Integer programming formulation for grouping clusters in ferry-assisted DTNs

K. Habibul Kabir, Masahiro Sasabe, and Tetsuya Takine
Department of Information and Communications Technology, Graduate School of Engineering,
Osaka University, Suita, Osaka, 565-0871, Japan.
habib@post.comm.eng.osaka-u.ac.jp, {sasabe, takine}@comm.eng.osaka-u.ac.jp

***Abstract*—Communication among isolated networks (clusters) in delay tolerant networks (DTNs) can be supported by a message ferry, which collects bundles from clusters and delivers them to a sink node. When there are lots of distant static clusters, multiple message ferries and sink nodes will be required. In this paper, we aim to make groups each of which consists of physically close clusters, a sink node, and a message ferry. Our main objective is minimizing the overall mean delivery delay of bundles in consideration of both offered load of clusters and distance between clusters and their sink nodes. We first model this problem as a nonlinear integer programming, based on the knowledge obtained in our previous work. Because it might be hard to solve this problem directly, we take two-step optimization approach based on linear integer programming, which yields an approximate solution of the problem. Through numerical results, we show the two-step optimization approach works well.**

***Keywords*-ferry-assisted DTN; grouping clusters; integer programming formulation;**

I. Introduction

The current end-to-end TCP/IP model does not adequately match with delay tolerant networks (DTNs) [2], [3], where there are no permanent end-to-end connections. Alternatively, a store-carry-forward message delivery scheme [2] and custody transfer mechanism [4] are used in those networks to assure reliable bundle transfers among nodes, where a bundle is the protocol data unit in DTNs. They perform a hop-by-hop reliable bundle transfer from a source node to its destination. To provide the opportunity of communication among isolated networks called *clusters*, Zhao et al. proposed message ferry schemes [11], [12], where a special mobile node proactively visits each cluster. This kind of networks can be applicable to sensor networking among physically distant regions and communications among rural areas without network infrastructure. In such situations, the system periodically collects information from isolated multi-cluster DTNs, where nodes can only communicate directly/indirectly with each other within a cluster through multi-hop communication. It is usually assumed that there exists a fixed base station called *sink node*, which serves as a connector to the Internet or to other sink nodes. In such a scenario, a message ferry helps the inter-cluster communication by acting as a mediator between each cluster and the outer world via the sink node. We call this scenario ferry-assisted multi-cluster DTNs.

If the arrival rates of bundles at clusters are different from each other and service times are not negligible, bundles in clusters with high arrival rate must wait for a long time to be delivered to the sink node, while less important visits to clusters with a few bundles also take place. In [6], we have already proposed a scheme to determine an optimal visiting order of a message ferry for one group, which minimizes the *mean delivery delay* of bundles, i.e., the average time interval from the generation of a bundle in a cluster to the completion of its delivery to the sink node. This optimization problem can be reduced to the minimization problem of the weighted mean waiting time in the conventional polling model of queueing theory [10]. The proposed visiting order is effective, especially when arrival rates of bundles in clusters and/or distances between clusters and the sink node are heterogeneous.

When there are lots of distant static clusters with heterogeneous offered load, there is potentially a drawback in designing a route using only one message ferry: The time spent for one cycle of the route increases with the number of clusters. This issue is the main concern of this paper. The whole system is divided into multiple groups, each of which consists of a sink node, clusters, and one message ferry. We assume that the sink node is constructed in one of clusters in each group. In what follows, we call the cluster with the sink node the *base cluster* and others group members. We further assume that high speed channels are available at the base cluster, so that the offered load of the base cluster is assumed to be excluded from the total offered load in each group. Note here that the total offered load handled by a message ferry should be less than one, and a moderate intensity, say 0.7 or less, is preferable. Moreover, for given number and positions of clusters, the total number of sink nodes (i.e., message ferries and groups) should be limited in order to suppress the introduction cost of the system.

Our main goal is making groups to minimize the mean delivery delay of bundles among groups. As mentioned above, we have already obtained the solution in the case of one group [6]: We have the explicit objective function that is a nonlinear function composed of arrival rate of bundles in clusters and distance between clusters and their sink nodes. Based on this knowledge, we first model our problem as a nonlinear integer programming. Due to the complexity of the objective function, however, it might be hard to solve this problem directly. Furthermore, this formulation may sacrifice the performance of lightly-loaded clusters with long distances from their base clusters, in order to minimize the overall mean delivery delay.

To tackle these problems, we introduce two-step optimization technique based on linear integer programming. In the first step, we find the minimum of longest distances between group members and their base clusters under the constraint that the offered load in each group is less than a predefined threshold (e.g., 0.7). For this purpose, we use a variant of the capacitated vertex p-center problem (CVPCP) in facility location problems [7], [9]. CVPCP tries to find locations of p capacitated facilities and assign customers to them in order to minimize the longest distance between facilities and their customers, when the locations and capacity of facilities, and locations and demand of customers are given. The first step optimization contributes to balancing the longest distance between a base cluster and its group members among groups.

The second optimization reconfigures the groups in order to minimize the overall mean delivery delay. Because the objective function in the original problem is an increasing function of the square root of the product of group member's arrival rates and distances from their base clusters, we consider minimizing the sum of those products under the constraint that the longest distance does not exceed the first step optimization result. We will give some numerical results to evaluate the characteristics of the obtained groups and how to find the optimal solution.

Besides, the intra-cluster communication can be efficiently handled by accumulating bundles from cluster members to a limited number of nodes called aggregators, with the help of the self-organized data aggregation technique in our previous work [5]. As a result, the message ferry needs to collect bundles only from the aggregators.

In summary, combining the current work with two of our previous works, we can comprehensively achieve effective data aggregation in ferry-assisted multi-clusters DTNs: 1) Making groups and determining base clusters, i.e., sink nodes according to the current work, 2) obtaining a visiting order for each group using the visiting order scheme [6], and 3) electing aggregators in each cluster based on the self-organized data aggregation technique [5].

The rest of this paper is organized as follows. Section II provides the problem formulation. In Section III, through numerical results, we demonstrate the characteristics of groups and explain how to find out the optimal grouping. Finally we conclude the paper in Section IV.

II. Problem formulation

A. Overview

Our goal is the development of a method for dividing clusters into several disjoint groups adequately in terms of the introduction cost and the total mean delivery delay. For each group, we select a base cluster, where a sink node is located, and we assign a message ferry. Recall that the sink node has a connection to the outer world and can directly handle the traffic generated in its base cluster via high speed channels, and the message ferry goes back and forth between the base cluster and other clusters in order to collect bundles. The optimal visiting order of clusters in a group, which results in the minimization of the total mean delivery delay is obtained according to [6].

In general, the total mean delivery delay of bundles decreases with the increase of the number of message ferries (which is equal to the number of groups). Therefore our problem is multi-objective. In order to restrain the introduction cost, it is preferable that the number of message ferries should be minimal within a range that the total mean delivery delay is allowable. Note that the mean delivery delay in each group of clusters has the following two features, because each group of clusters can be viewed as a polling model [6]. 1) The total offered load ρ handled by the message ferry should be moderate (e.g., $\rho \leq 0.7$) because the mean delivery delay is a nonlinear function of the total offered load ρ, which involves the factor $(1-\rho)^{-1}$. 2) Travel times between base cluster and group members linearly affect the mean delivery delay because they correspond to switchover times in the polling model.

Based on the above observation, we take the following approach. We first set the maximum allowable θ of offered load in each group and determine the lower bound K^*_{lower} of the number K of groups. We then attempt to solve a min-max integer program in order to minimize the total mean delivery delay of bundles. Note here that the number K of groups is first set to be K^*_{lower}, and if the program is not feasible, we add one to K and solve the program again. Repeating this procedure, we will have the solution with a minimum feasible $K = K^*_{\mathrm{min}}$ eventually.

B. Nonlinear integer programming formulation

We assume that there are V clusters labeled 1 to V in a certain geographical area. Let $\mathcal{V} = \{1, 2, \ldots, V\}$ denote the set of cluster indices. We define $\boldsymbol{d} = [d_{i,j}]$ $(i, j \in \mathcal{V})$ as a matrix of the message ferry's traveling time $d_{i,j}$ between cluster i and cluster j, where $d_{i,i}$ $(i \in \mathcal{V})$ is equal to zero. Also, let $\boldsymbol{\rho} = [\rho_i]$ $(i \in \mathcal{V})$ denote a vector of the offered load ρ_i of cluster i. We assume that transmission times of bundles at all clusters are independent and identically distributed (i.i.d.) according to a general distribution with mean h_i. Let λ_i $(i \in \mathcal{V})$ denote arrival rate of bundles of cluster i.

We first find the lower bound K^*_{lower} of the number K of groups. Suppose there exist K disjoint, non-empty group partitions for a given maximum allowable offered load θ in each group. Without loss of generality, we assume $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_N$. If cluster i for some $i > K$ is a base cluster of group k, there exists a cluster j $(j \leq K)$ of group k'. We then swap those two clusters; cluster j becomes a base cluster of group k and cluster i joins group k' as a group member. This swap yields another feasible group partition because it decreases the total offered load of group k' by $\rho_j - \rho_i$, and when $k \neq k'$, the total offered load of group k remains the same. Therefore, we consider only the case that clusters 1 to K are base clusters in discussing K^*_{lower} for a while. If a feasible partition of K groups is given, $\rho_{K+1} \leq \theta$ and $\rho_{K+1} + \rho_{K+2} + \cdots + \rho_V \leq K\theta$. We thus have

$$K^*_{\text{lower}} = \min_{1 \leq K \leq V} \{K;\ \rho_{K+1} \leq \theta, \sum_{i=K+1}^{V} \rho_i \leq K\theta\}.$$

Note that the minimum feasible number $K^*_{\min}$ of groups is not less than K^*_{lower}, i.e., $K^*_{\min} \geq K^*_{\text{lower}}$.

We define the set of base clusters as $\mathcal{K}$, where $|\mathcal{K}| = K$. Let $\mathcal{V}^{(k)}$ denote the set of clusters in group k, where $|\mathcal{V}^{(k)}| = V^{(k)}$. Let $\mathrm{E}[W^{(k)}_{\text{total}}]$ denote the overall mean delivery delay of bundles of group k. $\mathrm{E}[W^{(k)}_{\text{total}}]$ is defined as follows:

$$\mathrm{E}[W^{(k)}_{\text{total}}] = \frac{\displaystyle\sum_{i \in \mathcal{V}^{(k)} - \{k\}} \lambda_i \mathrm{E}[W^{(k)}_{\text{deliver},i}]}{\displaystyle\sum_{i \in \mathcal{V}^{(k)} - \{k\}} \lambda_i} \quad (k \in \mathcal{K}),$$

where the delivery delay $\mathrm{E}[W^{(k)}_{\text{deliver},i}]$ is the average time interval from the generation of a bundle of cluster i ($i \in \mathcal{V}^{(k)} - \{k\}$, $k \in \mathcal{K}$) to the completion of its delivery to the sink node in the base cluster at group k. The overall average weighted sum of total mean delivery delay of bundles of all groups becomes

$$\mathrm{E}[W_{\text{total}}] = \frac{\displaystyle\sum_{k \in \mathcal{K}} \lambda^{(k)}_{\text{total}} \mathrm{E}[W^{(k)}_{\text{total}}]}{\displaystyle\sum_{k \in \mathcal{K}} \lambda^{(k)}_{\text{total}}},$$

where $\lambda^{(k)}_{\text{total}} = \sum_{i \in \mathcal{V}^{(k)} - \{k\}} \lambda_i$.

Our main objective is to create groups of clusters in order to minimize $\mathrm{E}[W_{\text{total}}]$. Recall that minimization of $\mathrm{E}[W^{(k)}_{\text{total}}]$ can be obtained by optimizing the visiting order of the message ferry in each group k. The optimal visiting order of the message ferry can be achieved by adopting the minimization problem of conventional polling model as described in our previous work [6]. In [6], it is obtained that by ignoring constant factors and terms, the objective function of the minimization problem is reduced to

$$f^{(k)}(\boldsymbol{q}^{(k)}) = \sum_{i \in \mathcal{V}^{(k)} - \{k\}} \frac{\lambda_i}{q_i}, \tag{1}$$

where $\boldsymbol{q}^{(k)}$ is a vector of q_i ($i \in \mathcal{V}^{(k)} - \{k\}$), which is the mean number of visits at cluster i per unit time at group k ($k \in \mathcal{K}$), i.e.,

$$q_i = \frac{1 - \rho^{(k)}_{\text{total}}}{\displaystyle\sum_{j \in \mathcal{V}^{(k)} - \{k\}} \sqrt{2\lambda_j d_{k,j}}} \cdot \sqrt{\frac{\lambda_i}{2d_{k,i}}} \quad (i \in \mathcal{V}^{(k)} - \{k\}, k \in \mathcal{K}),$$

where $\rho^{(k)}_{\text{total}} = \sum_{i \in \mathcal{V}^{(k)} - \{k\}} \rho_i$. Therefore $f^{(k)}(\boldsymbol{q}^{(k)})$ in (1) is rewritten to be

$$f^{(k)}(\boldsymbol{q}^{(k)}) = \frac{\left(\displaystyle\sum_{i \in \mathcal{V}^{(k)} - \{k\}} \sqrt{2\lambda_i d_{k,i}}\right)^2}{1 - \rho^{(k)}_{\text{total}}}. \tag{2}$$

Based on the above discussion, in this paper the objective function of the grouping problem can be reduced to the minimization of weighted average of $f^{(k)}(\boldsymbol{q}^{(k)})$ among all groups:

$$\text{A:} \quad \text{minimize} \quad \frac{\displaystyle\sum_{k \in \mathcal{K}} \lambda^{(k)}_{\text{total}} f^{(k)}(\boldsymbol{q}^{(k)})}{\displaystyle\sum_{k \in \mathcal{K}} \lambda^{(k)}_{\text{total}}} \tag{3}$$

$$\text{subject to} \quad x_{i,j} \in \{0, 1\}, \quad \forall i, j \in \mathcal{V}, \tag{4}$$

$$\sum_{i \in \mathcal{V}} x_{i,j} = 1, \quad \forall j \in \mathcal{V}, \tag{5}$$

$$\sum_{i \in \mathcal{V}} x_{i,i} = K, \tag{6}$$

where $x_{i,j}$ ($i, j \in \mathcal{V}$) are decision variables such that

$$x_{i,j} = \begin{cases} 1, & \text{if } i = j \text{ and cluster } i \text{ is a base cluster,} \\ 1, & \text{if clusters } i \text{ and } j \text{ are in the same group and} \\ & \text{cluster } i \text{ is a base cluster,} \\ 0, & \text{otherwise.} \end{cases}$$

(4) and (5) ensure that cluster j is either a base cluster ($x_{j,j} = 1$) or a cluster member in the same group as base cluster i ($x_{i,j} = 1$ for $i \neq j$). (6) implies that there are K base clusters. Therefore $\mathcal{K}$ and $\mathcal{V}^{(k)}$ can be defined by $x_{i,j}$:

$$\mathcal{K} = \{i;\ x_{i,i} = 1\}, \qquad \mathcal{V}^{(k)} = \{j;\ x_{k,j} = 1\} \quad (k \in \mathcal{K}).$$

As (3) is a nonlinear function, it might be hard to solve Problem A with a straightforward method. Furthermore, the mean delivery delay of lightly-loaded group members with long distances from their base clusters may get large because the minimization of the overall mean delivery delay will be achieved at the sacrifice of the bad performance of such clusters.

To tackle these problems, we take a two-step approach based on linear integer programming. From the original objective function (3), we expect that achieving the following two characteristics leads to our objective: a) Reducing and balancing total offered load among groups under certain capacity limitation, and b) reducing and balancing the total traveling distances among groups. In the next subsection, we show this can be realized by a two-step optimization technique based on linear integer programming. Note that total offered load among groups can be reduced by selecting clusters with high offered load as base clusters.

C. Linear integer programming formulation for approximate solution: Two-step optimization

From (2), we observe that the offered load $\rho_{\text{total}}^{(k)}$ in each group should be moderate, say, 0.7 or less. We then introduce the upper bound threshold θ of the offered load in each group. The first step in our approach is a relief of lightly loaded clusters. More specifically, given $\mathcal{V}$, K, $\boldsymbol{d}$, and θ, we first try to find grouping where the longest distance between base clusters and their group members is minimized under the constraint of K and θ. This can balance the longest distance among groups. Note here that this kind of problem can be best studied by the capacitated vertex p-center problem in facility location problems [7], [9]. Hence, we can formulate the first step optimization as the following modified version of the capacitated vertex p-center problem.

$$
\begin{aligned}
\text{B:}\quad & \text{minimize} \quad W \\
& \text{subject to} \quad x_{i,j} \in \{0,1\}, \quad \forall i,j \in \mathcal{V}, \\
& \sum_{i \in \mathcal{V}} x_{i,j} = 1, \quad \forall j \in \mathcal{V}, \\
& \sum_{i \in \mathcal{V}} x_{i,i} = K, \\
& \sum_{j \in \mathcal{V}} \rho_j x_{i,j} - \rho_i x_{i,i} \le \theta x_{i,i}, \quad \forall i \in \mathcal{V}, \quad (7) \\
& \sum_{i \in \mathcal{V}} d_{i,j} x_{i,j} - W \le 0, \quad \forall j \in \mathcal{V}. \quad (8)
\end{aligned}
$$

Constraint (7) implies that for a base cluster i, the total offered load in its group is not greater than θ. Note here that for $i \in \mathcal{V}$ such that $x_{i,i} = 0$, both left and right hand sides of (7) are equal to zero. Constraint (8) ensures that the distance between a base cluster and group members in each group is not greater than W. Note that other constraints used in Problem B are the same as those used in Problem A. Recall that the initial value of K is set to be K_{lower}^* and is increased one by one to $K_{\min}^*$, where a feasible solution is found. The solution gives us base clusters and a group partition of clusters, which minimize the maximum distance between base clusters and their group members.

By solving Problem B, we obtain the minimum $W = W^*$, which provides the maximum allowable distance between group members and their base clusters. Under this constraint, we then try to minimize the mean delivery delay of bundles. Unfortunately, however, the objective function of the original problem is nonlinear and it might be difficult to solve it. We thus employ the following heuristics. From (2), we observe that the essential quantity in minimizing the mean delivery delay is $\sqrt{\lambda_i d_{k,i}}$ for group member i with base cluster k. Therefore, we reconfigure the groups to minimize the sum of

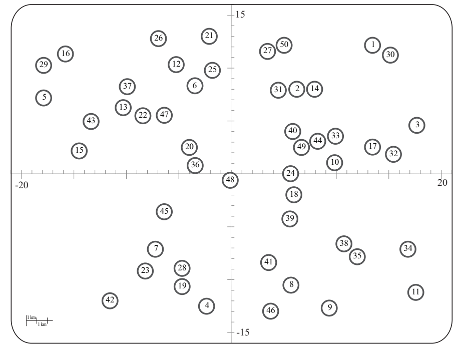

Figure 1. Random layout model with $V = 50$, where the numbers inside circles imply the cluster IDs.

$\sqrt{\lambda_i d_{k,i}}$ under the constraint of $\mathcal{K}$, θ, and W^*, where W^* is the solution of Problem B. The corresponding optimization problem is as follows.

$$
\begin{aligned}
\text{C:}\quad & \text{minimize} \quad \sum_{j \in \mathcal{V}} \sqrt{\lambda_j} \sum_{i \in \mathcal{V}} \sqrt{d_{i,j}} x_{i,j} \\
& \text{subject to} \quad x_{i,j} \in \{0,1\}, \quad \forall i,j \in \mathcal{V}, \\
& \sum_{i \in \mathcal{V}} x_{i,j} = 1, \quad \forall j \in \mathcal{V}, \\
& \sum_{i \in \mathcal{V}} x_{i,i} = K, \\
& \sum_{j \in \mathcal{V}} \rho_j x_{i,j} - \rho_i x_{i,i} \le \theta x_{i,i}, \quad \forall i \in \mathcal{V}, \\
& \sum_{i \in \mathcal{V}} d_{i,j} x_{i,j} - W^* \le 0, \quad \forall j \in \mathcal{V}.
\end{aligned}
$$

Note that in Problem C, the constraints are the same as those of Problem B except that $W = W^*$ is constant.

The remaining problem is finding optimal θ that satisfies (3). Given $x_{i,j}$ $(i, j \in \mathcal{V})$ by solving the two-step optimization problem, we can calculate (1) for each group. Therefore we can find the optimal θ as follows:

1) Set θ to be a maximum allowable offered load, e.g., 0.7.
2) Calculate the lower bound K_{lower}^* of the number K of groups according to the procedure in section II-B.
3) Find the minimum feasible number $K_{\min}^*$ of groups according to the procedure in section II-B.
4) With the help of line search technique [8], find the optimal $\theta = \theta^* \le 0.7$, which minimizes the value of the objective function of Problem A. Note that the finally obtained grouping also minimizes $\mathrm{E}[W_{\text{total}}]$.

III. Numerical results

We consider an area of 40 [km] $\times$ 30 [km], where fifty isolated clusters ($V = 50$) are randomly located, as illustrated in Fig. 1 and we then set $\boldsymbol{d} = [d_{ij}]$ $(i, j \in \mathcal{V})$ accordingly. For inter-cluster communications, we assume that each message ferry travels at a fixed speed of 10 m/s

Table I
SETTINGS OF ρ_i (V=50).

Case	ρ_1	ρ_2	ρ_3	$\cdots$	ρ_{50}	$\overline{\rho}$
Ascending	0.01	0.02	0.03	$\cdots$	0.50	0.255
Descending	0.50	0.49	0.48	$\cdots$	0.01	0.255

(i.e., 36 km/h). Table I illustrates two settings of $\boldsymbol{\rho}$ for heterogeneous and moderately loaded cases where ρ_i is assigned in an ascending order and a descending order with cluster IDs, and $\overline{\rho}$ is 0.255 in both cases. We assume that transmission times of bundles are i.i.d. according to an exponential distribution with mean $h_i = 1$ [s] ($i \in \mathcal{V}$). Since $\rho_i = \lambda_i h_i$ ($i \in \mathcal{V}$), the settings of λ_i ($i \in \mathcal{V}$) become identical to those of $\boldsymbol{\rho}$ in both cases. The total distance between base cluster k and its group members is denoted as $d^{(k)}_{\text{total}}$. By setting $\theta = 0.70$, we obtained $K^*_{\text{lower}} = K^*_{\min} = 12$ according to the procedures in section II-B. Therefore we fix $K = 12$ in the rest of this section. We also found that the minimum feasible θ_{lower} of θ is given by 0.62, so that we consider $\theta \in [0.62, 0.70]$.

We obtain the groups by solving the two-step optimization technique using CPLEX [1]. Recall that Problem B provides temporary groups by minimizing the longest distance W between base clusters and their group members, while Problem C reconfigures the groups and provides final results by minimizing the sum of $\sqrt{\lambda_i d_{k,i}}$ under the constraint of the allowable longest distance W^*. Next, we determine the optimal visiting order of the message ferry in each group according to [6]. Finally, we conduct the simulation experiments to obtain $\mathrm{E}[W^{(k)}_{\text{total}}]$ of group k and calculate $\mathrm{E}[W_{\text{total}}]$.

First, we observe the characteristics of grouping for different settings of θ in Table II for ascending case. To grasp how the grouping of clusters changes, we show the sum d_{total} of distances $d_{k,i}$ of group members from their base clusters. As we expected, there is some room to improve the performance, regardless of θ, and $\mathrm{E}[W_{\text{total}}]$ decreases in Step 2. Next, when θ decreases, d_{total} monotonically increases while the weighted average of $f^{(k)}(\boldsymbol{q}^{(k)})$, which is the objective function in the original problem, initially decreases but increases from a certain value of θ. This suggests that there is an optimal $\theta^* = 0.65$. We also observe that $\mathrm{E}[W_{\text{total}}]$ has the same tendency as the weighted average of $f^{(k)}(\boldsymbol{q}^{(k)})$ and the minimum $\mathrm{E}[W_{\text{total}}]$ is achieved at $\theta^* = 0.65$. Therefore, we can obtain the optimal θ by examining the weighted average of $f^{(k)}(\boldsymbol{q}^{(k)})$. Because of the limited search space for θ, this search does not require much computational overhead: θ should be not more than a moderate value, e.g., 0.7, and there will be the minimum feasible θ, θ_{lower}. Because we confirm the similar characteristics in descending case (Table III), we only focus on ascending case in what follows.

To examine the mean delivery delay in each group, we show $\rho^{(k)}_{\text{total}}$, $d^{(k)}_{\text{total}}$, and $\mathrm{E}[W^{(k)}_{\text{total}}]$ in Table IV ($\theta = 0.7$), Table V ($\theta = \theta^* = 0.65$), and Table VI ($\theta = \theta_{\text{lower}} = 0.62$).

Table II
d_{total}, $\mathrm{E}[W_{\text{total}}]$, AND WEIGHTED AVERAGE OF $f^{(k)}(\boldsymbol{q}^{(k)})$ ($K = 12$, ASCENDING CASE).

θ	d_{total} [km]		Weighted average	$\mathrm{E}[W_{\text{total}}]$ [s]	
	Step 1	Step 2	of $f^{(k)}(\boldsymbol{q}^{(k)})$	Step 1	Step 2
0.70	300.1	272.6	8,277.5	7,088.7	6,481.6
0.69	302.8	278.1	8,345.1	6,891.8	6,434.5
0.68	308.9	279.5	8,011.2	6,714.0	6,411.2
0.67	313.5	281.4	7,815.3	6,610.1	6,302.1
0.66	321.7	284.2	6,925.2	6,419.4	5,581.6
0.65	331.2	285.9	6,890.0	5,496.2	5,061.5
0.64	349.5	330.0	8,092.8	6,608.1	6,480.6
0.63	353.8	331.8	8,056.1	6,645.8	6,487.0
0.62	417.1	416.2	9,977.0	8,326.2	8,198.9

Table III
d_{total}, $\mathrm{E}[W_{\text{total}}]$, AND WEIGHTED AVERAGE OF $f^{(k)}(\boldsymbol{q}^{(k)})$ ($K = 12$, DESCENDING CASE).

θ	d_{total} [km]		Weighted average	$\mathrm{E}[W_{\text{total}}]$ [s]	
	Step 1	Step 2	of $f^{(k)}(\boldsymbol{q}^{(k)})$	Step 1	Step 2
0.70	257.2	239.7	7,489.4	6,109.1	5,782.2
0.69	274.6	248.2	7,382.1	5,959.9	5,600.1
0.68	290.4	254.1	6,920.5	5,812.6	5,550.1
0.67	299.0	260.3	6,625.1	5,632.5	5,391.7
0.66	304.9	262.8	6,762.7	5,401.1	5,036.4
0.65	308.9	281.5	6,370.8	5,178.6	4,935.2
0.64	347.8	317.0	7,756.6	6,237.6	5,946.5
0.63	349.7	349.7	8,290.2	7,663.5	7,328.7
0.62	414.5	392.3	9,425.0	7,761.9	7,588.9

We also present the obtained grouping for the optimal case ($\theta = 0.65$) in Fig. 2. From Table IV, if $\theta > \theta^*$, clusters with lower offered load, e.g., 14 and 32, can become base clusters, which results in higher average of $\rho^{(k)}_{\text{total}}$, i.e, 0.66. Note that the offered load of base cluster k is not included in $\rho^{(k)}_{\text{total}}$. In addition, $\rho^{(k)}_{\text{total}}$ is not well balanced: The maximum difference of $\rho^{(k)}_{\text{total}}$ among groups becomes 0.12. As a result, groups with high $\rho^{(k)}_{\text{total}}$ and large $d^{(k)}_{\text{total}}$ suffers high $\mathrm{E}[W^{(k)}_{\text{total}}]$, e.g., groups 37 and 38. If $\theta = \theta^*$ (Table V), the average and standard deviation of $\rho^{(k)}_{\text{total}}$ is improved, i.e., 0.64$\pm$0.02, with a small increase of d_{total}: $d_{\text{total}} = 272.6$ for $\theta = 0.7$ and $d_{\text{total}} = 285.9$ for $\theta = 0.65$. This can be achieved by selecting clusters with high $\rho^{(k)}_{\text{total}}$ in the dense region as in Fig. 2. If $\theta < \theta^*$ (Table VI), due to the severe bound of θ, clusters with high $\rho^{(k)}_{\text{total}}$ become base clusters regardless of their locations. As a result, $d^{(k)}_{\text{total}}$ steeply increases and thus $\mathrm{E}[W^{(k)}_{\text{total}}]$ becomes worse.

IV. CONCLUSION

In this paper, we focused on grouping clusters in ferry-assisted DTNs in order to minimize the mean delivery delay of bundles. We first modeled our problem as a nonlinear integer programming for exact solution. Due to the complexity of this problem, we further introduce two-step optimization technique based on linear integer programming for approximate solution. Through numerical results, we showed the two-step optimization can obtain optimal solution by setting

Table IV
RESULTS OF TWO-STEP OPTIMIZATION ($K = 12, \theta = 0.7$, ASCENDING CASE).

Base cluster ID	14	32	37	38	39	42	43	45	46	47	49	50
$\rho^{(k)}_{\text{total}}$	0.64	0.64	0.69	0.69	0.68	0.70	0.69	0.62	0.63	0.69	0.68	0.58
$d^{(k)}_{\text{total}}$ [km]	22.4	17.3	32.7	30.3	21.4	22.9	24.6	19.2	25.8	23.1	13.4	19.5
$\mathrm{E}[W^{(k)}_{\text{total}}]$ [s]	3,373.1	2,890.5	7,462.7	6,424.4	4,525.1	5,769.2	5,855.5	3,665.7	3,898.2	4,960.5	2,574.5	1,689.6

Table V
RESULTS OF TWO-STEP OPTIMIZATION ($K = 12, \theta = \theta^* = 0.65$, ASCENDING CASE).

Base cluster ID	33	35	37	39	42	43	44	46	47	48	49	50
$\rho^{(k)}_{\text{total}}$	0.64	0.64	0.64	0.64	0.58	0.65	0.63	0.64	0.64	0.65	0.64	0.63
$d^{(k)}_{\text{total}}$ [km]	34.9	35.2	20.9	18.8	24.4	28.2	17.7	26.6	15.7	15.9	8.5	39.1
$\mathrm{E}[W^{(k)}_{\text{total}}]$ [s]	5,373.8	4,568.3	3,036.2	2,414.4	3,091.9	4,272.5	3,015.2	3,385.0	2,353.2	2,929.5	1,001.2	3,128.1

Table VI
RESULTS OF TWO-STEP OPTIMIZATION ($K = 12, \theta = \theta_{\text{lower}} = 0.62$, ASCENDING CASE).

Base cluster ID	38	40	41	42	43	44	45	46	47	48	49	50
$\rho^{(k)}_{\text{total}}$	0.62	0.62	0.60	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62
$d^{(k)}_{\text{total}}$ [km]	32.2	21.4	27.1	36.6	37.3	23.9	42.6	34.6	22.1	57.2	38.7	42.5
$\mathrm{E}[W^{(k)}_{\text{total}}]$ [s]	4,755.9	2,928.3	4,658.6	4,867.4	5,605.6	3,859.9	6,093.8	5,523.8	3,289.3	8,596.9	4,938.5	5,717.9

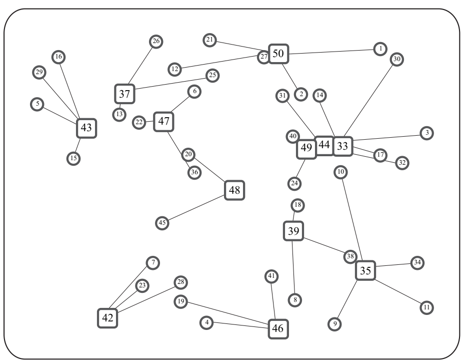

Figure 2. Optimal grouping obtained by the two-step optimization with $K = 12$ and $\theta = 0.65$ for the ascending case, where rectangles are base clusters and lines are drawn between base clusters and their group members.

θ adequately, which is realized using the original objective function as the stopping criterion.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Mutsunori Yagiura of Nagoya University for his valuable suggestions and comments. This research was supported by Global COE Program "Founding Ambient Information Society Infrastructure," MEXT, Japan, Grant-in-Aid for Young Scientists (B) of JSPS (No. 22700070), and Grant-in-Aid for Scientific Research (C) of JSPS (No. 22500056).

REFERENCES

[1] IBM ILOG CPLEX Optimizer, Available: http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/.

[2] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss, "Delay tolerant network architecture,," IETF RFC 4838 Draft..

[3] K. Fall, "A delay-tolerant network architecture for challenged internets," Proc. of *ACM SIGCOMM'03*, pp. 27–34, 2003.

[4] K. Fall and W. Hong, "Custody transfer for reliable delivery in delay tolerant networks," Intel Research Berkeley, Tech. Rep. IRB-TR-03-030, 2003.

[5] K.H. Kabir, M. Sasabe, and T. Takine, "Evolutionary game theoretic approach to self-organized data aggregation in delay tolerant networks," *IEICE Transactions on Communications*, vol. E93-B, pp. 490–500, 2010.

[6] K. H. Kabir, M. Sasabe, and T. Takine, "Optimal visiting order of isolated clusters in DTNs to minimize the total mean delivery time of bundles," *Numerical Algebra, Control and Optimization*, vol. 1, pp. 563–576, 2011.

[7] F. A. Özsoy and M. Ç. Pinar, "An exact algorithm for the capacitated vertex p-center problem," *Computers and Operations Research*, vol. 33, no. 5, pp. 1420–1436, 2006.

[8] R. L. Rardin, *Optimization in Operations Research*. Prentice Hall, 1998.

[9] M. Scaparra, S. Pallottino, and M. Scutellà, "Large-scale local search heuristics for the capacitated vertex p-center problem," *Networks*, vol. 43, no. 4, pp. 241–255, 2004.

[10] H. Takagi, "Queueing analysis of polling models: An update," *Stochastic Analysis of Computer and Communication Systems*, Elsevier, 1990.

[11] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," Proc. of *ACM MobiHoc'04*, pp. 187–198, 2004.

[12] W. Zhao, M. Ammar, and E. Zegura, "Controlling the mobility of multiple data transport ferries in a delay-tolerant network," Proc. of *INFOCOM'05*, pp. 1407–1418, 2005.