a single large cell or multiple smaller cells to serve these mobile hosts according to the number of mobile hosts, the locations of mobile hosts, and the bandwidth cost of each wireless technology specified by the network operators.
- Our mechanism is flexible since the bandwidth cost of each cell can be assigned with no restriction, and the cost of each cell can be adjusted dynamically. The flexible cost model enables the network operators to balance the load of wireless cells with the same or different technologies by adjusting the cost [25].
- Our mechanism is transparent to the video multicasting mechanisms. Each mobile host subscribes the video stream with the current video multicasting mechanisms after the mobile host selects the cell and the wireless technology according to our mechanism. We thereby require no modification on the current video multicasting mechanisms.
- Our mechanism requires no modification on the current wireless network infrastructures. The algorithm is implemented in only the mobile hosts, and the mobile hosts cooperatively select the cells and the wireless technologies.
- Our mechanism is adaptive. Our algorithm adapts to the change of the subscribers at each layer of a video stream, the change of locations of each mobile host, and the change of the bandwidth cost of each cell. In addition, our mechanism enables each mobile host to either automatically or manually choose the cell and the wireless technology. For example, when a mobile host manually chooses a cell, other nearby mobile hosts adaptively attach to the cell to reduce the total bandwidth consumption.

The rest of the paper is summarized as follows: We describe our assumption, present our ILP formulation, and show that the CTSP is NP-hard in Section 2. We propose algorithm LAGRANGE based on Lagrangean relaxation and the corresponding protocol in Section 3. We show our experimental results in Section 4. Finally, we conclude this paper in Section 5.

2 PROBLEM DESCRIPTION

In this paper, we consider the CTSP in the heterogeneous wireless networks for layer-encoded video multicasting. Video streams are delivered with one hop from the base station to each mobile host. The problem is to select the cells and the wireless technologies to multicast each layer of a video stream such that the total bandwidth cost is minimized. The problem is subject to the constraint that each mobile host can select at most one cell for each

available wireless technology. Moreover, to consider the heterogeneity of mobile hosts, the number of wireless technologies used by each mobile host can be different. In this paper, we concern only the wireless bandwidth consumption of each multicast group, and we assume that the networks have an abundant wireline bandwidth to connect to each cell. In addition, we assume the base station of a cell needs to multicast a layer of a video stream if at least one mobile host in the cell subscribes the layer from the cell. In this paper, a cell *covers* a mobile host if the mobile host is within the transmission range of the base station of the cell, and the cell in this case is a *covering cell* of the mobile host. A cell is a *candidate cell* if it covers at least one mobile host. A wireless technology is a *candidate technology* for a mobile host if the mobile host can access the Internet with the wireless technology. For simplicity, the selection of the cell for each mobile host means the selection of both the cell and the wireless technology in the rest of this paper.

The notation in this paper is summarized as follows:

- C : This is the set of cells in the heterogeneous wireless network.
- L : This is the set of layers of a layer-encoded video stream.
- M : This is the set of mobile hosts in the network.
- T : This is the set of wireless technologies.
- C_m : This is the set of cells that cover mobile host m and have good channel condition, $m \in M$, $C_m \subseteq C$; if mobile host m selects cell c manually, we let C_m contain only a cell c .
- C_t : This is the set of cells with wireless technology t , $t \in T$, $C_t \subseteq C$.
- L_m : This is the set of layers subscribed by mobile host m , $m \in M$, $L_m \subseteq L$.
- M_c : This is the set of mobile hosts covered by cell c , $c \in C$, $M_c \subseteq M$.
- T_m : This is the set of wireless technologies that can be used by mobile host m , $m \in M$, $T_m \subseteq T$.
- $b_{c,l}$: This is the bandwidth cost of cell c to deliver layer l of the video stream, $c \in C$, $l \in L$.
- t_m : This is the number of wireless technologies that can be used by mobile host m , $m \in M$.
- $\mu_{m,c,l}$: This is the Lagrange multiplier for mobile host m with respect to cell c and layer l , $m \in M$, $c \in C_m$, $l \in L_m$.

We use ILP to model the CTSP. The ILP formulation can find the optimal cells and technologies to multicast each layer of a video stream. The network operators can use our ILP formulation for network planning in the heterogeneous wireless networks. Our ILP formulation has the following variables:

- $\pi_{m,c,l}$: This is a binary variable; $\pi_{m,c,l}$ is one if mobile host m selects cell c for layer l of the video stream, $m \in M$, $c \in C_m$, $l \in L_m$.
- $\tau_{m,t}$: This is a binary variable; $\tau_{m,t}$ is one if mobile host m selects any cell with technology t , $t \in T_m$.
- $\phi_{m,c}$: This is a binary variable; $\phi_{m,c}$ is one if mobile host m selects cell c for any subscribed layer, $m \in M$, $c \in C_m$.

- $\sigma_{c,l}$: This is a binary variable; $\sigma_{c,l}$ is one if cell c delivers layer l of the video stream; namely, at least one mobile host subscribes layer l from cell c , $c \in C$, $l \in L$.

The objective function of our ILP formulation is given as follows:

$$\min \sum_{c \in C} \sum_{l \in L} b_{c,l} \times \sigma_{c,l}.$$

Although we do not maximize the number of subscribed layers for each mobile host in our problem, the objective function here minimizes the total wireless bandwidth consumption to deliver the current subscribed layers of the stream to the mobile hosts. Therefore, with the objective function, each mobile host is able to subscribe more layers with a fixed amount of wireless bandwidth and the networks thereby can support more video streams. Our ILP formulation includes the following constraints:

$$\sum_{c \in C_m} \pi_{m,c,l} = 1, \forall m \in M, \forall l \in L_m, \quad (1)$$

$$\pi_{m,c,l} \leq \phi_{m,c}, \forall m \in M, \forall c \in C_m, \forall l \in L_m, \quad (2)$$

$$\sum_{c \in C_t \cap C_m} \phi_{m,c} \leq 1, \forall m \in M, \forall t \in T_m, \quad (3)$$

$$\pi_{m,c,l} \leq \tau_{m,t}, \forall m \in M, \forall t \in T_m, \forall c \in C_t \cap C_m, \forall l \in L_m, \quad (4)$$

$$\sum_{t \in T_m} \tau_{m,t} \leq t_m, \forall m \in M, \quad (5)$$

$$\pi_{m,c,l} \leq \sigma_{c,l}, \forall m \in M, \forall c \in C_m, \forall l \in L_m. \quad (6)$$

Constraint (1) guarantees that each mobile host selects one cell for each subscribed layer. Constraints (2) and (3) enforce that each mobile host selects at most one cell for each available wireless technology. Constraints (4) and (5) state that the number of technologies used by a mobile host must be bounded by t_m . Each mobile host m can be associated with a different t_m to consider the heterogeneity of mobile hosts. Constraint (6) guarantees that a cell needs to multicast a layer of the video stream if any covered mobile host subscribes the layer from the cell. In addition to the above constraints, the problem is also subject to the integrality constraints to enforce that $\pi_{m,c,l}$, $\tau_{m,t}$, $\phi_{m,c}$, and $\sigma_{c,l}$ are all binary variables. We regard a set of selected cells that obey the above constraints as a *feasible solution* to the CTSP.

Previous works for layered video multicasting [7], [8], [9], [10], [11] assume that all layers of a stream are delivered in a single cell. In this paper, however, we show that utilizing different cells and technologies to deliver each layer of a video stream can reduce the bandwidth consumption. In addition, constraints (4) and (5) are formulated for the applications that use multiple multicast groups simultaneously, and we believe that video multicasting nowadays is the only application that is bandwidth demanding and conforms to the above requirement.

We show that the CTSP in the heterogeneous wireless networks for multicast communications is NP-hard because the *Minimum Set Cover* problem [26] is a special case of the CTSP. In the Minimum Set Cover problem, each set is assigned a cost and covers some elements. The problem is to select the sets with the minimum total cost such that every element is covered by at least one selected set.

Therefore, the Minimum Set Cover problem is the same as the CTSP if each mobile host subscribes only a single layer of the video stream and owns only one candidate wireless technology. In other words, the CTSP is more difficult than the Minimum Set Cover problem because our problem allows each mobile host to subscribe multiple layers and use multiple wireless technologies.

3 DESIGN OF ALGORITHM LAGRANGE

In this section, we propose algorithm LAGRANGE, which is based on Lagrangean relaxation on our ILP formulation proposed in Section 2. The algorithm can be implemented in the distributed manner on only mobile hosts. The algorithm adapts to the change of the subscribers of each layer, the change of the location of each mobile host, and the change of the bandwidth cost of each cell. In addition, the algorithm provides a lower bound on the total bandwidth cost of the optimal solution to the CTSP.

The algorithm relaxes a constraint of our ILP formulation and transfers the CTSP into the Lagrangean Relaxation Problem (LRP). The LRP owns a new objective function with the Lagrange multipliers and fewer constraints such that we can decompose the LRP into multiple subproblems, where each subproblem can be solved in a distributed manner by the mobile hosts. The mobile hosts in our algorithm cooperatively select the cells and the wireless technologies for each layer of a video stream according to the solutions to the subproblems. The solutions to the subproblems depend on the Lagrange multipliers of the LRP, and we iteratively adjust the multipliers to change the selected cells to iteratively reduce the total bandwidth consumption. In the rest of the section, we solve the CTSP with the following steps:

- Transfer the CTSP into the LRP.
- Decompose the LRP into multiple subproblems and solve each subproblem, respectively.
- Select the cell and the wireless technology for each subscribed layer of each mobile host according to the solutions to the subproblems.
- Reduce the total bandwidth cost of the selected cells by iteratively updating the cost, the Lagrange multiplier, of each cell for each subscribed layer of each mobile host.

3.1 Transferring the CTSP and Solving the LRP

Our algorithm relaxes constraint (6) in our ILP formulation to transfer the CTSP into the LRP, and the objective function of the LRP is shown as follows:

$$\begin{aligned} & \min \sum_{c \in C} \sum_{l \in L} b_{c,l} \times \sigma_{c,l} + \sum_{m \in M} \sum_{c \in C_m} \sum_{l \in L_m} \mu_{m,c,l} (\pi_{m,c,l} - \sigma_{c,l}) \\ & = \min \left[\sum_{c \in C} \sum_{l \in L} \left(b_{c,l} - \sum_{m: c \in C_m, l \in L_m} \mu_{m,c,l} \right) \sigma_{c,l} \right] \\ & \quad + \left[\sum_{m \in M} \sum_{c \in C_m} \sum_{l \in L_m} \mu_{m,c,l} \pi_{m,c,l} \right], \end{aligned}$$

where $\mu_{m,c,l}$ is the Lagrange multiplier, $\mu_{m,c,l} \geq 0$, $\forall m \in M$, $\forall c \in C_m$, $\forall l \in L_m$. The Lagrange multiplier $\mu_{m,c,l}$ is the cost

associated with cell c for layer l of mobile host m . The LRP includes the constraints (1)-(5) of our ILP formulation. Compared with the objective function of the CTSP, the objective function of the LRP owns a new term corresponding to the relaxed constraint (6). Intuitively, for any feasible solution to the LRP that contradicts the relaxed constraint, namely, $\pi_{m,c,l} > \sigma_{c,l}$, the objective function punishes the solution with a larger objective value. Besides, any feasible solution to the CTSP must also act as a feasible solution to the LRP since the set of constraints of the LRP is a subset of the constraints of the CTSP. Therefore, when we adopt the optimal solution to the CTSP as the feasible solution both to the LRP and the CTSP, the objective value of the LRP must be no more than the objective value of the CTSP because the new term in the objective function of the LRP is nonpositive. Therefore, the objective value of the optimal solution to the LRP must be no more than the objective value of the optimal solution to the CTSP. In other words, the optimal solution to the LRP provides a lower bound on the objective value of the optimal solution to the CTSP, which is the total bandwidth cost of the selected cells and the wireless technologies in the heterogeneous wireless networks.

We solve the LRP by decomposing the LRP into two subproblems. We divide the objective function and the constraints of the LRP into two parts, where each subproblem owns one part of the objective function and the constraints. The variables in the two subproblems are mutually independent such that we can solve each subproblem individually, and the solution to the LRP is just the combination of the solutions to the two subproblems.

The objective function of the first subproblem is the first part of the objective function in the LRP:

$$\min \sum_{c \in C} \sum_{l \in L} \left(b_{c,l} - \sum_{m: c \in C_m, l \in L_m} \mu_{m,c,l} \right) \sigma_{c,l}.$$

The first subproblem has only an integrality constraint enforcing that $\sigma_{c,l}$ must be a binary variable because $\sigma_{c,l}$ is not in constraints (1)-(5) of the LRP. Intuitively, each cell c in the subproblem is associated with a cost $b_{c,l}$ and a profit $\sum_{m: c \in C_m, l \in L_m} \mu_{m,c,l}$ for each layer l , and the subproblem is to minimize the *net cost* of each cell c for each layer l . Therefore, cell c needs to be selected for layer l in the subproblem if the cost is less than the profit. In other words, $\sigma_{c,l}$ is one in the solution to the subproblem if $b_{c,l}$ is less than $\sum_{m: c \in C_m, l \in L_m} \mu_{m,c,l}$.

The objective function of the second subproblem is the second part of the objective function in the LRP:

$$\min \sum_{m \in M} \sum_{c \in C_m} \sum_{l \in L_m} \mu_{m,c,l} \pi_{m,c,l}.$$

The second subproblem includes constraints (1)-(5) and the integrality constraints for the variables. In the subproblem, each mobile host m needs to select a cell c for each subscribed layer l , where each choice is associated with a nonnegative cost $\mu_{m,c,l}$. The subproblem is to minimize the total cost of the selected cells for all mobile hosts.

In the following, we first prove that the second subproblem is NP-hard:

Theorem 1. *The second subproblem of the LRP is NP-hard.*

Proof. We prove the theorem with reduction from *Maximum k -Facility Location* (MFL) problem, which is NP-hard [26]. The problem is given a set of vertices V in a complete graph, where each edge $e_{t,l}$ connecting vertices t and l is associated with a profit $p_{t,l}$. The problem is to find a subset of vertices $\bar{V} \subseteq V$ with $|\bar{V}| \leq k$ to maximize the following total profit:

$$\sum_{l \in \bar{V}} \max_{t \in \bar{V}} p_{t,l}.$$

For each instance in the MFL problem, we create the corresponding instance of our second subproblem as follows: The instance in our second subproblem has only one mobile host m . The instance contains a set of layers V of a video stream, and the set of technologies that can be accessed by m is also V . Each technology has only one cell that can be selected by m . For each layer l and the cell c with technology t , we assign Lagrange multiplier $\mu_{m,c,l}$ as follows:

$$\mu_{m,c,l} = \bar{q} - p_{t,l},$$

where \bar{q} is $\max_{u,v \in V} p_{u,v}$. Each host m can use only k technologies, namely, $t_m = k$.

We prove the theorem by showing that the total profit of any MFL instance is \bar{p} if and only if $|V| \times \bar{q} - \bar{p}$ is the total cost in the corresponding instance of our second subproblem. We first prove the sufficient condition. If the set of vertices selected in the MFL instance is \bar{V} , then we let mobile host m select the corresponding set of technologies \bar{V} . Therefore, for each layer l , the cell selected by mobile host m for layer l must have the following cost:

$$\min_{t: t \in \bar{V}, c \in C_t} \mu_{m,c,l} = \min_{t \in \bar{V}} \bar{q} - p_{t,l} = \bar{q} - \max_{t \in \bar{V}} p_{t,l}.$$

Therefore, the total cost of the instance in our second subproblem is given as follows:

$$\begin{aligned} \sum_{l \in V} \min_{c: t \in \bar{V}, c \in C_t} \mu_{m,c,l} &= \sum_{l \in V} \left(\bar{q} - \max_{t \in \bar{V}} p_{t,l} \right) \\ &= |V| \times \bar{q} - \sum_{l \in V} \max_{t \in \bar{V}} p_{t,l} = |V| \times \bar{q} - \bar{p}. \end{aligned}$$

The necessary condition can be proved similarly. \square

Although the second subproblem is NP-hard, the subproblem can be solved in a polynomial time if we assume the number of technologies that can be used by each mobile host is a small constant, irrelevant to the input size of the subproblem. Moreover, we assume that the number of covering cells for each mobile host is also a small constant. Based on the assumption, we solve the subproblem as follows: If the set of cells selected by a mobile host m is \tilde{C} in any solution, then m must select the cell $c_{m,l}^{\tilde{C}}$ with the minimum cost for each subscribed layer l , namely,

$$\tilde{c}_{m,l}^{\tilde{C}} = \arg \min_{c \in \tilde{C}} \mu_{m,c,l}.$$

Therefore, the total cost $\kappa_m^{\tilde{C}}$ of the selected cells for each mobile host m is given as follows:

$$\kappa_m^{\tilde{C}} = \sum_{l \in L_m} \mu_{m, c_{m,l}^{\tilde{C}}}$$

To find the optimal solution, we consider every possible set \tilde{C} , where $\tilde{C} \subseteq C_m$. Set \tilde{C} must follow two rules: First, the number of technologies used in \tilde{C} is no more than t_m . Second, each technology has at most one cell in \tilde{C} . Let \bar{C} denote the set of cells that induces the minimum cost, namely,

$$\bar{C} = \arg \min_{\tilde{C}} \left\{ \kappa_m^{\tilde{C}} \right\}.$$

Therefore, for each subscribed layer l , let $c_{m,l}^{\bar{C}}$ denote the cell selected by mobile host m ; in other words,

$$\pi_{m,c,l} \leftarrow \begin{cases} 1, & \text{if } c = c_{m,l}^{\bar{C}}, \\ 0, & \text{otherwise.} \end{cases}$$

The number of possible set \tilde{C} required to consider is $\sim O(|C_m|^{t_m})$, and we need $O(|L_m|t_m)$ comparisons to find $c_{m,l}^{\bar{C}}$ for all layers in L_m . Therefore, the time complexity of our algorithm for the second subproblem of each member m is $O(|L_m|t_m|C_m|^{t_m})$. We believe that the algorithm is suitable to be implemented in mobile hosts due to the following reasons: First, the algorithm is simple. For each set \tilde{C} , the algorithm simply finds the cell with the minimum cost, namely, the Lagrange multiplier, for each subscribed layer. Second, since each mobile host tends to be covered by a few cells and is able to access several wireless technologies, we can regard $|C_m|^{t_m}$ as a small constant, and our algorithm thereby can find the optimal solution to the second subproblem in reasonable time.

Since the variables to the first and second subproblems of the LRP are mutually independent, the solution to the LRP is just the combination of the solutions to the two subproblems. With the solution, algorithm LAGRANGE is able to find and improve the solution to the CTSP.

3.2 Finding and Improving the Solution to the CTSP

Although the first subproblem selects the cells with the minimum net cost, the selected cells may not be feasible to the CTSP because each mobile host is not guaranteed to be covered by at least one selected cell. Therefore, our algorithm selects the cells to deliver each layer of a video stream according to the solution to the second subproblem. Let $\tilde{\sigma}_{c,l}$ be a binary variable that represents if our algorithm selects cell c for the layer l of the video stream. Variable $\tilde{\sigma}_{c,l}$ as one if there exists at least one mobile host m selecting cell c in the second subproblem. In other words, variable $\pi_{m,c,l}$ is one in the second subproblem. The advantage of algorithm LAGRANGE is that the cell selection is performed only by the mobile host and is transparent to the video multicasting mechanisms. In other words, each mobile host subscribes a layer with the video multicasting mechanisms after it selects the cell and the wireless technology according to our algorithm. We thereby require no modification on the current video multicasting mechanisms. In addition, algorithm LAGRANGE is implemented in only the mobile host

and requires no modification on the current wireless network infrastructures.

Each mobile host m in algorithm LAGRANGE selects the cell c according to the cost $\mu_{m,c,l}$, the Lagrange multiplier, in the second subproblem. To reduce the total bandwidth consumption, we adjust the cost iteratively with the subgradient algorithm [12] and the solutions to the two subproblems of the LRP. Let $W(\mu)$ denote the objective function of the LRP in Section 3.1, where

$$\mu = (\mu_{m,c,l}, \forall m \in M, \forall c \in C_m, \forall l \in L_m).$$

The subgradient corresponding to the optimal solution of the LRP is denoted as

$$\nabla W(\mu) = (\partial W(\mu)/\partial \mu_{m,c,l}, \forall m \in M, \forall c \in C_m, \forall l \in L_m),$$

where

$$\frac{\partial W(\mu)}{\partial \mu_{m,c,l}} = \pi_{m,c,l} - \sigma_{c,l}.$$

The feasible solution to the CTSP obtained by algorithm LAGRANGE depends on the cost $\mu_{m,c,l}$, and the subgradient $\partial W(\mu)/\partial \mu_{m,c,l}$ indicates the direction of adjusting $\mu_{m,c,l}$ to find the better feasible solution to the CTSP, $\forall m \in M, \forall c \in C_m, \forall l \in L_m$. Algorithm LAGRANGE increases or decreases the cost $\mu_{m,c,l}$ at each iteration based on the subgradient $\partial W(\mu)/\partial \mu_{m,c,l}$. Algorithm LAGRANGE increases $\mu_{m,c,l}$ when $\pi_{m,c,l} - \sigma_{c,l}$ is positive. In other words, mobile host m selects cell c for layer l , but the cell is not selected in the first subproblem in this case. In contrast, algorithm LAGRANGE decreases $\mu_{m,c,l}$ when $\pi_{m,c,l} - \sigma_{c,l}$ is negative. In other words, mobile host m does not select cell c for layer l , but the cell is selected in the first subproblem in this case.

We explain the adjustment of the costs in an intuitive way as follows: Although the solution to the first subproblem may not find a feasible solution to the CTSP, the first subproblem provides an insight to select the cells with low bandwidth cost. The first subproblem tends to select the cells that cover more mobile hosts to save the wireless bandwidth because the cell tends to a own larger profit, where the profit of each cell c for layer l is $\sum_{m:c \in C_m} \mu_{m,c,l}$. Therefore, if mobile host m does not select cell c for layer l but the cell is selected in the first subproblem, we decrease $\mu_{m,c,l}$ such that the cell will more likely be chosen by m for layer l afterward. Therefore, the cost $\mu_{m,c,l}$, which is the Lagrange multiplier in the LRP, plays an important role to reduce the total bandwidth consumption.

Fig. 2 shows the details of our algorithm. Initially, algorithm LAGRANGE assigns a unit cost to each cell for each subscribed layer of each mobile host in step one, and each mobile host thereby can select any cell. Afterward, our algorithm iteratively reduces the total bandwidth cost of the selected cells and the wireless technologies. At each iteration, our algorithm first finds the solutions to the first subproblem and the second subproblem in steps two and three, respectively. The algorithm then adjusts the cost of each cell for each subscribed layer of each mobile host in step four such that we can select the cells and the wireless technologies with lower bandwidth cost at the next

```

1. Let  $\eta \leftarrow 1$ ,  $\mu_{m,c,l} \leftarrow 1$ ,  $\forall m \in M, \forall c \in C_m, \forall l \in L_m$ .
2. For each cell  $c$  and each layer  $l$ ,
   if  $b_{c,l} < \sum_{m:c \in C_m, l \in L_m} \mu_{m,c,l}$ , then let  $\sigma_{c,l} \leftarrow 1$ , else let  $\sigma_{c,l} \leftarrow 0$ ;
   end for.
3. For each mobile host  $m$ ,
   for every possible set  $\tilde{C}$ ,
     let  $c_{m,l}^{\tilde{C}} \leftarrow \arg \min_{c \in \tilde{C}} \mu_{m,c,l}$ ,  $\forall l \in L_m$ ; let  $\kappa_m^{\tilde{C}} \leftarrow \sum_{l \in L_m} \mu_{m,c_{m,l}^{\tilde{C}},l}$ ;
   end for;
   let  $\bar{C} \leftarrow \arg \min_{\tilde{C}} \{\kappa_m^{\tilde{C}}\}$ ; let  $\pi_{m,c,l} \leftarrow \begin{cases} 1, & \text{if } c = c_{m,l}^{\bar{C}}, \\ 0, & \text{otherwise,} \end{cases} \forall l \in L_m$ ;
   end for;
   let  $\tilde{\sigma}_{c,l} \leftarrow \begin{cases} 1, & \text{if there is a mobile host } m \text{ with } \pi_{m,c,l} = 1; \\ 0, & \text{otherwise.} \end{cases}$ 
4. If  $\eta \leq N$  and  $\left[ \sum_{c \in C} \sum_{l \in L} b_{c,l} \times \tilde{\sigma}_{c,l} \right] - \left[ \sum_{c \in C} \sum_{l \in L} b_{c,l} \times \sigma_{c,l} + \sum_{m \in M} \sum_{c \in C_m} \sum_{l \in L_m} \mu_{m,c,l} (\pi_{m,c,l} - \sigma_{c,l}) \right] > \theta$ , then let  $\eta \leftarrow \eta + 1$ ,
   let  $\gamma \leftarrow \sum_{m \in M} \sum_{c \in C_m} \sum_{l \in L_m} (\pi_{m,c,l} - \sigma_{c,l})^2$ , let  $\mu_{m,c,l} \leftarrow \max \left\{ 0, \mu_{m,c,l} + \frac{\varepsilon \times (\pi_{m,c,l} - \sigma_{c,l})}{\gamma} \right\}$ ,  $\forall m \in M, \forall c \in C_m, \forall l \in L_m$ ;
   if any Lagrange multiplier is adjusted, then return to step 2;
   end if.

```

Fig. 2. Algorithm LAGRANGE.

iteration. Algorithm LAGRANGE stops 1) when the number of iterations η is larger than a threshold N , 2) when our algorithm can no longer adjust the cost, or 3) when the difference of the total bandwidth cost of our solution and the lower bound on the total bandwidth cost of the optimal solution is within a threshold θ . We enforce that the algorithm can perform a limited number of iterations N to find the trade-off between the quality of the solution and the computational time. The parameter ε in step five dominates the modification of the cost at each iteration in the subgradient algorithm. The solution with a larger ε improves faster, but the solution with a smaller ε converges to a better solution. In this paper, we thereby reduce ε as the improvement of the selected cells becomes smaller.

Consider Fig. 3 as an illustrative example. The bandwidth cost of each UMTS and each Wi-Fi cell is 3.5 and 2.5 to deliver each layer of a video stream. Each mobile host can use both wireless technologies. For simplicity, the video stream here has only two layers, and the bandwidth allocated for each UMTS and Wi-Fi cell can support one and two layers, respectively. The solid line connecting a mobile host and a cell represents that the mobile host selects the cell for Layer 1, the base layer, of a video stream. The long dash line indicates that the mobile host selects the cell for Layer 2, the enhancement layer, of the video stream. Each mobile host selects the cell for each subscribed layer according to the solution to the second subproblem of the LRP. In the second subproblem of the LRP, each cell is associated with a cost for each subscribed layer of each mobile host, and the cost is shown beside each line in Fig. 3. In addition, the first subproblem of the LRP provides an insight for bandwidth-efficient cells, where each selected cell selected is presented by a solid circle.

Initially, we set the cost of each cell as one. Therefore, only the UMTS cell is selected in the first subproblem for Layer 1 because the net cost of the cell is negative in Fig. 3a.

In other words, the bandwidth cost of the cell is less than the profit, which is the sum of costs for the four mobile hosts. In the second subproblem, mobile hosts A and B subscribe the two layers from cells Wi-Fi 1 and Wi-Fi 2, and mobile hosts C and D subscribe the base layer from cell UMTS. Therefore, the two Wi-Fi cells need to deliver both layers, and the UMTS cell needs to multicast the base layer of the video stream. In other words, variables $\tilde{\sigma}_{\text{Wi-Fi1,1}}$, $\tilde{\sigma}_{\text{Wi-Fi2,1}}$, $\tilde{\sigma}_{\text{Wi-Fi1,2}}$, $\tilde{\sigma}_{\text{Wi-Fi2,2}}$, and $\tilde{\sigma}_{\text{UMTS,1}}$ are one in Fig. 3a, and the total bandwidth cost is 13.5.

The first subproblem encourages the mobile hosts to select the UMTS cells for the layer since the cell is selected in the first subproblem for Layer 1 of the video stream. However, mobile hosts A and B do not select the UMTS cell in the second subproblem. Therefore, algorithm LAGRANGE reduces $\mu_{\text{A,UMTS,1}}$ and $\mu_{\text{B,UMTS,1}}$ to 0.8 accordingly. On the contrary, the two Wi-Fi cells are not selected in the first subproblem, but mobile host A and B selects the two Wi-Fi cells in the second subproblem. Therefore, algorithm LAGRANGE increases $\mu_{\text{A,Wi-Fi1,1}}$, $\mu_{\text{A,Wi-Fi1,2}}$, $\mu_{\text{B,Wi-Fi2,1}}$, and $\mu_{\text{B,Wi-Fi2,2}}$ to 1.2 accordingly, and Fig. 3b shows the new costs at the beginning of the second iteration. At the second iteration, both mobile hosts A and B select the UMTS cell for Layer 1 of the video stream in the second subproblem because the cost for the UMTS cell is smaller than the cost for the Wi-Fi cell. Similarly, both mobile hosts A and B select cell Wi-Fi 1 for Layer 2 of the video streams. Therefore, variables $\tilde{\sigma}_{\text{UMTS,1}}$ and $\tilde{\sigma}_{\text{Wi-Fi1,2}}$ are one in Fig. 3b, and the total bandwidth cost is 6, identical to the optimal solution in the example. Algorithm LAGRANGE shows two advantages here: First, it selects the bandwidth-efficient wireless technology according to the bandwidth cost of each cell and the number of covered mobile hosts. It uses the UMTS cell to serve all mobile hosts for Layer 1 of the video stream. Second, algorithm LAGRANGE clusters the mobile hosts to reduce the

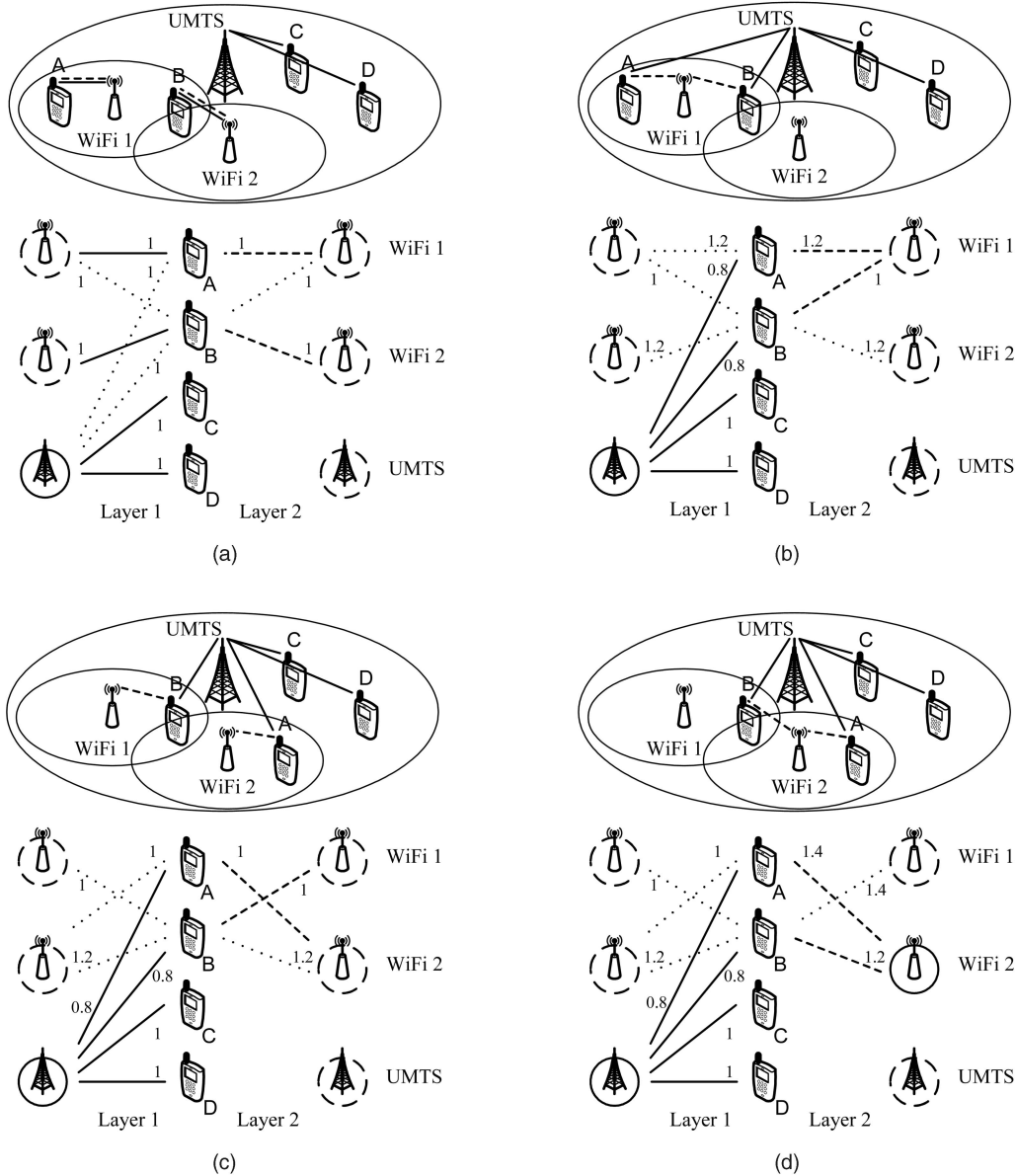


Fig. 3. An example of algorithm LAGRANGE. (a) Iteration 1 (cost = 13.5). (b) Iteration 2 before A moves (cost = 6). (c) Iteration 2 after A moves (cost = 8.5). (d) Iteration 4 (cost = 6).

number of cells for each layer of the video stream. It selects cell Wi-Fi 1 to multicast Layer 2 of the video stream to mobile hosts A and B, instead of using the two Wi-Fi cells.

Figs. 3c and 3d show that algorithm LAGRANGE adapts to the mobility of users. We assume that mobile host A moves to cell Wi-Fi 2 during the second iteration, and the total bandwidth cost increases to 8.5. The cost of cell Wi-Fi 2 for Layer 2 of mobile host A is set as one in Fig. 3c. For Layer 1 of the video stream, each cell selected in the first subproblem is also selected in the second subproblem, and each cell that is not selected in the first subproblem is not selected in the second subproblem either. Therefore, the cost of each cell for Layer 1 of each mobile host remains the same. For Layer 2 of the video stream, since cells Wi-Fi 1 and Wi-Fi 2 are not selected in the first subproblem, algorithm LAGRANGE increases the cost of Wi-Fi 2 for mobile host A and the cost of Wi-Fi 1 for mobile host B.

Therefore, the costs $\mu_{A,Wi-Fi2,2}$ and $\mu_{B,Wi-Fi1,1}$ become 1.2 at the beginning of the third iteration. The third iteration only increases the costs $\mu_{A,Wi-Fi2,2}$ and $\mu_{B,Wi-Fi1,1}$ to 1.4, and Fig. 3d shows the new costs at the beginning of the fourth iteration. At this iteration, mobile host B selects cell Wi-Fi 2 for Layer 2 of the video stream since the cost $\mu_{B,Wi-Fi2,2}$ is smaller than $\mu_{B,Wi-Fi1,1}$. Therefore, we need only Wi-Fi 2 to deliver Layer 2 of the video stream. The total bandwidth cost is 6, which is identical to the optimal solution in the example. In addition, algorithm LAGRANGE stops at this iteration because we can no longer increase or decrease the cost of every cell.

The above example shows that algorithm LAGRANGE can reduce the bandwidth consumption to multicast a video stream in the heterogeneous wireless networks. This example also indicates that algorithm LAGRANGE adapts to change of the locations of mobile hosts. Each mobile host

in the algorithm is able to individually select the cell for each subscribed layer because each mobile host is able to solve the corresponding two subproblems individually. Therefore, algorithm LAGRANGE is suitable to be implemented in a distributed manner, and we design a protocol for the algorithm in Section 3.3.

3.3 Protocol Design

The proposed algorithm LAGRANGE can be implemented in either a centralized or a distributed manner. In the centralized manner, the operator uses a server to find the solution with our proposed algorithm. The centralized approach requires no additional communication between the members or between the members and the base stations to find the solution. However, it suffers from the scalability problem because the number of servers is proportional to the number of video streams. Therefore, in this paper, we propose a protocol to implement the algorithm in a distributed manner. The protocol can be implemented in the video subscription software in mobile hosts, and the network operator requires no server to provide the services. In this paper, we assume that the subscription of the layers is controlled in mobile hosts to ensure the scalability to support more members and streams, just like the previous works. Therefore, our protocol informs each mobile host of the identity of the cell and technology for each subscribed layer, and we require no modification on the existing video multicast mechanisms.

In the following, we present our protocol based on algorithm LAGRANGE. We assume the wireless networks support multicast between mobile hosts. Each mobile host in our protocol periodically exchanges the costs with other mobile hosts in the same cell to solve the two subproblems of the LRP. Each mobile host selects the cell and technology to subscribe each layer of the video stream from the solutions of the two subproblems. Each cell needs to deliver a layer of the video stream if at least one mobile host in the cell subscribes the layer from the cell.

To reduce the protocol overhead, we divide the exchange of the costs into two modes: Unsuppress and Suppress. Our protocol inserts multiple Suppress exchanges between two Unsuppress exchanges of the costs. In the Unsuppress mode, each mobile host sends the costs via multicast to other mobile hosts in each covering cell. In the Suppress mode, however, each mobile host suppresses sending a cost to other mobile hosts if the difference of the current cost and the cost that was sent to other mobile hosts is within a threshold. Each mobile host m in our protocol stores the following information for each covering cell c : 1) The cost that was sent to other mobile hosts for each layer l in L_m . 2) The profit of the cell for each layer l in the first subproblem of the LRP. The profit is the sum of the total costs of all mobile hosts in the cell, namely, $\sum_{\bar{m}: c \in C_{\bar{m}}, l \in L_{\bar{m}}} \mu_{\bar{m}, c, l}$. A mobile host obtains the profit by exchanging the Lagrange multipliers with other mobile hosts in the cell.

Each mobile host needs to access each covering cell only during the exchange of multipliers. In other times, each mobile host accesses only the selected cell to receive the video stream. In our protocol, the exchange of multipliers is

periodically initiated by sending a trigger message via multicast to all mobile hosts. The trigger message indicates the mode of the exchange of the costs and can be sent by the video server. In the Unsuppress exchange of multipliers, each mobile host obtains the profit of a covering cell after it receives the costs from other mobile hosts in the cell. In the Suppress exchange of multipliers, each mobile host finds the profit of a covering cell with the profit stored in the mobile host and the differences of the costs sent by other mobile hosts in the cell. Afterward, each mobile host solves the two subproblems of the LRP and selects the cell to subscribe each layer of the video stream.

Our protocol supports the mobility of users. When a mobile host hands over to a new cell, it sends the cost that is initially set as one to other mobile hosts in the new cell during the exchange of the costs. If the exchange of the costs is in the Suppress mode, the mobile host also sends a negative cost to the old cell to reduce the total profit. Our protocol is resilient to the message loss. If any message with the cost in the Suppress mode is lost, other mobile hosts can still obtain the correct profit of the cell after an Unsuppress exchange of the costs. If a mobile host keeps moving back and forth between adjacent cells, our algorithm can handle this situation by reducing the duration between two iterations. The amount of time required by our algorithm to find the best solution depends on the length of the time between two iterations.

Our protocol supports the dynamic group membership. For a new member, it distributes and collects the Lagrange multipliers in the way as the member hands over into the cell. For a member that unsubscribes a video stream, the behavior of the member in a cell is just like the way that the member hands over out of the cell.

Current video multicast mechanisms mainly focus on adaptively adding and dropping a layer according to the available bandwidth to avoid quality fluctuation. The current mechanisms adopt buffering techniques to synchronize the playback of each layer at mobile hosts. Each mobile host in the mechanisms uses a state machine to ensure that the addition or drop of a layer is in the progressive manner. The state machine also enforces that each mobile host must first subscribe the base layer when it would like to receive a video stream. In addition, the state machine includes timers and hysteresis algorithms to avoid the ping-pong effect of the quality fluctuation between layers. In contrast, our approach is orthogonal to them and needs to adopt one of them to subscribe a video stream.

4 EXPERIMENTAL STUDIES

In this section, we present our simulation results. As we know, there is no related algorithm for the CTSP in the previous works. Therefore, we compare algorithm LAGRANGE with two other algorithms that can represent the reasonable user behaviors. In the first algorithm RAND, each mobile host randomly selects a cell to subscribe all layers of a video stream. In the second algorithm LOCAL, each mobile host individually selects the wireless technology with the minimum bandwidth cost to subscribe all layers of the video stream because the mobile host tends to

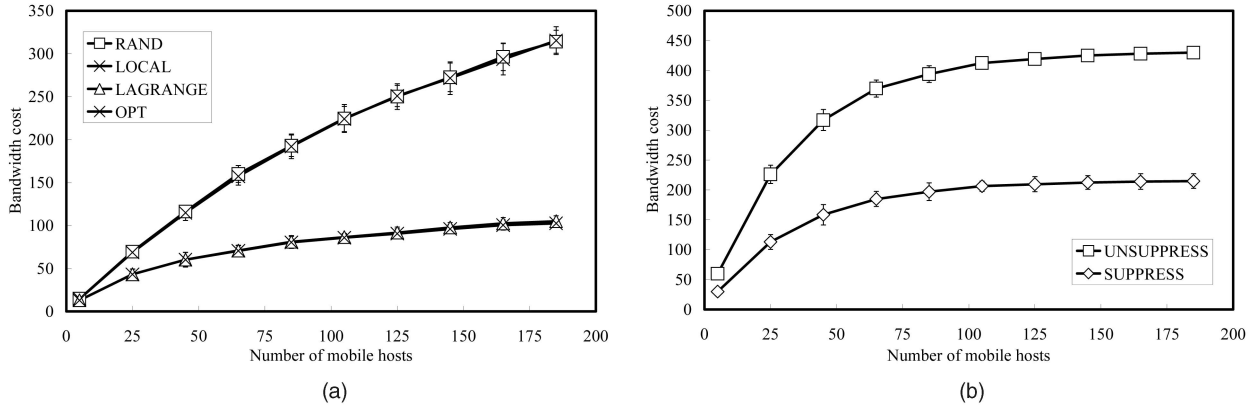


Fig. 4. Simulation results of homogeneous wireless networks. (a) The bandwidth cost to deliver the video stream. (b) The bandwidth cost of CONTROL.

spend the least monetary cost in this case. The mobile host selects the cell with the minimum distance to the base station if there are multiple cells with the technology covering the mobile host. In addition to RAND and LOCAL, we compare our solution with the optimal solution OPT obtained by CPLEX [27] with our ILP formulation in small networks. We also compare the total bandwidth cost of our solution with the lower bound LB on the total bandwidth cost of the optimal solution in large networks, where we obtain the lower bound by solving the LRP in Section 3.

To test the performance of our algorithm in different scenarios, we change the following parameters:

1. number of mobile hosts that subscribe the video stream to consider the scalability of our algorithm and protocol,
2. number of wireless technologies that can be used by each mobile host to consider the heterogeneity of users,
3. the bandwidth cost of each wireless technology, and
4. transmission range of a base station.

For each wireless technology, the size of the overlapping area of adjacent cells changes when the transmission range of a base station is different. We measure 100 samples in each scenario with 95 percent confidence interval.

The performance metrics in our simulation are listed as follows: 1) Bandwidth cost induced by each algorithm. We measure the total bandwidth cost and the bandwidth cost of each wireless technology to deliver a video stream. 2) Control overhead of our protocol, denoted as CONTROL. We measure the bandwidth cost of the cells to exchange the costs in order to compare CONTROL with the bandwidth costs to send the video stream. The number of control messages in our protocol is proportional to the number of cells to exchange the costs and the number of mobile hosts in a cell. In our simulation, we compare the overhead of our protocol in only Unsuppress mode, denoted as UNSUPPRESS, with the overhead of our protocol in both Unsuppress and Suppress modes, denoted as SUPPRESS. In the SUPPRESS exchange of the costs, there are five Suppress exchanges of the costs between two Unsuppress exchanges. In our simulation, the solution

obtained by algorithm LAGRANGE adopts the SUPPRESS exchange of the costs.

In our simulation, we adopt the probability-based mobility model widely used in the previous works. Initially, we distribute all mobile hosts uniformly at random in the networks in a $10 \text{ km} \times 10 \text{ km}$ service area. Afterward, each mobile host randomly decides to stay or move at each minute. If the mobile host decides to move, it randomly chooses the direction and the speed from 0 to 1 km per minute. Our simulation adopts the cellular network infrastructure, which has been widely used in the experiments of related works. In addition, we decide the locations of Wi-Fi cells randomly because we assume that Wi-Fi cells need to cover only hotspot areas and, as we know, there is no related work on the distribution of hotspot areas. The mobility model in this paper is very similar to the Random Walk model that has been widely adopted in the literature. However, we assume that mobile hosts sometimes can stop for a while in this paper to model the behaviors of pedestrians or vehicles in a city. For algorithm LAGRANGE, we run one iteration every 30 seconds. The simulation time takes 1,000 minutes.

We first compare the solutions obtained by our algorithms with the optimal solutions obtained by CPLEX with our ILP formulation. Because solving large ILP problems is computational infeasible, here, we simulate the scenario that the video stream is not layer encoded, and the network operators own only one wireless technology. The network has 144 Wi-Fi cells, where the bandwidth cost of each cell is three. Fig. 4a presents the bandwidth costs of algorithms LAGRANGE, RAND, LOCAL, and OPT. Fig. 4a shows that LAGRANGE outperforms RAND and LOCAL because LAGRANGE clusters the mobile hosts to reduce the number of cells to multicast the video stream. Mobile hosts in LAGRANGE jointly reduce the total bandwidth consumption by exchanging the costs. On the contrary, each mobile host in RAND and LOCAL selects the cells independently. Fig. 4b shows that our protocol in Suppress mode can reduce the number of cells to exchange the costs. Therefore, the above results indicate that our algorithm and protocol can reduce the total bandwidth consumption even when the video stream is not layer encoded, and the network operators own only one wireless technology.

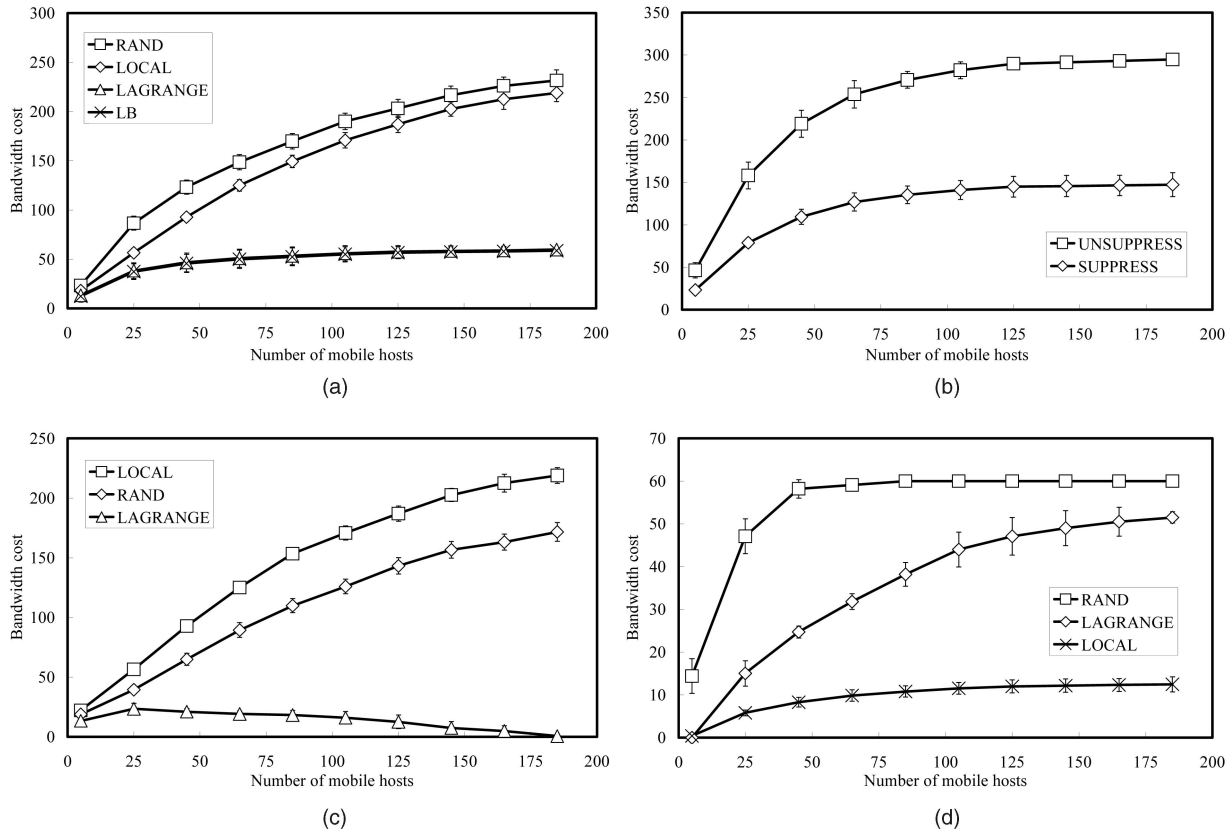


Fig. 5. Simulation results of heterogeneous wireless networks. (a) The total bandwidth cost of both Wi-Fi cells and UMTS cells. (b) The bandwidth cost of CONTROL. (c) The bandwidth cost of Wi-Fi cells. (d) The bandwidth cost of UMTS cells.

Our protocol is based on the proposed algorithm LAGRANGE, which needs to exchange the Lagrange multipliers in not only the cells that deliver the stream, but also the cells that cover the mobile hosts but do not multicast the stream. The mobile hosts need to exchange the multipliers in the latter cells to reduce the bandwidth consumption in a distributed manner. Therefore, more cells are counted in the results of CONTROL. However, each mobile host only accesses the latter cells during the exchange of multipliers.

In the following, we test our algorithm in the heterogeneous wireless networks. The network has 144 Wi-Fi cells and four UMTS cells, where the bandwidth cost of each Wi-Fi cell and UMTS cell is 3 and 15, respectively, for each layer of the video stream. Fig. 5a presents the total bandwidth cost to multicast the video stream. Algorithm LOCAL is slightly better than RAND because the mobile hosts in the LOCAL prefer to use the Wi-Fi cells. However, mobile hosts in RAND waste the network bandwidth because multiple mobile hosts in the same cell may choose the different wireless technologies even when only one technology is required here. In contrast, algorithm LAGRANGE outperforms both RAND and LOCAL, and the solutions obtained by LAGRANGE and LB are very close.

Figs. 5c and 5d compare the total bandwidth consumption in each wireless network. Algorithms RAND and LOCAL select more Wi-Fi cells as the number of mobile hosts increases. In contrast, algorithm LAGRANGE selects fewer Wi-Fi cells but more UMTS cells because a single

UMTS cell induces less bandwidth cost than multiple Wi-Fi cells as the number of mobile hosts increases. Algorithm RAND uses more UMTS cells than LAGRANGE because each mobile host selects the cell individually, regardless of the selections of other mobile hosts. Although algorithm LAGRANGE uses more UMTS cells than LOCAL, the total bandwidth cost of LAGRANGE is much less than the cost of LOCAL, as shown in Fig. 5a. The reason is that LAGRANGE has the abilities to cluster the mobile hosts and to select the proper wireless technology according to the number and locations of mobile hosts.

We then test our algorithm when the number of wireless technologies that can be used by each mobile host is different. Some mobile hosts would like to improve the video quality by subscribing more layers with both wireless technologies, whereas others may concern power consumption and thereby subscribe a limited number of layers with a single wireless technology. In the following, the video has four layers here, and the bandwidth allocated for each UMTS and Wi-Fi cell can support two and four layers. We change the proportion of mobile hosts that can use both wireless technologies in the networks. Fig. 6a shows that LAGRANGE uses more UMTS cells as more mobile hosts can use both technologies. However, Fig. 6b shows that the total bandwidth cost decreases as more mobile hosts can use both technologies. The reason is that more mobile hosts subscribe Layers 1 and 2 of the video stream from the UMTS cells, and fewer Wi-Fi cells are required to multicast the two layers. Therefore, these cells save the bandwidth of

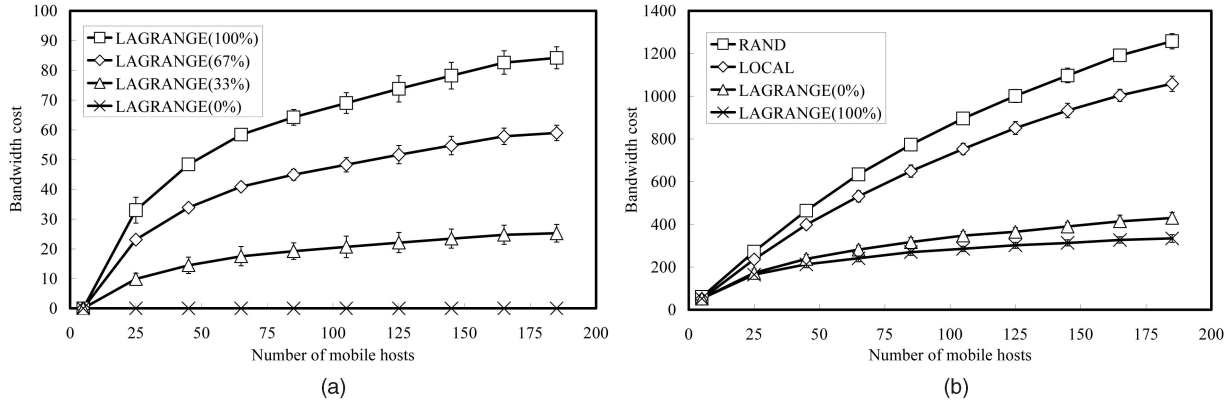


Fig. 6. Simulation results of different percentages of mobile hosts that can use the both wireless technologies. (a) The bandwidth cost of UMTS cells. (b) The total bandwidth cost of both Wi-Fi and UMTS cells.

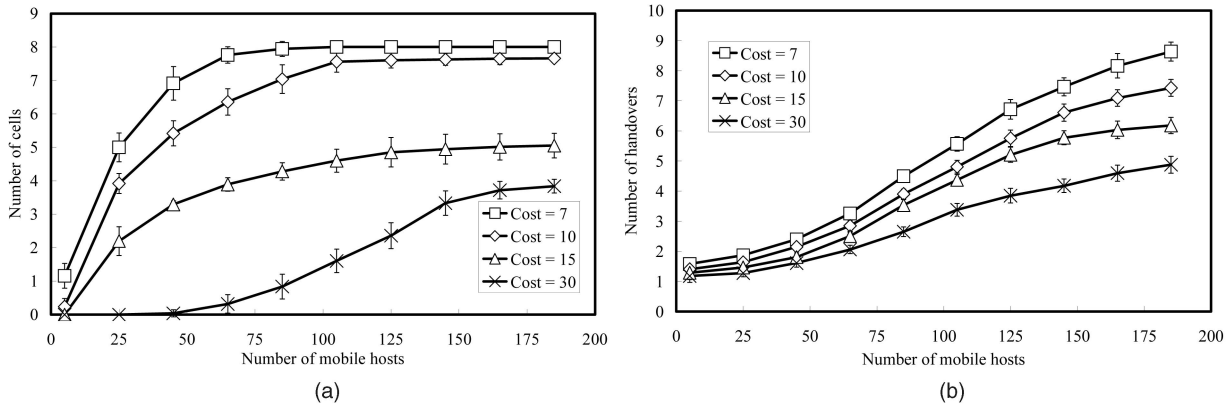


Fig. 7. Simulation results of different bandwidth costs assigned to each UMTS cells. (a) The number of used UMTS cells. (b) The number of handovers for each mobile host.

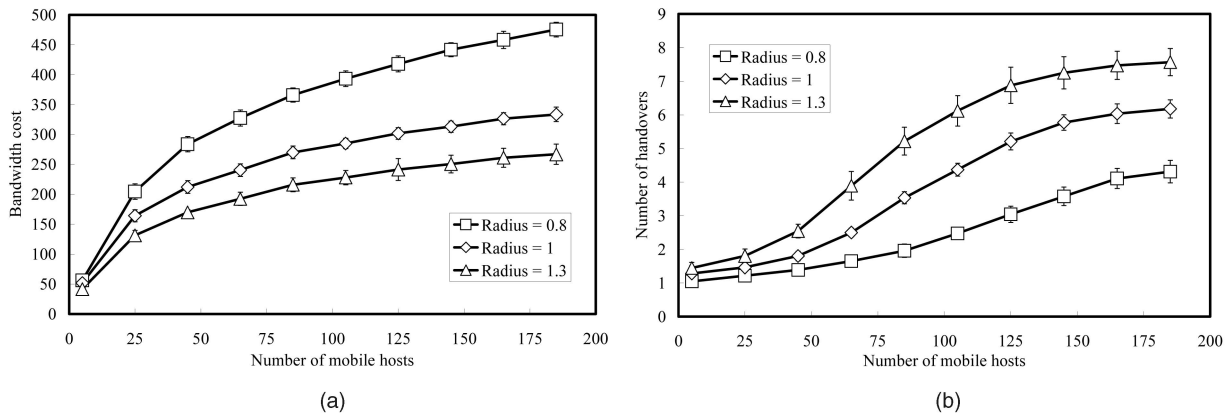


Fig. 8. Simulation results of different radii of a cell with the ratio respect to the original radius. (a) The total bandwidth cost of both Wi-Fi and UMTS cells. (b) The number of handovers for each mobile host.

the two layers and have the ability to multicast more layers to the mobile hosts to improve the quality of the video stream.

The network operators can change the bandwidth cost of each cell to adjust the distribution of traffic. Fig. 7a shows the total number of cells to multicast the video stream with different costs to the UMTS cells. With a larger cost assigned to each UMTS cell, algorithm LAGRANGE uses fewer UMTS cells in Fig. 7a. Therefore, the network operators can increase the bandwidth cost of a cell when

the cell is congested. In addition, Fig. 7b shows that a larger cost for UMTS cells also leads to fewer handovers for each mobile host.

Fig. 8a shows the total bandwidth cost when we change the radius of a cell with the ratio with respect to the original radius. As the radius increases, each mobile host is covered by more cells, and more mobile hosts can be clustered together to use fewer cells. Therefore, algorithm LAGRANGE requires less network bandwidth to multicast the video stream. However, our algorithm in this case

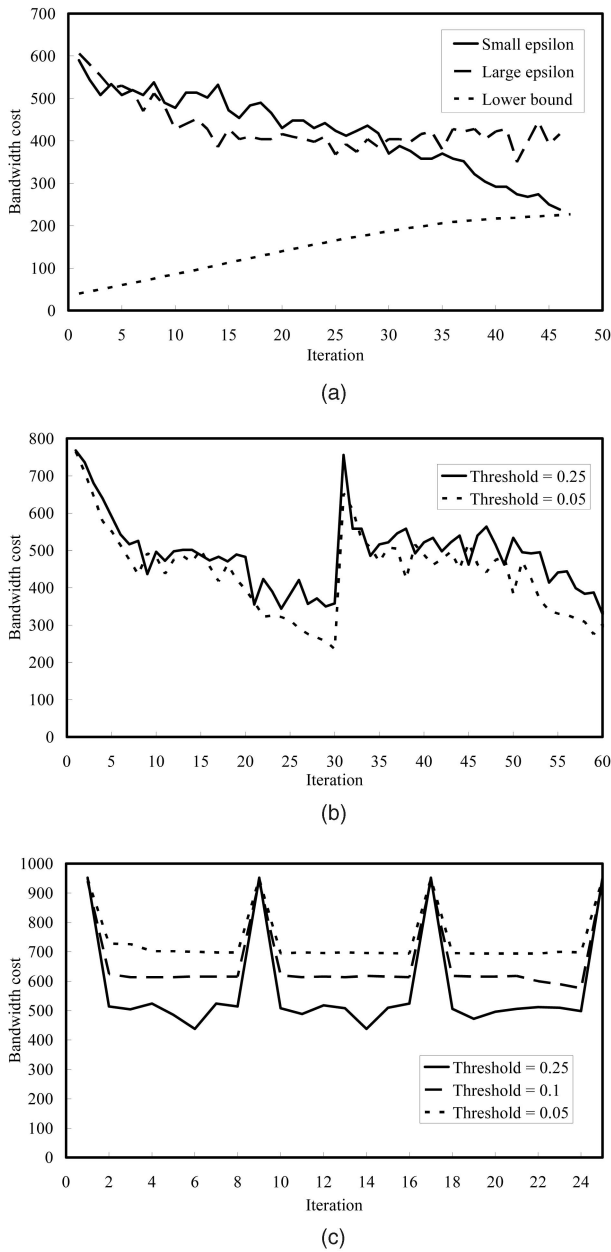


Fig. 9. Transient behavior of algorithm LAGRANGE. (a) The bandwidth cost of LAGRANGE. (b) The bandwidth cost of LAGRANGE. (c) The bandwidth cost of CONTROL.

requires more iterations to find the solution because each mobile host has more candidate cells for each subscribed layer. Therefore, each mobile host needs more handovers in this case, as shown in Fig. 8b.

Fig. 9a shows the transient behavior of our algorithm. With a large ε , that is, $\varepsilon = 10$ in this case, the bandwidth cost converges faster to 400 with only 14 iterations, whereas a small ε , that is, $\varepsilon = 1$ in this case, requires 29 iterations. However, our algorithm with a small ε can obtain a better solution that approaches the lower bound. Our algorithm sometimes obtains a solution slightly worse than the solutions of the nearby iterations. The logic behind searching toward a locally worse direction is to avoid being trapped in a locally optimal solution in order to find a better solution later. Our algorithm, a large ε in Fig. 9a, obtains a

solution at the 43rd iteration, which is better than the one at the 14th iteration.

Fig. 9b shows the transient behavior of our protocol based on the proposed algorithm. There are nine Suppress exchanges of the costs between two Unsuppress exchanges of the costs in our simulation. We change the threshold in the Suppress mode of our protocol. Each mobile host suppresses sending the cost if the difference of the current cost and the previous sent cost is within the threshold. Here, we let all mobile hosts move at iteration 30 to show the effect of mobility in our algorithm. Algorithm LAGRANGE iteratively reduces the total bandwidth consumption, as shown in Fig. 9b. After the mobile hosts move, algorithm LAGRANGE consumes more bandwidth. However, algorithm LAGRANGE reduces the total bandwidth consumption iteratively after the mobile hosts exchange the costs. Fig. 9c shows the control overhead of our protocol. Our protocol induces more overhead as the threshold increases in Fig. 9c. However, our protocol in this case also converges faster toward the optimal solution, as shown in Fig. 9b.

5 CONCLUSION

In this paper, we propose a new mechanism to select the cells and the wireless technologies for layer-encoded video multicasting in heterogeneous wireless networks. Each mobile host in our mechanism can select a different cell to subscribe each layer of the video stream, and each cell can multicast only a subset of layers of the video stream to reduce the bandwidth consumption. We formulate CTSP in the heterogeneous wireless networks as an optimization problem. We use ILP to model the problem. The network operators can use the ILP formulation to find the optimal solutions for network planning. We show that the problem is NP-hard and design an algorithm LAGRANGE, which is based on Lagrangean relaxation on our ILP formulation. Algorithm LAGRANGE iteratively converges toward the optimal solutions and can be implemented in the distributed manner in only the mobile hosts. Our mechanism requires no change of the current video multicasting mechanisms and the current wireless network infrastructures. Our algorithm is adaptive to the change in the subscribers at each layer and the change of the location of each mobile host. Our algorithm is flexible such that network operators can balance the load of the wireless cells with the same or different wireless technologies.

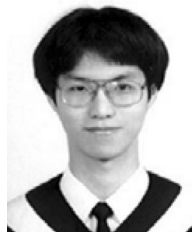
ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for their helpful comments to improve this paper.

REFERENCES

- [1] W. Kumwilaisak et al., "A Cross-Layer Quality-of-Service Mapping Architecture for Video Delivery in Wireless Networks," *IEEE J. Selected Areas in Comm.*, vol. 21, no. 10, pp. 1685-1698, Dec. 2003.
- [2] M. Chen and A. Zakhor, "Rate Control for Streaming Video over Wireless," *IEEE Wireless Comm.*, vol. 12, no. 4, pp. 32-41, Aug. 2005.

- [3] Y. Liang and B. Girod, "Network-Adaptive Low-Latency Video Communication over Best-Effort Networks," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 1, pp. 72-81, Jan. 2006.
- [4] W. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 301-317, Mar. 2001.
- [5] J. Ohm, "Advances in Scalable Video Coding," *Proc. IEEE*, vol. 93, no. 1, pp. 42-56, Jan. 2005.
- [6] J.-L. Huang, W.-C. Peng, and M.-S. Chen, "SOM: Dynamic Push-Pull Channel Allocation Framework for Mobile Data Broadcasting," *IEEE Trans. Mobile Computing*, vol. 5, no. 8, pp. 974-990, Aug. 2006.
- [7] S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-Driven Layered Multicast," *Proc. ACM SIGCOMM*, vol. 26, no. 4, pp. 117-130, 1996.
- [8] B. Vickers, C. Albuquerque, and T. Suda, "Source-Adaptive Multilayered Multicast Algorithms for Real-Time Video Distribution," *IEEE/ACM Trans. Networking*, vol. 8, no. 6, pp. 720-733, Dec. 2000.
- [9] T. Kim and M. Ammar, "Optimal Quality Adaptation for MPEG-4 Fine-Grained Scalable Video," *Proc. IEEE INFOCOM*, vol. 1, pp. 641-651,
- [10] J. Liu, B. Li, Y. Hou, and I. Chlamtac, "On Optimal Layering and Bandwidth Allocation for Multisession Video Broadcasting," *IEEE Trans. Wireless Comm.*, vol. 3, no. 2, pp. 656-667, Mar. 2004.
- [11] S. Zhao, Z. Xiong, and X. Wang, "Optimal Resource Allocation for Wireless Video over CDMA Networks," *IEEE Trans. Mobile Computing*, vol. 4, no. 1, pp. 56-67, Jan./Feb. 2005.
- [12] G.L. Nemhauser and L.A. Wosley, "Integer and Combinatorial Optimization," *Wiley-Interscience Series in Discrete Mathematics and Optimization*, 1999.
- [13] X. Yong, D. Harrison, S. Kalyanaraman, K. Ramachandran, and A. Venkatesan, "Accumulation-Based Congestion Control," *IEEE/ACM Trans. Networking*, vol. 13, no. 1, pp. 69-80, Feb. 2005.
- [14] I. Paschalidis and Y. Liu, "Pricing in Multiservice Loss Networks: Static Pricing, Asymptotic Optimality and Demand Substitution Effects," *IEEE/ACM Trans. Networking*, vol. 10, no. 3, pp. 425-438, June 2002.
- [15] C. Beard and V. Frost, "Prioritized Resource Allocation for Stressed Networks," *IEEE/ACM Trans. Networking*, vol. 9, no. 5, pp. 618-633, Oct. 2001.
- [16] M. Chiang, "Balancing Transport and Physical Layers in Wireless Multihop Networks: Jointly Optimal Congestion Control and Power Control," *IEEE J. Selected Areas in Comm.*, vol. 23, no. 1, pp. 104-116, Jan. 2005.
- [17] Y. Zhang, O. Yang, and H. Liu, "A Lagrangean Relaxation and Subgradient Framework for the Routing and Wavelength Assignment Problem in WDM networks," *IEEE J. Selected Areas in Comm.*, vol. 22, no. 9, pp. 1752-1765, Nov. 2004.
- [18] E.W.M. Wong, A.K.M. Chan, and T.-S.P. Yum, "Analysis of Rerouting in Circuit-Switched Networks," *IEEE/ACM Trans. Networking*, vol. 8, no. 3, pp. 419-427, June 2000.
- [19] K.-C. Lee and V.O.K. Li, "A Wavelength Rerouting Algorithm in Wide-Area All-Optical Networks," *IEEE J. Lightwave Technology*, vol. 14, no. 6, pp. 1218-1229, June 1996.
- [20] A. Donner, M. Berioli, and M. Werner, "MPLS-Based Satellite Constellation Networks," *IEEE J. Selected Areas in Comm.*, vol. 22, no. 3, pp. 438-448, Apr. 2004.
- [21] K. Pahlavan et al., "Handoff in Hybrid Mobile Data Networks," *IEEE Personal Comm.*, vol. 7, no. 2, pp. 34-47, Apr. 2000.
- [22] J. McNair, I.F. Akyildiz, and M. Bender, "An Inter-System Handoff Technique for the IMT-2000 System," *Proc. IEEE INFOCOM*, vol. 1, pp. 208-216, 2000.
- [23] J. McNair and F. Zhu, "Vertical Handoffs in Fourth-Generation Multinetwork Environments," *IEEE Wireless Comm.*, vol. 11, no. 3, pp. 8-15, June 2004.
- [24] Y. Pan, M. Lee, J.B. Kim, and T. Suda, "An End-to-End Multipath Smooth Handoff Scheme for Stream Media," *IEEE J. Selected Areas in Comm.*, vol. 22, no. 4, pp. 653-663, May 2004.
- [25] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World," *IEEE J. Selected Areas in Comm.*, vol. 20, no. 4, pp. 756-767, May 2002.
- [26] M.R. Gary and D.S. Johnson, *Computer and Intractability: A Guide to the Theory of NP-Hardness*. W.H. Freeman, 1979.
- [27] CPLEX Mathematical Programming Optimizer, <http://www.ilog.com/products/cplex/>, 2007.



De-Nian Yang received the BS and PhD degrees from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, in 1999 and 2004, respectively. He is currently a postdoctoral researcher in military service in the Department of Electrical Engineering, National Taiwan University. His research interests include network planning, multicasting, and Quality of Service in wireless networks. He is a member of the IEEE.



Ming-Syan Chen received the BS degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, and the MS and PhD degrees in computer, information, and control engineering from the University of Michigan, Ann Arbor, Michigan, in 1985 and 1988, respectively. He is currently a professor and the chairman of the Graduate Institute of Communication Engineering, a professor in the Electrical Engineering Department, and also a professor in the Computer Science and Information Engineering Department, National Taiwan University, Taipei, Taiwan. He was a research staff member at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, from 1988 to 1996. His research interests include database systems, data mining, mobile computing systems, and multimedia networking, and he has published more than 180 papers in his research areas. In addition to serving as a program committee member in many conferences, he served as an associate editor of the *IEEE Transactions on Knowledge and Data Engineering* from 1997 to 2001, is currently on the editorial board of the *Very Large Databases Journal*, *Knowledge and Information Systems Journal*, *Journal of Information Science and Engineering*, and the *International Journal of Electrical Engineering*, and is also a guest coeditor for the *IEEE Transactions on Knowledge and Data Engineering* on a special issue for data mining in December 1996. He was a distinguished visitor of the IEEE Computer Society for Asia-Pacific from 1998 to 2000. He served as the program chair of the Pacific Area Knowledge Discovery and Data Mining (PAKDD '02), program vice chair of the IEEE International Conference on Distributed Computing Systems (ICDCS) 2005 and the International Conference on Parallel Processing (ICPP) 2003, program vice chair of Very Large Databases Conference (VLDB '02), and also the general chair and program chair of several other conferences. He was a keynote speaker on Web data mining at the International Computer Congress in Hong Kong in 1999, a tutorial speaker on Web data mining at the International Conference on Database Systems for Advanced Applications (DASFAA '99) and on parallel databases at the 11th IEEE International Conference on Data Engineering (ICDE) in 1995. He holds or has applied for 18 US patents and seven ROC patents in the areas of data mining, Web applications, interactive video playback, video server design, and concurrency and coherency control protocols. He is a recipient of the National Science Council (NSC) Distinguished Research Award and the K.-T. Li Research Penetration Award for his research work and the Outstanding Innovation Award from IBM Corporate for his contribution to a major database product. He also received numerous awards for his research, teaching, inventions, and patent applications. He is a fellow of the IEEE and the ACM.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.