AUTHOR QUERY FORM

Dear Author,

During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.

Many thanks for your assistance.

<table>
<thead>
<tr>
<th>Query References</th>
<th>Query</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AUTHOR: The reference citations have been renumbered from Reference 1 (originally Reference 24) onward, and the reference list has been renumbered accordingly. Please confirm that this is OK.</td>
<td></td>
</tr>
</tbody>
</table>
10

Cost Optimization for Scalable Communication in Wireless Networks with Movement-Based Location Management

Keqin Li

10.1 INTRODUCTION

To effectively and efficiently deliver network services to mobile users in a wireless communication network, a key component of the network, that is, a dynamic location management scheme, should be carefully designed, and its cost should be thoroughly analyzed. There are two essential tasks in location management, namely, location update (location registration) and terminal paging (call delivery). Location update is the process for a mobile terminal to periodically notify its current location to a network so that the network can revise the mobile terminal’s location profile in a location database. Terminal paging is the process for a network to search a mobile terminal by sending polling signals based on the information of its last reported location so that an incoming phone call can be routed to the mobile terminal.

Both location update and terminal paging consume significant communication bandwidth of a wireless network, battery power of mobile terminals, memory space
in location registers and databases, and computing time at base stations. Therefore, both location update cost and terminal paging cost should be minimized. However, there is a trade-off between the cost of location update and the cost of terminal paging. More location updates reduce the cost of terminal paging, while insufficient location update increases the cost of terminal paging. A dynamic location management scheme has the capability to adjust its location update and terminal paging strategies based on the mobility pattern and incoming call characteristics of a mobile terminal, such that the combined cost of location update and terminal paging is minimized.

Analysis and minimization of the total location management cost for a single mobile terminal has been studied extensively by many researchers. However, there are two problems in existing research. First, cost of location update (measured by the number of location updates and related to resource consumption in base stations) and cost of terminal paging (measured by the number of cells in a paging area (PA) and related to communication bandwidth) are different in nature and difficult to be unified. Second, cost minimization should be carried out at the network level for multiple heterogeneous mobile terminals which simultaneously exist in a network and share network resources, not just for a single mobile terminal, so that network-wide performance optimization and network scalability can be considered. Hence, the traditional way of finding the best movement threshold that minimizes the total location management cost for a single mobile terminal might not be interesting and might not make much sense.

In this chapter, we consider the problem of cost optimization in wireless communication networks with movement-based location management. Our approach is to minimize the total cost of terminal paging of all mobile terminals with a constraint on the total cost of location update of all mobile terminals and to minimize the total cost of location update with a constraint on the total cost of terminal paging. These problems are formulated as multivariable optimization problems and are solved numerically. This approach allows network administration to freely control one of the total cost of location update and the total cost of terminal paging of all mobile terminals in a wireless network and to minimize the other. (This is not the case for a single mobile terminal, for which the determination of one cost also determines the other.) This is an effective way to handle cost trade-off, as in many other computing and communication systems. Such network-wide cost control and optimization is extremely important for scalable communication in a wireless network when the number of mobile terminals increases. Our approach also eliminates the issue of cost unification of location update and terminal paging since the two kinds of cost are separated. To the best of our knowledge, such cost optimization in an entire wireless communication network with heterogeneous mobile terminals has not been considered before.

The rest of the chapter is organized as follows. In Section 10.2, we provide necessary background information on dynamic location management in wireless communication networks. In Section 10.3, we discuss cost measure and optimization for a single mobile terminal. In Section 10.4, we define and solve the terminal paging cost minimization problem subject to constraint on location update cost. In Section 10.5, we define and solve the location update cost minimization problem subject to constraint on terminal paging cost. In Section 10.6, we present a numerical example. In Section 10.7, we summarize the chapter.
10.2 BACKGROUND INFORMATION

Our model of a wireless communication network follows that in Reference 1. A wireless communication network has the common hexagonal cell configuration or mesh cell configuration. In the hexagonal cell structure (see Fig. 10.1), cells are hexagons of identical size and each cell has six neighbors. In the mesh cell structure (see Fig. 10.2), cells are squares of identical size and each cell has eight neighbors. Throughout the chapter, we let $q$ be a constant such that $q = 3$ for the hexagonal cell configuration and $q = 4$ for the mesh cell configuration. By using the constant $q$, the hexagonal cell configuration and the mesh cell configuration can be treated in a unified way. For instance, we can say that each cell has $2q$ neighbors without mentioning the particular cell structure. The network is homogeneous in the sense that the behavior of a mobile terminal is statistically the same in all the cells.

Let $s$ be the cell registered by a mobile terminal in the last location update. The cells in a wireless network can be divided into rings, where $s$ is the center of the network and is called ring 0. The $2q$ neighbors of $s$ constitute ring 1. In general, the neighbors of all the cells in ring $r$, except those neighbors in rings $r - 1$ and $r$, constitute ring $r + 1$. For all $r \geq 0$, the cells in ring $r$ have distance $r$ to $s$. For all $r \geq 1$, the number of cells in ring $r$ is $2qr$. Notice that the rings are defined with respect to $s$. When a mobile terminal updates its location to another cell $s'$, $s'$ becomes the center of the network, and ring $r$ consists of the $2qr$ cells whose distance to $s'$ is $r$.

A mobile terminal $u$ constantly moves from cell to cell. Such movement also results in movement from ring to ring. Let the sequence of cells visited by $u$ before
the next phone call be denoted as $s_0, s_1, s_2, \ldots, s_d, \ldots$, where $s_0 = s$ is $u$’s last registered cell (not the cell in which $u$ received the previous phone call) and considered as $u$’s current location. There are three location update methods proposed in the current literature, namely, the distance-based method, the movement-based method, and the time-based method.

- In the **distance-based location update method**, location update is performed as soon as $u$ moves into a cell $s_j$ in ring $d$, where $d$ is a distance threshold; that is, the distance of $u$ from the last registered cell $s$ is $d$, such that $s_j$ is registered as $u$’s current location. It is clear that $j \geq d$, that is, it takes at least $d$ steps for $u$ to reach ring $d$.

- In the **movement-based location update method**, location update is performed as soon as $u$ has crossed cell boundaries for $d$ times since the last location update, where $d$ is a movement threshold. It is clear that the sequence of registered cells for $u$ is $s_d, s_2d, s_3d, \ldots$.

- In the **time-based location update method**, location update is performed every $\tau$ units of time, where $\tau$ is a time threshold, regardless of the current location of $u$.

In all dynamic location management schemes, a current PA consists of rings 0, 1, 2, \ldots, $d-1$, where $d$ is some value appropriately chosen. We say that such a PA
has radius $d$. Since the number of cells in ring $r$ is $2qr$, for all $r \geq 1$, the total number of cells in a PA is $qd^2 - qd + 1$. It should be noticed that a PA is defined with respect to the current location of a mobile terminal and is changed whenever a mobile terminal updates its location. The radius $d$ of a PA can be adjusted in accordance with various cost and performance considerations. On the other hand, the location and size of a cell are fixed in a wireless network.

We will consider two different call handling models:

- In the call plus location update (CPLU) model, the location of a mobile terminal is updated each time a phone call arrives; that is, in addition to distance-based or movement-based or time-based location updates, the arrival of a phone call also initiates location update and defines a new PA. This causes the original location update cycle of a mobile terminal being interrupted.
- In the call without location update (CWLU) model, the arrival of a phone call has nothing to do with location update; that is, a mobile terminal still keeps its original location update cycles.

Our cost analysis and optimization will be conducted for both models.

For a random variable $T$, we use $E(T)$ to represent the expectation of $T$ and $\lambda_T = E(T)^{-1}$. The probability density function (pdf) of $T$ is $f_T(t)$. There are two important random variables in the study of dynamic location management. The intercall time $T_c$ is defined as the length of the time interval between two consecutive phone calls. The cell residence time $T_s$ is defined as the time a mobile terminal stays in a cell before it moves into a neighboring cell. The quantity $\rho = \lambda_T / \lambda_{T_c}$ is the call-to-mobility ratio.

The cost of dynamic location management contains two components, that is, the cost of location update and the cost of terminal paging. The cost of location update is proportional to the number of location updates. If there are $X_u$ location updates between two consecutive phone calls, the cost of location update is $\Delta_u X_u$, where $\Delta_u$ is a constant. Since $X_u$ is a random variable, the location update cost is actually calculated as $\Delta_u (E(X_u))$. The cost of terminal paging is proportional to the number of cells paged. If a PA has radius $d$, the cost of paging is $\Delta_p (qd^2 - qd + 1)$, where $\Delta_p$ is a constant.

Dynamic location management is per-terminal based. A mobile terminal is specified by $f_T(t)$ and $f_T(t)$, where $f_T(t)$ is the call pattern and $f_T(t)$ is the mobility pattern. Since a location update method determines the location update cost and a terminal paging method determines the terminal paging cost, for given a mobile terminal, we need to find a balanced combination of a location update method and a terminal paging method, such that the total location management cost for the mobile terminal is minimized.

Dynamic location management has been studied extensively by many researchers. The performance of movement-based location management schemes has been investigated in References 2–9. The performance of distance-based location management schemes has been studied in References 3 and 10–14. The performance of time-based location management schemes has been considered in References 3 and 15–17. Terminal paging methods with low cost and time delay have been studied by several researchers [2, 12, 18–22]. Others studied were reported in References 23–32.
number of cells has been treated as an optimization problem that is solved by using bioinspired methods such as simulated annealing, neural networks, and genetic algorithms [33–38]. The reader is also referred to the surveys in References 39–42 (Chapter 15) and 43 (Chapter 11).

10.3 COST MEASURE AND OPTIMIZATION FOR A SINGLE USER

Minimization of total location management cost for a single mobile terminal has been studied extensively. Traditionally, the performance measure for a single mobile user is the cost of location management (including the cost of location update and the cost of terminal paging) between two consecutive phone calls. However, different mobile terminals have different intercall times. Our performance measure in this chapter is the cost of location management per unit of time so that the cost of different users can be added.

In this chapter, we will focus on the movement-based location update method. Our investigation in this chapter is based on our recent work in Reference 1, where we successfully developed closed-form expressions of location update cost for a mobile terminal under both CPLU and CWLU models.

The following theorem gives the cost of location management per unit of time for a single mobile user under the CPLU model.

**Theorem 10.1** If \( f_{T(t)} \) has an Erlang distribution, the total cost of location management per unit of time under the CPLU model is

\[
M_{\text{CPLU}}(d) = \lambda_T \Delta_u \left( \left( \frac{\rho}{2} + 1 \right) \frac{1}{d} + \frac{\rho}{2} \right) + \lambda_T \Delta_p (qd^2 - qd + 1),
\]

where \( \rho = \frac{\lambda_T}{\lambda_T} \), for any probability distribution \( f_{T(t)} \).

**Proof:** It was proven in Reference 1 that if \( f_{T(t)} \) has an Erlang distribution, the total cost of location management between two consecutive phone calls under the CPLU model is

\[
\Delta_u \left( \frac{1}{pd} + \frac{d+1}{2d} \right) + \Delta_p (qd^2 - qd + 1),
\]

for any probability distribution \( f_{T(t)} \). Since there are \( \lambda_T \) phone calls per unit of time, we get the theorem.

The following theorem gives the cost of location management per unit of time for a single mobile user under the CWLU model.

**Theorem 10.2** The total cost of location management per unit of time under the CWLU model is

\[
M_{\text{CWLU}}(d) = \lambda_T \frac{\Delta_u}{d} + \lambda_T \Delta_p (qd^2 - qd + 1)
\]

for any probability distributions \( f_{T(t)} \) and \( f_{T(t)} \).
Proof: It was proven in Reference 1 that for any probability distributions \( f_{T(t)} \) and \( f_{T(s)} \), the total cost of location management between two consecutive phone calls under the CWLU model is

\[
\frac{\Delta_u}{\rho d} + \Delta_p (qd^2 - qd + 1).
\]

Since there are \( \lambda_T \) phone calls per unit of time, we get the theorem.

Let us consider the derivative of \( M_{CPLU}(d) \) in Theorem 10.1 for the CPLU model:

\[
\frac{\partial M_{CPLU}(d)}{\partial d} = -\lambda_T \Delta_u \left( \frac{\rho}{2} + 1 \right) \frac{1}{d^2} + \lambda_T \Delta_p (2d - 1) = 0;
\]

that is,

\[
-\Delta_u \left( \frac{\rho + 2}{2\rho} \right) 1d^2 + \Delta_p (2d - 1) = 0,
\]

which yields

\[
2d^3 - d^2 - y_{CPLU} = 0,
\]

where

\[
y_{CPLU} = \left( \frac{\rho + 2}{2\rho} \right) \left( \frac{\Delta_u}{\Delta_p} \right).
\]

It can be verified that \( 2d^3 - d^2 \geq 0 \) if \( d \geq 1/2 \), and \( 2d^3 - d^2 \) is an increasing function in the range \([1/2, \infty)\). Thus, for \( y \geq 0 \), there is a unique solution: \( d \in [1/2, \infty) \), which satisfies \( 2d^3 - d^2 = y \). Furthermore, \( d \geq 1 \) if and only if \( y \geq 1 \). To ensure that there is a solution of \( d \geq 1 \), we assume that \( y_{CPLU} \geq 1 \). The above equation of \( d \) can be solved using basic algebra, and the solution is given below. We define a function \( g(y) \) as

\[
g(y) = \left( \frac{1}{4} \left( y + \frac{1}{54} + \sqrt{y^2 + \frac{y}{27}} \right) \right) + \left( \frac{1}{4} \left( y + \frac{1}{54} - \sqrt{y^2 + \frac{y}{27}} \right) \right) + \frac{1}{6}.
\]

We have the following theorem.

**Theorem 10.3** The optimal value of the movement threshold which minimizes \( M_{CPLU}(d) \) is either \( \lfloor d \rfloor \) or \( \lceil d \rceil \), where \( d = g(y_{CPLU}) \), whichever minimizes \( M_{CPLU}(d) \).

Let us consider the derivative of \( M_{CWLU}(d) \) in Theorem 10.2 for the CWLU model:

\[
\frac{\partial M_{CWLU}(d)}{\partial d} = -\lambda_T \Delta_u \frac{1}{d^2} + \lambda_T \Delta_p (2d - 1) = 0;
\]

that is,

\[
-\frac{\Delta_u}{\rho d^2} + \Delta_p (2d - 1) = 0.
\]
which yields
\[ 2d^3 - d^2 - y_{CWLU} = 0, \]
where
\[ y_{CWLU} = \frac{1}{\rho q} \left( \frac{\Delta_u}{\Delta_p} \right). \]

Again, we assume that \( y_{CWLU} \geq 1 \). The above equation of \( d \) can be solved using basic algebra, and the solution is given below.

**Theorem 10.4** The optimal value of the movement threshold which minimizes \( M_{CWLU}(d) \) is either \( \lfloor d \rfloor \) or \( \lceil d \rceil \) where \( d = g(y_{CWLU}) \), whichever minimizes \( M_{CWLU}(d) \).

### 10.4 COST OPTIMIZATION WITH LOCATION UPDATE CONSTRAINT

In this section, we consider the terminal paging cost minimization problem with constraint on location update cost.

#### 10.4.1 The CPLU Model

Assume that there are \( n \) heterogeneous mobile terminals \( u_1, u_2, \ldots, u_n \) in a wireless communication network. Each \( u_i \) has \( \lambda_{T_{ic}}, \lambda_{T_{is}}, \rho_i = \lambda_{T_{ic}}/\lambda_{T_{is}}, \) and
\[ y_i = \left( \frac{\rho_i + 22}{\rho q} \right) \left( \frac{\Delta_u}{\Delta_p} \right), \]
where \( 1 \leq i \leq n \). We are going to minimize the total cost of terminal paging of all mobile terminals per unit of time, which is defined as a function of \( d_1, d_2, \ldots, d_n \), that is,
\[ P_{CPLU}(d_1, d_2, \ldots, d_n) = \Delta_p \sum_{i=1}^{n} \lambda_{T_{ic}}(qd_i^2 - qd_i + 1), \]
with a constraint on the total cost of location update of all mobile terminals per unit of time; that is,
\[ U_{CPLU}(d_1, d_2, \ldots, d_n) = \Delta_u \sum_{i=1}^{n} \lambda_{T_{is}} \left( \frac{\rho_i}{2} + 1 \right) \left( \frac{1}{d_i} \right) = U, \]
where
\[ \frac{1}{2} \Delta_u \sum_{i=1}^{n} \lambda_{T_{is}} = U_1 < U \leq U_2 = \Delta_u \sum_{i=1}^{n} \lambda_{T_{is}}(\rho_i + 1). \]
Notice that the lower bound $U_1$ for $\tilde{U}$ is the total cost of location update of the $n$ mobile terminals per unit of time when all the $d_i$’s are infinity. The upper bound $U_2$ for $\tilde{U}$ is the total cost of location update of the $n$ mobile terminals per unit of time when all the $d_i$’s are 1. For mathematical convenience, we treat the $d_i$’s as continuous variables in the range $[1/2, \infty)$.

To minimize $P_{CPLU}(d_1, d_2, \ldots, d_n)$, we establish a Lagrange multiplier system:

$$\nabla P_{CPLU}(d_1, d_2, \ldots, d_n) = \alpha \nabla U_{CPLU}(d_1, d_2, \ldots, d_n),$$

where $\alpha$ is a Lagrange multiplier. The above equation implies that

$$\frac{\partial P_{CPLU}(d_1, d_2, \ldots, d_n)}{\partial d_i} = \frac{\partial U_{CPLU}(d_1, d_2, \ldots, d_n)}{\partial d_i},$$

that is,

$$\Delta_y \lambda_{r,q} (2d_i - 1) = -\alpha \Delta_y \lambda_{r,q} \left( \frac{\rho_i}{2} + 1 \right) \frac{1}{d_i^2},$$

for all $1 \leq i \leq n$. From the last equation, we get

$$2d_i^2 - d_i^2 - Y_i = 0,$$

where

$$Y_i = -\alpha y_i = -\alpha \left( \frac{\rho_i + 2}{2\rho_q} \right) \left( \frac{\Delta_u}{\Delta_p} \right),$$

which implies that

$$d_i = g(Y_i).$$

Substituting $d_1, d_2, \ldots, d_n$ into the constraint

$$U_{CPLU}(d_1, d_2, \ldots, d_n) = \tilde{U},$$

we obtain

$$\Delta_u \sum_{i=1}^{n} \lambda_{r,q} \left( \frac{\rho_i}{2} + 1 \right) \frac{1}{g(Y_i)} + \frac{\rho_i}{2} = \tilde{U}.$$

What remains to be done is to solve the above equation of $\alpha$.

It is easy to see that the left-hand side of the last equation is an increasing function of $\alpha$ in the range $(-\infty, 0]$. To find $\alpha$, we need to find $\alpha_1$ and $\alpha_2$, such that $\alpha$ is guaranteed in the range $[\alpha_1, \alpha_2]$. We notice that

$$g(y) > \frac{\sqrt{y}}{2}.$$
Therefore, we have

\[
\Delta_u \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{1}{g(Y_i)} + \frac{\rho_i}{2} \right) < \Delta_u \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2} + 1 \sqrt{\frac{2}{Y_i}} + \frac{\rho_i}{2} \right).
\]

We further let

\[
\Delta_u \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{2}{\sqrt{\alpha}} \frac{2}{\rho_i + 2} + \frac{\rho_i}{2} \right) = \bar{U};
\]

that is,

\[
\Delta_u \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{2}{\sqrt{\alpha}} \frac{2}{\rho_i + 2} + \frac{\rho_i}{2} \right) = \bar{U},
\]

or

\[
1 = \alpha \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{2}{\sqrt{\alpha}} \frac{2}{\rho_i + 2} + \frac{\rho_i}{2} \right) = \bar{U} - U_i.
\]

The last equation implies that if

\[
\alpha = \alpha_i = \left( \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{2}{\sqrt{\alpha}} \frac{2}{\rho_i + 2} + \frac{\rho_i}{2} \right) \right)^{-1},
\]

we have

\[
\Delta_u \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{1}{g(Y_i)} + \frac{\rho_i}{2} \right) < \bar{U}.
\]

We also notice that \(g(0) = 1/2 < 1\). Hence, if \(\alpha = \alpha_2 = 0\), we have

\[
\Delta_u \sum_{i=1}^{n} \lambda_{T_{i,j}} \left( \frac{\rho_i}{2+1} \frac{1}{g(Y_i)} + \frac{\rho_i}{2} \right) > U_2 \geq \bar{U}.
\]

The classical bisection method can be employed to search for a numerical solution to \(\alpha\) in the range \([\alpha_i, \alpha_2]\).

10.4.2 The CWLU Model

The terminal paging cost minimization problem with constraint on location update cost for the CWLU model is described and solved in a way similar to that for the CPLU model. Each \(u_i\) has \(\lambda_{T_{i,j}}, \lambda_{T_{i,j}}\), and \(\rho_i = \lambda_{T_{i,j}} / \lambda_{T_{i,j}}\). Let

\[
y_i = \frac{1}{\rho_i g(Y_i)} \left( \frac{\Delta_u}{\Delta_p} \right).
\]
We are going to optimize the total cost of terminal paging of all mobile terminals per unit of time; that is,

\[ P_{\text{CWLU}}(d_1, d_2, \ldots, d_n) = \Delta_p \sum_{i=1}^{n} \lambda_{t,i} (qd_i^2 - qd_i + 1), \]

subject to the constraint

\[ U_{\text{CWLU}}(d_1, d_2, \ldots, d_n) = \Delta_u \sum_{i=1}^{n} \lambda_{t,i} d_i = \bar{U}, \]

where

\[ 0 = U_1 < \bar{U} \leq U_2 = \Delta_u \sum_{i=1}^{n} \lambda_{t,i}. \]

Notice that the lower bound \( U_1 \) is the total cost of location update of the \( n \) mobile terminals per unit of time when all the \( d_i \)'s are infinity. The upper bound \( U_2 \) is the total cost of location update of the \( n \) mobile terminals per unit of time when all the \( d_i \)'s are 1.

To minimize \( P_{\text{CWLU}}(d_1, d_2, \ldots, d_n) \), we establish a Lagrange multiplier system,

\[ \nabla P_{\text{CWLU}}(d_1, d_2, \ldots, d_n) = \alpha \nabla U_{\text{CWLU}}(d_1, d_2, \ldots, d_n), \]

where \( \alpha \) is a Lagrange multiplier. The above equation implies that

\[ \frac{\partial P_{\text{CWLU}}(d_1, d_2, \ldots, d_n)}{\partial d_i} = \alpha \frac{\partial U_{\text{CWLU}}(d_1, d_2, \ldots, d_n)}{\partial d_i}; \]

that is,

\[ \Delta_p \lambda_{t,i} q(2d_i - 1) = -\alpha \Delta_u \lambda_{t,i} \frac{1}{d_i^2}, \]

for all \( 1 \leq i \leq n \). From the last equation, we get

\[ 2d_i^3 - d_i^2 - Y_i = 0, \]

where

\[ Y_i = -\alpha Y_i = -\frac{\alpha}{\rho q} \left( \frac{\Delta_u}{\Delta_p} \right), \]

which implies that

\[ d_i = g(Y_i). \]
Substituting \( d_1, d_2, \ldots, d_n \) into the constraint
\[
U_{\text{CWLU}}(d_1, d_2, \ldots, d_n) = \bar{U},
\]
we obtain
\[
\Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i}}{g(Y_i)} = \bar{U},
\]
which gives an equation of \( \alpha \).

To solve the above equation of \( \alpha \), we need to find \( \alpha_1 \) and \( \alpha_2 \), such that \( \alpha \) is guaranteed in the range \([\alpha_1, \alpha_2]\). By using the fact that \( g(y) > \sqrt{y} / 2 \), we obtain
\[
\Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i}}{g(Y_i)} < \Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i} \sqrt{2}}{Y_i}.
\]
Let
\[
\Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i}}{g(Y_i)} = \bar{U};
\]
that is,
\[
\Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i} \sqrt{2}}{Y_i} = \bar{U}.
\]
The last equation implies that if
\[
\alpha = \alpha_1 = - \left( \Delta_u \bar{U} \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i} \sqrt{2 \rho_q \cdot \Delta u}}{\Delta u} \right),
\]
we have
\[
\Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i}}{g(Y_i)} < \bar{U}.
\]
It is clear that if \( \alpha = \alpha_2 = 0 \), we have
\[
\Delta_u \sum_{i=1}^{n} \frac{\hat{\lambda}_{r,i}}{g(Y_i)} > U_2 \geq \bar{U}.
\]
Again, we can use the classical bisection method to search for a numerical solution to \( \alpha \) in the range \([\alpha_1, \alpha_2]\).

### 10.5 COST OPTIMIZATION WITH TERMINAL PAGING CONSTRAINT

We now consider the location update cost minimization problem with constraint on terminal paging cost.
10.5.1 The CPLU Model

Assume that there are \( n \) heterogeneous mobile terminals \( u_1, u_2, \ldots, u_n \) in a wireless communication network. Each \( u_i \) has \( \lambda_{Ti}, \lambda_{ti}, \rho_i = \lambda_{Ti}/\lambda_{ti} \), and

\[
y_i = \left( \frac{\rho_i + 2}{2 \rho_i q} \right) \left( \frac{\Delta_u}{\Delta_p} \right),
\]

here \( 1 \leq i \leq n \). We are going to minimize the total cost of location update of all mobile terminals per unit of time, which is defined as a function of \( d_1, d_2, \ldots, d_n \); that is,

\[
U_{\text{CPLU}}(d_1, d_2, \ldots, d_n) = \Delta_u \sum_{i=1}^{n} \lambda_{Ti,i} \left( \left( \frac{\rho_i + 1}{2 d_i} \right) + \frac{\rho_i}{2} \right),
\]

with a constraint on the total cost of terminal paging of all mobile terminals per unit of time; that is,

\[
P_{\text{CPLU}}(d_1, d_2, \ldots, d_n) = \Delta_p \sum_{i=1}^{n} \lambda_{Ti,i} \left( qd_i^2 - qd_i + 1 \right) = \bar{P},
\]

where

\[
\Delta_p \sum_{i=1}^{n} \lambda_{Ti,i} = P_1 \leq \bar{P} < P_2 = \infty.
\]

Notice that the lower bound \( P_1 \) for \( \bar{P} \) is the total cost of terminal paging of the \( n \) mobile terminals per unit of time when all the \( d_i \)'s are 1. The upper bound \( P_2 \) for \( \bar{P} \) is the total cost of terminal paging of the \( n \) mobile terminals per unit of time when all the \( d_i \)'s are infinity.

To minimize \( U_{\text{CPLU}}(d_1, d_2, \ldots, d_n) \), we establish a Lagrange multiplier system:

\[
\nabla U_{\text{CPLU}}(d_1, d_2, \ldots, d_n) = \beta \nabla P_{\text{CPLU}}(d_1, d_2, \ldots, d_n),
\]

where \( \beta \) is a Lagrange multiplier. The above equation implies that

\[
\frac{\partial U_{\text{CPLU}}(d_1, d_2, \ldots, d_n)}{\partial d_i} = \beta \frac{\partial P_{\text{CPLU}}(d_1, d_2, \ldots, d_n)}{\partial d_i},
\]

that is,

\[
-\Delta_u \lambda_{Ti,i} \left( \frac{\rho_i + 1}{2 d_i^2} \right) = \beta \Delta_p \lambda_{Ti,i} q(2d_i - 1),
\]

for all \( 1 \leq i \leq n \). From the last equation, we get

\[
2d_i^2 - d_i - Y_i = 0,
\]
where

\[ Y_i = \frac{y_i}{\beta} = -\frac{1}{\beta} \left( \frac{\rho_i + 2}{2 \rho_i q} \right) \left( \Delta_u \right), \]

which implies that

\[ d_i = g(Y_i). \]

Substituting \( d_1, d_2, \ldots, d_n \) into the constraint

\[ P_{\text{CPLU}}(d_1, d_2, \ldots, d_n) = \tilde{P}, \]

we obtain

\[ \Delta_p \sum_{i=1}^{n} \lambda_{i} \left( q(g(Y_i))^2 - qg(Y_i) + 1 \right) = \tilde{P}. \]

What remains to be done is to solve the above equation of \( \beta \).

It is easy to see that the left-hand side of the last equation is an increasing function of \( \beta \) in the range \((-\infty, 0)\). To find \( \beta \), we need to find \( \beta_1 \) and \( \beta_2 \), such that \( \beta \) is guaranteed in the range \([\beta_1, \beta_2]\). We notice that \( g(y) \leq \sqrt{y} \) if \( y \geq 1 \). Therefore, we have

\[ \Delta_p \sum_{i=1}^{n} \lambda_{i} \left( q(g(Y_i))^2 - qg(Y_i) + 1 \right) < \Delta_p \sum_{i=1}^{n} \lambda_{i} \left( q\sqrt{Y_i^2} + 1 \right). \]

We further let

\[ \Delta_p \sum_{i=1}^{n} \lambda_{i} \left( q\sqrt{Y_i^2} + 1 \right) = \tilde{P}, \]

that is,

\[ \Delta_p \sum_{i=1}^{n} \lambda_{i} \left( q\sqrt{\left( -\frac{1}{\beta} \frac{\rho_i + 2}{2 \rho_i q} \right)^2} \right) + 1 = \tilde{P}, \]

or

\[ \frac{1}{\sqrt{(-\beta)^2}} \sum_{i=1}^{n} \lambda_{i} \left( \frac{\rho_i + 2}{2 \rho_i q} \right)^2 \left( \frac{\Delta_u}{\Delta_p} \right)^2 = \frac{\tilde{P} - \tilde{P}}{q\Delta_p}. \]
The last equation implies that if
\[ \beta = \beta_1 = -\left(\sum_{i=1}^{n} \lambda_{t,i} \left(\frac{\rho + 2 \cdot \Delta}{2 \rho q \cdot \Delta_p}\right)^2\right)^{3/2}, \]
we have
\[ \Delta_p \sum_{i=1}^{n} \lambda_{t,i} (q(g(Y_i))^2 - qg(Y_i)+1) < \tilde{P}. \]

As for \( \beta_i \), we need to choose \( \beta_i \), which is sufficiently close to 0, such that
\[ \Delta_p \sum_{i=1}^{n} \lambda_{t,i} (q(g(Y_i))^2 - qg(Y_i)+1) \geq \tilde{P}. \]

Thus, we must have \( \beta \in [\beta_1, \beta_2] \), which can be found by using the bisection method.

10.5.2 The CWLU Model

The location update cost minimization problem with constraint on terminal paging cost for the CWLU model is formulated and solved in a way similar to that for the CPLU model. Each \( u_i \) has \( \lambda_{t,i}, \lambda_{t,i}, \) and \( \rho = \lambda_{t,i}/\lambda_{t,i} \). Let
\[ y_i = \frac{1}{\rho q} \frac{\Delta_p}{\Delta_p}. \]

We are going to optimize the total cost of location update of all mobile terminals per unit of time; that is,
\[ U_{\text{CWLU}}(d_1, d_2, \ldots, d_n) = \Delta_p \sum_{i=1}^{n} \frac{\lambda_{t,i}}{d_i}, \]
subject to the constraint
\[ P_{\text{CWLU}}(d_1, d_2, \ldots, d_n) = \Delta_p \sum_{i=1}^{n} \lambda_{t,i} (qd_i^2 - qd_i + 1) = \tilde{P}, \]
where
\[ \Delta_p \sum_{i=1}^{n} \lambda_{t,i} = P_1 \leq \tilde{P} < P_2 = \infty. \]

Notice that the lower bound \( P_1 \) is the total cost of terminal paging of the \( n \) mobile terminals per unit of time when all the \( d_i \)'s are 1. The upper bound \( P_2 \) is the total cost of terminal paging of the \( n \) mobile terminals per unit of time when all the \( d_i \)'s are infinity.
To minimize $U_{CWLU}(d_1, d_2, \ldots, d_n)$, we establish a Lagrange multiplier system:

$$\nabla U_{CWLU}(d_1, d_2, \ldots, d_n) = \beta \nabla P_{CWLU}(d_1, d_2, \ldots, d_n),$$

where $\beta$ is a Lagrange multiplier. The above equation implies that

$$\frac{\partial U_{CWLU}(d_1, d_2, \ldots, d_n)}{\partial d_i} = \beta \frac{\partial P_{CWLU}(d_1, d_2, \ldots, d_n)}{\partial d_i},$$

that is,

$$-\Delta_p \lambda_{r,i} \frac{1}{d_i} = \beta \Delta_p \lambda_{r,i} q(2d_i - 1),$$

for all $1 \leq i \leq n$. From the last equation, we get

$$2d_i^2 - d_i - Y_i = 0,$$

where

$$Y_i = -\frac{y_i}{\beta} = -\frac{1}{\beta} \left( \frac{1}{\rho q} \left( \frac{\Delta_p}{\Delta_p} \right) \right),$$

which implies that

$$d_i = g(Y_i).$$

Substituting $d_1, d_2, \ldots, d_n$ into the constraint

$$P_{CWLU}(d_1, d_2, \ldots, d_n) = \tilde{P},$$

we obtain

$$\Delta_p \sum_{i=1}^{n} \lambda_{r,i} (q(g(Y_i))^2 - qg(Y_i) + 1) = \tilde{P},$$

which gives an equation of $\beta$.

To solve the above equation of $\beta$, we need to find $\beta_i$ and $\beta_2$, such that $\beta$ is guaranteed in the range $[\beta_1, \beta_2]$. By using the fact that $g(y) \leq \sqrt[2]{y}$, we obtain

$$\Delta_p \sum_{i=1}^{n} \lambda_{r,i} (q(g(Y_i))^2 - qg(Y_i) + 1) < \Delta_p \sum_{i=1}^{n} \lambda_{r,i} \left( q\sqrt{Y_i^2} + 1 \right).$$

Let

$$\Delta_p \sum_{i=1}^{n} \lambda_{r,i} \left( q\sqrt{Y_i^2} + 1 \right) = \tilde{P}.$$
that is,
\[
\Delta_p \sum_{i=1}^{n} \lambda_{T,i} \left( q \sqrt{\frac{1}{\beta p} \frac{1}{\rho q} \frac{\Delta u}{\Delta p}} \right)^2 + 1 = \bar{P}.
\]

The last equation implies that if
\[
\beta = \beta_1 = -\left( \sum_{i=1}^{n} \lambda_{T,i} \left( \frac{1}{\rho q} \frac{\Delta u}{\Delta p} \right)^2 \right) / \left( \bar{P} - P \right)^{3/2},
\]
we have
\[
\Delta_p \sum_{i=1}^{n} \lambda_{T,i} \left( q(g(Y)_i) \right)^2 - qg(Y_i) + 1 < \bar{P}.
\]

As for \( \beta_2 \), we need to choose \( \beta_2 \), which is sufficiently close to 0, such that
\[
\Delta_p \sum_{i=1}^{n} \lambda_{T,i} \left( q(g(Y)_i) \right)^2 - qg(Y_i) + 1 \geq \bar{P}.
\]

A numerical solution to \( \beta \) can be found by using the bisection method.

### 10.6 NUMERICAL DATA

In this section, we demonstrate numerical data.

We consider \( n = 9m \) mobile terminals classified into nine categories:

- few calls with low mobility: \( \lambda_{T,1} = 1, \lambda_{T,2} = 10, \rho = 0.1 \).
- few calls with moderate mobility: \( \lambda_{T,1} = 1, \lambda_{T,2} = 30, \rho = 0.03 \).
- few calls with high mobility: \( \lambda_{T,1} = 1, \lambda_{T,2} = 100, \rho = 0.01 \).
- moderate calls with low mobility: \( \lambda_{T,1} = 5, \lambda_{T,2} = 10, \rho = 0.5 \).
- moderate calls with moderate mobility: \( \lambda_{T,1} = 5, \lambda_{T,2} = 30, \rho = 0.17 \).
- moderate calls with high mobility: \( \lambda_{T,1} = 5, \lambda_{T,2} = 100, \rho = 0.05 \).
- much calls with low mobility: \( \lambda_{T,1} = 20, \lambda_{T,2} = 10, \rho = 2 \).
- much calls with moderate mobility: \( \lambda_{T,1} = 20, \lambda_{T,2} = 30, \rho = 0.67 \).
- much calls with high mobility: \( \lambda_{T,1} = 20, \lambda_{T,2} = 100, \rho = 0.2 \).

The number of mobile terminals in each category is \( m = 10 \).

We set the parameters \( \Delta_p \) and \( \Delta_u \) as \( \Delta_p = 1 \) and \( \Delta_u = 10, 20, 30, 40 \).

In Figures 10.3–10.6, we display the impact of location update constraint on terminal paging cost optimization for the CPLU model with \( q = 3 \) in Figure 10.3, the CWLU model with \( q = 3 \) in Figure 10.4, the CPLU model with \( q = 4 \) in Figure 10.5, and the CWLU model with \( q = 4 \) in Figure 10.6. Each curve shows the terminal paging cost \( P_{\text{CPLU}}(d_1, d_2, \ldots, d_n) \) or \( P_{\text{CWLU}}(d_1, d_2, \ldots, d_n) \) (divided by \( 10^4 \)) versus \( U \).
The interval \([U_1, U_2]\) shifts to the right on the \(\bar{U}\)-axis and the length of the interval increases as \(\Delta_n\) increases. It is observed that in the range \([U_1 + (U_2 - U_1)/3, U_2]\), \(P_{\text{CWLU}}(d_1, d_2, \ldots, d_n)\) or \(P_{\text{CPLU}}(d_1, d_2, \ldots, d_n)\), increases smoothly as \(\bar{U}\) decreases. However, further reduction of \(\bar{U}\), that is, more location update constraint, causes dramatic increase in \(P_{\text{CPLU}}(d_1, d_2, \ldots, d_n)\) or \(P_{\text{CWLU}}(d_1, d_2, \ldots, d_n)\) due to the quadratic terminal paging cost.
In Figures 10.7–10.10, we display the impact of terminal paging constraint on location update cost optimization for the CPLU model with $q = 3$ in Figure 10.7, the CWLU model with $q = 3$ in Figure 10.8, the CPLU model with $q = 4$ in Figure 10.9, and the CWLU model with $q = 4$ in Figure 10.10. Each curve shows the location update cost $U_{\text{CPLU}}(d_1, d_2, \ldots, d_n)$ or $U_{\text{CWLU}}(d_1, d_2, \ldots, d_n)$ (divided by $10^4$) versus $\tilde{P}$ (divided by $10^4$). It is observed that as $\Delta_u$ increases, $U_{\text{CPLU}}(d_1, d_2, \ldots, d_n)$ or $U_{\text{CWLU}}(d_1, d_2, \ldots, d_n)$ decreases smoothly.
FIGURE 10.7. Optimization with terminal paging constraint (CPLU, \( q = 3 \)).

FIGURE 10.8. Optimization with terminal paging constraint (CWLU, \( q = 3 \)).
**FIGURE 10.9.** Optimization with terminal paging constraint (CPLU, \(q = 4\)).

**FIGURE 10.10.** Optimization with terminal paging constraint (CWLU, \(q = 4\)).
10.7 CONCLUDING REMARKS

We have emphasized the motivation and significance of network-wide cost control and optimization for scalable communication in a wireless network with heterogeneous mobile terminals. Our approach is to minimize one of the total cost of location update and the total cost of terminal paging of all mobile terminals in a wireless network by fixing the other. Our cost minimization problems are formulated as multivariable optimization problems and are solved numerically for the CPLU model and the CWLU model. Our approach eliminates the issue of cost unification of location update and terminal paging since the two kinds of cost are separated. The model and method developed in this chapter provide an effective way to handle cost trade-off in large-scale wireless communication networks.

REFERENCES


