# Impact of In-Network Aggregation on Target Tracking Quality Under Network Delays

Wei Wei, Ting He, Chatschik Bisdikian, Dennis Goeckel, Bo Jiang, Lance Kaplan, and Don Towsley

***Abstract*—In this paper, we investigate how in-network aggregation approach impacts the target tracking quality in multi-hop wireless sensor networks under network delays. Specifically, we use the mean squared error (MSE) of the target location estimate to quantify the target tracking quality, and investigate how in-network aggregation affects the MSE. To obtain insights without being obscured by onerous mathematical details, we assume a Brownian motion mobility model for the target, Gaussian measurement noise for the sensors, and independent per-hop delays. Under the above assumptions, we first propose an aggregation scheme that preserves a sufficient statistic for optimal tracking under data aggregation at the intermediate nodes and arbitrary network delays. We then analytically study the impact of aggregation in three increasingly more complicated scenarios: single task tracking with only transmission delay, single task tracking with both transmission delay and queueing delay at intermediate nodes, and multi-task tracking. Our results demonstrate that in-network aggregation improves tracking quality in all three scenarios. Furthermore, our analysis provides guidelines on how to choose aggregation parameters in practice.**

***Index Terms*—In-network aggregation, Target tracking, Brownian motion, and Mean squared error.**

## I. Introduction

IN-NETWORK aggregation, i.e., aggregating packets at intermediate nodes enroute to the sink(s), has been proposed to reduce resource consumption in wireless sensor networks [1]. The main idea is that aggregating packets inside network through local computation (e.g., fusing sensor readings related to the same event) reduces the amount of data to be transmitted inside the network, and hence reduces energy consumption and network bandwidth usage. Furthermore, from the data sink's point of view, in-network aggregation reduces the amount of data to be processed, and makes further data processing more efficient [2]. Various issues related to in-network aggregation have been studied in the literature, including how to design routing protocols to provide efficient in-network aggregation [3], [4], [5], [6], [7], [8], [9], how to design aggregation functions [10], [11], [12], and how to represent data [13], [14], [15].

In-network aggregation can lead to various tradeoffs in resource efficiency, accuracy, timeliness and granularity of the data [1]. In this paper, we study the impact of in-network aggregation on target tracking quality. Existing work on target tracking mostly ignores the impact of network dynamics, especially the random delays due to propagation and queueing, on the efficiency of in-network aggregation. While it is known that appropriate aggregation schemes can achieve optimal tracking accuracy (e.g., by computing the sufficient statistic) without network delay, it remains open whether the same holds when there are substantial network delays: aggregation induces extra delays at intermediate nodes (to wait for inputs from upstream nodes), and the delays translate into additional tracking errors due to target movement. Intuitively, one may expect this phenomenon to be more prominent when delays exhibit a higher level of randomness.

In this paper, we investigate the impact of in-network aggregation on target tracking quality in the presence of random network delays due to data transmission and queuing. Specifically, we aim to answer the following questions: *Can optimal target location estimates be achieved in the presence of aggregation inside the network? Does aggregation improve or reduce target tracking quality in the presence of (possibly random) network delays, and how much is the improvement or reduction?*

To answer the above questions, we quantify the target tracking quality through the mean squared error (MSE), and investigate how in-network aggregation affects the MSE. To obtain insights without being obscured by onerous mathematical details, we consider a simple model where target movement follows 1-D Brownian motion and sensor measurement noise follows a Gaussian distribution, and analyze the impact of in-network aggregation on target tracking quality in the case that network resources are used for a single tracking task, and where multiple tasks share the network resources. Our main contributions are as follows.

- In one-shot single task tracking, we prove that for general tree topologies and general network delay distributions, the aggregation approach always improves the tracking quality.
- In periodic single task tracking where measurement nodes track a target periodically, we investigate the asymptotic behavior of MSE, and show that aggregation can significantly outperform non-aggregation.
- In multiple task tracking, we show that the aggregation approach leads to superior tracking quality relative to not using aggregation provided that aggregation parameters are carefully chosen. Our analysis also provides guidelines on choosing the optimal parameter in practice.

In related work, the studies of [16] and [17] evaluate energy consumption and tracking accuracy of several data

Manuscript received March 1, 2012; revised December 29, 2012.

W. Wei, B. Jiang, and D. Towsley are with Department of Computer Science, University of Massachusetts, Amherst, MA 01003 (e-mail: weiwei@cs.umass.edu).

T. He and C. Bisdikian are with IBM T. J. Watson Research Center, 19 Skyline Drive, Hawthorne, NY 10532, USA.

D. Goeckel is with the Electrical and Computer Engineering Department, University of Massachusetts, Amherst, MA 01003.

L. Kaplan is with the Army Research Laboratory, Adelphi, MD 20783.

Digital Object Identifier 10.1109/JSAC.2013.130416.

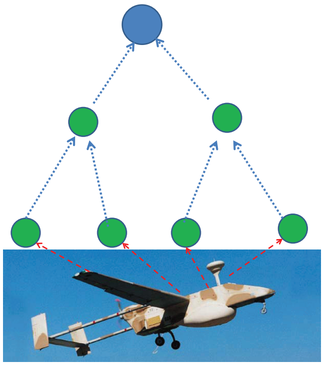


Fig. 1. Illustration of target tracking based on measurements from a sensor network: the measurements are taken simultaneously at time 0 by the sensors that detect the target, and are forwarded to the sink node along a routing tree.

aggregation schemes for single and multiple target tracking, respectively. The evaluation assumes ideal target detection and communication links. Contrary to these studies, we develop an analytical framework to study the impact of data aggregation on target tracking quality in the presence of network delays and measurement noise. Several studies consider target localization and tracking in sensor networks [18], [19], [20], [21], [22]. They, however, do not consider data aggregation. Several studies provide theoretical treatment on in-network aggregation: Giridhar and Kumar [23] provide a theoretical framework for in-network computation, Balister et al. [24] designs policies for in-network function computation with minimum energy consumption subject to a latency constraint; Banerjee et al. [25] develop joint aggregation, routing and scheduling algorithms. These studies differ in scope from ours. To the best of knowledge, our work is the first that accounts for random network delays in studying the impact of data aggregation on target tracking quality in sensor networks. Finally, the general topic of quality of information (of which quality of tracking estimation is a subcase) has been presented in [26].

The remainder of the paper is organized as follows. Section II presents the problem setting. Section III analyzes the impact of data aggregation on target tracking quality assuming a single tracking task. Sections IV considers periodic single task tracking. Section V considers the scenario with multiple tracking tasks. Finally Section VI concludes the paper and presents future work.

## II. Problem Setting

Consider a target tracking application using a sensor network, as illustrated in Fig. 1. The sensor network contains a set of sensor nodes deployed in a field. A mobile target moves inside the field. At a certain point in time, the group of sensors close to the target measure the location of the target and send the measurements to the sink through a network whose topology is a tree. We assume all intermediate nodes in the tree aggregate or do not aggregate the packets from their child nodes. The former is referred to as *in-network aggregation* (or simply *aggregation*), and the latter is referred to as *non-aggregation*. After receiving the measurements, the sink estimates the location of the moving target using the measurements.

Due to packet delays in the sensor network, the time when the sink uses the measurements to infer the target location differs from the time when the measurements are taken. In addition, the sensor measurements have errors. These factors may affect the quality of the target tracking. The goal of this paper is to quantitatively and qualitatively investigate the impact of aggregation on target tracking quality.

### *A. Assumptions*

For ease of exposition, we assume the target moves along a straight line (extending this mobility model to 2-D or 3-D scenarios is straightforward). For analytical tractability, we assume the target moves according to Brownian motion. As we shall see, this mobility model allows us to obtain closed-form results of the target location estimates, which provide valuable insights on the impact of network and sensor measurement qualities, and in-network aggregation on tracking quality. Specifically, let $\theta(t)$ denote the target location at time $t$. The location at time 0, $\theta(0) = \theta_0$, is unknown. Under the Brownian motion model, for $t_2 > t_1$, $\theta(t_2) - \theta(t_1)$ follows a Gaussian distribution $N(0, c(t_2 - t_1))$, where $c$ is a positive real number, which scales the volatility of the Brownian motion, and is assumed to be known.

We assume the sensor clocks are synchronized so that they take measurements simultaneously. Without loss of generality, we assume each of $n$ sensors makes a measurement of the target location at time 0. For each sensor, the measurement error follows a Gaussian distribution with zero mean, and the measurement errors from different sensors are independent. Specifically, let $z_i$ denote the measurement of sensor $i$, $i = 1 \ldots, n$. We assume $z_i$ follows a Gaussian distribution with mean $\theta_0$ and variance $\sigma_i^2$, where $\theta_0$ is unknown and $\sigma_i$'s are known. Last, we assume a measurement is encapsulated in a packet (hence we use packets and measurements interchangeably in the rest of the paper).

Sensor measurements are sent to the sink via the intermediate nodes along a routing tree (the sink is the root of the tree, and the $n$ sensors are the leaves of the tree). For simplicity, we assume per-hop delays along the routing tree are independent random variables and there is no packet loss; considering correlated per-hop delays and packet losses is left as future work.

### *B. Target scenarios*

We consider both single task and multiple task tracking. In single task tracking, all the measurements are about a single tracking task, and the network resource is used only by this tracking task. In multiple task tracking, the network resource is shared by multiple tasks.

*1) Single task tracking:* Within this scenario, we consider both one-shot tracking where a target is tracked only once, and periodic tracking where a target is tracked periodically in multiple rounds. In either case, when using the aggregation approach, all intermediate nodes perform aggregation. More specifically, an intermediate node waits for the packets from all of its children to arrive, includes all measurements into a single packet, and then forwards the combined measurement data to its parent. When using the non-aggregation approach, none of the intermediate nodes perform aggregation, i.e., each intermediate node simply forwards a packet to its parent immediately after receiving the packet.

We next briefly describe target location estimation for one-shot tracking; target estimation for periodic tracking is deferred to Section IV. Let random variable $T$ denote the time by which all the measurements reach the sink (i.e., $T$ is the time when the last measurement reaches the sink since measurements are made at time 0). Note that, since intermediate nodes perform different operations in the aggregation and non-aggregation schemes, $T$ may take different values under these two schemes even under the same network conditions. For simplicity, we assume the time required to aggregate measurements is negligible compared to the network delays; considering non-negligible aggregation delays is left as future work.

When the sink receives all the packets from its children at time $T$, it estimates the location of the target at time $T$, i.e., it estimates $\theta(T)$. Based on the sensor measurements (taken at time 0), the minimum variance unbiased estimator (MVUE) for location estimate of $\theta_0$ is

$$\frac{\sum_{i=1}^{n} z_i/\sigma_i^2}{\sum_{i=1}^{n} 1/\sigma_i^2},$$

with an MSE of

$$\frac{1}{\sum_{i=1}^{n} 1/\sigma_i^2}.$$

By the property of Brownian motion, the conditional probability distribution function $f(\theta(T) \mid \theta(0), T)$ is Gaussian. Therefore, the MVUE for $\theta(T)$ is

$$\frac{\sum_{i=1}^{n} z_i/\sigma_i^2}{\sum_{i=1}^{n} 1/\sigma_i^2}, \tag{1}$$

with an MSE of

$$\frac{1}{\sum_{i=1}^{n} 1/\sigma_i^2} + cT, \tag{2}$$

where $c > 0$ is the Brownian motion volatility parameter.

Since $T$ is a random variable, the expectation of the MSE in (2) is

$$\frac{1}{\sum_{i=1}^{n} 1/\sigma_i^2} + cE[T]. \tag{3}$$

Throughout the paper, we use the MSE of the target location estimate (3) to characterize target tracking quality.

*2) Multiple task tracking:* In multiple task tracking, the measurements (packets) for the multiple tracking tasks are sent through a routing tree, jointly using the network resources. When using aggregation, an intermediate node aggregates the measurements from the same tracking task (e.g., the multiple tracking tasks are differentiated with IDs). The sink obtains the target location estimate for each tracking task, based on the measurements for that task. We describe the assumptions under multiple tracking tasks in more detail in Section V.

### C. Aggregation

Consider the following aggregation scheme that specifies the operations of the leaf and intermediate nodes. A leaf node $j$ sends the pair $(z_j, \sigma_j^2)$ to its parent, where $z_j$ is its measurement of the target location and $\sigma_j^2$ is the variance of the measurement. An intermediate node $i$ sends $(z_i, \sigma_i^2)$ to its parent, where $z_i$ and $\sigma_i^2$ are calculated as

$$z_i = \frac{\sum_{j\in C(i)} z_j/\sigma_j^2}{\sum_{j\in C(i)} 1/\sigma_j^2} \tag{4}$$

$$\sigma_i^2 = \frac{1}{\sum_{j\in C(i)} 1/\sigma_j^2}, \tag{5}$$

where $C(i)$ is the set of the children of node $i$.

Assuming Brownian motion mobility model for the target and Gaussian measurement noise for the sensors, the above aggregation scheme preserves a sufficient statistic for the root to make an optimal estimate (in the sense of MVUE) of $\theta(T)$, i.e., the target location at time $T$. The above statement is established as follows.

First, we show that $\sum_{i=1}^{n} z_i/\sigma_i^2$ is a sufficient statistic for estimating $\theta(T)$ based on $\mathbf{z} = (z_i)_{i=1}^n$. Given $\theta(T)$, each sensor measurement $z_i$ can be written as $z_i = \theta(T) - \Delta + W_i$, where $\Delta := \theta(T) - \theta(0)$ is the change in target location from time 0 to time $T$, and $W_i$ the measurement noise at sensor $i$. Under our model, we know that $\mathbf{z}$ given $\theta(T)$ is Gaussian distributed, with mean $\mu = \mathbf{1}\theta(T)$ and covariance matrix $\Sigma = cT \cdot \mathbf{1} + \mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)$, where $cT \cdot \mathbf{1}$ denotes an $n \times n$ matrix with all elements equal to $cT$, and $\mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ a diagonal matrix with diagonal elements $\sigma_1^2, \ldots, \sigma_n^2$. Therefore, the likelihood function of $\theta(T)$ is

$$p(\mathbf{z}|\theta(T)) = c_1 \exp\left(-\frac{1}{2}(\mathbf{z}-\mu)^T \Sigma^{-1} (\mathbf{z}-\mu)\right), \tag{6}$$

where $c_1$ is a constant independent of $\mathbf{z}$ and $\theta(T)$. Rewriting (6) shows that a transformation of the likelihood function $c_2(\log p(\mathbf{z}|\theta(T)) - \log c_1)$ is in the form of

$$\begin{aligned} & f_1(\mathbf{z}) + f_2(\theta(T)) \\ & + \theta(T)\left(\sum_{i\neq j} \frac{z_i + z_j}{\sigma_i^2\sigma_j^2} - 2\sum_i \frac{z_i}{\sigma_i^2}(1 + \sum_{j\neq i}\frac{1}{\sigma_j^2})\right) \\ & = f_1(\mathbf{z}) + f_2(\theta(T)) - 2\theta(T)\sum_i \frac{z_i}{\sigma_i^2}, \end{aligned} \tag{7}$$

where $c_2$ is also a constant (i.e., independent of $\mathbf{z}$ and $\theta(T)$), $f_1(\mathbf{z})$ a function independent of $\theta(T)$, and $f_2(\theta(T))$ a function independent of $\mathbf{z}$. By the Fisher-Neyman factorization theorem, the sufficient statistic of $\theta(T)$ is $\sum_{i=1}^{n} z_i/\sigma_i^2$.

Next, we show that the proposed aggregation scheme (4)–(5) preserves this sufficient statistic. We prove this through induction on the depth of the routing tree. Let $l$ denote the depth of the tree. When $l = 1$, the statement is true based on the definition of the aggregation scheme. We now show that if the statement is true when the depth of the tree is at most $l$, then it is also true when the depth of the tree is $l + 1$. For simplicity, we first consider the case where the root has two children: one aggregating measurements from a tree of depth at most $l$ with $n - m$ sensors, and the other from a tree of depth at most $l$ with $m$ sensors ($m > 1$). Let $(z_a, \sigma_a^2)$ be the aggregation result from the $n-m$ sensors, and $(z_b, \sigma_b^2)$ be the aggregation result from the $m$ sensors. Then

$$z_a = \frac{\sum_{i=1}^{n-m} \frac{z_i}{\sigma_i^2}}{\sum_{i=1}^{n-m} \frac{1}{\sigma_i^2}}, \tag{8}$$

$$\sigma_a^2 = \frac{1}{\sum_{i=1}^{n-m} \frac{1}{\sigma_i^2}}, \tag{9}$$

and

$$z_b = \frac{\sum_{i=n-m+1}^{n} \frac{z_i}{\sigma_i^2}}{\sum_{i=n-m+1}^{n} \frac{1}{\sigma_i^2}}, \tag{10}$$

$$\sigma_b^2 = \frac{1}{\sum_{i=n-m+1}^{n} \frac{1}{\sigma_i^2}}. \tag{11}$$

Hence, the final aggregation results at the root $(z_r, \sigma_r^2)$ can be calculated as

$$z_r = \frac{z_a\sigma_b^2 + z_b\sigma_a^2}{\sigma_a^2 + \sigma_b^2}, \tag{12}$$

$$\sigma_r^2 = \frac{\sigma_a^2\sigma_b^2}{\sigma_a^2 + \sigma_b^2}. \tag{13}$$

Substituting (8) to (11) into (12) and (13), we have

$$z_r = \frac{\sum_{i=1}^{n} \frac{z_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}, \tag{14}$$

$$\sigma_r^2 = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}, \tag{15}$$

which preserves the sufficient statistic. In fact, $z_r$ already gives the MVUE estimate of $\theta(T)$. When the root has more than two children, by following a similar procedure as above, we can prove that the aggregation from all of the children provides the same results as in (14) and (15).

Recall that target tracking quality is represented by the MSE of the target location estimate in (3). Since the above aggregation scheme preserves optimality in the sense of MVUE, when using this scheme, the first term in (3) is the same as that when not using aggregation. Therefore, the target tracking quality when using and not using aggregation only differs in the second term of (3), i.e., $E[T]$.

## III. One-Shot Single Task Tracking

In this section, we consider one-shot single task tracking in a sensor network. As we showed in Section II-C, the target tracking qualities when using the aggregation and non-aggregation approaches only differ in $E(T)$, i.e., the expected delay for the measurements to reach the sink. To provide some intuition, we first obtain $E(T)$ using these two approaches for several examples. We then present the main results, and results when incorporating queuing delays.

### A. Motivating examples

The examples are based on a single tracking task and exponential delay distributions; a more formal treatment in more generalized settings is deferred to later sections.

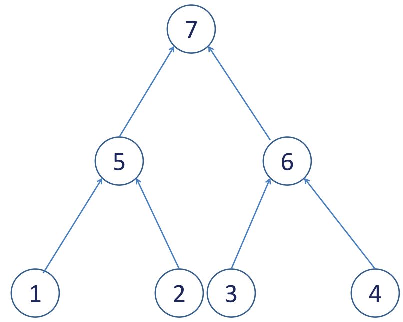



Fig. 2. A binary tree example to illustrate the impact of aggregation on target tracking quality.

As described in Section II, when using aggregation, an intermediate node needs to wait for the packets from all its children to aggregate them together. Therefore, one may intuitively think that aggregation leads to larger delays for the measurements to reach the sink, and hence an inferior target tracking quality. As we shall see, perhaps surprisingly, this is not the case. In the following, we first describe results for a simple three-level binary-tree example (i.e., depth of 3) with seven nodes, and then describe large-scale simulation results in trees with more nodes and larger depths. Theorem 1 in Section III-B then provides a formal statement that generalizes such behavior to many scenarios of interest.

*1) A binary tree example:* In this example, we assume that four sensors, $1, 2, 3,$ and $4$, measure the target location simultaneously at time $0$, as shown in Fig. 2. The measurements from sensors $1$ to $4$ are sent to the sink (node $7$) via intermediate nodes $5$ and $6$. For sensor $i$, the measurement error follows a Gaussian distribution $N(0, \sigma_i^2)$, $i = 1, 2, 3, 4$. Let $X_i$ be the delay on the link from node $i$ to its parent. In this example, for simplicity, we assume the delays for the links are i.i.d. and each follows an exponential distribution with mean $1/\eta$. Let $f_X(x)$ and $F_X(x)$ denote respectively the probability density distribution (PDF) and cumulative distribution function (CDF) of the link delay.

**Aggregation case.** When using aggregation, the measurements from nodes $1$ and $2$ are aggregated at node $5$ on their route to the sink. Similarly, the measurements from nodes $3$ and $4$ are aggregated at node $6$. Let $T_i$ be the cumulative delay to reach node $i$ from leaf nodes. We have

$$T_5 = \max\{X_1, X_2\},\ T_6 = \max\{X_3, X_4\}.$$

The cumulative distribution function (CDF) of $T_5$ is

$$\begin{aligned} F_{T_5}(x) &= P(T_5 \le x) = P(\max\{X_1, X_2\} \le x) \\ &= P(X_1 \le x)P(X_2 \le x) \\ &= (P(X \le x))^2 \\ &= (1 - e^{-\eta x})^2 \end{aligned}$$

Similarly, we can calculate the CDF of $T_6$, which is the same as the CDF of $T_5$.

Next, we calculate the CDF of $T_7$, i.e., the delay to reach the sink (node 7).

$$\begin{aligned}
F_{T_7}(x) &= P(T_7 \le x) \\
&= P(\max\{X_5 + T_5, X_6 + T_6\} \le x) \\
&= (P(X_5 + T_5 \le x))^2 \\
&= \left(\int_0^x F_{T_5}(y) f_X(x-y)dy\right)^2 \\
&= \left(\int_0^x (1-e^{-\eta y})^2 \eta e^{-\eta(x-y)}dy\right)^2 \\
&= \left((e^{2\eta x} - 2x\eta e^{\eta x} - 1)e^{-2\eta x}\right)^2 \\
&= (1 - 2\eta x e^{-\eta x} - e^{-2\eta x})^2
\end{aligned}$$

Hence the expected cumulative delay to reach the sink is

$$\begin{aligned}
E_{T_7}(x) &= \int_0^\infty (1 - F_{T_7}(x))\, dx \\
&= \int_0^\infty \left(1 - (1 - 2\eta x e^{-\eta x} - e^{-2\eta x})^2\right) dx \\
&= \frac{119}{36}\frac{1}{\eta} \approx 3.31\frac{1}{\eta}.
\end{aligned}$$

That is, in this example, when using aggregation, the expected cumulative delay to reach the sink is approximately 3.31 times the mean one-hop delay.

**Non-aggregation case** In the absence of aggregation, each measurement travels two links to reach the sink. The sink infers the target location and calculates the MSE of the target location after receiving all the four measurements. We calculate the CDF of the delay, $T_7$, to reach the sink as

$$\begin{aligned}
F_{T_7}(x) &= P(T_7 \le x) \\
&= (P(X_1 + X_5 \le x))^4 \\
&= \left(\int_0^x F_{X_1}(y) f_{X_5}(x-y)dy\right)^4 \\
&= \left(\int_0^x (1-e^{-\eta y})\eta e^{-(x-y)}dy\right)^4 \\
&= (1 - e^{-\eta x} - \eta x e^{-\eta x})^4.
\end{aligned}$$

Hence

$$\begin{aligned}
E_{T_7}(x) &= \int_0^\infty (1 - F_{T_7}(x))\, dx \\
&= \int_0^\infty \left(1 - (1 - e^{-\eta x} - \eta x e^{-\eta x})^4\right) dx \\
&= \frac{12259}{3456}\frac{1}{\eta} \approx 3.55\frac{1}{\eta}.
\end{aligned}$$

Comparing the results for aggregation and non-aggregation cases, we observe the aggregation approach outperforms the non-aggregation approach. This is surprising since one would conjecture that the delay incurred at an intermediate node in order to aggregate all measurements from its children would slow down the aggregation approach, and hence lead to worse target tracking quality.

*2) Large-scale simulation results:* We next use large-scale simulation to further compare the cumulative delays to reach the sink when using and not using aggregation. Specifically, we construct binary trees of larger depths (with 3 to 6 levels). The individual link delays follow exponential distributions with mean 1. For each tree, we obtain the expected cumulative delays to reach the sink through $10,000,000$ simulation runs. Table I lists the results (the 95% confidence intervals are tight and hence are not presented). We observe that, for all topologies, the cumulative delays under the aggregation approach are shorter than those under the non-aggregation approach. We also explored non-binary tree topologies (in particular, the number of children can vary from 2 to 5), and obtained similar results. Next, we compare the performance of aggregation and non-aggregation approaches in a general tree topology assuming a general delay distribution, and formally prove that indeed aggregation leads to better target tracking qualities.

### B. Main Theorem

For simplicity, we assume that there is no queuing delay at the intermediate nodes in the non-aggregation case. That is, when multiple packets arrive at an intermediate node simultaneously, they can be forwarded to the parent independently without incurring any queuing delays. Section III-C relaxes this assumption and considers the scenario where a packet needs to be queued at an intermediate node when previous packets have not been transmitted by the intermediate node.

Specifically, consider a routing tree. Inside the tree, consider a parent node $P$ with $\alpha$ children, indexed $1, 2, \ldots, \alpha$. Let $X_i$ be the delay between node $i$ and the parent node $P$. When using aggregation, let $T_i$ be the cumulative delay to reach node $i$ from the leaves, $i = 1, \ldots, \alpha$. Then we have the following relationship

$$T_P = \max_{i=1,\ldots,\alpha}\{T_i + X_i\}. \tag{16}$$

For any given tree, the above stochastic relationship can be applied recursively to obtain the cumulative delay to reach the root. The CDF of the cumulative delay to reach the parent node, $P$, based on the CDF of the cumulative delay and the per-hop delay distribution of each child node can be calculated as

$$\begin{aligned}
F_P(x) &= \prod_{i=1}^{\alpha} F_{T_i + X_i}(x) \\
&= \prod_{i=1}^{\alpha} \int_0^x F_{T_i}(y) f_{X_i}(x-y)dy
\end{aligned} \tag{17}$$

Applying the above formula recursively, we can obtain the CDF of the cumulative delay to reach the root. Therefore, we can obtain the expected cumulative delay to reach the root when using aggregation.

The theorem below shows that under a general tree topology and a general per-hop delay distribution, the expected cumulative delay of the aggregation scheme is smaller than that under non-aggregation.

TABLE I
THE EXPECTED CUMULATED DELAYS FOR THE MEASUREMENTS TO REACH THE SINK FROM MONTE CARLO SIMULATIONS FOR BINARY TREES.

| | aggregation | non-aggregation |
|---|---|---|
| 7 sensors, 3 levels | 3.31 | 3.55 |
| 15 sensors, 4 levels | 5.31 | 5.84 |
| 31 sensors, 5 levels | 7.44 | 8.25 |
| 63 sensors, 6 levels | 9.67 | 10.73 |

**Theorem 1.** *Assuming per-hop delays are independent, the aggregation scheme generates a smaller expected cumulative delay than a non-aggregation approach.*

*Proof:* Let $n$ denote the number of sensors. Let $E_i$ denote the set of links on the path from the $i$-th sensor to the sink, and $E = \bigcup_{1 \le i \le n} E_i$ the set of all links in the routing tree.

In the non-aggregation scheme, for $e \in E_i$, let $X_{e,i}$ be the single-hop delay experienced on link $e$ by a packet originated from the $i$-th sensor. The total delay from the $i$-th sensor to the sink is $Y_i = \sum_{e \in E_i} X_{e,i}$. Note that $X_{e,i}$ and $X_{e',i'}$ are independent if $i \ne i'$ or $e \ne e'$, so the $Y_i$'s are also independent. The cumulative delay for all of the measurements to reach the sink $T = \max_{1 \le i \le n} Y_i$. For example, in Fig. 2, the total delay from sensor 1 to the sink is $Y_1 = X_{(1,5),1} + X_{(5,7),1}$ and $T = \max\{Y_1, Y_2, Y_3, Y_4\}$.

In the aggregation scheme, for $e \in E$, let $\hat{X}_e$ be the single-hop delay experienced by all *aggregated* packets through link $e$. The total delay from the $i$-th sensor to the sink is $\hat{Y}_i = \sum_{e \in E_i} \hat{X}_e$. Note that $\hat{X}_e$ and $\hat{X}_{e'}$ are independent if $e \ne e'$. Since $\hat{X}_e$ and $X_e$ are identically distributed for all $e \in E_i$, so are $\hat{Y}_i$ and $Y_i$. Furthermore, it can be shown that the cumulative delay for all of the measurements to reach the sink $\hat{T} = \max_{1 \le i \le n} \hat{Y}_i$, which is a consequence of the basic fact that, for a random variable $Z$, $\max_i\{X_i\} + Z = \max_i\{X_i + Z\}$. For example, in Fig. 2,

$$\begin{aligned}\hat{T} &= \max\{\max\{\hat{X}_{(1,5)}, \hat{X}_{(2,5)}\} + \hat{X}_{(5,7)}, \\ &\qquad \max\{\hat{X}_{(3,6)}, \hat{X}_{(4,6)}\} + \hat{X}_{(6,7)}\} \\ &= \max\{\hat{X}_{(1,5)} + \hat{X}_{(5,7)}, \hat{X}_{(2,5)} + \hat{X}_{(5,7)}, \\ &\qquad \hat{X}_{(3,6)} + \hat{X}_{(6,7)}, \hat{X}_{(4,6)} + \hat{X}_{(6,7)}\} \\ &= \max\{\hat{Y}_1, \hat{Y}_2, \hat{Y}_3, \hat{Y}_4\}.\end{aligned}$$

Being independent, the random variables $\hat{X}_e$ $(e \in E)$ are positively associated by Theorem 2.1 in [27]. Clearly, each $Y_i$ is a nondecreasing function of the $\hat{X}_e$'s. By Theorem 5.1 in [27],

$$\begin{aligned}\mathbb{P}(\hat{T} \le y) &= \mathbb{P}(\hat{Y}_1 \le y, \dots, \hat{Y}_n \le y) \\ &\ge \prod_{i=1}^{n} \mathbb{P}(\hat{Y}_i \le y) \\ &= \prod_{i=1}^{n} \mathbb{P}(Y_i \le y) = \mathbb{P}(T \le y).\end{aligned}$$

It then follows that

$$\mathbb{E}\hat{T} = \int_0^\infty \mathbb{P}(\hat{T} > y)dy \le \int_0^\infty \mathbb{P}(T > y)dy = \mathbb{E}T.$$

■

**Remark III.1.** *Our proof in Theorem 1 assumes that the per-hop delays when using and not using aggregation are the same. This is conservative since the traffic reduction through aggregation can further reduce delays in the network (e.g., through shorter media access time due to less contention). Therefore, we expect even more delay reduction when using aggregation in practice.*

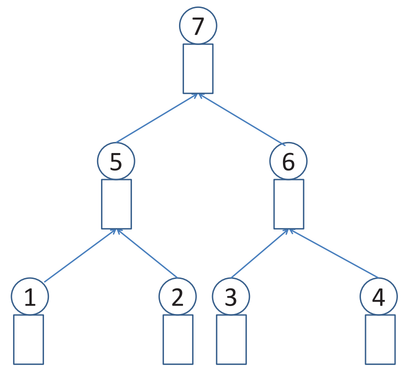



Fig. 3. Illustration of the setting for single tracking task with queueing delays at the intermediate nodes.

### *C. Incorporation of Queuing Delays*

We now relax the assumption in Section III-B and consider the scenario where there is a buffer at each intermediate node (we assume the buffer is of sufficient size so that no packet will be lost at the node), and packets can be queued up in the buffer, as illustrated in Fig. 3. Therefore, a one-hop delay contains two components: the transmission delay (i.e., the actual delay for a packet to reach its parent after being transmitted) and queuing delay (i.e., the delay for a packet to wait in the buffer before being transmitted). Again, we assume that transmission delays on the hops (links) are independent of each other. We then have the following result comparing the aggregation and non-aggregation approaches under general tree topologies and general delay distributions.

**Corollary III.1.** *For the scenario with queuing delays, assuming per-hop delays are independent, the aggregation approach generates a shorter expected cumulative delay than a non-aggregation approach.*

*Proof:* In a given setting (fixed tree topology, per-hop delays), the cumulative delay under the aggregation scheme when considering queuing delays at the intermediate nodes is the same as that when not considering queuing delays (since in either case an intermediate node needs to wait for the packets from all its children and then aggregate them, and there is no other traffic for either case), while in the non-aggregation case, considering queuing delays leads to larger expected cumulative delay since the packets might be queued at intermediate nodes. We know non-aggregation already leads to larger expected cumulative delay than aggregation in the non-queueing setting (see Theorem 1). When considering queuing delays, the expected cumulative delay under non-aggregation is larger than that under aggregation. ■

### D. Analysis Under Constant Delays

Next, we discuss under what conditions aggregation achieves the best performance. For aggregation, the tree topology and the packet flow form a fork-join queueing network, where packets from all the children nodes are joined at a parent node. By applying results from fork-join queueing networks [28], we know that when per-hop delay is a constant (i.e., each hop has the same constant delay since we assume per-hop delays follow the same distribution), the expected cumulative delay is minimized. That is, given link delays with arbitrary distributions, $E[T]$ can be lower bounded by $E[T']$ for a system where the link delays are constant with values equal to the expected link delays of the first system. We next use a simple example to illustrate the benefits of the aggregation approach compared to the non-aggregation approach under constant transmission delays.

Consider a complete binary tree with $l$ levels. The number of leaf nodes of the tree is therefore $2^l$. Suppose the transmission delay on each hop is a constant, denoted as $d$. For the aggregation approach, the delay for the root node to receive all the measurements is $ld$. For the non-aggregation case, before time $ld$, the root node receives no measurements; starting from $ld$, the root node receives 2 measurements every $d$ time units. Therefore, the total delay for the root to receive all the measurements is $2^l/2d + (l-1)d = 2^{l-1}d + (l-1)d$.

The above analysis demonstrates for this example that the MSE increases linearly with the depth of the tree under aggregation, while it increases exponentially under non-aggregation, highlighting the benefits of aggregation. We can further determine good stopping times (i.e., when the sink should estimate the target location without waiting for any additional sensor measurements) for both approaches. It is easy to see that for aggregation, since all of the measurements arrive at time $ld$, the optimal stopping time is $ld$. For non-aggregation, the MSE at time $(l+x-1)d$, $x = 1, 2, \cdots$, is

$$\frac{\sigma^2}{2x} + (l+x-1)cd.$$

It is easy to show that the optimal value of $x$ is $(\frac{\sigma}{\sqrt{2cd}} + l - 1)d$, which indicates that under the non-aggregation approach, lower measurement quality and deeper tree lead to a larger optimal stopping times, while higher volatility of the target motion leads to a shorter optimal stopping times.

## IV. Periodic Tracking

In Section III-C, we have shown that aggregation outperforms non-aggregation in one-shot single task tracking when considering queuing delays. In periodic tracking where sensors take measurements periodically, measurements can be queued up at an intermediate node, leading to increasing queuing delays over time. Therefore, an interesting question is: what are the asymptotic behaviors of aggregation and non-aggregation approaches in periodic tracking when considering queuing delays? We next answer this question. Our investigation focuses on a scenario where links have the same constant transmission delay, motivated by the observation that the aggregation approach achieves the best performance under this scenario (see Section III-C).

Suppose that each sensor takes measurements periodically with period $\tau$. Accordingly, we divide time into intervals of length $\tau$, called *rounds*; assume measurements are taken at the *beginning* of each round. When the root receives all the measurements taken in round $m$ ($m = 1, 2, \ldots$), it estimates the target location at the beginning of this round. Let $\sigma_m^2$ denote the MSE for estimating the target location in round $m$, which also gives the target's variance given measurements up to round $m$. At the beginning of round $m+1$, the target's variance increases to $\sigma_m^2 + c\tau$ due to its mobility, which establishes a prior for the estimation in round $m+1$. Let $\sigma_0^2$ denote the aggregate noise variance of all the measurements, i.e., $\frac{1}{\sigma_0^2} = \sum_{i=1}^n \frac{1}{\sigma_i^2}$. Since both the prior and the measurement noise are Gaussian, the MSE for the $(m+1)$-th round is

$$\sigma_{m+1}^2 = \frac{(\sigma_m^2 + c\tau)\sigma_0^2}{\sigma_m^2 + \sigma_0^2 + c\tau},$$

where $c$ is the volatility parameter of the target mobility model.

### A. Asymptotic Behavior

When the period is too short (i.e., measurements are taken too frequently), packets can be queued indefinitely inside the network, leading to infinite queuing delays. We instead consider the scenario where the queuing delays inside the network are bounded. The following theorem summarizes the asymptotic behavior of $\sigma_m^2$.

**Theorem 2.** *When the tracking period is sufficiently long so that the queuing delays in the network are bounded, $\sigma_m^2$ converges as $m \to \infty$, and*

$$\lim_{m\to\infty} \sigma_m^2 = -\frac{1}{2}c\tau + \frac{1}{2}\sqrt{c\tau(c\tau + 4\sigma_0^2)}. \quad (18)$$

*Proof:* We prove the convergence by proving that

$$f(x) \quad = \quad \frac{\sigma_0^2(x + c\tau)}{x + \sigma_0^2 + c\tau}$$

is a contraction mapping.

Note

$$\begin{aligned} & |f(x) - f(y)| \\ = \quad & \sigma_0^2 \left| \frac{(x+c\tau)(y+c\tau+\sigma_0^2) - (y+c\tau)(x+c\tau+\sigma_0^2)}{(x+c\tau+\sigma_0^2)(y+c\tau+\sigma_0^2)} \right| \\ = \quad & \frac{\sigma_0^4}{(x+c\tau+\sigma_0^2)(y+c\tau+\sigma_0^2)} \cdot |x-y|, \end{aligned}$$

where

$$0 < \frac{\sigma_0^4}{(x+c\tau+\sigma_0^2)(y+c\tau+\sigma_0^2)} < 1.$$

This implies that $f(x)$ is a contraction mapping and hence has a fixed point. Therefore, $\sigma_m^2$ converges as $m \to \infty$.

Next, we calculate $\tilde{\sigma}^2 = \lim_{m\to\infty} \sigma_m^2$. Let $m \to \infty$, we have

$$\tilde{\sigma}^2 \quad = \quad \frac{\sigma_0^2(\tilde{\sigma}^2 + c\tau)}{\tilde{\sigma}^2 + \sigma_0^2 + c\tau}.$$

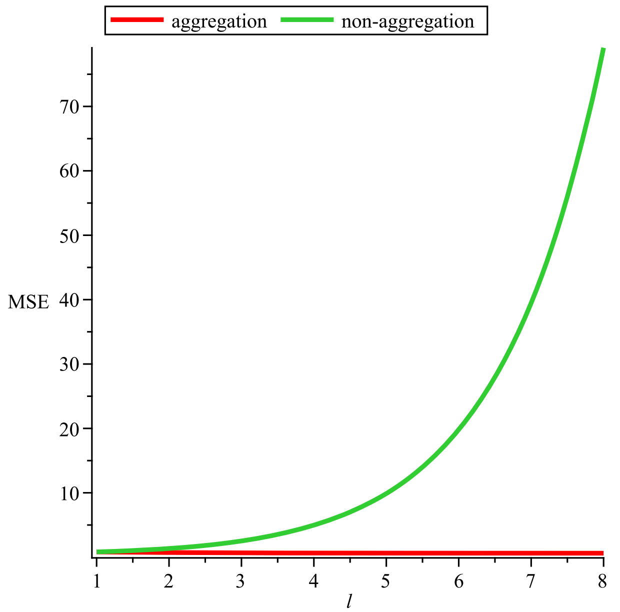



Fig. 4. Optimal asymptotic MSE for aggregation and non-aggregation under periodic tracking with binary tree topology.

Hence,

$$\tilde{\sigma}^4 + c\tau\tilde{\sigma}^2 - c\tau\sigma_0^2 = 0.$$

By solving the equation, we have

$$\tilde{\sigma}^2 = -\frac{1}{2}c\tau + \frac{1}{2}\sqrt{c\tau(c\tau + 4\sigma_0^2)}.$$

■

Based on this theorem, it can be shown that $\lim_{m\to\infty}\sigma_m^2$ is an increasing function of $\tau$. Therefore, the optimal $\tau$ is the minimum value so that queue lengths inside the network are finite. Although we have only considered estimating the target location at the beginning of each round, the above analysis applies to arbitrary points of inception, with an asymptotic MSE of $-\frac{1}{2}c\tau + \frac{1}{2}\sqrt{c\tau(c\tau + 4\sigma_0^2)} + cT$, where $T$ is the time from the beginning of each round $m$ to the corresponding inception point (i.e., the time to estimate the target location for, using measurements up to round $m$).

### B. Periodic Tracking under Aggregation and Non-aggregation

We next consider periodic tracking under aggregation and non-aggregation approaches when measurements are routed along a complete binary tree. As mentioned before, we assume the transmission delay on each hop is a constant, denoted as $d$. To achieve the the optimal asymptotic MSE, we need to set the measurement period, $\tau$, to be $d$ when using aggregation. This is because of pipelining. When $n > l$, the sink will receive one set of measurements from all of the sensors in each round. Therefore, we can set the measurement period to be $d$ without overflowing the network. This is not the case for non-aggregation, where we need to set $\tau$ to $2^{l-1}d$. This is because it takes $2^{l-1}$ rounds for the sink to receive the measurements from all of the sensors. Since $d \ll 2^{l-1}d$ when $l$ is large, the optimal MSE for aggregation is much smaller than that of non-aggregation, indicating that aggregation can dramatically improve the tracking quality for periodic tracking.

We now use an example to illustrate the tracking quality under aggregation and non-aggregation. Suppose $c = d = \sigma^2 = 1$, where $\sigma^2$ is the measurement MSE of each leaf node. Since there are $2^{l-1}$ sensors that take measurements in a complete binary tree of $l$ layers, $\sigma_0^2 = \sigma^2/2^{l-1}$. Fig. 4 plots the optimal asymptotic MSEs versus $l$ for both aggregation and non-aggregation approaches. Observe that for non-aggregation, the optimal asymptotic MSE increases exponentially with $l$, while for aggregation, it is a constant (independent of $l$).

Compared with one-shot tracking, aggregation provides more benefit under periodic tracking due to pipelining. Moreover, the performance of aggregation is independent of degree of a complete tree. If we apply the above analysis to a complete tree with degree $d_0$, the optimal $\tau$ for aggregation is still $d$, while the optimal $\tau$ for non-aggregation becomes $d_0^{l-1}d$. Furthermore, for any routing tree with an arbitrary topology, the optimal $\tau$ for aggregation is $d$ despite the topology difference.

In summary, when the per-hop transmission delay is a constant, under periodic tracking, the optimal MSE when using aggregation is independent of the tree topology, while the optimal MSE when not using aggregation depends heavily on the topology. When the sensor network is deployed in a battlefield or in an ad hoc manner, topology control is not always possible. In such scenarios, tracking quality when not using aggregation can be significantly inferior to that when using aggregation.

In closing this section, it should be noted that what was presented can serve as a basis for future investigation under more advanced aggregation process. For example, sensors may preconstruct multiple aggregation trees and dynamically select one such tree for each measurement based on network conditions (e.g., queue sizes at intermediate nodes) and also which other measurements they have intersected for improved quality of target location estimation. System analysis for such time-shared, interleaved multi-tree aggregation may be performed along the same lines as presented earlier. It is anticipated that the performance characteristics per aggregation tree will be similar as before, calculated for the selected subset of measurements.

Specifically, the expected per-tree delay remains the same for the case of independent per-hop delays, while its value is reduced for the case involving queueing delays due to reduced load. In both cases, the overall expected cumulative delay is a weighted average of the per-tree delays, with weights proportional to the time shares of using the trees. While quantitative analysis of the delays will require a separate study, we point out that the qualitative comparison between aggregation scheme and non-aggregation scheme remains valid for each tree and also for the overall delay, given that both schemes follow the same time-sharing schedule.

## V. Multiple Tracking Tasks

In this section, we consider a sensor network used for multiple tasks. For instance, the sensor network is used to track the location of a target as well to keep track of environmental characteristics (e.g., temperature, humidity, wind

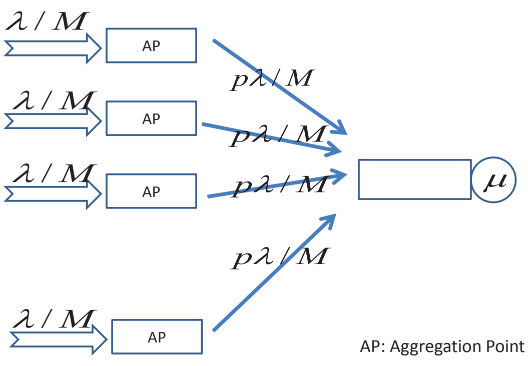



Fig. 5. Setting of multiple tracking tasks.

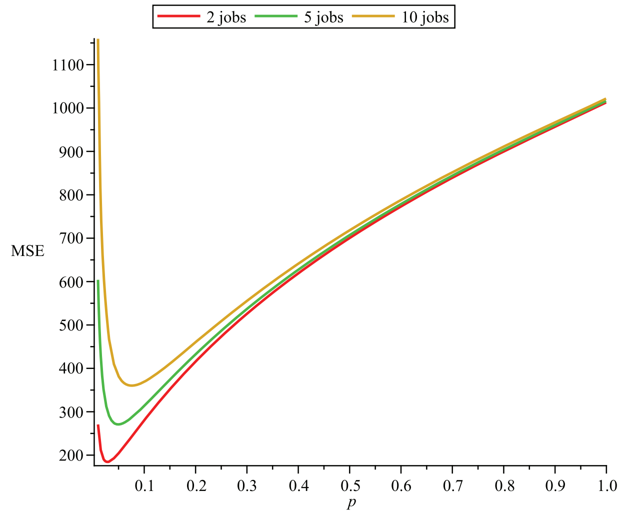



Fig. 6. Multiple tracking tasks, MSE versus $p$, where $M = 2, 5, 10$, $\lambda = 0.9, \mu = 1, c = 1$, and $\sigma^2 = 1000$.

level, etc.). The packets corresponding to the multiple tasks share the resources of the sensor network. In the following, we consider specifically the target tracking task, and investigate target tracking quality in the presence of other tasks under aggregation and non-aggregation.

### A. Network Model

In this multiple task setting, it is intractable to model each link individually. We therefore adopt a simplified model that models the entire network as an $M/M/1$ queue, as illustrated in Fig. 5. We assume there are an infinite number of sensors, and, for each task, all of the measurements are made at the same time at the sensors, but are sent to an *aggregation point* according to a Poisson process with rate $\lambda/M$, where $M$ is the number of simultaneous tracking tasks. Measurements for different tasks are sent to different aggregation points to satisfy that only packets that belong to the same tracking task can be aggregated together. The Poisson arrival process models the delays that packets experience while being transmitted in the network to reach the aggregation point.

Now consider specifically the aggregation point for the tracking task. We assume all sensor measurement errors follow the same Gaussian distribution with mean 0 and variance $\sigma^2$. A packet of the tracking task is aggregated with previous packets that have not been sent out, and with probability $p$, the aggregated measurements are sent to the network (i.e. the $M/M/1$ queue). This implies that, on average, $1/p$ packets are aggregated together. Note that the non-aggregation approach is a special case of the above model with $p = 1$ (since $p = 1$, a packet is sent out immediately, and does not aggregate with other packets). Following the above model, each aggregation point sends out packets to the $M/M/1$ queue following a Poisson process with rate $p\lambda/M$, where the processing delay for a packet at the queue models the delay the packet encounters after its aggregation point to reach the sink. Combining the packets from the $M$ tracking tasks, the aggregate arrival rate at the $M/M/1$ queue is $p\lambda$.

### B. Target Location Estimate

Under the above assumptions, the probability that $n$ aggregated measurements are sent out to the $M/M/1$ queue by a tracking job is $p(1-p)^{n-1}$. The average time to accumulate these $n$ measurements is $nM/\lambda$. The average time to wait in the queue is $1/(\mu - p\lambda)$ according to the property of $M/M/1$ queue, where $\mu$ is the service rate of the queue. Therefore, following (3), when the sink receives the measurements, the MSE of the target location estimate is calculated as

$$\begin{aligned} &\sum_{n=1}^{\infty} p(1-p)^{n-1}(cnM/\lambda + c/(\mu - p\lambda) + \sigma^2/n) \\ = \quad & c/(\mu - p\lambda) + \sum_{n=1}^{\infty} p(1-p)^{n-1}(cnM/\lambda + \sigma^2/n) \\ = \quad & \frac{c}{\mu - p\lambda} + \frac{cM}{p\lambda} - \frac{\sigma^2 p \ln p}{1-p}. \end{aligned} \tag{19}$$

Note that the second term in (19) increases linearly with $M$, indicating that tracking quality decreases linearly with the number of tracking tasks. Further note that the first and third terms in (19) increase with $p$, while the second term decreases with $p$, indicating that for a fixed $M$ the optimal $p$ is somewhere between 0 and 1. Specifically, the optimal $p$, denoted as $p^*$, can be obtained by solving

$$\frac{c\lambda}{(\mu - p\lambda)^2} - \frac{cM}{p^2\lambda} - \frac{\sigma^2(1 + \ln p)}{1-p} - \frac{\sigma^2 p \ln p}{(1-p)^2} = 0. \tag{20}$$

Recall that $p = 1$ corresponds to non-aggregation. Therefore, the MSE when using aggregation with the optimal $p$ outperforms that when not using aggregation. Fig. 6 plots an example, where $p$ is varied from 0 to 1, and $M$ is set to 2, 5 or 10. We observe that, for the same $p$, the MSE indeed increases with $M$, and for the same $M$, the optimal $p$ is indeed between 0 and 1.

The above analysis, although conducted under simplifying assumptions, provides useful guidance on parameter selection when designing aggregation approaches for multiple tracking tasks. For instance, when measurements errors have variance

$\sigma^2$, for a network that has capacity $\mu$, average arrival packet rate $\lambda$, and serves $M$ simultaneous tracking tasks, our results indicate that aggregating every $1/p^*$ measurements together for each tracking task is a reasonable choice in order to achieve good tracking quality.

## VI. Conclusions and Future Work

In this paper, we have investigated how commonly used in-network aggregation approaches impact the target tracking quality in multi-hop wireless sensor networks. Specifically, we use the MSE of target location estimates to quantify the target tracking quality, and investigate how in-network aggregation affects the MSE. We started with proposing an aggregation scheme that preserves sufficient statistic for making an optimal estimate, and then explored the impact of aggregation in several examples to provide intuition. We then analytically studied the impact of aggregation in three increasingly more complicated scenarios.

As future work, we plan to study the impact of various delays (e.g., delays to aggregate packets and delays to differentiate different tracking tasks) on target tracking quality, and scenarios with correlated per-hop delays and packet losses.

## Acknowledgement

This work was sponsored in part by the National Science Foundation under grants EEC-0313747, and by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

## References

[1] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: A survey," *IEEE Wireless Commun.*, vol. 14, no. 2, pp. 70 – 87, 2007.

[2] I. Solis and K. Obraczka, "In-network aggregation trade-offs for data collection in wireless sensor networks," *International J. Sensor Networks*, vol. 1, Sept. 2006.

[3] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Trans. Netw.*, vol. 11, pp. 2–16, Feb. 2002.

[4] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 660–670, Oct. 2002.

[5] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Trans. Parallel Distrib. Syst.*, vol. 13, pp. 924–935, Sept. 2002.

[6] Y. Yao and J. Gehrke, "Query processing for sensor networks," in *ACM CIDR*, (Asilomar, CA, US), Jan. 2003.

[7] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: a tiny aggregation service for ad-hoc sensor networks," in *USENIX OSDI*, (Boston, MA, US), Dec. 2002.

[8] S. Nath, P. B. Gibbons, Z. R. Anderson, and S. Seshan, "Synopsis diffusion for robust aggregation in sensor networks," in *ACM SenSys*, (Baltimore, MD, US), Nov. 2004.

[9] A. Manjhi, S. Nath, and P. B. Gibbons, "Tributaries and deltas: Efficient and robust aggregation in sensor network stream," in *ACM SIGMOD*, (Baltimore, MD, US), June 2005.

[10] E. Cohen and H. Kaplan, "Spatially-decaying aggregation over a network: Model and algorithms," in *ACM SIGMOD*, (Paris, France), June 2004.

[11] A. Sharaf, J. Beaver, A. Labrinidis, and K. Chrysanthis, "Balancing energy efficiency and quality of aggregate data in sensor networks," *The VLDB J.*, vol. 13, pp. 384–403, Dec. 2004.

[12] T. He, B. M. Blum, J. A. Stankovic, and T. Abdelzaher *ACM Trans. Embedded Computing Systems*, vol. 3, pp. 426–457, May 2004.

[13] N. Shrivastava, C. Buragohain, D. P. Agrawal, and S. Suri, "Medians and beyond: New aggregation techniques for sensor networks," in *ACM SenSys*, (Baltimore, MD, US), Nov. 2004.

[14] A. Bezenchek, M. Rafanelli, and L. Tininini, "A data structure for representing aggregate data," in *IEEE SSDBM*, (Stockholm, Sweden), June 1996.

[15] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Mag.*, vol. 21, pp. 80–94, Sept. 2004.

[16] C. Lageweg, J. Janssen, and M. Ditzel, "Data aggregation for target tracking in wireless sensor networks," in *European Conference on Smart Sensing and Context, Lecture Notes in Computer Science*, vol. 4272, pp. 15–24, 2006.

[17] M. Ditzel, C. Lageweg, J. Janssen, and A. Theil, "Multi-target data aggregation and tracking in wireless sensor networks," *J. Networks*, vol. 3, Jan. 2008.

[18] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, and et al., "A line in the sand: A wireless sensor network for target detection, classification, and tracking," *Computer Networks, Special Issue on Military Communications Systems and Technologies*, vol. 46, pp. 605–634, July 2004.

[19] B. Horling, R. Vincent, R. Mailler, J. Shen, R. Becker, K. Rawlins, and V. Lesser, "Distributed sensor network for real time tracking," in *Proc. of International conference on Autonomous agents*, (New York, NY, USA), pp. 417–424, 2001.

[20] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration for target tracking," *IEEE Signal Processing Mag.*, vol. 19, pp. 61–72, Mar. 2002.

[21] D. Li, K. Wong, Y. H. Hu, and A. Sayeed, "Detection, classification and tracking of targets in distributed sensor networks," *IEEE Signal Processing Mag.*, vol. 19, no. 2, pp. 17–29, 2002.

[22] S. Pattem, S. Poduri, and B. Krishnamachari, "Energy-quality tradeoffs for target tracking in wireless sensor networks," in *Information Processing in Sensor Networks (IPSN)*, pp. 32–46, Apr. 2003.

[23] A. Giridhar and P. R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Commun. Mag.*, vol. 44, pp. 98 – 107, April 2006.

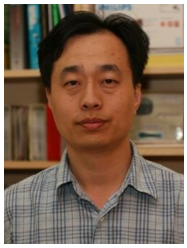
[24] P. Balister, B. Bollobas, A. Anandkumar, and A. Willsky, "Energy-latency tradeoff for in-network function computation in random networks," in *Proc. IEEE INFOCOM*, 2011.

[25] S. Banerjee, P. Gupta, and S. Shakkottai, "Towards a queueing-based framework for in-network function computation," in *IEEE International Symposium on Information Theory (ISIT)*, (Saint Petersburg, Russia), July 31-Aug 5 2011.

[26] C. Bisdikian, L. Kaplan, M. Srivastava, D. Thornley, D. Verma, and R. Young, "Building principles for a quality of information specification for sensor information," in *Information Fusion, 2009. FUSION '09. 12th International Conference on*, pp. 1370 –1377, july 2009.

[27] J. Esary, F. Proschan, and D. Walkup, "Association of random variables with applications," in *The Annals of Mathematical Statistics*, vol. 38, pp. 1466–1474, Oct. 1967.

[28] F. Baccelli, W. A. Massey, and D. Towsley, "Acyclic fork-join queueing networks," tech. rep., Amherst, MA, USA, 1987.



**Wei Wei** received his B.S. degree in Applied Mathematics from Peking University, China in 1992, and M.S. degree in Statistics from Texas A & M University in 2000. He then received M.S. degrees in Computer Science and Applied Mathematics, and a Ph.D. in Computer Science from the University of Massachusetts, Amherst in 2004, 2004, and 2006 respectively. He is currently a Senior Research Fellow at the Computer Science Department, the University of Massachusetts, Amherst. He was a Research Scientist at United Technologies Research Center and a visiting Assistant Professor at University of Connecticut. His current research interests include modeling, data mining and statistical inference in computer networks (including cellular data networks), and social networks.

**Ting He** is a Research Staff Member in the Wireless Networking Resarch group at IBM T.J. Watson Research Center, Hawthorne, NY. She received the Ph.D. degree from the School of Electrical and Computer Engineering, Cornell University, in 2007 and the B.S. degree in Computer Science from Peking University, China, in 2003. At IBM, Ting works as a task lead under the International Technology Alliance (ITA) program funded by US ARL and UK MoD, the ARRA program funded by NIST, and other research projects on network science and technologies. Her research aims at applying classic signal processing techniques to network design and analysis. Her recent research focuses on controlled mobility in hybrid communication networks, network optimization in coalition networks, and modeling and scheduling in cloud computing networks. Her general research interests include detection and estimation theory, stochastic control, statistical signal processing, information theory, and network security.

**Chatschik Bisdikian** (PhD, FIEEE) holds a Ph.D. degree in EE from the Univ. of Connecticut, USA. He is a Research Staff Member with the IT & Wireless Convergence department group at IBMs T. J. Watson Research Center, in Hawthorne, NY, USA. He has been with IBM Research since 1989 and over the years he has worked in numerous projects covering a variety of research topics in communications, networking, pervasive computing, IPTV services, computer system management, sensor networks, with special emphasis in the areas of quality and value of information (QoI and VoI) for sensor networks.

He has authored over 120 peer-reviewed papers, holds several patents in the aforementioned areas, and co-authored the book Bluetooth Revealed (Prentice Hall). He has served as the Editor-in-Chief of IEEE Network Magazine and currently serves in the editorial board of the ACM Computing Surveys and Pervasive and Mobile Computing journals. He is currently serving as General Vice-Chair for the IEEE PerCom13 conference and was: the Technical Program Chair for IEEE PerCom09; general co-chair of the 2010 and 2011 IEEE Intl Workshop on Information Quality and Quality of Service (IQ2S) and the 2008 IEEE Workshop on Quality of Information (QoI) for Sensor Networks (QoISN). He received the 2010 Best Paper award at IEEE RTSS2010 for the paper titled Quality Tradeoffs in Object Tracking with Duty-Cycled Sensor Networks for 2002 best tutorial award from IEEE Communications Society for his paper titled An Overview of the Bluetooth Wireless Technology. He has been involved with the development of the Bluetooth specification from its early stages and has served as vice-chair of the IEEE 802.15.1 task group that developed a standard for personal area networks adapted from the Bluetooth specification.

He is an IEEE Fellow, a member of the Academy of Distinguished Engineers and Hall of Fame of the School of Engineering of the Univ. of Connecticut, and a 1995 finalist of the Eta Kappa Nu Honor Societys Outstanding Young Electrical Engineer Award Program.

**Dennis Goeckel** split time between Purdue University and Sundstrand Corporation from 1987-1992, receiving his BSEE from Purdue in 1992. From 1992-1996, he was a National Science Foundation Graduate Fellow and then Rackham Pre-Doctoral Fellow at the University of Michigan, where he received his MSEE in 1993 and his Ph.D. in 1996, both in Electrical Engineering with a specialty in Communication Systems. In September 1996, he joined the Electrical and Computer Engineering department at the University of Massachusetts, where he is currently a Professor. His current research interests are in the areas of communication systems and wireless network theory.

Dr. Goeckel was the recipient of a 1999 CAREER Award from the National Science Foundation for "Coded Modulation for High-Speed Wireless Communications". He was a Lilly Teaching Fellow at UMass-Amherst for the 2000-2001 academic year and received the University of Massachusetts Distinguished Teaching Award in 2007. He is a Fellow of the IEEE.

**Bo Jiang** received his B.S. in Electronic Engineering from Tsinghua University in 2006, and his M.S. in Electrical and Computer Engineering and in Applied Mathematics from University of Massachusetts Amherst in 2008 and 2012, respectively. He is currently a Ph.D. Candidate in the Department of Computer Science at University of Massachusetts Amherst. His research interests include network science, network modeling and analysis.

**Lance M. Kaplan** received the B.S. degree with distinction from Duke University, Durham, NC, in 1989 and the M.S. and Ph.D. degrees from the University of Southern California, Los Angeles, in 1991 and 1994, respectively, all in Electrical Engineering. From 1987-1990, Dr. Kaplan worked as a Technical Assistant at the Georgia Tech Research Institute. He held a National Science Foundation Graduate Fellowship and a USC Dean's Merit Fellowship from 1990-1993, and worked as a Research Assistant in the Signal and Image Processing Institute at the University of Southern California from 1993-1994. Then, he worked on staff in the Reconnaissance Systems Department of the Hughes Aircraft Company from 1994-1996. From 1996-2004, he was a member of the faculty in the Department of Engineering and a senior investigator in the Center of Theoretical Studies of Physical Systems (CTSPS) at Clark Atlanta University (CAU), Atlanta, GA. Currently, he is a researcher in the Networked Sensing and Fusion branch of the U.S. Army Research Laboratory. Dr. Kaplan serves as Editor-In-Chief for the IEEE Transactions on Aerospace and Electronic Systems (AES). In addition, he also serves on the Board of Governors of the IEEE AES Society and on the Board of Directors of the International Society of Information Fusion. He is a three time recipient of the Clark Atlanta University Electrical Engineering Instructional Excellence Award from 1999- 2001. His current research interests include signal and image processing, automatic target recognition, information/data fusion, and resource management.

**Don Towsley** holds a B.A. in Physics (1971) and a Ph.D. in Computer Science (1975) from University of Texas. From 1976 to 1985 he was a member of the faculty of the Department of Electrical and Computer Engineering at the University of Massachusetts, Amherst. He is currently a Distinguished Professor at the University of Massachusetts in the Department of Computer Science. He has held visiting positions at IBM T.J. Watson Research Center, Yorktown Heights, NY; Laboratoire MASI, Paris, France; INRIA, Sophia-Antipolis, France; AT&T Labs - Research, Florham Park, NJ; and Microsoft Research Lab, Cambridge, UK. His research interests include networks and performance evaluation.

He has served as Editor-in-Chief of IEEE/ACM Transactions on Networking and on the editorial boards of Journal of the ACM, and IEEE Journal on Selected Areas in Communications, and has previously served on numerous other editorial boards. He was Program Co-chair of the joint ACM SIGMETRICS and PERFORMANCE '92 conference and the Performance 2002 conference. He is a member of ACM and ORSA, and Chair of IFIP Working Group 7.3.

He has received the 2007 IEEE Koji Kobayashi Award, the 2007 ACM SIGMETRICS Achievement Award, the 2008 ACM SIGCOMM Achievement Award, the 1998 IEEE Communications Society William Bennett Best Paper Award, and numerous best conference/workshop paper awards. Last, he has been elected Fellow of both the ACM and IEEE.