Minimizing the Maximum Delay for Reaching Consensus in Quorum-Based Mutual Exclusion Schemes

Tatsuhiro Tsuchiya, Member, IEEE, Masatoshi Yamaguchi, and Tohru Kikuno, Member, IEEE

Abstract—The use of quorums is a well-known approach to achieving mutual exclusion in distributed computing systems. This approach works based on a coterie, a special set of node groups where any pair of the node groups shares at least one common node. Each node group in a coterie is called a quorum. Mutual exclusion is ensured by imposing that a node gets consensus from all nodes in at least one of the quorums before it enters a critical section. In a quorum-based mutual exclusion scheme, the delay for reaching consensus depends critically on the coterie adopted and, thus, it is important to find a coterie with small delay. In [5], Fu introduced two related measures called max-delay and mean-delay. The former measure represents the largest delay among all nodes, while the latter is the arithmetic mean of the delays. She proposed polynomial-time algorithms for finding max-delay and mean-delay optimal coteries when the network topology is a tree or a ring. In this paper, we first propose a polynomial-time algorithm for finding max-delay optimal coteries and, then, modify the algorithm so as to reduce the mean-delay of generated coteries. Unlike the previous algorithms, the proposed algorithms can be applied to systems with arbitrary topology.

Index Terms—Quorums, coteries, communication delay, mutual exclusion, distributed systems.

1 INTRODUCTION

The distributed mutual exclusion problem is to guarantee that at most one computing node can enter a critical section at a time. This problem is widely recognized as one of the most fundamental problems in distributed computing since it arises in various kinds of distributed systems. For example, consider a system in which its nodes share an exclusive resource. If two or more nodes try to access the resource, a conflict occurs. To avoid such a situation, mutual exclusion has to be ensured.

The use of coteries is known as an elegant approach to mutual exclusion in distributed systems. A coterie is a special set of node groups such that any two node groups have at least one node in common (intersection property) and no node group is a superset of any other node group (minimality property) [4]. Node groups in a coterie are called quorums. Given a coterie, mutual exclusion can be achieved as follows: Before entering the critical section, a node has to acquire permission from all the nodes in at least one quorum. On the other hand, each node is allowed to give its permission to at most one node. By the intersection property, it is guaranteed that no more than one node can enter the critical section at a time.

Since the performance or robustness of such a mutual exclusion scheme depends critically on the coterie adopted by the scheme, many researchers have studied methods of designing coteries that optimize various objective functions [7], [8], [9], [14], [15]. In addition to availability and message complexity [12], the communication delay needed for achieving quorum consensus is also recognized as an important factor. Especially for systems requiring short response time, such as replicated database systems, minimizing the delay for reaching consensus is a very significant task.

Recently, the notions of max-delay and mean-delay of coteries have been introduced by Fu [5]. The max-delay of a coterie is the maximum of the delays among all nodes, while the mean-delay is the average. Fu has shown that there must be a delay-optimal coterie in a special subset of coteries, called nondominated (ND) coteries. Based on this result, she has proposed polynomial-time algorithms to find max-delay optimal and mean-delay optimal coteries for systems with special topologies, such as trees and rings. Since the number of ND coteries in such networks is very small, the algorithms can efficiently find delay-optimal coteries by enumerating ND coteries.

However, finding delay-optimal coteries on general graphs has been left as an open problem. The difficulty of this problem is mainly due to the fact that enumerative approaches are impractical for systems with general topology. This is because the number of ND coteries explodes when the number of nodes exceeds only five [4]. Recently, Bioch and Ibaraki developed an enumeration method based on Boolean algebra [1]. According to [1], however, the maximum size of the system that the method was able to handle was only seven nodes.

In this paper, we consider the problem of finding max-delay optimal coteries in systems with arbitrary
topology. We propose two algorithms to solve this problem. These algorithms take a completely different approach from Fu’s in the sense that they do not use the notion of ND coteries. The first algorithm we present finds a max-delay optimal coterie on an arbitrary network with time complexity $O(n^3 \log n)$, where $n$ is the number of nodes. The second algorithm is a modification of the first algorithm. By incorporating heuristics, this algorithm finds a max-delay optimal coterie with smaller mean-delay than that found by the first algorithm. The time complexity of the second algorithm is also $O(n^3 \log n)$.

The remainder of the paper is organized as follows: In Section 2, we explain the system model and basic notions about coteries. In Section 3, we propose an algorithm to find max-delay optimal coteries. In Section 4, we extend the proposed algorithm in order to generate max-delay optimal coteries with smaller mean-delay than that found by the first algorithm. In Section 5, we show the results obtained by applying the proposed algorithms to several sample systems. Finally, we conclude this paper with a brief summary in Section 6.

2 PRELIMINARIES

We adopt the system model used in [5]. We consider a distributed system modeled by an edge-weighted undirected graph $G = (V, E, w)$, where $V = \{v_1, v_2, \ldots, v_n\}$ is a set of $n$ nodes and $E$ is a set of edges. We assume that $G$ is connected and has neither self-loop nor parallel edges. Each edge $e \in E$ is assigned a positive real number $w(e) > 0$ as the weight of the edge. Fig. 1 illustrates an example of a system. The number attached to each edge represents its weight.

For two distinct nodes $v_i$ and $v_j$, the virtual distance $\text{dist}(v_i, v_j) = \text{dist}(v_j, v_i)$ between the two nodes is defined as the length of the shortest path on $G$, where the length of a path is the sum of the weights assigned to the edges in the path. The virtual distance represents the communication delay between the two nodes. In Fig. 1, for example, $\text{dist}(v_1, v_6)$ is 5.6. We assume that, for any node $v_i$, the virtual distance $\text{dist}(v_i, v_i)$ is zero.

Given a subset $V'$ of $V$, we assume that a node $v_i$ can communicate with all members of $V'$ with delay $\max_{v \in V'} \{\text{dist}(v, v)\}$. We let $\text{delay}(v_i, V')$ denote this delay, i.e.,

$$\text{delay}(v_i, V') = \max_{v \in V'} \{\text{dist}(v, v)\},$$

where $V'$ is a nonempty subset of $V$.

In the following, we give the formal definition of a coterie. The notion of coterie was first introduced by Garcia-Molina and Barbara [4].

**Definition 1 (Coterie).** A coterie $C$ is a set of nonempty subsets of $V$ such that the following conditions hold:

1. (Intersection property) $Q \cap Q' \neq \emptyset$ for any pair $Q, Q' \in C$.
2. (Minimality property) There is no pair $Q, Q' \in C$ such that $Q \subset Q'$.

Each element in a coterie is called a quorum. Given a coterie, mutual exclusion can be achieved using some distributed mutual exclusion algorithms [2], [9], [13]. In the following, we use the term node group to refer to a nonempty subset of $V$.

**Example 1.** Let the set of all nodes $V$ be $\{v_1, v_2, v_3, v_4, v_5, v_6\}$. Now, consider the following sets of node groups, $C_1, C_2, C_3$, and $C_4$.

$$C_1 = \{\{v_1\}\}$$
$$C_2 = \{\{v_2, v_4\}, \{v_2, v_5\}, \{v_4, v_5\}\}$$
$$C_3 = \{\{v_1, v_2, v_3\}, \{v_4, v_5, v_6\}\}$$
$$C_4 = \{\{v_1\}, \{v_1, v_2, v_3\}\}.$$

Among the four sets of node groups, $C_1$ and $C_2$ are coteries. $C_3$ is not a coterie because it does not satisfy the intersection property. $C_4$ is not a coterie either because it does not satisfy the minimality property.

Adopting definitions from [5], we now introduce the notions of the delay of a node in a coterie and the max-delay and the mean-delay of a coterie. The delay of node $v_i$ in coterie $C$, or $\text{Delay}(v_i, C)$, is given by

$$\text{Delay}(v_i, C) = \min_{Q \in C} \{\max_{v \in Q} \{\text{dist}(v, v)\}\} = \min_{Q \in C} \{\text{delay}(v_i, Q)\},$$

and the max-delay and the mean-delay of coterie $C$ are given by

$$\max-\text{delay}(C) = \max_{v \in V} \{\text{Delay}(v, C)\}$$
$$\text{mean-} \text{delay}(C) = \frac{1}{|V|} \sum_{v \in V} \text{Delay}(v, C).$$

Intuitively, the delay of a node in a coterie is the communication delay for achieving consensus when the node accesses its nearest quorum. The max-delay of a coterie is the maximum of the delays among all nodes and the mean-delay of a coterie is the arithmetic mean of the delays of all nodes.

**Example 2.** Consider the system shown in Fig. 1. Let $C = \{\{v_2, v_4\}(=Q_1), \{v_2, v_5\}(=Q_2), \{v_4, v_5\}(=Q_3)\}$ be a coterie in the system. Now, take the node $v_1$ as an example. Among the three quorums, $Q_2$ is the nearest quorum from $v_1$, and $\text{Delay}(v_1, C)$ is
Begin
/* Step 1. Calculate the virtual distance between every pair of nodes */
  Execute an algorithm such as [6] or [10].

/* Step 2. Find \( \min \) by binary search */
  Sort \( \{\text{dist}(v_i, v_j) \mid v_i, v_j \in V\} \) in ascending order and store in \( a_1, a_2, \ldots, a_m \)
  \( il := 1, iu := m \)
  While \((il \neq iu)\) do
      Begin
          \( i := [(il + iu)/2] \)
          If Intersection_Check\(_i\) = true then \( iu := i \) else \( il := i + 1 \)
      End
  \( D_i := \{v \in V \mid \text{dist}(v_i, v) \leq a_i\} \) for all \( v_i \in V \) /* \( \min = a_i, D_i = NB_i(\min) \) */

/* Step 3. Remove all superset from \( NB_i(\min)'s \). */
  \( tmp := \) a list of \( D_i's \)
  For all pairs \( v_i, v_j \in V \) do
      Begin
          If \( D_i \subseteq D_j \) then remove \( D_j \) from \( tmp \)
          If \( D_j \subset D_i \) then remove \( D_i \) from \( tmp \)
      End
  \( C_{opt} := \) a set of node groups in \( tmp \) /* max-delay optimal coterie */
End

Function Intersection_Check\(_i\)(\( r \))
Begin
  \( D_i := \{v \in V \mid \text{dist}(v_i, v) \leq r\} \) for all \( v_i \in V \)
  For all pairs \( v_i, v_j \in V \) do
      If \( D_i \cap D_j = \emptyset \) then return(false)
      return(true)
End

Fig. 2. The proposed algorithm for obtaining a max-delay optimal coterie.

\[
delay(v_1, Q_2) = \text{dist}(v_1, v_2) = 4.1.
\]

For \( v_2, v_3, v_4, v_5, \) and \( v_6, \) the nearest quorums are \( Q_1, Q_2, Q_1, Q_3, \) and \( Q_3, \) respectively. Since \( Delay(v_1, C) = 4.1, \)
\( Delay(v_2, C) = 2.5, \)
\( Delay(v_3, C) = 2.2, \)
\( Delay(v_4, C) = 2.5, \)
\( Delay(v_5, C) = 2.6, \)
and \( Delay(v_6, C) = 2.0, \)
\( \max - delay(C) \) and \( \text{mean-delay}(C) \) are 4.1 and 2.65, respectively.

Now, a max-delay optimal coterie is defined as follows:

**Definition 2 (Max-delay optimal coterie).** Let \( UC \) be the set of all coteries. A coterie \( C \in UC \) is said to be max-delay optimal if and only if
\[
\max - delay(C) = \min_{C' \in UC} \{\max - delay(C')\}.
\]

3 Finding Max-Delay Optimal Coteries

As mentioned before, the problem we consider in this paper is to find a max-delay optimal coterie for a given \( G = (V, E, w) \).

Suppose that virtual distance \( \text{dist}(v_i, v_j) \) for any \( v_i, v_j \in V \) has been computed from \( G \). In fact, this is possible in polynomial time by using some previously proposed algorithms (e.g., [10]). We can then restate the problem by explicitly specifying a quorum from which each node gets consensus.
Problem 1. Given \( V \) and virtual distance \( \text{dist}(v_i, v_j) \) for any \( v_i, v_j \in V \), find an \( n \)-tuple of node groups \( (N_1, N_2, \ldots, N_n) \) that minimizes \( \max_{v \in V} \{ \text{delay}(v, N_i) \} \), under the conditions that 1) for any two groups \( N_i \) and \( N_j \), \( N_i \cap N_j \neq \emptyset \), and 2) for any two groups \( N_i \) and \( N_j \), \( N_i \not\subset N_j \).

Suppose \( (N_1, N_2, \ldots, N_n) \) is an optimal solution to Problem 1 and let \( C \) be the set of all \( N_i \)s \( (i = 1, 2, \ldots, n) \). Note that the number of elements in \( C \) is not necessarily \( n \) since \( N_i \) may be equal to \( N_j \) for some \( i, j(\neq i) \). Then, \( C \) is obviously a max-delay optimal coterie.

In order to solve this problem, we consider a more tractable problem, removing Condition 2) from Problem 1.

Problem 2. Given \( V \) and virtual distance \( \text{dist}(v_i, v_j) \) for any \( v_i, v_j \in V \), find an \( n \)-tuple of node groups \( (N_1, N_2, \ldots, N_n) \) that minimizes \( \max_{v \in V} \{ \text{delay}(v, N_i) \} \), under the condition that 1) for any two node groups \( N_i \) and \( N_j \), \( N_i \cap N_j \neq \emptyset \).

Lemma 1. Suppose that \( (N_1, N_2, \ldots, N_n) \) is an optimal solution to Problem 2. Let \( C \) be the set of \( N_i \)s such that no other \( N_j(i \neq j) \) is a proper subset of \( N_i \), i.e.,

\[
C = \{ N_i | N_i \in N \text{ such that } \forall N_j \in N - \{N_i\}, N_j \not\subset N_i \},
\]

where \( N \) is the set of all \( N_i \)s \( (i = 1, 2, \ldots, n) \). (Note that the number of elements in \( N \) is not necessarily \( n \) since \( N_i \) may be equal to \( N_j \) for some \( i, j(\neq i) \).) Then, \( C \) is a max-delay optimal coterie.

Proof. By definition, it is clear that \( C \) is a coterie. If \( N_i \) is not in \( C \), then there is another node group \( N_j \) in \( C \) such that \( N_j \subset N_i \). Now, consider another \( n \)-tuple \( (M_1, M_2, \ldots, M_n) \) such that if \( N_i \in C \), then \( M_i = N_i \); otherwise, \( M_i \) is equal to a node group \( N_j \) in \( C \) such that \( N_j \subset N_i \). Then, since \( M_i \subseteq N_i \) holds for any \( i \), \( \text{delay}(v, M_i) \leq \text{delay}(v, N_i) \) also holds for any \( i \). It is then clear that

\[
\max_{v \in V} \{ \text{delay}(v, M_i) \} \leq \max_{v \in V} \{ \text{delay}(v, N_i) \}.
\]

In addition, since \( (N_1, N_2, \ldots, N_n) \) is an optimal solution to Problem 2,

\[
\max_{v \in V} \{ \text{delay}(v, M_i) \} \geq \max_{v \in V} \{ \text{delay}(v, N_i) \}.
\]

Hence, \( \max_{v \in V} \{ \text{delay}(v, M_i) \} = \max_{v \in V} \{ \text{delay}(v, N_i) \} \). This implies that \( (M_1, M_2, \ldots, M_n) \) is also an optimal solution to Problem 2.

Since Problem 2 differs from Problem 1 only in that it does not impose Condition 2) on solutions, the inequality \( \max_{v \in V} \{ \text{delay}(v, M_i) \} \leq \max_{v \in V} \{ \text{delay}(v, L_i) \} \) holds if \( (L_1, L_2, \ldots, L_n) \) is an optimal solution to Problem 1. Notice that \( (M_1, M_2, \ldots, M_n) \) satisfies Conditions 1) and 2) in Problem 1. Hence, it is also an optimal solution to Problem 1. By definition, \( C \) is equal to the set of \( M_i \)s, thus, \( C \) is a max-delay optimal coterie.

Proof. The proof is clear from Lemma 1 and Lemma 2.

Theorem 1 leads to an algorithm for finding a max-delay optimal coterie. Fig. 2 shows this algorithm. It consists of three steps.

In Step 1, the virtual distance between every pair of nodes is calculated. This can be done, for example, by using Floyd’s classical algorithm with time complexity \( O(n^3) \) [6]. (There are also faster algorithms, such as [10].)

In Step 2, \( \min \) and \( NB_i(min) \)s are calculated. At the beginning of this step, elements of \( \{ \text{dist}(v_i, v_j) | v_i, v_j \in V \} \), \( NB_i(min) \cap NB_j(min) \neq \emptyset \). Then, the following lemma holds.

Lemma 2. \( (NB_1(min), NB_2(min), \ldots, NB_n(min)) \) is an optimal solution to Problem 2.

Proof. (By contradiction) If

\[
(NB_1(min), NB_2(min), \ldots, NB_n(min))
\]

is not an optimal solution to Problem 2, there must exist another \( n \)-tuple \( (N_1, N_2, \ldots, N_n) \) such that \( N_i \cap N_j \neq \emptyset \) for any \( i, j \) and \( \min \max_{v \in V} \{ \text{delay}(v, N_i) \} \).

Let

\[
d = \max_{v \in V} \{ \text{delay}(v, N_i) \}.
\]

Then, since \( N_i \subseteq NB(d) \) holds for any \( i \), \( NB(d) \cap NB_i(min) \neq \emptyset \) for any \( i, j \). Because

\[
\max_{v \in V} \{ \text{delay}(v, NB_i(min)) \} = d,
\]

\( (NB_1(d), NB_2(d), \ldots, NB_n(d)) \) is also an optimal solution to Problem 2. Since \( d < \min \), this is a contradiction to the definition of \( \min \).

Let \( C_{opt} \) be the set of \( NB_i(min)s \) such that no \( NB_i(min)(i \neq j) \) is a proper subset of \( NB_i(min) \), i.e.,

\[
C_{opt} = \{ NB_i(min)|NB_i(min) \in NB \text{ such that } \forall NB_j(min) \in NB - \{NB_i(min)\}, NB_i(min) \not\subset NB_j(min)\},
\]

where \( NB \) is the set of all \( NB_i(min)s \) \( (i = 1, 2, \ldots, n) \). Then, we obtain the following theorem.

Theorem 1. \( C_{opt} \) is a max-delay optimal coterie.

Proof. The proof is clear from Lemma 1 and Lemma 2.

Problem 2'. Reduce the sizes of \( D_i \)'s */

POOL := \{ (D_i, v_j) | 1 \leq i \leq n, v_j \in D_i \}

While (POOL \( \neq \emptyset \)) do

Begin

/* Select one tuple \( (D_i, v_j) */

From POOL, choose \( (D_i, v_j) \) such that

\[
\text{dist}(v_i, v_j) = \max_{(D_i, v_j) \in \text{POOL}} \{ \text{dist}(v_i, v_j) \}.
\]

(If their is more than one such tuple, choose a tuple with the largest value of \( |D_i| \))

POOL := POOL - \( (D_i, v_j) \)

/* Check whether \( (D_i, v_j) \) can be removed or not */

If \( \forall D_k \neq D_i \) \( |D_k \cap \{v_j\} \neq \emptyset \) then \( D_i := D_i - \{v_j\} \)

End

Fig. 3. Step 2’ of the modified algorithm.
that is, all candidates for $\min$, are sorted in ascending order and stored in $a_1, a_2, \cdots$. Using this data structure and Function Intersection_Check, $\min$ can be obtained by a binary search. Function Intersection_Check checks whether $NB_i(r) \cap NB_j(r) \neq \emptyset$ holds for any $i, j$ or not.

Fig. 4. Sample systems.

<table>
<thead>
<tr>
<th>System</th>
<th>Original Algorithm</th>
<th>Modified Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>${v_1, v_2, v_3}, {v_2, v_5, v_6, v_7}$</td>
<td>${v_2, v_3}, {v_2, v_4}, {v_2, v_6}$</td>
</tr>
<tr>
<td>2</td>
<td>${v_1, v_2, v_5}, {v_5, v_6, v_7}$</td>
<td>${v_1, v_5}, {v_1, v_6}, {v_5, v_6}$</td>
</tr>
<tr>
<td>3</td>
<td>${v_3, v_7, v_8, v_9}, {v_1, v_2, v_3, v_4, v_5}$</td>
<td>${v_3, v_4}, {v_1, v_3, v_5}, {v_1, v_4, v_7}$</td>
</tr>
<tr>
<td></td>
<td>${v_1, v_2, v_3, v_4, v_5}$</td>
<td>${v_3, v_7, v_8}, {v_4, v_5, v_7}, {v_4, v_5, v_8}$</td>
</tr>
<tr>
<td></td>
<td>${v_1, v_4, v_6, v_7, v_8}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>${v_4, v_5, v_6, v_7, v_8}$</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>${v_7, v_{10}, v_{11}}, {v_4, v_5, v_9, v_{11}}$</td>
<td>${v_5, v_7}, {v_5, v_{11}}, {v_7, v_{11}}$</td>
</tr>
<tr>
<td>5</td>
<td>${v_7, v_8, v_{10}}, {v_1, v_3, v_4, v_7, v_8, v_9}$</td>
<td>${v_2, v_8}, {v_3, v_8}, {v_7, v_8}$</td>
</tr>
<tr>
<td></td>
<td>${v_1, v_2, v_3, v_5, v_6, v_7, v_8}$</td>
<td>${v_2, v_3, v_7}$</td>
</tr>
<tr>
<td></td>
<td>${v_2, v_3, v_5, v_6, v_7, v_8, v_9}$</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>${v_1, v_2, v_3, v_7}$</td>
<td>${v_2, v_7}, {v_4, v_7}, {v_6, v_7}, {v_7, v_{14}}$</td>
</tr>
<tr>
<td></td>
<td>${v_5, v_6, v_7, v_8, v_{10}, v_{11}, v_{13}, v_{14}}$</td>
<td>${v_2, v_7}, {v_4, v_7}, {v_6, v_7}, {v_7, v_{14}}$</td>
</tr>
<tr>
<td></td>
<td>${v_6, v_7, v_8, v_{10}, v_{11}, v_{12}, v_{14}}$</td>
<td>${v_2, v_4, v_6, v_{14}}$</td>
</tr>
<tr>
<td></td>
<td>${v_4, v_5, v_6, v_7, v_8, v_{10}, v_{11}, v_{12}}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>${v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_{10}, v_{11}, v_{14}}$</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1
Max-Delay Optimal Coteries
The time complexity of Step 2 is \( O(n^3) \) because each if statement takes \( O(n^2) \) time and the for loop is iterated \( \frac{n(n-1)}{2} \) times. Consequently, it is seen that the time complexity of this algorithm is \( O(n^3 \log n) \).

**Example 3.** Consider the system shown in Fig. 1. The proposed algorithm works as follows: In Step 1, the virtual distance between every pair of nodes is calculated. Then, in Step 2, the values of \( \min \) and \( NB_i(min) \)s are calculated. First, candidates for \( \min \) are sorted as follows:

\[
1.5, 1.8, 2.0, 2.1, 2.2, 2.5, 2.6, 3.6, 4.1, 4.3, 4.5, 5.6.
\]

Using a binary search, the value of \( \min \) is calculated. In this case, \( \min = 3.6 \). Then, each \( NB_i(min) \) is calculated and stored in \( D_i \) as follows:

\[
\begin{align*}
D_1 &= \{v_1, v_2, v_3\} \\
D_2 &= \{v_1, v_2, v_3, v_4\} \\
D_3 &= \{v_1, v_2, v_3, v_5, v_6\} \\
D_4 &= \{v_2, v_4, v_5, v_6\} \\
D_5 &= \{v_3, v_4, v_5, v_6\} \\
D_6 &= \{v_3, v_4, v_5, v_6\}.
\end{align*}
\]

In Step 3, we obtain a max-delay optimal coterie by removing all supersets from the \( NB_i(min) \)s. In this case, the max-delay optimal coterie is

\[
\{\{v_1, v_2, v_3\}, \{v_2, v_4, v_5, v_6\}, \{v_5, v_1, v_5, v_6\}\}.
\]

Obviously, the max-delay of the coterie is 3.6, while its mean-delay is 2.533.

### 4 Extension for Reducing Mean-Delay

The algorithm presented in the previous section can successfully find a max-delay optimal coterie \( C_{opt} \) for any given system \( G \). However, there may be other max-delay optimal coteries whose mean-delay is smaller than \( C_{opt} \). In order to find such a max-delay optimal coterie with smaller mean-delay, we propose a new algorithm by modifying the original algorithm.

Specifically, we insert a new step (referred to as Step 2') into the original algorithm between Step 2 and Step 3. At the end of Step 2, \( NB_1(min), NB_2(min), \ldots, NB_n(min) \) have been obtained. Each \( NB_i(min) \) is stored in \( D_i \) in the algorithm shown in Fig. 2. In Step 2', we examine the nodes in \( D_1, D_2, \ldots, D_n \) one by one, in a certain order, and remove those nodes whose removal does not destroy the property that every pair of \( D_i \) and \( D_j \) intersect. Even with such an additional stage that reduces the sizes of the \( D_i \)s, it is guaranteed that the coterie obtained is max-delay optimal and its mean-delay is not larger than \( C_{opt} \). The following theorem ensures this.

**Theorem 2.** Let \( (N_1, N_2, \ldots, N_n) \) be an \( n \)-tuple of node groups such that \( N_i \cap N_j = \emptyset \) for any \( v_i, v_j \in V \) and \( N_i \subseteq NB_i(min) \) for any \( v_i \in V \). Let \( C \) be the set of \( N_i \)s such that no other \( N_j(i \neq j) \) is a proper subset of \( N_i \), i.e.,

\[
C = \{N_i|N_i \subseteq N^i \text{ such that } \forall N_j \in N - \{N_i\}, N_j \nsubseteq N_i\},
\]

where \( N^i \) is the set of all \( N_i \)s (\( i = 1, 2, \ldots, n \)). Then, \( C \) is a max-delay optimal coterie and mean-delay(\( C \)) \leq \text{mean-delay}(C_{opt}).

**Proof.** Since \( N_i \subseteq NB_i(min) \) for any \( i \), \( \text{delay}(v_i, N_i) \leq \text{delay}(v_i, NB_i(min)) \) holds for any \( i \). By Lemma 2, \( (NB_1(min), NB_2(min), \ldots, NB_n(min)) \) is an optimal solution to Problem 2. Hence, \( \max_{v_i \in V} \{\text{delay}(v_i, N_i)\} = \min \) and \( (N_1, N_2, \ldots, N_n) \) is also an optimal solution to Problem 2. By Lemma 1, it is then clear that \( C \) is a max-delay optimal coterie. By definition, for any quorum \( Q \in C_{opt} \), there is a \( Q' \in C \) such that \( Q' \subseteq Q \). Note that if \( Q' \subseteq Q \), then, for any node \( v \), \( \text{delay}(v, Q') \leq \text{delay}(v, Q) \). Hence, \( \text{mean-delay}(C) \leq \text{mean-delay}(C_{opt}) \).
In order to determine which nodes are checked and removed, we do the following: Let the tuple \((D_i, v_j)\) denote a node \(v_j\) in \(D_i\). (There are then \(\sum_{v_i \in V} |D_i|\) tuples that have to be considered.) We choose \((D_i, v_j)\) with the largest value of \(\text{dist}(v_i, v_j)\) first, since this is most intuitive. If there is more than one tuple with the largest value, a node is chosen from \(D_i\) such that \(j_{D_i} = \max\). Fig. 3 shows Step 2' of the modified algorithm. (Steps 1, 2, and 3 are omitted since

Fig. 5. The details of Step 2' of the modified algorithm.
they are exactly the same as the original algorithm shown in Fig. 2.)

Since the total number of tuples considered is less than \( n^2 \), the while loop in Step 2' is repeated no more than \( n^2 \) times. It takes \( O(n) \) time to select one tuple during each iteration of the while loop by sorting the nodes in each \( D_i \) according to their distances from \( v_i \). Unlike in the first algorithm, whether or not the intersection property continues to hold after removal of the selected tuple can be determined in \( O(n) \) time by maintaining the number of nodes shared by every pair of \( D_i \)'s. (The details of Step 2' are shown in the Appendix.) Hence, the time complexity of this step is \( O(n^3) \). Since the original algorithm (the one without Step 2') is of \( O(n^3 \log n) \) time complexity, the modified algorithm is also of \( O(n^3 \log n) \) time complexity.

Example 4. Consider the system shown in Fig. 1. Since Steps 1 and 2 of the modified algorithm are the same as the original algorithm, each \( D_i \) in the modified algorithm has been set to \( N B_i \) at the end of Step 2, as shown in the previous example. Thus, there are 24 tuples that have to be considered in this case.

Step 2' reduces the sizes of the \( D_i \)'s as follows: First, among the 24 tuples, tuple \( (D_1, v_j) \) is selected such that the value of \( \text{dist}(v_i, v_j) \) is the largest. In this case, \( (D_3, v_6) \) and \( (D_6, v_3) \) have the greatest value, \( 3.6 = \text{dist}(v_3, v_6) = \text{dist}(v_6, v_3) \). Since \( |D_3| = 5 \) is larger than \( |D_6| = 4 \), \( D_3, v_6 \) is chosen first. Because \( D_3 \) = \( \{v_6\} \), i.e., \( \{v_1, v_2, v_3, v_4, v_5, v_6\} \) has at least one common node with each of \( D_1, D_2, D_3, D_5, \) and \( D_6, v_6 \) is removed from \( D_3 \). Then, the \( D_i \)'s become as follows:

\[
D_1 = \{v_1, v_2, v_3\} \\
D_2 = \{v_1, v_2, v_3, v_4\} \\
D_3 = \{v_1, v_2, v_3, v_5\} \\
D_4 = \{v_2, v_4, v_5, v_6\} \\
D_5 = \{v_1, v_4, v_5, v_6\} \\
D_6 = \{v_1, v_4, v_5, v_6\}.
\]

Next, \( (D_6, v_3) \) is chosen for checking. If \( v_3 \) were removed from \( D_6 \), \( D_6 \) would no longer intersect with \( D_1 \). Thus, this node is not removed. \( (D_4, v_5) \) is chosen next since \( \text{dist}(v_4, v_5) \) is the largest among the remaining unchecked tuples. This process is repeated until all tuples have been checked. Consequently, the \( D_i \)'s become as follows:

\[
D_1 = \{v_2, v_3\} \\
D_2 = \{v_2, v_3\} \\
D_3 = \{v_2, v_3\} \\
D_4 = \{v_2, v_6\} \\
D_5 = \{v_3, v_6\} \\
D_6 = \{v_3, v_6\}.
\]

A max-delay optimal coterie is derived from the \( D_i \)'s in Step 3. In this case, the coterie is \( \{\{v_2, v_3\}, \{v_2, v_6\}, \{v_3, v_6\}\} \). The mean-delay of this coterie is 2.433. (Recall that the mean-delay of the coterie obtained by the original algorithm was 2.533.)

5 Evaluation

Using the C language, we coded the original algorithm and the modified algorithm. For Step 1, we adopted Floyd’s algorithm [6]. We took a collection of networks from [3], as shown in Fig. 4, and used them to represent the topologies of sample systems. Random weights were assigned to the edges. For each system in Fig. 4, we executed the programs on a SUN Ultra 1 workstation and obtained max-delay optimal coteries. In all cases, the running time needed for generating a max-delay optimal coterie was less than 0.1 second.

Table 1 shows the max-delay optimal coteries obtained by the two algorithms. Table 2 summarizes the max-delays and mean-delays of the coteries. These results clearly show that max-delay optimal coteries generated by the modified algorithm have much smaller mean-delays than those generated by the original algorithms. For System 6, for example, the mean-delay of the coterie generated by the modified algorithm is about 35 percent smaller than that of the coterie generated by the original algorithm.

6 Conclusions

In this paper, we have proposed two algorithms to find max-delay optimal coteries for systems with arbitrary topology. The first algorithm finds a max-delay optimal coterie on an arbitrary network with \( O(n^3 \log n) \) time complexity, where \( n \) is the number of nodes. The second algorithm is a modification of the first algorithm. By incorporating heuristics, this algorithm finds a max-delay optimal coterie with smaller mean-delay than the original algorithm. The time complexity of the second algorithm is also \( O(n^3 \log n) \). Through case studies, we have shown that this modification can lower the mean-delay effectively.

The modified algorithm is an approximation method for solving the problem of finding a max-delay optimal coterie such that its mean-delay is minimized. Unfortunately, we have not yet developed an optimal algorithm to solve this problem nor do we know the complexity of this problem. Further, how to find max-delay optimal coteries in arbitrary networks is still left as an open problem. In addition, evaluation of the proposed algorithms using measures other than delay, such as availability or load [11], also needs further study.

Appendix

Fig. 5 shows the details of Step 2' of the modified algorithm.

Acknowledgments

The authors wish to thank Professor Ichiro Suzuki of the University of Wisconsin-Milwaukee for his useful comments. We also express our gratitude to the anonymous reviewers for useful comments, especially on the complexity of the proposed algorithms. This work is in part supported by a grant from Inamori Foundation.
REFERENCES


