Throughput Capacity of Mobility-assisted Data Collection in Wireless Sensor Networks

Wang Liu†‡, Jianping Wang‡, Guoliang Xing§, and Liusheng Huang‡
† Department of Computer Science, University of Science and Technology of China, Hefei, China
‡ Department of Computer Science, City University of Hong Kong, Hong Kong
§ Department of Computer Science and Engineering, Michigan State University, MI, USA
Email:wangliu@mail.ustc.edu.cn, jianwang@cityu.edu.hk‡, glxing@msu.edu§, lshuang@ustc.edu.cn†

Abstract—Recently, mobility-assisted data collection has been proposed to prolong network lifetime. However, the upper bound of throughput capacity in such mobility-assisted data collection models has not been studied. In this work, we first derive the upper bound of throughput capacity when a mobile sink is available for data collection. Given the traveling speed of the mobile sink and the required delay deadline, we analyze the necessary conditions (e.g., optimal number of clusters, minimum buffer size at each cache, and minimum traveling distance) to achieve the maximum throughput capacity. Our analysis shows that \( \frac{4W}{v} \) per-node throughput capacity can be achieved at a low traveling speed, which is 3 times of the throughput capacity in a static WSN. We further extend the analysis to the case where multiple mobile devices can assist in data collection. We show that 4 mobile relays are enough to achieve the upper bound of throughput capacity, \( O\left(\frac{W}{v}\right) \). Finally, we derive the throughput capacity under existing mobility-assisted data collection models.

Keywords: WSN, Mobility, throughput capacity, network lifetime

I. INTRODUCTION

Wireless Sensor Networks (WSNs) hold the promise of revolutionizing the way that we observe and interact with the physical world in a wide range of application domains. A WSN consists of many source sensors which send the sensing data to the sink through multi-hop communications. In such a many-to-one communication mode, per-node throughput capacity, which is referred to as the maximum throughput that a sensor node can send to the sink in WSNs, is limited due to the interference (Without ambiguity, we use throughput capacity and per-node throughput capacity interchangeably in this paper).

As studied in [1], the per-node throughput capacity in a static flat WSN with \( n \) nodes is \( \frac{W}{n} \) when \( W \) is the data rate of transmission and we assume that the interference range is twice of the transmission range. Furthermore, as shown in [1], the throughput capacity per node can be higher by using clustering if the cluster heads have larger transmission range/rate and high energy to communicate with the sink directly. To relax such requirements on cluster heads, an alternative approach is to use a mobile sink which travels to each cluster head to collect data. Thus, the cluster head can communicate with the mobile sink without requiring larger transmission range/rate. In this paper, we focus on deriving the upper bound of throughput capacity in a WSN with a mobile sink.

Applications in WSNs often require data to be delivered within a certain deadline. For instance, a user may issue the following sliding-window query: “sample seismic data every 10s and archive at sink every 10 minutes”, where the deadline is 10 minutes. On the other hand, the battery of the mobile sink may only last for certain time and need to be recharged periodically. Due to limited storage at the robomote node, the mobile sink needs to dump collected data to a static server periodically. All these factors will require the mobile sink to collect data from all clusters within a deadline.

Rendezvous-based data collection with delay requirements in a WSN has been studied in [2] where a few nodes in a large sensing field serve as rendezvous points (RPs) and source nodes form into different clusters with RPs as the cluster heads. RPs buffer data sent (possibly through multiple hops) from source nodes. The mobile sink periodically visit the RPs and pick up the cached data within the required deadline, \( T \). The use of RPs enables the mobile sink to collect a large volume of data at a time without traveling a long distance, which achieves high data bandwidth and low communication delay at the same time. The work in [2] focuses on the selection of RPs along a data routing tree to minimize the energy consumption. Given the traveling speed and delay deadline, the maximum achievable throughput capacity of such a cache-based data collection assisted by a mobile sink has not been studied.

In this paper, we address the following fundamental issues related to throughput capacity of mobility-assisted data collection in large-scale WSNs with mobile devices.

• If only one mobile sink is available, given the traveling speed \( v \) and deadline \( T \), what is upper bound of throughput capacity? what is the optimal number of clusters to achieve such an upper bound of throughput capacity? With more clusters, each cluster has less number of sensors, which means less interference within each cluster, thus, throughput capacity could be larger. However, on the other hand, larger number of clusters means that the mobile sink will spend more time on traveling and less time on data collection, which limits throughput capacity. Thus, the number of clusters must be carefully chosen in order to meet the delay deadline.

978-1-4244-5113-5/09/$25.00 ©2009 IEEE
• If several mobile devices are available, how many mobile devices are enough to achieve the upper bound of throughput capacity $O(W)$? If there are several mobile devices available, they can be used as mobile relays to collect data and send to a static sink. Suppose that the static sink only has one radio to communicate with the mobile relays, the mobile relays have to take turns to dump their collected data to the static sink. If too many mobile relays are used, the time portion assigned to each mobile relay to send data to the sink will be small, which means that the time portion for each cluster to transmit data to the mobile relay is small. Therefore, too many mobile relays may not contribute to higher throughput capacity.

This paper aims to provide firm answers to the above fundamental issues. We systematically analyze the impact of various parameters on throughput capacity and provide necessary conditions to achieve the upper bound of throughput capacity. Our contributions can be summarized as follows.

• We prove that maximum 4 clusters will be enough to achieve the upper bound of throughput capacity where the number of nodes $k$ ($2 \leq k \leq 4$) is determined according to the traveling speed and delay deadline. Given traveling speed of $v$ and transmission range of $r$, we show that $\frac{3W}{4n}$ per-node throughput capacity can be achieved when $v \in \left[ \frac{2W}{3r}, \frac{W}{r} \right]$ and the number of clusters is 3, which is 3 times of the throughput capacity in a static WSN.

• We prove that 4 mobile relays are enough to achieve the per-node throughput capacity of $O(\frac{W}{n})$ which is also the upper bound of throughput capacity in any WSN without traveling speed and/or delay deadline constraints.

• We derived the throughput capacity of existing mobility-assisted data collection models. Our results provide important insights into the design of mobility-assisted sensor networks under different data collection models.

The rest of the paper is organized as follows. We review the literature work in Section II. The network model is given in Section III. We define our problem in Section IV. We analyze the upper bound of throughput capacity with one mobile sink in Section V. We derive the upper bound of throughput capacity with one static sink and $k$ mobile relays in Section VI. We discuss the throughput capacity under other mobility-assisted data collection models in Section VII. We present the simulation results in Section VIII and conclude the paper in Section IX.

II. RELATED WORK

In [3], Gupta and Kumar studied the throughput capacity of a random wireless network, where static nodes are randomly placed in the network and each node sends data to a randomly chosen destination. The throughput capacity per node is shown to be $\Theta(\frac{W}{\sqrt{nlogn}})$, where $n$ is the number of nodes in the network and $W$ is the radio capacity. It is assumed that all nodes are mobile nodes in [4] and they can coordinateably move to reduce the interference, thus, it can achieve $O(W)$ throughput capacity per node at the cost of unbounded delay and buffer requirement. The work in [1] showed that the per-node throughput capacity in a static flat WSN with $n$ nodes is $\frac{W}{n}$ when the interference range is twice of the transmission range. It is also shown in [1] that the throughput capacity per node can be higher by using clustering if the cluster heads have larger transmission range/rate and high energy to communicate with the sink directly.

Recent work has exploited controlled mobility to achieve the balance between delay, buffer requirement and throughput capacity. Controlled mobility is studied in [5], [6], [7], [8] to enhance the connectivity of sparse ad hoc networks and reduce the energy consumption of WSNs. We can classify different controlled mobility schemes according to the role of the mobility nodes played into the following categories:

• Mobile sink without cache. Under this category, the mobile nodes visit source nodes and gather data from them via either one-hop or multi-hop communication. If the mobile nodes visits all source nodes, the communication between source nodes and the mobile sink is one-hop communication. Shah et al. [9] studies the performance of mobile nodes based on the random mobility model. Several heuristics are proposed in [10], [11] to schedule the movement of mobile nodes such that the source nodes can be visited before buffer overflow. While this approach minimizes the network energy consumption by avoiding multi-hop wireless transmissions, it incurs high latency when collecting data from large sensing fields due to the slow speed of mobile nodes.

An alternative approach is to partition the network into clusters and the mobile sink will visit each cluster at a RP. When a mobile sink visits a cluster, all source nodes in the cluster communicate with the mobile sink through multi-hop communication. It is shown in [12], [13] that the optimal path of the mobile sinks is the perimeter of the sensing field. However, the average network energy consumption in this approach is high as nodes must communicate with the mobile sink(s) through multi-hop routes. Moreover, as the mobile sinks often change their paths dynamically, additional overhead is incurred in maintaining efficient routing topology.

• Mobile sink with cache. The cache nodes are used to collect data when the mobile sink is traveling. Thus, network throughput capacity can be improved. In [14], [15], [16], source nodes send data to the nodes close to the traveling path of the mobile sink waiting for the mobile sink to collect. The RP selection and traveling path are studied in [11], [10] to minimize the energy consumption on the network paths from sources to caches while ensuring that the caches can be visited by mobile sinks within temporal constraints.

In the existing cache-based mobility-assisted data collection models, there is only one cache node in each cluster. Therefore, in each cluster, the cache node will consume
more energy and becomes the bottlenecks of the network lifetime. In this paper, we propose a new cache node selection approach where the network is divided into $k$ clusters and the mobile sink only visits a RP in each cluster. All one-hop neighbors of the RP will become the cache nodes. Give the speed of the mobile sink and the temporal constraint, we derive the optimal $k$ and traveling path to maximize the network throughput.

- Static sink with mobile relays. In [17], there is a static base station with $k$ mobile relays in the network. Under such an approach, $k$ mobile relays collect data along $k$ rings. Such $k$ mobile relays then form a multi-hop communication path to relay data to the static sink. Since the mobile relay closest to the static sink will be the bottleneck, the throughput capacity is not maximized. In this paper, we will analyze the minimum number of mobile relays to achieve the upper bound of throughput capacity, $O(\frac{m}{n})$.

III. NETWORK MODEL

In this section, we discuss the necessary conditions to achieve the upper bound of throughput capacity. As shown in [1], clustering can improve per-node throughput capacity of WSNs. In this paper, we study the upper bound of throughput capacity in a large scale WSN with a mobile sink. Practical mobile sensor platforms are only capable of slow-speed movement. The typical speed of mobile sensor systems (e.g., Networked Infomechanical Systems [20], Packbot [21] and Robomote [13]) is about 0.2-2 m/s. The low traveling speed of the mobile sink indicates that the mobile sink should travel along the shortest trail to visit all clusters in order to catch the deadline and spend more time on collecting data.

We now discuss the minimum traveling distance that a mobile sink has to travel. In a cluster, the 1-hop neighbors of the cache nodes need to relay data from all other sensors. Such nodes are the most heavily loaded nodes in a cluster. Thus, in order to achieve the maximum throughput capacity, we need to assure that transmissions from such nodes to their cache nodes will not be interfered by the transmissions from other clusters. As shown in Fig. 1, $s_1$ and $s_2$ are 1-hop neighbors of the cache nodes in two different clusters. The distance between $s_1$ and $s_2$ must be at least $2r$, where $r$ is the transmission range, in order to guarantee that $s_1$ and $s_2$ can transmit to their cache nodes simultaneously. Since $s_1$ is 2 hops away from its RP and $s_2$ is also 2 hops away from its RP, the minimum distance between any two RPs must be $6r$. Thus, if a WSN is divided into $k$ clusters, the traveling distance of the mobile sink will be at least $6kr$.

With the necessary conditions of cache nodes and minimum traveling distance, we will derive the upper bound of throughput capacity in the following sections. Fig. 2 gives a network model satisfying the above necessary conditions where the mobile sink travels along a regular triangle with edge length $6r$ and all 1-hop neighbors of each RP are cache nodes.

IV. PROBLEM DESCRIPTION

Suppose that $n$ sensors are uniformly deployed in a sensing field with area of $A$ and these sensors are grouped into $k$ clusters. Let the traveling speed of the mobile sink be $v$ m/s. Let the delay deadline be $T$. Given the speed $v$, transmission range $r$, and delay deadline $T$, we aim to derive the optimal $k$ such that the throughput capacity per-node is maximized. The symbols used in this paper are given in Table I.

If the mobile sink can travel fast and spend enough time on collecting data, the throughput capacity is determined by
the allowable maximum data collection delay.

the area of the sensing field.

the average number of cache nodes in each cluster.

the waiting time of mobile sink in each RP.

the volume of data received by all cache nodes

the amount of time that the mobile sink collects

data from each cluster in every $T$ time period.

the velocity of the mobile sink.

the traveling distance.

the average number of nodes in each cluster.

maximum per-node throughput capacity

radio capacity.

maximum per-node throughput capacity for a given

the total number of nodes in the network.

each cache node can spend

collecting data from the cache nodes in each cluster. Thus,

throughput capacity is limited either by contention constraint

and harvesting constraint, there exists an optimal $x$
such that data sent to the cache nodes in $T - x$ seconds can
be picked up by the mobile sink in $x$ seconds at each cycle.

If the traveling speed of the mobile sink is fast enough,
when the mobile sink arrives at the RP of a cluster, it will
wait for some time such that more data can be sent to the

and the mobile sink then spends $x$ second on collecting data from the cache nodes in each cluster. Let the

waiting time of the mobile sink at each cluster be $\Delta$ in each
circle, we have

$$k(x + \Delta) + \frac{l}{v} = T, \quad x > 0, \Delta \geq 0$$

Relaxing the harvesting constraint, let us consider the max-

imum per-node throughput capacity in a cluster under the

contention constraint. Suppose that there is a virtual node at
each RP. According to [1], if all one-hop nodes relay their data
to the virtual node and the per-node throughput capacity is $\frac{Wx}{4}$,
which means that the virtual node is only busy in $\frac{1}{4}$ of the time
at any time interval. In a time interval of $T - x$, the virtual
node is receiving data in $\frac{4}{x}$ seconds. Thus, the one-hop
nodes spend $\frac{(T-x)}{x}$ seconds on sending and $\frac{3(T-x)}{x}$ seconds on
receiving. The volume of data received by the virtual node is

$W(T-x)$. In other words, the one-hop nodes receives $W(T-x)$
data in $\frac{3(T-x)}{x}$ seconds. Thus, the total throughput at one-hop
nodes is $\frac{W}{x}$.

In our proposed cache-based data collection model, there is
no such a virtual node and the one-hop nodes directly relay
their data to the mobile sink. Therefore, the one-hop nodes
will spend $T - x$ seconds in each $T$ time period on receiving
data and the amount of data received by these nodes would be:

$$D = \frac{W}{3} \times (T - x) = \frac{W(T - x)}{3}$$

Thus, per-node throughput capacity under the contention con-
straint is

$$U(k,v,l) \leq \frac{Dk}{Tn} = \frac{W(T-x)k}{3Tn} = \frac{W - \frac{k}{v} + \frac{k}{v} + \frac{k}{v}}{3n}$$

The total amount of data that the mobile sink can collect
from $k$ clusters in $x$ seconds will be $Wxk$. Thus, the maximum
throughput capacity will be $\frac{Wxk}{Tn}$. Therefore, the throughput
capacity under the harvesting constraint is

$$U(k,v,l) \leq \frac{Wzk}{Tn} = \frac{W}{n}(1 - \frac{l}{vT} - \frac{k\Delta}{T})$$

Let $U_1(k,v,l) = \frac{W(k(1 + \frac{k}{v} + \frac{k}{v})}{3n}$ and $U_2(k,v,l) = \frac{W}{n}(1 - \frac{l}{vT} - \frac{k\Delta}{T})$. Let $U_{max}(k,v)$ be the maximum per-node through-
put capacity for given $k$ and $v$, then we have

$$U_{max}(k,v) = \max \min \{U_1(k,v,l), U_2(k,v,l)\}$$

Let $U_{max}(v)$ be the maximum per-node throughput capacity
for a given $v$, we have

$$U_{max}(v) = \max_k U_{max}(k,v)$$

<table>
<thead>
<tr>
<th>Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
</tr>
<tr>
<td>$n$</td>
</tr>
<tr>
<td>$W$</td>
</tr>
<tr>
<td>$k$</td>
</tr>
</tbody>
</table>
| $x$ | the amount of time that the mobile sink collects

data from each cluster in every $T$ time period. |
| $v$ | the velocity of the mobile sink. |
| $l$ | the traveling distance. |
| $N$ | the average number of nodes in each cluster. |
| $N_0$ | the average number of cache nodes in each cluster. |
| $A$ | the area of the sensing field. |
| $D$ | the volume of data received by all cache nodes

of a cluster in $T$. |
| $U(k,v,l)$ | per-node throughput capacity for a given $k$, $v$, and $l$. |
| $U_{max}(k,v)$ | maximum per-node throughput capacity

for a given $k$ and $v$. |
| $U_{max}(v)$ | maximum per-node throughput capacity for a given $v$. |
| $\Delta$ | the waiting time of mobile sink in each RP. |
Our objective is to derive the optimal $k$ such that $U_{\text{max}}(v)$ is maximized for a given $v$ and $T$.

V. UPPER BOUND OF THROUGHPUT CAPACITY WITH A MOBILE SINK

Since our aim is to derive the upper bound of per-node throughput capacity, we assume that the mobile sink visits all clusters with minimum traveling distance $l = 6kr$ where such an assumption is achievable by selecting appropriate RPs such that they are exactly $6r$ apart. Therefore, we can redefine $U_1(k, v) = \frac{W(1 - \frac{6kr}{vT} - \frac{k\Delta}{T})}{n}$ and $U_2(k, v) = \frac{W}{n}(1 - \frac{6kr}{vT} - \frac{k\Delta}{T})$ in the discussion of this section.

We first show that when $k$ is larger than 4, the network cannot achieve maximum per-node throughput capacity since the advantages of having more clusters are offset by decreased throughput capacity caused by the harvesting constraint. We have the following lemma.

Lemma 1: When $k \geq 4$, $U_{\text{max}}(k, v) \leq U_{\text{max}}(4, v)$.

Proof. When $k \geq 4$, we have $\frac{k\Delta}{T} \geq 1$.

\[ U_2(k, v) = \frac{W}{n}(1 - \frac{6kr}{vT} - \frac{k\Delta}{T}) \]

\[ \leq \frac{W}{n}(k - 1 + \frac{6kr}{vT} + \frac{k\Delta}{T}) = U_1(k, v) \]

which means that per-node throughput capacity is limited by the harvesting constraint when $k \geq 4$, i.e., the mobile sink cannot collect all data sent to the cache nodes in a cluster during its harvesting time interval. Since

\[ U_2(k, v) = \frac{W}{n}(1 - \frac{6kr}{vT} - \frac{k\Delta}{T}) \quad (7) \]

$U_2(k, v)$ decreases with the increase of $k$. We have $U_2(k, v) \leq U_2(4, v)$ and $U_{\text{max}}(k, v) = \max U_2(k, v) < \max U_2(4, v) = U_{\text{max}}(4, v)$ when $k \geq 4$. Therefore, the lemma holds.

With Lemma 1, in order to derive the upper bound of throughput capacity, we only need to discuss the cases of $k = 2, 3, 4$, and $k = 4$. For the case of $k = 4$, if $v > \frac{24r}{T}$, $U_{\text{max}}(4, v) = \frac{W}{n}(T - \frac{24r}{vT} - k\Delta)$ with $l = 24r$. If $v \leq \frac{24r}{T}$, the mobile sink will spend all time on traveling and cannot harvest any data. Thus, the per-node throughput is minimum. We now consider the case of $k = 2$. We have the following lemma.

Lemma 2: When $k = 2$,

\[ U_{\text{max}}(2, v) = \begin{cases} \frac{W}{n}(1 - \frac{12r}{vT}), & v \in \left(\frac{12r}{T}, \frac{24r}{T}\right) \\ \frac{W}{2}, & v \in \left(\frac{24r}{T}, \infty\right) \end{cases} \quad (8) \]

Proof. Since $U_1(2, v)$ increases with the increase of $\frac{l}{v}$ ($l = 12r$) and $U_2(2, v)$ decreases with the increase of $\frac{l}{v}$. Then $U_{\text{max}}(2, v)$ can be achieved when $U_1(2, v) = U_2(2, v)$, which is:

\[ \frac{W}{3nT}(T + \frac{12r}{v} + k\Delta) = \frac{W}{nT}(T - \frac{12r}{v} - k\Delta) \quad (9) \]

which implies that $U_{\text{max}}(2, v)$ can achieve the maximum capacity as long as $\frac{12r}{v} + k\Delta = \frac{4nT}{v} = \frac{2}{T}$.

The upper bound of throughput capacity under a given $v$ and the optimal $k$ to achieve the upper bound of throughput capacity

<table>
<thead>
<tr>
<th>$v$</th>
<th>$U_{\text{max}}(v)$</th>
<th>$k$</th>
<th>$\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\left(\frac{12r}{T}, \frac{24r}{T}\right)$</td>
<td>$\frac{W}{n}(1 - \frac{12r}{vT})$</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>$\left(\frac{24r}{T}, \infty\right)$</td>
<td>$\frac{W}{n}(1 - \frac{12r}{vT})$</td>
<td>3</td>
<td>$\frac{1}{4}(\frac{T}{r} - \frac{1}{2})$</td>
</tr>
<tr>
<td>$\left(\frac{24r}{T}, \infty\right)$</td>
<td>$\frac{W}{n}(1 - \frac{12r}{vT})$</td>
<td>3</td>
<td>$\frac{1}{4}(\frac{T}{r} - \frac{1}{2})$</td>
</tr>
<tr>
<td>$\left(\frac{24r}{T}, \infty\right)$</td>
<td>$\frac{W}{n}(1 - \frac{12r}{vT})$</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Theorem 1: $U_{\text{max}}(v) = \max_{2 \leq k \leq 4}\{U_{\text{max}}(k, v)\}$.

Proof. The theorem follows directly from Lemma 1, Lemma 2 and Lemma 3.

We summarize the relationship between $v$, $U_{\text{max}}(v)$, $k$, $\Delta$ in Table II. From Table II, we can see that when the traveling speed is slow, it is not worth to let the mobile sink travel to too many clusters. When the traveling speed is moderate, e.g., $v \in \left[\frac{12r}{T}, \frac{96r}{T}\right]$, the upper bound of throughput capacity is $\frac{3W}{4n}$, which is 3 times of that in a static WSN.

When $v < \frac{12r}{T}$, the mobile sink will stay at one position and collect data from its one-hop cache nodes. It will achieve $\frac{W}{4n}$ per-node throughput capacity which is same as the per-node throughput capacity achieved in a WSN with a static sink.

In the above analysis, we assume that the buffer size at each cache node is large enough. In what follows, we analyze the minimum buffer size requirement to achieve the maximum throughput capacity.

Lemma 4: To achieve the maximum throughput capacity, the minimum memory requirement in cache nodes is $M_0 = \frac{A}{k\pi r^2}TU_{\text{max}}(v)$.
Proof. In a cluster, there are $\frac{n}{A}$ sensors and $N_0 = \frac{n^2}{A}$ of them are cache nodes. Let $M$ be the buffer size of each cache. We have

$$U_{\text{max}}(v)T\frac{n}{k} \leq M\frac{\pi r^2}{A}$$

Thus,

$$M \geq A\frac{2}{k\pi r^2}TU_{\text{max}}(v)$$

Therefore, the minimum buffer size $M_0$ is $A\frac{2}{\pi r^2}TU_{\text{max}}(v)$. \[ \square \]

VI. THROUGHPUT CAPACITY WITH A STATIC SINK AND $k$ MOBILE RELAYS

In the above analysis, we have derived the upper bound of throughput capacity with a mobile sink under the given traveling speed and delay deadline, which is less than $\frac{W}{n}$. From Table II, we can see that, in order to achieve per-node throughput capacity of $\frac{W}{n}$, the traveling speed must be infinitely fast.

If multiple mobile relays are available, it can compensate the limitation of low traveling speed since each mobile relay only needs to travel at most the distance between a RP and the static sink, which could be less than the distance of $6r$. Thus, the upper bound of throughput capacity $\frac{W}{n}$ may be achieved. In this section, we derive the minimum number of mobile relays and the minimum traveling speed requirement to achieve per-node throughput capacity of $\frac{W}{n}$.

Since we aim to derive the minimum number of mobile relays to achieve such an upper bound of throughput capacity, we assume that each mobile relay moves back and forth between the static sink and the RP in each cluster to collect data. In order to achieve per-node throughput capacity of $\frac{W}{n}$, the mobile sink has to be busy all the time, i.e., after one mobile relay dumps its collected data to the static sink, another mobile sink arrives at the position which is one-hop away from the static sink and starts to send it collected data to the static sink, and so on. Fig. 3 illustrates such a data collection model.

![Fig. 3. Data collection with k mobile relays](image)

Let $l$ be the distance between each RP and the location that mobile relay can start to transmit data to the sink. In each $T$, the mobile relay spends $\frac{2l}{v}$ seconds on traveling, and the time spent on collecting data from the cache nodes is the same as the time it relays the data to the base station. Suppose that a mobile relay spends $x$ seconds in each $T$ time period on collecting data, then we have $T = \frac{2l}{v} + 2x$. According to (3), in each cluster the cache nodes collect data in $T - x$ seconds and the volume of data collected will be $\frac{W}{n}(T - x) = \frac{W}{n}(x + \frac{2l}{v})$. The volume of data that a mobile relay can collect within $x$ seconds is $Wx$. Thus, all data collected by caches during $x + \frac{2l}{v}$ seconds can be forward to the mobile relay in $x$ time interval when the mobile relay stays at the RP of the cluster.

$$\frac{W}{3}(x + \frac{2l}{v}) \geq Wx$$

Thus, we have

$$\frac{l}{v} \geq x \tag{12}$$

Since $T = \frac{2l}{v} + 2x$, we have, $x \leq \min\{\frac{T}{4}, \frac{l}{v}\}$. The throughput capacity is:

$$U(k, v, l) = \frac{Wkx}{nT} = \frac{Wk(T - \frac{2l}{v})}{2n} \tag{13}$$

$$= \frac{Wkx}{nT} \tag{14}$$

$$\leq \frac{Wk}{4n} \tag{15}$$

From (15), at least $k \geq 4$ mobile relays are needed to achieve the upper bound of throughput capacity $\frac{W}{n}$. We now prove the upper bound of throughput capacity can be achieved when $k = 4$ with an affordable traveling speed. When $k = 4$ and $U(k, v, l) = \frac{W}{n}$, we require $x = \frac{T}{4} = \frac{l}{4}$, thus:

$$v = \frac{4l}{T} \tag{16}$$

As we known, the minimum distance between any two RPs is $6r$, then we have following minimum traveling distance $l$ for mobile relays:

$$(l + r)\sin\left(\frac{\pi}{k}\right) = 3r$$

Thus, $\[ l = \frac{3r}{\sin\left(\frac{\pi}{k}\right)} - r \] \tag{17}$

When $k = 4$, we have $l = (3\sqrt{2} - 1)r$, $x = \frac{l}{4}, v = \frac{4(3\sqrt{2} - 1)r}{3}$, and the throughput capacity is $\frac{W}{n}$ which is upper bound of throughput capacity.

VII. THROUGHPUT CAPACITY UNDER EXISTING MOBILITY-ASSISTED DATA COLLECTION MODELS

In the previous sections, we derived the upper bound of throughput capacity for the case of a mobile sink and the case of one static sink and multiple mobile relays. The upper bound of throughput capacity can be achieved when each pair of RPs are exactly $6r$ apart and all 1-hop nodes of the RP in each cluster are cache nodes. In this section, we analyze
the throughput capacity under existing mobility-assisted data collection models. For the existing data collection models, we can classify them according to the number of clusters and cache nodes used in the data collection models and derive the achievable throughput capacity under each category.

A. One mobile sink with \( n \) clusters

The mobility-assisted data collection model proposed in [22] assumes that a mobile sink visits every node in the network. Thus, every node sends data to the mobile sink only when the mobile sink visits it. The data collection model is equivalent to having \( n \) clusters where there is only one node in each cluster. Under such an approach, the mobile sink can collect data from each sensor node for \( \frac{T-n}{n} \) seconds in each \( T \) period. Thus, the throughput capacity can be derived as

\[
U(n, v, l) = \frac{W(T - \frac{l}{v})}{nT} \tag{18}
\]

where \( l = 0.7R\sqrt{\pi n t} \) as shown in [22] and the traveling speed \( v \) must be faster than \( 0.7R\sqrt{\pi n t} \) which is hard to satisfy in a large scale WSN.

B. One mobile sink without any cache node

The mobility-assisted data collection model proposed in [9] assumes that the mobile sink travels to \( k \) RPs to collect data where there is no cache node. Under such an approach, the network is clustered into \( k \) clusters where each cluster is centered at a RP. When the mobile sink travels to a RP, it serves as the cluster head and all nodes in the cluster starts to transmit data to the sink. The time that the mobile sink can spend on collecting data in each cluster will be \( \frac{T-k}{k} \). During the time that the mobile sink collects data, the per-node throughput capacity in a cluster [1] is \( \frac{W}{4nT} \). Thus, throughput capacity can be derived as

\[
U(k, v, l) = \frac{W}{4nT}k\frac{T-n}{T} < \frac{W}{4n} \tag{19}
\]

which is less than the per-node throughput capacity in a WSN with a static sink.

C. One mobile sink with one cache node in each cluster

The mobility-assisted data collection model proposed in [2] assumes that the network is divided into \( k \) clusters where there is a cache node in each cluster. The mobile sink collects data from those cache nodes periodically. According to [1], the throughput capacity in each cluster is \( \frac{W}{4T} \), where \( N \) is the number of nodes in each cluster. We first show that under such a model, when \( k \) is larger than 5, the network cannot achieve the maximum per-node throughput capacity even when the mobile sink travels along the shortest trail between any two RPs.

**Lemma 5**: When \( k \geq 5 \), \( U_{\text{max}}(k, v) \leq U_{\text{max}}(5, v) \).

**Proof**: When \( k \geq 5 \), we have \( \frac{k-1}{vT} \geq 1 \).

\[
U_2(k, v) = \frac{W}{nT}(1 - \frac{l}{vT} - \frac{k\Delta}{T}) \leq \frac{W}{nT}(1 - \frac{l}{vT} - \frac{5\Delta}{T}) \tag{20}
\]

On the other hand,

\[
U_2(k, v) = \frac{W}{nT}(1 - \frac{l}{vT} - \frac{k\Delta}{T}) \leq \frac{W}{nT}(1 - \frac{6kr}{vT} - \frac{k\Delta}{T}) \tag{21}
\]

Thus, \( U_2(k, v) \) decreases with the increase of \( k \). We have \( U_2(k, v) \leq U_2(5, v) \) and \( U_{\text{max}}(k, v) = \max U_2(k, v) < \max U_2(5, v) = U_{\text{max}}(5, v) \) when \( k \geq 5 \). Therefore, the lemma holds.

With Lemma 5, in order to derive the maximum throughput capacity under such a data collection model, we need to compute the maximum throughput capacity for the cases of \( k = 2, 3, 4, 5 \) respectively. From Lemma 5, when \( k = 5 \) and \( v > \frac{30r}{l} \), \( U_{\text{max}}(5, v) = \frac{W}{nT}(T - 30r) \) with \( l = 30r \). We now discuss the cases of \( k = 2, 3, 4 \).

**Lemma 6**: When \( k = 2 \),

\[
U_{\text{max}}(2, v) = \begin{cases} \frac{W}{nT}(1 - \frac{12r}{vT}), & v \in (\frac{12r}{v}, \frac{20r}{v}) \\ \frac{3W}{5n}, & v \in [\frac{20r}{v}, \infty) \end{cases} \tag{22}
\]

**Proof**: Since \( U_1(2, v) \) increases with the increase of \( \frac{1}{v} \) (\( l = 12r \)) and \( U_2(2, v) \) decreases with the increase of \( \frac{1}{v} \). Then \( U_{\text{max}}(2, v) \) can be achieved when \( U_1(2, v) = U_2(2, v) \), that is:

\[
\frac{W}{4nT}(T + \frac{l}{v} + k\Delta) = \frac{W}{nT}(T - \frac{l}{v} - k\Delta) \tag{23}
\]

which implies that \( U_{\text{max}}(2, v) \) can achieve the maximum capacity as long as \( \frac{l}{v} + k\Delta = \frac{2k\Delta - l}{v} \).

- When \( v \in (\frac{12r}{v}, \frac{20r}{v}) \), \( \frac{l}{v} > \frac{12r}{v} \). Then, \( U_2(2, v) < U_1(2, v) \) and \( U_{\text{max}}(2, v) = \max U_2(2, v) = \frac{W}{nT}(T - \frac{12r}{v}) \).

- When \( v \in [\frac{20r}{v}, \infty) \), given \( v \), to achieve \( U_1(2, v) = U_2(2, v) \), we let \( \Delta = \frac{l}{v} - \frac{12r}{v} \), therefore, \( U_{\text{max}}(2, v) = \max U_1(2, v) = \max U_2(2, v) = \frac{3W}{5n} \).

Thus, the lemma holds.

**Lemma 7**: When \( k = 3 \),

\[
U_{\text{max}}(3, v) = \begin{cases} \frac{W}{nT}(1 - \frac{18r}{v}), & v \in (\frac{18r}{v}, \frac{45r}{v}) \\ \frac{3W}{5n}, & v \in [\frac{45r}{v}, \infty) \end{cases} \tag{24}
\]

**Proof**: Similar to Lemma 6, \( U_{\text{max}}(3, v) \) can be achieved when \( U_1(3, v) = U_2(3, v) \), that is:

\[
\frac{W}{4nT}((k-1) + \frac{l}{vT} + k\Delta) = \frac{W}{nT}(1 - \frac{l}{vT} - k\Delta) \tag{25}
\]

which implies that \( U_{\text{max}}(3, v) \) can achieve the maximum capacity as long as \( \frac{l}{v} + k\Delta = \frac{5\Delta - l}{v} \). The lemma can
be proved similarly to Lemma 6. Due to the space limit, we omit the proof.

Lemma 8: When \( k = 4 \),

\[
U_{\text{max}}(4, v) = \left\{ \begin{array}{ll}
\frac{W}{n} (1 - \frac{24r}{vT}), & v \in (\frac{24r}{vT}, \frac{120r}{vT}) \\
\frac{4W}{n}, & v \in [\frac{120r}{vT}, \infty) \end{array} \right.
\]  

(26)

Proof. Similar to Lemma 6, \( U_{\text{max}}(4, v) \) can be achieved when \( U_1(4, v) = U_2(4, v) \), that is:

\[
W \left( (k - 1) + \frac{l}{vT} + \frac{k\Delta}{T} \right) = \frac{W}{n} \left( 1 - \frac{l}{vT} - \frac{k\Delta}{T} \right)
\]  

(27)

which implies that \( U_{\text{max}}(4, v) \) can achieve the maximum capacity as long as \( \frac{l}{v} + k\Delta = \frac{24r}{vT} \). The lemma can be proved similarly to Lemma 6. Due to the space limit, we omit the proof.

Theorem 2: \( U_{\text{max}}(v) = \max_{2k \leq k \leq 5} \{U_{\text{max}}(k, v)\} \).

Proof. The theorem follows directly from Lemma 5, Lemma 6 and Lemma 7, Lemma 8.

We summarize the relationship between \( v, U_{\text{max}}(v), k, l \) in Table III.

<table>
<thead>
<tr>
<th>( v )</th>
<th>( U )</th>
<th>( k )</th>
<th>( \Delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{W}{n} (1 - \frac{24r}{vT}) )</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \frac{W}{n} (1 - \frac{120r}{vT}) )</td>
<td>2</td>
<td>( \frac{24r}{vT} )</td>
<td>0</td>
</tr>
<tr>
<td>( \frac{4W}{n} )</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \frac{W}{n} (1 - \frac{120r}{vT}) )</td>
<td>3</td>
<td>( \frac{24r}{vT} )</td>
<td>( \frac{120r}{vT} )</td>
</tr>
<tr>
<td>( \frac{W}{n} (1 - \frac{24r}{vT}) )</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \frac{W}{n} (1 - \frac{120r}{vT}) )</td>
<td>4</td>
<td>( \frac{24r}{vT} )</td>
<td>( \frac{120r}{vT} )</td>
</tr>
<tr>
<td>( \frac{W}{n} (1 - \frac{24r}{vT}) )</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE III

MAXIMUM THROUGHPUT CAPACITY WITH ONLY ONE CACHE IN EACH CLUSTER FOR A GIVEN \( v \) AND THE OPTIMAL NUMBER OF CLUSTERS TO ACHIEVE THE MAXIMUM THROUGHPUT CAPACITY

D. One static sink and multiple mobile relays without cache nodes

The mobility-assisted data collection model proposed in [17] assumes that there is a static sink and \( k \) mobile relays. \( k \) mobile relays collect data along \( k \) rings. They then form a multi-hop communication to relay data to the static sink. Since the mobile relay which is closest to the static sink will be the bottleneck, the per-node throughput capacity is:

\[
U \leq \frac{k}{2k - 1} \frac{W}{n}
\]  

(28)

which is less than the upper bound of throughput capacity \( \frac{W}{n} \).

We summarize the throughput capacity under different mobility-assisted data collection models in Table IV where for the case of one mobile sink, given \( v \), the upper bound of throughput capacity is listed in the columns with all one-hop nodes of the RP at each cluster are cache nodes. It can be verified that for a given \( v \), the throughput capacity in such columns is larger than that in existing mobility-assisted data collection model.

<table>
<thead>
<tr>
<th>Mobility</th>
<th>Clusters &amp; Caches</th>
<th>( v )</th>
<th>( U )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static sink</td>
<td>no clustering, no cache</td>
<td>0</td>
<td>( \frac{W}{n} )</td>
</tr>
<tr>
<td>Mobile sink</td>
<td>( k ) clusters, no cache</td>
<td>( &gt; 0 )</td>
<td>( &lt; \frac{W}{n} )</td>
</tr>
<tr>
<td>Mobile sink</td>
<td>( k ) clusters and one cache node at each cluster</td>
<td>( \frac{W}{n} (1 - \frac{24r}{vT}) )</td>
<td>( \frac{W}{n} (1 - \frac{120r}{vT}) )</td>
</tr>
<tr>
<td>Mobile sink</td>
<td>( k ) clusters and all one-hop nodes of the RP at each cluster are cache nodes</td>
<td>( \frac{W}{n} (1 - \frac{24r}{vT}) )</td>
<td>( \frac{W}{n} (1 - \frac{120r}{vT}) )</td>
</tr>
<tr>
<td>Mobile sink</td>
<td>( n ) clusters, no cache nodes</td>
<td>( O\left( \frac{nT}{r} \right) )</td>
<td>( \frac{W}{n} )</td>
</tr>
<tr>
<td>Mobile relay</td>
<td>( k ) clusters and no cache nodes</td>
<td>( O\left( \frac{Wk}{vT} \right) )</td>
<td>( \frac{W}{n} )</td>
</tr>
<tr>
<td>Mobile relay</td>
<td>( 4 ) clusters, all one-hop nodes of the RP at each cluster are cache nodes</td>
<td>( \frac{W}{n} )</td>
<td>( \frac{W}{n} )</td>
</tr>
</tbody>
</table>

TABLE IV

THROUGHPUT CAPACITY UNDER DIFFERENT MOBILITY-ASSISTED DATA COLLECTION MODELS

VIII. SIMULATION RESULTS

The upper bound of throughput capacity is derived under a proposed data collection model where all 1-hop nodes of each RP are cache nodes and the distance between any two RPs are exactly \( 6r \). In this section, we conduct simulation under such a data collection model to verify how close the achievable throughput capacity is to our derived upper bound of throughput capacity.

In our simulation, 3000 source nodes are uniformly deployed in a circular sensing field with radius of \( R = 20m \). The sensing range \( r \) is set to be \( 1m \). The delay deadline is set to be \( T = 2 \) minutes. The traveling speed \( v \) varies in the range of \([0.1m/s, 1m/s]\). The radio parameters are set according to the data sheet of the CC1000 radio on Mica2 motes [23]. Radio bandwidth is 40 Kbps and transmission power is 4 dbm with the current consumption of 11.6 mA. The size of each packet is 30 bytes.

Given \( v \) and \( T \), we select \( k \) RPs and determine the traveling distance \( l \) according to our analysis in Section 4. Load balanced routing protocol is applied in each cluster to determine a cache node for each source node. In the first \( \frac{k-1}{k}T + \frac{l}{k} \) seconds of each \( T \) time period, all source nodes will keep transmitting packets to their correspondent cache nodes. Let \( D_i \) be the amount of data that the mobile sink can collect from cluster \( i \). We calculate the per-node throughput obtained in the simulation as:

\[
U_s = \frac{\sum_{i=1}^{k} D_i}{Tn}
\]
Let $U_t$ be the derived throughput capacity. We are interested in how close $U_s$ is to $U_t$. Given $v$, we are also interested in the impact of $k$ on the throughput as well as the impact of the buffer size at each cache node on the throughput.

Let $U_0$ be the throughput per-node capacity in a static WSN. Figure 4 shows that the per-node throughput $U_s$, the theoretical throughput capacity $U_t$ and $U_0$ when $v$ changes in the range of $[0.1m/s, 1m/s]$. We can see that cache-based mobility-assisted data collection can significantly improve the throughput compared with that in a static WSN. We can also see that the per-node throughput is very close to the theoretical value we derived in Section 4.

In our proposed cache-based mobility-assisted data collection model, there are multiple cache nodes in each cluster where there is only one cache-node in each cluster under the mobility-assisted data collection model proposed in [2]. We show the advantages of using multiple cache nodes in each cluster in Fig. 5. We can see that the throughput capacity in our proposed data collection model is consistently better than that in the mobility-assisted data collection model proposed in [2].

Figure 6 depicts the throughput capacity under various $k$ with the change of $v$. We can see that when $v \in [0.1m/s, 0.3m/s]$, the throughput for $k = 2$ excels the throughput achieved for other values of $k$, when $v \in [0.3m/s, 0.8m/s]$, the throughput for $k = 3$ is the best, and when $v \in [0.8m/s, 1m/s]$, the throughput for $k = 4$ is the best. The results in Fig. 6 are consistent with our analysis. From Fig. 6, we can also see that the throughput for $k = 4$ is consistently better than that for $k = 5$.

Figure 7 shows the minimum buffer size requirement to achieve the maximum throughput. Let $N_i$ be the number of cache nodes in cluster $i$. Then the necessary buffer size will be $M_s = \frac{1}{k} \sum_{i=1}^{k} \frac{D_i}{N_i}$. As we can see from Fig. 7, the buffer size is very close to our derived minimum buffer size to achieve the maximum throughput capacity under different $v$. The maximum buffer size is less than 50k bytes. Mica2 motes have 128k flash memory which is enough to satisfy our memory requirement on cache nodes.

Overall, the simulation results validate our analysis on the optimal $k$ and minimum buffer size to achieve the maximum throughput capacity. The simulation results also demonstrate that the derived throughput capacity is achievable as the throughput obtained through the simulation is very close to...
our theoretical value.

IX. CONCLUSION

In this paper, we systematically studied the throughput capacity of mobility-assisted data collection in WSNs. Given the traveling speed of the mobile sink and the required delay deadline, we derived the optimal number of clusters, traveling distance, and memory size at each cache node in order to achieve the upper bound of throughput capacity. We also proved that 4 mobile relays with moderate traveling speed are enough to achieve the upper bound of throughput capacity \( O\left(\frac{\sqrt{n}}{n}\right) \) in a WSN. Besides deriving the upper bound of throughput capacity in mobility-assisted WSN, we also propose a mobility-assisted data collection model where each pair of RPs are within the minimum distance and all one-hop neighbors of the RP in each cluster are cache nodes. Such a mobility-assisted data collection model can achieve the derived upper bound of throughput capacity. At last, we analyzed the maximum throughput capacity of existing mobility-assisted data collection model. Our results provide important insights into the design of mobility-assisted sensor networks under different data collection models.

ACKNOWLEDGEMENT

This work is partially supported by the grant from the Research Grants Council of the Hong Kong Special Administrative Region, China under Project CityU 121107.

REFERENCES


