# Relationship between Approximability and Request Structures in the Minimum Certificate Dispersal Problem

Izumi, Tomoko

Izumi, Taisuke

Ono, Hirotaka

Wada, Koichi

https://hdl.handle.net/2324/14761

出版情報：Lecture Notes in Computer Science. 5609, 2009-07. Springer
バージョン：
権利関係：

# Relationship between Approximability and Request Structures in the Minimum Certificate Dispersal Problem[⋆]

Tomoko IZUMI[1], Taisuke IZUMI[2], Hirotaka ONO[3], and Koichi WADA[2]

[1] College of Information Science and Engineering, Ritsumeikan University,
Kusatsu, 525-8577 Japan.
izumi-t@fc.ritsumei.ac.jp
[2] Graduate School of Engineering, Nagoya Institute of Technology, Nagoya, 466-8555, Japan.
{t-izumi, wada}@nitech.ac.jp
[3] Graduate School of Information Science and Electrical Engineering, Kyushu University,
Fukuoka, 819-0395, Japan.
ono@csce.kyushu-u.ac.jp

**Abstract.** Given a graph $G = (V, E)$ and a set $R \subseteq V \times V$ of requests, we consider to assign a set of edges to each node in $G$ so that for every request $(u, v)$ in $R$ the union of the edge sets assigned to $u$ and $v$ contains a path from $u$ to $v$. *The Minimum Certificate Dispersal Problem* (MCD) is defined as one to find an assignment that minimizes the sum of the cardinality of the edge set assigned to each node. In this paper, we give an advanced investigation about the difficulty of MCD by focusing on the relationship between its (in)approximability and request structures. We first show that MCD with general $R$ has $\Theta(\log n)$ lower and upper bounds on approximation ratio under the assumption $P \neq NP$, where $n$ is the number of nodes in $G$. We then assume $R$ forms a clique structure, called *Subset-Full*, which is a natural setting in the context of the application. Interestingly, under this natural setting, MCD becomes to be 2-approximable, though it has still no polynomial time approximation algorithm whose factor better than $677/676$ unless $P = NP$. Finally, we show that this approximation ratio can be improved to 3/2 for undirected variant of MCD with Subset-Full.

## 1 Introduction

*Background and Motivation.* Let $G = (V, E)$ be a directed graph and $R \subseteq V \times V$ be a set of ordered pairs of nodes, which represents requests about reachability between two nodes. For given $G$ and $R$, we consider an assignment of a set of edges to each node in $G$. The assignment satisfies a request $(u, v)$ if the union of the edge sets assigned to $u$ and $v$ contains a path from $u$ to $v$. *The Minimum Certificate Dispersal Problem* (MCD) is the one to find the assignment satisfying all requests in $R$ that minimizes the sum of the cardinality of the edge set assigned to each node.

[⋆] This work is supported in part by KAKENHI no. 19700058, 21500013 and 21680001, Asahi-glass Foundation, Inamori Foundation and Hori Information Science Promotion Foundation.

This problem is motivated by a requirement in public-key based security systems, which are known as a major technique for supporting secure communication in a distributed system [3, 5–8, 10, 11]. The main problem of the systems is to make each user's public key available to others in such a way that its authenticity is verifiable. One of well-known approaches to solve this problem is based on public-key certificates. A public-key certificate contains the public key of a user $v$ encrypted by using the private key of a user $u$. Any user who knows the public key of $u$ can use it to decrypt the certificate from $u$ to $v$ for obtaining the public key of $v$. All certificates issued by users in a network can be represented by a certificate graph: Each node corresponds to a user and each directed edge corresponds to a certificate. When a user $w$ has communication request to send messages to a user $v$ securely, $w$ needs to know the public key of $v$ to encrypt the messages with it. To compute $v$'s public-key, $w$ uses a set of certificates stored in $w$ and $v$ in advance. Therefore, in a certificate graph, if a set of certificates stored in $w$ and $v$ contains a path from $w$ to $v$, then the communication request from $w$ to $v$ is satisfied. In terms of cost to maintain certificates, the total number of certificates stored in all nodes must be minimized for satisfying all communication requests.

While, from the practical aspect, MCD should be handled in the context of distributed computing theory, its inherent difficulty as an optimization problem is not so clear even in centralized settings: Jung et al. discussed MCD with a restriction of available paths in [8] and proved that the problem is NP-hard. In their work, to assign edges to each node, only the restricted paths which are given for each request is allowed to be used. In [11], MCD, with no restriction of available paths, is proved to be also NP-hard even if the input graph is strongly connected. Known results about the complexity of MCD are actually only these NP-hardness. This fact yields a theoretical interest of revealing the (in)approximability of MCD. As for the positive side, MCD is polynomially solvable for bidirectional trees, rings and Cartesian products of graphs [11].

This paper also investigates how the request structures affect the difficulty of MCD. As seen above, MCD is doubly structured in a sense: One structure is the graph $G$ itself and the other is the request structure $R$. We would like to investigate how the tractability of MCD changes as the topology of $R$ changes. On MCD, our interest here is to investigate whether the hardness (of approximation) of MCD depends on the restrictions about $R$. This is a natural question not only from the theoretical viewpoint but also from the practical viewpoint, because, in public-key based security systems, a set of requests should have a certain type of structures. For example, it is reasonable to consider the situation in which a set of nodes belonging to a certain community should have requests between each other in the community. This situation is interpreted that $R$ forms a clique structure. Thus the following question arises: If $R$ forms a clique, can the approximability of MCD be improved?

*Our Contribution.* In this paper, we investigate the approximability of MCD from the perspective how the structure of $R$ affects the complexity of MCD. We classify the set $R$ of requests according to the elements of $R$: $R$ is *subset-full* if for a subset $V'$ of $V$, $R$ consists of all reachable pairs of nodes in $V'$, and $R$ is *full* if the subset $V'$ is equal to $V$. Note that Subset-Full corresponds to the situation that $R$ forms a clique. Table 1 summarizes the results in this paper.

**Table 1.** Approximability / Inapproximability shown in this paper

| | Restriction on request | | |
|---|---|---|---|
| | Arbitrary | Subset-Full | Full |
| Inapproximability | $\Omega(\log n)$ | 677/676 | open |
| | 261/260 (for bidirectional graphs) | | |
| Approximation ratio | $O(\log n)$ | 2 | 2 [11] |
| | | 3/2 (for undirected graphs) | |

Here we review our contribution. We first consider the general case: We show that if we have no restriction about $R$, a lower bound on approximation ratio for MCD is $\Omega(\log n)$ and an upper bound is $O(\log n)$, where $n$ is the number of nodes. Namely, the lower and upper bounds coincide as $\Theta(\log n)$ in terms of order. Moreover, it is proved that we can still obtain the inapproximability $\Omega(1)$ of MCD even when the graph class is restricted to bidirectional graphs. As the second half of the contribution, for subset-full requests, we show that the lower bound of approximation ratio for MCD is 677/676 and the upper bound is 2. The upper bound is proved by a detailed analysis of the algorithm MinPivot , which is proposed in [11]. While Zheng et al. have shown that MinPivot achieves approximation ratio 2 with full requests, we can obtain the same approximation ratio by a different approach even when the set of requests is subset-full. In addition, by extending the approach, it is also shown that MinPivot guarantees 3/2 approximation ratio for MCD of the undirected variant with subset-full requests.

The remainder of the paper is organized as follows. In Section 2, we define the Minimum Certificate Dispersal Problem (MCD). Section 3 presents inapproximability of MCD with general $R$ and one with Subset-Full. The upper bound of MCD with general $R$ and one with Subset-Full are shown in Sections 4 and 5 respectively. Section 6 concludes the paper. All the proofs are omitted due to space limitation.

## 2 Minimum Certificate Dispersal Problem

Let $G = (V, E)$ be a directed graph, where $V$ and $E$ are the sets of nodes and edges in $G$ respectively. An edge in $E$ connects two distinct nodes in $V$. The edge from node $u$ to $v$ is denoted by $(u, v)$. The numbers of nodes and edges in $G$ are denoted by $n$ and $m$, respectively (i.e., $n = |V|, m = |E|$). A sequence of edges $p(v_0, v_k) = (v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k)$ is called a *path* from $v_0$ to $v_k$ of length $k$. A path $p(v_0, v_k)$ can be represented by a sequence of nodes $p(v_0, v_k) = (v_0, v_1, \ldots, v_k)$. For a path $p(v_0, v_k)$, $v_0$ and $v_k$ are called the source and destination of the path respectively. The length of a path $p(v_0, v_k)$ is denoted by $|p(v_0, v_k)|$. For simplicity, we treat a path as the set of edges on the path when no confusion occurs. A shortest path from $u$ to $v$ is the one whose length is the minimum of all paths from $u$ to $v$, and the distance from $u$ to $v$ is the length of a shortest path from $u$ to $v$, denoted by $d(u, v)$.

A *dispersal* $D$ of a directed graph $G = (V, E)$ is a family of sets of edges indexed by $V$, that is, $D = \{D_v \subseteq E | v \in V\}$. We call $D_v$ a local dispersal of $v$. A local dispersal $D_v$ indicates the set of edges assigned to $v$. The *cost* of a dispersal $D$, denoted by $c.D$, is the sum of the cardinalities of all local dispersals in $D$ (i.e., $c.D = \Sigma_{v \in V} |D_v|$). A request is a reachable ordered pair of nodes in $G$. For a request $(u, v)$, $u$ and $v$ are called the

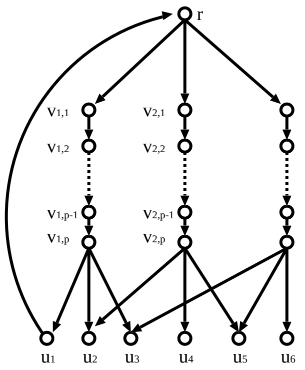

**Fig. 1.** Reduction for general case (from SET-COVER)

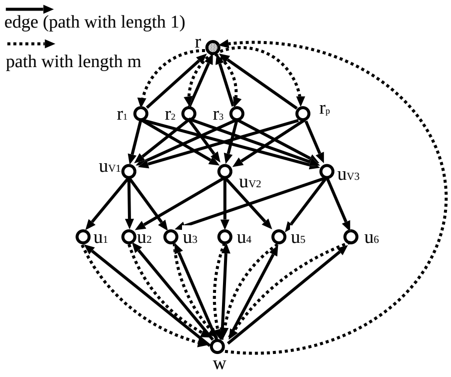

**Fig. 2.** Reduction for Subset-Full (from VERTEX-COVER)

source and destination of the request respectively. A set $R$ of requests is *subset-full* if there exists a subset of $V$ such that $R$ consists of all reachable pairs of nodes in $V'$ ($i.e.$, $R = \{(u, v)|u$ is reachable to $v$ in $G$, $u, v \in V' \subseteq V\}$), and $R$ is *full* if the subset $V'$ is equal to $V$. We say a dispersal $D$ of $G$ *satisfies* a set $R$ of requests if a path from $u$ to $v$ is included in $D_u \cup D_v$ for any request $(u, v) \in R$.

The *Minimum Certificate Dispersal Problem (MCD)* is defined as follows:

**Definition 1 (Minimum Certificate Dispersal Problem (MCD)).**
*INPUT: A directed graph $G = (V, E)$ and a set $R$ of requests.*
*OUTPUT: A dispersal $D$ of $G$ satisfying $R$ with minimum cost.*

The minimum among costs of dispersals of $G$ that satisfy $R$ is denoted by $c_{min}(G, R)$. For short, the cost $c_{min}(G, R)$ is also denoted by $c_{min}(G)$ when $R$ is full. Let $D^{Opt}$ be an optimal dispersal of $G$ which satisfies $R$ (i.e., $D^{Opt}$ is one such that $c.D^{Opt} = c_{min}(G, R)$).

In this paper, we deal with MCD for undirected graphs in Section 5.3. For an undirected graph $G$, the edge between nodes $u$ and $v$ is denoted by $(u, v)$ or $(v, u)$. When an edge $(u, v)$ is included in a local dispersal $D_v$, the node $v$ has two paths from $u$ to $v$ and from $v$ to $u$.

## 3 Inapproximability

It was shown in [11] that MCD for strongly connected graphs is NP-hard by a reduction from the VERTEX-COVER problem. In this section, we provide another proof of NP-hardness of MCD for strongly connected graphs, which implies a stronger inapproximability. Here, we show a reduction from the SET-COVER problem. For a collection $C$ of subsets of a finite universal set $U$, $C' \subseteq C$ is called a *set cover* of $U$ if every element in $U$ belongs to at least one member of $C'$. Given $C$ and a positive integer $k$, SET COVER is the problem of deciding whether a set cover $C' \subseteq C$ of $U$ with $|C'| \leq k$ exists.

The reduction from SET-COVER to MCD is as follows: Given a universal set $U = \{1, 2, \ldots, n\}$ and its subsets $S_1, S_2, \ldots, S_m$ and a positive integer $k$ as an instance $\mathcal{I}$

of SET-COVER, we construct a graph $G_{\mathcal{I}}$ including gadgets that mimic (a) elements, (b) subsets, and (c) a special gadget: (a) For each element $i$ of the universe set $U = \{1, 2, \ldots, n\}$, we prepare an element gadget $u_i$ (it is just a vertex); let $V_U$ be the set of element vertices, i.e., $V_U = \{u_i \mid i \in U\}$. (b) For each subset $S_j \in C$, we prepare a directed path $(v_{j,1}, v_{j,2}, \ldots, v_{j,p})$ of length $p - 1$, where $p$ is a positive integer used as a parameter. The end vertex $v_{j,p}$ is connected to the element gadgets that correspond to elements belonging to $S_j$. For example, if $S_1 = \{2, 4, 5\}$, we have directed edges $(v_{1,p}, u_2)$, $(v_{1,p}, u_4)$ and $(v_{1,p}, u_5)$. (c) The special gadget just consists of a base vertex $r$. This $r$ has directed edges to all $v_{j,1}$'s of $j = 1, 2, \ldots, m$. Also $r$ has an incoming edge from each $u_i$. See Figure 1 as an example of the reduction, where $S_1 = \{1, 2, 3\}, S_2 = \{2, 4, 5\}$ and $S_3 = \{3, 5, 6\}$. We can see that $G_{\mathcal{I}}$ is strongly connected. The set $R$ of requests contains the requests from the base vertex $r$ to all element vertices $u_i$, i.e., $R = \{(r, u_i) \mid u_i \in V_U\}$.

We can show the following, although we omit the proof because it is straightforward: (i) If the answer of instance $\mathcal{I}$ of SET-COVER is yes, then $c_{min}(G, R) \le pk + n$. (ii) Otherwise, $c_{min}(G, R) \ge p(k+1)+n$. About the inapproximability of SET-COVER, it is known that SET-COVER has no polynomial-time approximation algorithm with factor better than $0.2267 \ln n$, unless $P = NP$ [1]. From this inapproximability, we obtain a *gap-preserving reduction* [2] as follows:

**Lemma 1.** *The above construction of $G_{\mathcal{I}}$ is a gap-preserving reduction from SET-COVER to MCD for strongly connected graphs such that*

(i) *if $OPT_{SC}(\mathcal{I}) \le g(\mathcal{I})$, then $c_{min}(G, R) \le p \cdot g(\mathcal{I}) + n$,*
(ii) *if $OPT_{SC}(\mathcal{I}) \ge g(\mathcal{I}) \cdot c \ln n$, then $c_{min}(G, R) \ge (p \cdot g(\mathcal{I}) + n)\left(c \ln n - \frac{cn \ln n - n}{p \cdot g(\mathcal{I}) + n}\right)$,*

*where $OPT_{SC}(\mathcal{I})$ denotes the optimal value of SET-COVER for $\mathcal{I}$ and $c = 0.2267$.*

By taking $p$ large so as to satisfy $p \cdot g(I) + n = n^{1+\alpha}$ for $\alpha > 0$, we have the following:

**Theorem 1.** *There exists no $(0.2267(1 + \alpha)^{-1} \ln |V| - \varepsilon)$ factor approximation polynomial time algorithm of MCD for strongly connected graphs unless $P = NP$, where $\alpha$ and $\varepsilon$ are arbitrarily small positive constants.*

We can obtain some inapproximability result for bidirectional graphs, by slightly modifying the graph $G_{\mathcal{I}}$, though we omit the details.

**Theorem 2.** *There exists no $(261/260 - \varepsilon)$ factor approximation polynomial time algorithm of MCD for bidirectional graphs unless $P = NP$, where $\varepsilon$ is an arbitrarily small positive constant.*

Again we consider another reduction from VERTEX-COVER for graphs with degree at most 4, in which we embed an instance to MCD problem with a subset-full request structure. As well as the reduction from SET-COVER, we prepare (a) edge gadgets, (b) vertex gadgets, and (c) special gadgets. The reduction from VERTEX-COVER to MCD with subset-full requests is as follows: Given $G = (V, E)$ with degree at most 4 and a positive integer $k$ as an instance $\mathcal{I}$ of VERTEX-COVER, where $V = \{1, 2, \ldots, n\}$ is the vertex set and $E = \{e_1, e_2, \ldots, e_m\}$ is the edge set, we construct an MCD graph $G'_{\mathcal{I}}$. (a) For each edge $e_i$ in $E$, we prepare an $m$-length directed path $(u_i, u_{i,1}, \ldots, u_{i,m-1}, w)$

and $(w, u_i)$ as an edge gadget, where $w$ is a common vertex among edge gadgets. (b) For each vertex $j \in V$, we prepare a vertex $u_j^V$ as a vertex gadget. If $j$ is connected with edge $e_i$, we add directed edges $(u_j^V, u_i)$. For example, if $e_4 = \{2, 3\}$, we have directed edges $(u_2^V, u_4)$, $(u_3^V, u_4)$. Note that each $u_i$ has exactly two incoming edges from vertex gadgets. (c) The special gadgets consist of $p$ base vertices $r_1, r_2, \ldots, r_p$ and one root vertex $r$. Each $r_j$ and $r$ are connected by path $(r, r_{j,1}, \ldots, r_{j,m-1}, r_j)$ and edge $(r_j, r)$. Also, each $r_i$ has directed edges to all $u_j^V$'s of $j = 1, 2, \ldots, m$. Furthermore, we prepare an $m$-length directed path from $w$ to $r$, i.e., $(w, w_1, \ldots, w_{m-1}, r)$. See Figure 2 as an example of the reduction, in which we have $e_2 = \{1, 2\}, e_3 = \{1, 3\}$ and $e_5 = \{2, 3\}$. We can see that $G'_{\mathcal{I}}$ is strongly connected. The set $R'$ of requests are defined as $R' = R_{a,a} \cup R_{a,c} \cup R_{c,c}$, where $R_{a,a} = \{(u_i, u_j) \mid i, j = 1, 2, \ldots, m, \text{ and } i \neq j\}$, $R_{a,c} = \{(u_i, r_j), (r_j, u_i) \mid i = 1, \ldots, m\}$ and $R_{c,c} = \{(r_i, r_j) \mid i, j = 1, 2, \ldots, p, \text{ and } i \neq j\}$.

**Lemma 2.** *Let $p = m$. The above construction of $G'_{\mathcal{I}}$ and $R'$ is a gap-preserving construction from VERTEX-COVER with degree at most* 4 *to MCD with subset-full requests for strongly connected graphs such that:*

(i) *If $OPT_{VC}(\mathcal{I}) = g(\mathcal{I})$, then $c_{min}(G'_{\mathcal{I}}, R') \leq m(g(\mathcal{I}) + 3m + 3)$.*
(ii) *If $OPT_{VC}(\mathcal{I}) > c \cdot g(\mathcal{I})$, then $c_{min}(G'_{\mathcal{I}}, R') > m(g(\mathcal{I}) + 3m + 3)(c - \frac{(3m+3)(c-1)}{g(\mathcal{I})+3m+3})$,*

*where $OPT_{VC}(\mathcal{I})$ denotes the optimal value of VERTEX-COVER for $\mathcal{I}$ and $c = 53/52$.*

The constant $c = 53/52$ represents an inapproximability bound for VERTEX-COVER with degree at most 4 under the assumption $P \neq NP$ [4]. From this lemma and $4g(\mathcal{I}) \geq m$, we obtain the following theorem:

**Theorem 3.** *There exists no $(677/676 - \varepsilon)$ factor approximation polynomial time algorithm of MCD with subset-full requests for strongly connected graphs unless $P = NP$, where $\varepsilon$ is an arbitrarily small positive constant.*

## 4 Approximability

In the previous section, we show that it is difficult to design a polynomial time approximation algorithm of MCD whose factor is better than $(0.2267(1 + \alpha)^{-1} \ln n - \varepsilon)$, even if we restrict that the input graph is strongly connected. In this section, in contrast, we show that MCD has a polynomial time approximation algorithm whose factor is $O(\log n)$, which is applicable for general graphs. This implies that we clarify an optimal approximability / inapproximability bound in terms of order under the assumption $P \neq NP$.

The idea of $O(\log n)$-approximation algorithm is based on formulating MCD as a *submodular set cover problem* [9]: Let us consider a finite set $N$, a nonnegative cost function $c_j$ associated with each element $j \in N$, and non-decreasing submodular function $f : 2^N \mapsto Z^+$. A function $f$ is called *non-decreasing* if $f(S) \leq f(T)$ for $S \subseteq T \subseteq N$, and is called *submodular* if $f(S) + f(T) \geq f(S \cap T) + f(S \cup T)$ for $S, T \subseteq N$. For a subset $S \subseteq N$, the cost of $S$, say $c(S)$, is $\sum_{j \in S} c_j$.

By these $f$, $c$ and $N$, the submodular set cover problem is formulated as follows: **[Minimum Submodular Set Cover (SSC)]**

$$\min\left\{\sum_{j\in S} c_j : f(S) = f(N)\right\}.$$

It is known that the greedy algorithm of SSC has approximation ratio $H(\max_{j\in N} f(j))$ where $H(i)$ is the $i$-th harmonic number if $f$ is integer-valued and $f(\emptyset) = 0$ [9]. Note that $H(i) < \ln i + 1$.

We here claim that our problem is considered a submodular set cover problem. Let $N = \bigcup_{u\in V}\{x_{e,u} \mid e \in E\}$. Intuitively, $x_{e,u} \in S \subseteq N$ represents that the local dispersal of $u$ contains $e \in E$ in $S$, i.e., $e \in D_u$. For $S \subseteq N$, we define $d_S(u, v)$ as the distance from $u$ to $v$ under the setting that each edge $e \in D_u \cup D_v$ of $S$ has length 0 otherwise 1. That is, if all edges are included in $D_u \cup D_v$ of $S$, then $d_S(u, v) = 0$. If no edge is included in $D_u \cup D_v$ of $S$, then $d_S(u, v)$ is the length of a shortest path from $u$ to $v$ of $G$. Let $f(S) = \sum_{(u,v)\in R}(d_\emptyset(u, v) - d_S(u, v))$. This $f$ is integer-valued and $f(\emptyset) = 0$. In the problem setting of MCD, we can assume that for any $(u, v) \in R$, $G$ has a (directed) path from $u$ to $v$. (Otherwise, we have no solution). Then the condition $f(N) = f(S)$ means that all the requests are satisfied. Also cost $c$ reflects the cost of MCD.

Then we have the following lemma:

**Lemma 3.** *Function $f$ defined as above is a non-decreasing submodular function.*

Notice that $f$ can be computed in polynomial time.

By these, MCD is formulated as a submodular set cover problem. Since we have $\max_{x_{e,u}\in N} f(\{x_{e,u}\}) \le |R| \max_{u,v} d_\emptyset(u, v) \le n^3$, the approximation ratio of the greedy algorithm is $O(\log n)$. We obtain the following.

**Theorem 4.** *There is a polynomial time algorithm with approximation factor $O(\log n)$ for MCD.*

## 5 Approximation Algorithm for Subset-Full

Zheng et al. have proposed a polynomial-time algorithm for MCD, called MinPivot, which achieves approximation ratio 2 for strongly connected graphs when a set $R$ of requests is full. In this section, we show that even when $R$ is subset-full, MinPivot achieves approximation ratio 2 for strongly connected graphs. Moreover, we show that MinPivot is a 3/2-approximation algorithm for MCD of the undirected variant with subset-full requests.

### 5.1 Algorithm MinPivot

A pseudo-code of the algorithm MinPivot is shown in Algorithm 1[4]. For the explanation of the algorithm, we define $\mathcal{P}(u, v)$ as the minimum-cardinality set of edges that constitute a round-trip path between $u$ and $v$ on $G$.

[4] Although the original MinPivot is designed to work for any set of requests, we here show a simplified one because we focus on the case when $R$ is subset-full.

**Algorithm 1** MinPivot $(G = (V, E), R)$

```
$V' := \{v, w \in V | (v, w) \in R\}$
for all $u \in V$ do
    for all $v \in V'$ do
        $D_v := \mathcal{P}(u, v)$ and $D(u) := \{D_v \mid v \in V\}$
    end for
end for
output $\min_{u \in V}\{c.D(u)\}$ and its $D(u)$.
```

In dispersals returned by MinPivot , one node is selected as a *pivot*. Each request is satisfied by a path via the selected pivot. The algorithm works as follows: It picks up a node $u$ as a candidate of the pivot. Then, for nodes $v, w$ in each request $(v, w) \in R$, MinPivot stores a round-trip path between $v$ (reps. $w$) and the pivot $u$ in $D_v$ (resp. $D_w$) such that the sum of edges included in the round-trip path is minimum. Since there is a path from $v$ to $w$ via the pivot $u$ in $D_v \cup D_w$ for each request $(v, w)$, the dispersal satisfies $R$. For every pivot candidate, the algorithm MinPivot computes the corresponding dispersal and returns the minimum-cost one among all computed dispersals.

In [11], the following theorem is proved.

**Theorem 5.** *For a strongly connected graph $G$,* MinPivot *is a 2-approximation algorithm for MCD on $G$ with a full request. It completes in $O(n^7)$ time for a strongly connected graph and in $O(nm)$ time for an undirected graph.*

### 5.2 Proof of 2-approximation for Strongly Connected Graphs

In this subsection, we prove the following theorem.

**Theorem 6.** *For a strongly connected graph $G$ and a subset-full request $R$,* MinPivot *is a 2-approximation algorithm.*

We first introduce several notations used in the proof: The set of nodes included in requests in $R$ is denoted by $V_R$, that is, $V_R = \{u, v \mid (u, v) \in R\}$. Let $x$ be a node in $V_R$ with the minimum local dispersal in $D^{Opt}$ (i.e., $|D_x^{Opt}| = \min\{|D_v^{Opt}| \mid v \in V_R\}$). When there is more than one node with the minimum local dispersal, $x$ is defined as one of them chosen arbitrarily. In the following argument, we can consider only the case of $|D_x^{Opt}| > 0$: If $|D_x^{Opt}|$ is zero, any node in $V_R$ must have two paths from/to $x$ in its local dispersal to satisfy the requests for $x$. Then, the optimal solution is equivalent to that computed by MinPivot whose pivot candidate is $x$, which implies that MinPivot returns an optimal solution. Let $D^{MP}$ denote an output of the algorithm MinPivot. The following proposition clearly holds.

**Proposition 1.** *For a dispersal $D$, if there exists a node $u$ such that the local dispersal $D_v$ of any node $v$ in $V_R$ contains a round-trip path between $v$ and $u$, then $c.D^{MP} \leq c.D$.*

The idea of the proof is that we construct a feasible dispersal $D$ with cost at most $2 \cdot c.D^{Opt}$, which satisfies the condition shown in Proposition 1. It follows that the cost of the solution by MinPivot is bounded by $2 \cdot c.D^{Opt}$. We construct the dispersal $D$ from $D^{Opt}$ by additionally giving the minimum-size local dispersal to all nodes in $V_R$. More precisely, for every node $v \in V_R$, $D_v = D_v^{Opt} \cup D_x^{Opt}$.

Theorem 6 is easily proved from the following lemma and Proposition 1.

**Lemma 4.** *In the dispersal $D$ constructed in the above way, every node $v$ in $V_R$ has a round-trip path between $v$ and $x$ in $D_v$. In addition, $c.D \leq 2 \cdot c.D^{Opt}$ is satisfied.*

### 5.3 Proof of 3/2-approximation for Undirected Graphs

In this subsection, we prove that the approximation ratio of MinPivot is improved for MCD of the undirected variant. That is, we prove the following theorem.

**Theorem 7.** *For an undirected graph $G$ and a subset-full request $R$,* MinPivot *is a 3/2-approximation algorithm.*

In the proof, we take the same approach as the one of Theorem 6: We construct a dispersal $D$ with cost at most $\frac{3}{2} \cdot c.D^{Opt}$, which satisfies the condition in Proposition 1. Since Proposition 1 also clearly holds in undirected graphs, it follows that the cost of the solution by MinPivot is bounded by $\frac{3}{2} \cdot c.D^{Opt}$. In the proof of Theorem 6, we show that when all the edges in $D_x^{Opt}$ are added to the local dispersal of every node in $V_R$, the cost of the dispersal $D$ is at most twice as much as that of the optimal dispersal. Our proof of Theorem 7 is based on the idea that we construct a dispersal $D$ by adding each edge in $D_x^{Opt}$ to at most $|V_R|/2$ local dispersals.

In what follows, we show the construction of $D$. We define a rooted tree $T$ from an optimal dispersal $D^{Opt}$. To define $T$, we first assign a *weight* to each edge: To any edge in $D_x^{Opt}$, the weight zero is assigned. All the other edges are assigned the weight one. A rooted tree $T = (V, E_T)$ ($E_T \subseteq E$) is defined as a shortest path tree with root $x$ (in terms of weighted graphs) that spans all the nodes in $V_R$. Let $p_T(u, v)$ be the shortest path from a node $u$ to $v$ on the tree $T$. The weight of a path $p(u, v)$ is defined by the total weight of the edges on the path and denoted by $w.p(u, v)$. For each node $v$, let $p_T(v, v) = \phi$ and $w.p_T(v, v) = 0$. From the construction of the tree $T = (V, E_T)$, we obtain that $\sum_{v \in V_R} w.p_T(x, v) < c.D^{Opt}$.

For each edge $e$ in $D_x^{Opt}$, let $C(e)$ be the number of nodes from which path to the node $x$ on $T$ includes the edge $e$: $C(e) = |\{v \in V_R \mid e \in p_T(x, v)\}|$. The construction of the desired dispersal depends on whether any edge $e$ in $D_x^{Opt}$ satisfies $C(e) \leq |V_R|/2$ or not.

In the case that $C(e) \leq |V_R|/2$ holds for any edge $e$ in $D_x^{Opt}$, the dispersal $D'$ is constructed in the following way: $D' = \{D'_v \mid v \in V\}$, where $D'_v = p_T(x, v)$ for node $v$ in $V_R$, and $D'_v = \phi$ for node $v$ in $V \setminus V_R$.

**Lemma 5.** $c.D' \leq \frac{3}{2} \cdot c.D^{Opt}$

We consider the case that there is an edge such that $C(e) > |V_R|/2$. Let $T_v$ be the subtree of $T$ induced by node $v$ and all of $v$'s descendants, and $V(T_v)$ be a set of nodes in $T_v$. The set of edges in $D_x^{Opt}$ such that $C(e) > |V_R|/2$ is denoted by $\hat{D}_x^{Opt}$. Let $y$ be the node farthest from $x$ of those adjacent to some edge in $\hat{D}_x^{Opt}$. A dispersal $D''$ is constructed such that every node in $V_R$ has a path from itself to node $y$: $D'' = \{D''_v \mid v \in V\}$, where $D''_v = p_T(y, v)$ for node $v$ in $V_R \cap V(T_y)$, $D''_v = p_T(x, v) \cup p_T(x, y)$ for node $v$ in $V_R \setminus V(T_y)$, and $D''_v = \phi$ for node $v$ in $V \setminus V_R$.

**Lemma 6.** $c.D'' \leq \frac{3}{2} \cdot c.D^{Opt}$

From Lemmas 5 and 6, Theorem 7 is proved.

## 6 Concluding remarks

In this paper, we investigate the (in)approximability of MCD from a perspective of how topological structures of $R$ affect the complexity of MCD. While the approximability bound of MCD for a general setting of $R$ is evaluated as $\Theta(\log n)$ under the assumption $P \neq NP$, MCD for Subset-Full is 2-approximable though it is still inapproximable within a small constant factor unless $P = NP$. The complexity of MCD for Full, which is a special case of Subset-Full, is still open. We actually conjecture that MinPivot returns an optimal solution for MCD with Full; if it is correct, we will obtain an interesting contrast similar to the relation between Minimum Steiner Tree and Minimum Spanning Tree.

## References

1. N. Alon, D. Moshkovitz, and S. Safra. Algorithmic construction of sets for k-restrictions. *ACM Transactions on Algorithms*, 2(2):153–177, April 2006.
2. S. Arora and C. Lund. Hardness of approximation. In D. Hochbaum, editor, *Approximation Algorithms for NP-hard problems*, pages 399–446. PWS publishing company, 1995.
3. S. Capkun, L. Buttyan, and J.-P. Hubaux. Self-organized public-key management for mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 2(1):52–64, March 2003.
4. M. Chlebík and J. Chlebíková. Complexity of approximating bounded variants of optimization problems. *Theoretical Computer Science*, 354(3):320–338, 2006.
5. M. G. Gouda and E. Jung. Certificate dispersal in ad-hoc networks. In *in Proceeding of the 24th International Conference on Distributed Computing Systems (ICDCS'04)*, pages 616–623, March 2004.
6. M. G. Gouda and E. Jung. Stabilizing certificate dispersal. In *in Proceeding of the 7th International Symposium on Self-Stabilizing Systems (SSS'05)*, pages 140–152, October 2005.
7. J. Hubaux, L. Buttyan, and S. Capkun. The quest for security in mobile ad hoc networks. In *in Proceeding of the 2nd ACM international symposium on Mobile ad hoc networking and computing (Mobihoc'01)*, pages 146–155, October 2001.
8. E. Jung, E. S. Elmallah, and M. G. Gouda. Optimal dispersal of certificate chains. In *in Proceeding of the 18th International Symposium on Distributed Computing (DISC'04)*, pages 435–449, October 2004.
9. L. A. Wolsey. An analysis of the greedy algorithm for the submodular set covering pr blem. *Combinatorica*, 2(4):385–393, 1982.
10. H. Zheng, S. Omura, J. Uchida, and K. Wada. An optimal certificate dispersal algorithm for mobile ad hoc networks. *IEICE Transactions on Fundamentals*, E88-A(5):1258–1266, May 2005.
11. H. Zheng, S. Omura, and K. Wada. An approximation algorithm for minimum certificate dispersal problems. *IEICE Transactions on Fundamentals*, E89-A(2):551–558, February 2006.