Data Fusion Improves the Coverage of Wireless Sensor Networks

Abstract—Recently, large-scale wireless sensor networks have been increasingly available for mission-critical applications such as object detection and tracking. Coverage is an important performance measure that characterizes how well a surveillance field is monitored by a sensor network. Due to the high deployment cost, it is crucial to predict and understand the expected coverage performance of large-scale sensor networks. However, most existing theoretical studies are based on simplistic sensing models such as the boolean model that does not capture the stochastic nature of sensing. In this paper, we explore the fundamental limits of coverage based on a data fusion model that can improve detection performance by aggregating noising measurements of individual sensors. We derive several network density bounds for achieving full coverage for both static and mobile sensor networks and compare against the existing results based on the boolean sensing model. We show that the same level of coverage can be achieved with significantly lower network density by exploiting the collaboration among sensors in data fusion. Our results help understand the limitations of the existing results based on the boolean sensing model and provide important insights into the design of large-scale sensor network based on data fusion models.

I. INTRODUCTION

Recent years have witnessed the deployments of wireless sensor networks in a class of mission-critical applications such as security surveillance, object detection and tracking [13], [19]. Many of these applications involve a large number of sensors that are deployed in a vast geographical area. Due to the high cost of deployment, it is important to predict and understand the expected performance of these large-scale sensor networks. A critical performance measure is sensing coverage that characterizes how well a surveillance field is monitored by a sensor network. To this end, many recent works aimed to analyze the coverage performance of large-scale sensor networks.

Most existing studies on the coverage problem [2], [16], [28], [24], [34], [31], [32], [33] are based on simplistic sensing models. In particular, the sensing region of a sensor is often modeled as a disk with radius \( r \) centered at the position of sensor. \( r \) is referred to as the sensing range. A sensor detects the targets (events) within its sensing range with a probability of one. Although this boolean sensing model allows a geometric treatment to the coverage problem, it fails to capture the stochastic nature of sensing, such as false alarms caused by noise and probabilistic detectability. Moreover, each sensor performs detection independently and do not take advantage of possible collaboration among the sensors.

The aforementioned complexities of sensing have been the research focus of the data fusion community for over three decades [26]. A key advantage of data fusion is to improve the system detection performance by jointly considering the noisy measurements of multiple sensors. Early work on data fusion focused on analyzing optimal fusion strategies for small-scale powerful sensor networks (e.g., a handful of radars). Recent work [8], [19], [9] has considered the properties of wireless sensor networks such as spatial distribution of sensors. In practice, many sensor network systems designed for target detection, tracking and classification [9], [13], [19] have employed some kind of data fusion schemes. However, the fundamental performance limits of large-scale wireless sensor networks designed based on data fusion models has not been systematically analyzed.

In this paper, we develop a theoretical framework to study the fundamental limits of coverage of large-scale wireless sensor networks under the data fusion model. We adopt the following data fusion model. When a possible target appears, the sensors within a distance of \( R \) (referred to as the fusion range) from the target position measure signal energy and the detection decision is made by comparing the sum of individual sensor measurements against a predefined threshold. Based on this model, we characterize the coverage of a physical location by \((\alpha, \beta)\), i.e., the probability of detecting an event occurring at the location must be above \(\beta\) while the false alarm rate must be below \(\alpha\). This definition of coverage captures the requirement of mission-critical surveillance applications. For instance, it is natural for a user to require that “the probability of detecting any intruder is no lower than 95% and no more than 5% of the network reports are false alarms”.

In this paper, we first derive the minimum network density that is required to achieve the desired coverage \((\alpha, \beta)\) specified by the application. We then study the condition to guarantee the full coverage over a large region. We prove that, when sensors are deployed according to a 2-dimensional Poisson point process, the ratio of sensor densities (denoted by \(\rho\) and \(\rho_b\), respectively) required to achieve full coverage under the data fusion and boolean models satisfies \(\rho/\rho_b = 2r^2/R^2\). As fusion range \(R\) can increase with network density and is much larger than sensing range \(r\), \(\rho\) is significantly smaller than \(\rho_b\). Moreover, when the fusion range is optimized with respect to network density, we have \(\rho = O(\sqrt{\rho_b})\).

Recent work [30], [29], [4] shows that mobility can be introduced to trade with network density in achieving coverage. In such a scheme, randomly distributed mobile sensors can relocate themselves to fill coverage holes in the initial network deployment. In this paper, we extend a mobile relocation algorithm [30] to the data fusion model. We show that the
mobile sensor density required to achieve full coverage is no greater than that under the boolean model, i.e., \( \rho \leq \rho_0 \). Moreover, the ratio between \( \rho_0 \) and \( \rho \) increases with SNR. At the same time, the maximum mobile moving distance under the fusion model remains the same order as under the boolean model.

These results show that the boolean model is not efficient in providing coverage and does not achieve the full potential of a sensor network. Consequently, the resulted network deployments yield substantial over-provision of sensing quality. On the other hand, by exploiting the collaboration among sensors, the data fusion model can achieve the same level of coverage using significantly fewer sensors. In addition to the advantage of network density reduction, our results can also be applied to prolong network lifetime. An effective energy conservation approach for large networks is to maintain sufficient network coverage via only a fraction of on-duty sensors while all other sensors are switched to sleep (or a low duty-cycle). The data fusion model reduces the number of necessary on-duty sensors by taking advantage of the collaboration among them, resulting in longer network lifetime.

The rest of this paper is organized as follows. Section II reviews related work. Section III and IV introduce the background and the definition of our problem, respectively. In Section V, we derive coverage under the data fusion and boolean sensing models. We study the network densities for achieving full coverage for static and mobile sensor networks in Section VI and VII-C, respectively. We present simulation results in Section VIII and conclude the paper in Section IX.

II. RELATED WORK

As one of the most fundamental problems in wireless sensor networks, the coverage problem has attracted significant research attention. Previous work falls into two basic categories, namely coverage maintenance algorithms/protocols and analysis of coverage performance of large-scale networks.

Most existing coverage maintenance algorithms/protocols [3], [31], [32], [33] are designed based on the boolean sensing model. The effectiveness of these schemes largely relies on the assumption of circular sensing range and deterministic sensing capability of sensors. Several recent proposals have adopted probabilistic sensing models [14], [23], [1]. Different from our focus on analyzing the fundamental limits of coverage in sensor networks, these studies aim to devise algorithms/protocols for coverage maintenance.

Theoretical analysis of the coverage of large-scale sensor networks has been conducted in [2], [20], [16], [28], [24], [34]. These work aim to study the asymptotic behavior of coverage with respect to network density and sensing range. In [16], [28], [24], [34], critical conditions for ensuring k-coverage are derived for various sensor deployment strategies (e.g., grid, random and Poisson deployments). The boolean model is adopted in all these work except [20] in which the coverage is analyzed based on both the boolean model and a general sensing model accounting for signal decay. In this paper, we compare our results derived under a data fusion model against the results from [2], [20].

Recent work [30], [29], [4] has exploited mobility to reduce the density of sensors in achieving coverage. In such a scheme, randomly distributed mobile sensors can relocate themselves to fill coverage holes in the initial network deployment. Wang et al. [30] proposed a sensor relocation algorithm that bounds the movement distance of mobile sensors. In this paper, we extend this algorithm to a data fusion model. Our analysis shows that data fusion results in lower density of sensors without increasing the movement distance of mobile sensors.

There is a vast literature on stochastic signal detection based on multi-sensor data fusion. Early work [26] focuses on devising optimal fusion rules for single-hop wired sensor networks (e.g., those composed of a handful of radars). Recent work on data fusion [8], [19], [9] have considered the properties of wireless sensor networks such as spatial node distribution and limited sensing/communication capability. Moreover, various data fusion schemes have been implemented in practice for target detection, tracking and classification [19], [9] using wireless sensor networks.

Niu et al. [22] derived the relation between average system detection probability and network density in large random networks under a data fusion model. Different from their work that studies the average detection performance, our paper focuses on the asymptotic behavior of full coverage (i.e., the performance of detecting any target that appears in a region). Moreover, we study the detection performance of mobile sensor networks while the issue of mobility is not addressed in [22].

III. PRELIMINARIES

In this section, we describe the preliminaries of our work, which include the target energy model, the sensor measurement model and the data fusion model.

A. Target and Sensor Measurement Models

Sensors detect targets by measuring the energy of signals, e.g., acoustic signal, emitted by targets. The energy attenuates with the distance from the source. Suppose the distance between sensor \( i \) and the signal source is \( r_i \). The signal energy measured by sensor \( i \), \( s_i \), follows the following attenuation model:

\[
s_i = \frac{a}{1 + r_i^b}
\]

where \( a \) is the original energy emitted by the signal source, \( b \) is a decaying factor which is typically from 2 to 5. In particular, \( b = 2 \) for acoustic signal [25]. We assume \( b = 2 \) in the rest of this paper. The model in (1) avoids singularity when \( r_i \) approaches 0, and approximates a power law decay when \( r_i \) is large. This model has also been used in the literature of data fusion [22]. Sensor \( i \)'s measurement, \( y_i \), is contaminated by Gaussian noise \( n_i \).

\[
y_i = s_i + n_i
\]
Data fusion [27] is a widely used technique for improving the performance of detection systems. A sensor network that employs data fusion is often organized into clusters. Each cluster head is responsible for making a final decision regarding the presence of a target by fusing the information gathered by member nodes in the cluster. There exist two basic data fusion schemes, namely, value fusion and decision fusion.

In a value fusion scheme [6], each sensor sends its signal energy measurements to the cluster head, which makes a decision based on the energy measurements received from all member nodes. In a decision fusion scheme, each sensor makes a local decision (0 or 1) based on its measurements and sends its decision to the cluster head, which makes a final decision according to the local decisions. Value fusion often yields more accurate detection decisions than decision fusion as all the information gathered by a node is considered in the decision making.

In this paper, we adopt a value fusion scheme as follows. Each cluster head makes the detection decision by comparing the sum of all measurements against a threshold $T$. Suppose there are $N$ member sensors in the cluster and $Y$ denotes the sum of all measurements, i.e., $Y = \sum_{i=1}^{N} y_i$. If $Y \geq T$, the cluster head decides 1; otherwise, it decides 0.

### IV. Problem Definition and Summary of Contributions

**A. Notations**

We define the following notations in this paper.

- $\phi(x)$, $\Phi(x)$ and $Q(x)$ represent the probability density function (PDF), the cumulative distribution function (CDF) and the complementary CDF of the standard normal distribution, respectively. Formally, $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $\Phi(x) = \int_{-\infty}^{x} \phi(t) dt$, $Q(x) = 1 - \Phi(x)$.
- $Q^{-1}(x)$ represents the inverse function of $Q(x)$. For a given false alarm rate $\alpha$ and detection probability $\beta$, we use $\alpha'$ and $\beta'$ to represent $Q^{-1}(\alpha)$ and $Q^{-1}(\beta)$, respectively.
- $f_{\text{Poi}}(k|\lambda)$ and $F_{\text{Poi}}(x|\lambda)$ represent the PDF and CDF of the Poisson distribution with mean $\lambda$. Formally, $f_{\text{Poi}}(k|\lambda) = \frac{e^{-\lambda} \lambda^k}{k!}$ and $F_{\text{Poi}}(x|\lambda) = \sum_{k=0}^{\lfloor x \rfloor} f_{\text{Poi}}(k|\lambda)$.
- $\mathbb{E}[x]$ and $\text{Var}[x]$ represent the mean and variance of random variable $x$.
- We use $O(g(x))$, $\Omega(g(x))$ and $\Theta(g(x))$ to represent the asymptotic upper, lower and tight bounds of function $g(x)$\(^1\).

**B. Network Model**

For any physical point $p$, we define the fusion range as the circle of radius $R$ centered at $p$. In order to detect whether a target is present at point $p$, all the sensors within $R$ from $p$ participate in the data fusion. $R$ can be used as an input parameter of a clustering algorithm. Sensors in the network may self-organize into clusters during initialization or form a cluster dynamically around a possible target [5]. In this work, we are interested in the fundamental performance limits of coverage under the data fusion model and the design of clustering algorithms is not the focus of this paper.

Sensors are deployed according to a 2-dimensional stationary Poisson point process of density $\rho$. The number of sensors that falls within fusion range $R$ follows the Poisson distribution with mean equal $\rho \pi R^2$, i.e., $N \sim \text{Poi}(\rho \pi R^2, \rho \pi R^2)$. We assume the density of sensors is high. This assumption is motivated by the fact that most wireless sensors in practice have very limited sensing capability because of cost and power

\[^1\] $f(x) = O(g(x))$ means that $\exists c$ and $\exists x_0$ such that $f(x) \leq c g(x)$ for $x \geq x_0$; 2) $f(x) = \Omega(g(x))$ means that $\exists c$ and $\exists x_0$ such that $f(x) \geq c g(x)$ for $x \geq x_0$; 3) $f(x) = \Theta(g(x))$ means that $f(x) \in O(g(x)) \cap \Omega(g(x))$. 

False alarm rate and detection probability are two major metrics to characterize the performance of a detection system. False alarm rate (denoted by $P_F$) is the probability of making a positive decision when the target is actually absent, and detection probability (denoted by $P_D$) is the probability that a target is correctly detected. Under the aforementioned value fusion scheme, $P_F$ and $P_D$ are given by:

$$P_F = \Pr(Y > T|\mathcal{H}_0), \quad P_D = \Pr(Y > T|\mathcal{H}_1)$$

where $\mathcal{H}_0$ and $\mathcal{H}_1$ represent that the target is absent and present, respectively.
constraints. For instance, vehicle detection experiments based on Mica2 motes showed that four Mica2 motes deployed in a 10 m × 10 m region yields a detection probability of only 40% when the false alarm rate is required to be below 5% [10].

When the density ρ is large enough, we use the normal distribution \( N(\rho \pi R^2, \rho \pi R^2) \) to approximate the Poisson distribution \( \text{Poi}(\rho \pi R^2, \rho \pi R^2) \). In practice, this approximation is accurate when \( \rho \pi R^2 > 20 \) [21]. This condition can be easily met in our problem. Suppose fusion range \( r^2 \) is 20 meters and Mica2 motes are deployed in the same density as in the aforementioned experiments (i.e., 4 motes per 10 m × 10 m area), the number of sensors within the fusion range is about 50.

Fusion range is an important design parameter of our data fusion scheme. As signal energy decays with distance, fusion range lower-bounds the “quality” of information that is fused at a cluster head. We note that similar distance-based fusion schemes have been employed by surveillance sensor networks in practice [7].

For a physical point \( p \), the number of sensors within the fusion range of \( p \) is represented by \( N(p) \). For conciseness, we use \( N \) for \( N(p) \) when the point of interest is clear.

C. Problem Definition

Definition 1 (\((\alpha, \beta)\)-coverage): Given two real number \( \alpha \in [0, 0.5] \) and \( \beta \in [0.5, 1] \), a point \( p \) is covered if the false alarm rate\(^2\) \( P_F \) and detection probability \( P_D(p) \) when a target is present at point \( p \) satisfy

\[
P_F \leq \alpha, \quad P_D(p) \geq \beta
\]

The coverage of a region \( X \), \( c(X) \), is defined as the fraction of points that are covered.

The full coverage of a region refers to the scenario where the coverage of the region approaches 1, i.e., the false alarm rate is below \( \alpha \) and the probability of detecting a target present at any location of a region is above \( \beta \).

We note that the above definition of coverage is consistent with the requirement of mission-critical surveillance applications [10], [27], [9] in which a small false alarm rate (\( \ll 0.5 \)) and a high detection probability (\( \geq 0.5 \)) are often desired. For instance, it is natural for a user to require that “the probability of detecting any intruder is no lower than 95% and no more than 5% of network reports are false alarms”.

D. Main Contributions

The main contributions of this paper are as follows:

1) We derive the coverage under the fusion model as \( c = 1 - F_{p_0}(\Gamma(R) \rho \pi R^2), \) where \( \Gamma(R) = \Theta(\frac{R^2}{m_{\text{opt}}}) \). We also establish the relation between the optimal fusion range and network density as \( \frac{m_{\text{opt}}}{m_{\text{opt}}} = \Theta(\sqrt{\rho}) \). These results can be used to compute the optimal fusion range

2The reliable sensing range of a single Mica2 mote is about 9 meters [11]. The fusion range can be much larger than the sensing range of a node, as detailed later in this paper.

3\( P_F \) is the probability of making a position decision when the target is absent, which is hence not dependent on any location.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Fixed Fusion Range ( R )</th>
<th>Optimal Fusion Range ( R )</th>
<th>SNR (( \delta = \frac{a}{\sigma} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Sensors</td>
<td>( \frac{m}{p_0} = \frac{2\pi a^2}{R^2} )</td>
<td>( \rho = \Theta(\sqrt{p_0}) )</td>
<td>( \frac{m}{p_0} = \Theta(\frac{1}{\delta}) )</td>
</tr>
<tr>
<td>Mobile Sensors</td>
<td>( \rho \leq p_0 )</td>
<td>( \rho = \Theta(\sqrt{p_0}) )</td>
<td>( \frac{m}{p_0} = \Theta(\frac{1}{\delta}) )</td>
</tr>
</tbody>
</table>

TABLE I

NETWORK DENSITIES TO ACHIEVE FULL COVERAGE UNDER THE DATA FUSION AND BOOLEAN MODELS.

and the sufficient density to ensure the desirable level of coverage.

2 We study the fundamental limits of full coverage. We prove that, the ratio of network densities under the fusion and boolean models satisfy \( \frac{m}{p_0} = \frac{2\pi a^2}{R^2} \) where \( r \) is the sensing range of each sensor. When the fusion range \( R \) is optimized with respect to network density, \( \rho = \Theta(\sqrt{p_0}) \). These results show that the fusion model leads to significant density reduction compared to the boolean model while achieving full coverage.

3 We extend a mobile relocation approach [30] to the data fusion model. We show that the mobile sensor density required to achieve full coverage is no greater than that under the boolean model, i.e., \( \rho \leq p_0 \), while the maximum mobile moving distances under the two models remain the same order.

4 We study the impact of SNR on the network density under full coverage. We prove that \( \frac{m}{p_0} = \Theta(\delta) \) when the fusion range is fixed. This result shows that the coverage performance of the boolean and fusion models are comparable only when SNR is very large. However, when the fusion range is optimized with respect to network density, \( \rho = \Theta(\sqrt{p_0}) \). That is, the performance gap between the fusion and boolean models increases with SNR. Similarly, when sensors are mobile, \( \frac{m}{p_0} = \Theta(\frac{1}{\delta}) \).

In summary, Table IV-D compares the network densities required to achieve full coverage under the data fusion and boolean models.

V. COVERAGE UNDER DATA FUSION AND BOOLEAN MODELS

In this section, we first derive the closed-form expression of coverage under the fusion model. We then establish the condition for the optimal fusion range that maximizes the coverage. Finally, we derive the coverage under an extended boolean sensing model that accounts for signal decay and probabilistic detection.

A. Derivation of Coverage under the Data Fusion Model

We first study the condition for the coverage of an arbitrary point in the region. We have the following lemma.

Lemma 1: Consider an arbitrary point \( p \) in a network uniformly distributed in a large region. Denote \( s_i \) as the measurement of an arbitrary sensor within the fusion range
of p. The sufficient condition to ensure the coverage of p is:

$$N \geq \left( \frac{\alpha \sigma - \beta}{\mu_a} \right)^2$$  \hspace{1cm} (3)$$

where N is the number of sensors within R from the position of p, $\mu_a = E[s_i]$ and $\sigma^2_s = \text{Var}[s_i]$.

Proof: When no target is present, the sum of all measurements of sensors within the fusion range of p follows a normal distribution (as the noise measured by each sensor follows a normal distribution),

$$Y|H_0 = \sum_{i=1}^{N} n_i \sim N(N\mu, N\sigma^2)$$

Thus, the false alarm rate can be calculated by

$$P_f = Pr(Y > T)|H_0 = Q\left( \frac{T - N\mu}{\sqrt{N}\sigma} \right)$$  \hspace{1cm} (4)$$

where T is the detection threshold. To meet the upper bound of the false alarm rate, we set $P_f = \alpha$. According to (4), the detection threshold can be derived as $T = N\mu + \sqrt{N}\sigma\alpha'$. When a target is present at point p, $Y|H_1 = \sum_{i=1}^{N} s_i + \sum_{i=1}^{N} n_i$. The measurement of each sensor, $s_i$, is independent and identically-distributed (i.i.d.). As network density is high (as assumed in Section IV-B), $\sum_{i=1}^{N} s_i$ approaches the normal distribution according to the Central Limit Theorem (CLT). Formally, $\sum_{i=1}^{N} s_i \sim N(N\mu_s, N\sigma^2_s)$. Accordingly,

$$Y|H_1 \sim N(N\mu_s + N\mu, N\sigma^2 + N\sigma^2)$$

The detection probability at point p is

$$P_D = Pr(Y > T)|H_1 = Q\left( \frac{T - N\mu_s - N\mu}{\sqrt{N}\sigma} \right)$$

$$= Q\left( \frac{N\mu + \sqrt{N}\sigma\alpha' - N\mu_s - N\mu}{\sqrt{N}\sigma + \sigma^2_s} \right)$$

$$= Q\left( \frac{\sigma\alpha' - \sqrt{N}\mu_s}{\sqrt{\sigma^2 + \sigma^2_s}} \right)$$  \hspace{1cm} (5)$$

According to (5), in order to meet the performance requirement $P_D \geq \beta$, N must be lower-bounded. It can be easily verified that the condition of N is given by (3).

Because sensor deployment follows a stationary Poisson point process, the fraction of the area being covered equals the probability that an arbitrary point is covered by the network. Therefore, we can derive the coverage provided by a random network based on Lemma 1.

**Theorem 1:** When sensors are deployed according to a 2-dimensional Poisson point process of density $\rho$, the coverage is given by

$$c = 1 - F_{\rho\mu}(\Gamma(R)|\rho\pi R^2)$$  \hspace{1cm} (6)$$

where

$$\Gamma(R) = \left( \frac{\alpha' \rho \pi R^2 - \beta' \sqrt{\rho \pi R^2 \ln(1 + R^2) + \frac{\pi R^2}{2}}}{\ln(1 + R^2)} \right)^2$$

Proof: We first derive the mean value and variance of $s_i$, the measurement of an arbitrary sensor i within the fusion range of point p. Suppose the coordinates of point p and sensor i are denoted by $(x_p, y_p)$ and $(x_i, y_i)$, respectively. As only the sensors within the fusion range participate in the data fusion, the PDF of $(x_i, y_i)$ is

$$f(x_i, y_i) = \frac{1}{\pi R^2}, \quad (x_i - x_p)^2 + (y_i - y_p)^2 \leq R^2$$

The CDF and PDF of the distance between sensor i and point p, $r_i$, are

$$F(r_i) = \int_0^{2\pi} d\theta \int_0^{r_i} \frac{1}{\pi R^2} \cdot r dr = \frac{r_i^2}{R^2}, \quad 0 \leq r_i \leq R$$

$$f(r_i) = \frac{dF(r_i)}{dr_i} = \frac{2r_i}{R^2}, \quad 0 \leq r_i \leq R$$

According to (1), the derivative of $r_i$ with respect to $s_i$ is

$$\frac{dr_i}{ds_i} = \frac{a}{2} \left( \frac{a}{s_i} - 1 \right)^{-\frac{1}{2}} \frac{1}{s_i}$$

Accordingly, the PDF of $s_i$ is

$$f(s_i) = f(r_i) \frac{dr_i}{ds_i} = \frac{a}{2R^2} \cdot \frac{1}{s_i} \cdot \frac{a}{1 + R^2} \leq s_i \leq a$$

The mean value of $s_i$ is

$$\mu_s = \int_0^{a} s f(s_i) ds_i = \frac{a^2}{2R^2} \ln(1 + R^2)$$  \hspace{1cm} (7)$$

And variance of $s_i$ is

$$\sigma^2_s = \int_0^{a} (s - \mu_s)^2 f(s_i) ds_i = \frac{a^2}{1 + R^2} - \frac{a^2}{R^2} \ln^2(1 + R^2)$$

We can obtain the condition to ensure the coverage of point p by substituting $\mu_s$ and $\sigma^2_s$ in (3). One can easily verify that the condition is $N \geq \Gamma(R)$. As N is a Poisson random variable, the expected coverage is

$$c = Pr(\text{point p is covered})$$

$$= Pr(N \geq \Gamma(R))$$

$$= 1 - F_{\rho\mu}(\Gamma(R)|\rho\pi R^2)$$

According to Theorem 1, one can compute the network density that achieves the desirable level of coverage under the fusion model. When the network density is high, the normal distribution is an excellent approximation to the Poisson distribution. Therefore, coverage c can be approximated by

$$c = 1 - \Phi\left( \frac{\Gamma(R) - \rho\pi R^2}{\sqrt{\rho\pi R^2}} \right)$$  \hspace{1cm} (8)$$

Figure 3 plots the coverage computed under both distributions. From Figure 3, we can see that the coverage initially increases with fusion range R, and approaches 1 when R is between 2 to 8, but decreases to 0 eventually. Intuitively, as the fusion range increases, more sensors contribute to the data fusion resulting in better sensing quality. However, as the fusion range becomes very large, the aggregate noise starts to cancel out the benefit because the target signal decreases exponentially with the distance from the target. In other words, the measurements of sensors far away from the target contain no useful information and hence fusing them leads to lower detection performance. An important problem is how to choose the fusion range $\hat{R}$ to maximize the coverage for a given network density. We study the optimal fusion range in the next section.
In the boolean sensing model, each sensor has a sensing range \( r \). A location is covered if it lies within at least one sensor’s sensing range. The coverage under the boolean model has been studied based on the stochastic geometry theory \([20]\), the stochastic nature of sensing. In the following, we describe an extended boolean model that accounts for signal decay and probabilistic detection.

For any sensor, we choose an appropriate detection threshold (as detailed later) such that 1) its false alarm rate is no greater than \( \alpha \); and 2) the probability of detecting any target present in its sensing range \( r \) is no lower than \( \beta \). Under this model, if any point in a region falls within the sensing range \( r \) of at least one sensor (e.g., the region is regarded as being covered under the boolean model), the probability of detecting any target present in the region is no lower than \( \beta \). At the same time, the false alarm rate of the network is no greater than \( \alpha \).

The relation between the optimal fusion range \( R_{\text{opt}} \) and node density is \( R_{\text{opt}} = \Theta(\sqrt{\rho}) \).

According to Theorem 2, the optimal fusion range \( R_{\text{opt}} \) increases with network density \( \rho \). Figure 4 plots the optimal fusion range (solved from (27)) versus the network density.

**B. Optimal Fusion Range**

Eq. (6) shows that the coverage depends on \( \Gamma(R) \). We first study the properties of \( \Gamma(R) \). We have the following lemma.

**Lemma 2:** The upper and lower bounds of \( \Gamma(R) \) are

\[
\Delta_1 = \frac{\alpha' - \beta'}{\delta} \sqrt{1 + \frac{1}{\beta^2}}, \quad \Delta_2 = \frac{\alpha' - \beta'}{\delta}
\]

The tight bound of \( \Gamma(R) \) is

\[
\Gamma(R) = \Theta\left(\frac{R^4}{\ln^2(1 + R^2)}\right)
\]

Based on the bounds of \( \Gamma(R) \), we can find the relation between the optimal fusion range and network density. We have the following theorem (the proof is given in Appendix B).

**Theorem 2:** The relation between the optimal fusion range \( R_{\text{opt}} \) and node density is

\[
R_{\text{opt}} = \Theta(\sqrt{\rho})
\]

The tight bound of \( \Gamma(R) \) is

\[
\Gamma(R) = \Theta\left(\frac{R^4}{\ln^2(1 + R^2)}\right)
\]

The upper and lower bounds of \( \Gamma(R) \) are

\[
\Delta_1 = \frac{\alpha' - \beta'}{\delta} \sqrt{1 + \frac{1}{\beta^2}}, \quad \Delta_2 = \frac{\alpha' - \beta'}{\delta}
\]

Therefore, the definition of coverage under the boolean model. This extension allows us to study the limitation of the boolean model on providing the probabilistic coverage. We now discuss how to choose the detection threshold of each sensor.

For sensor \( i \), its detection decision is to test the following hypotheses:

\[
H_0 : y_i \sim N(\mu, \sigma^2) \\
H_1 : y_i \sim N(\mu + s_i, \sigma^2)
\]

The optimal decision rule of above hypothesis testing is the Likelihood Ratio Test \([27]\), in which sensor \( i \) compares its measurement \( y_i \) against a detection threshold \( t \). If \( y_i \) exceeds \( t \), sensor \( i \) makes a positive decision; otherwise, it makes a negative decision. Thus, the false alarm rate \( P_F \) and detection probability \( P_D \) are given by

\[
P_F = \Pr(y_i > t | H_0) = Q\left(\frac{t - \mu}{\sigma}\right)
\]

\[
P_D = \Pr(y_i > t | H_1) = Q\left(\frac{t - \mu - s_i}{\sigma}\right)
\]

To satisfy the requirements of coverage, we set \( P_F = \alpha \) and \( P_D \geq \beta \). The detection threshold \( t \) can then be solved as

\[
t = \mu + \sigma \alpha'
\]

By replacing \( t \) and \( s_i \) (defined by (1)) in (13), we have:

\[
r \leq \sqrt{\frac{\delta}{\alpha' - \beta'}} - 1 = \sqrt{\frac{1}{\Delta_2} - 1}
\]

where \( \Delta_2 \) is defined in (9). Note that the requirements of detection performance should satisfy \( \alpha' - \beta' < \delta \). Eq. (14) shows that the sensing range of a sensor depends on SNR and the user requirements. SNR depends on the property of target, noise level of environment and the sensitivity of sensors. As a result, the sensing range of a sensor varies. In the vehicle detection experiments based on Mica2 motes \([10]\), the average SNR is about 50. Thus, if \( \alpha = 5\% \) and \( \beta = 50\% \), \( r \) is about 5.4 meters and if the required \( \beta \) increases to 95\%, \( r \) decreases to 3.8 meters. When the noise level of the environment increases (e.g., due to wind), \( r \) can be reduced significantly.

**VI. NETWORK DENSITY FOR ACHIEVING FULL COVERAGE**

In this section, we explore the fundamental limits of full coverage under both data fusion and boolean sensing models. In particular, we compare the minimum network densities under the two sensing models when the coverage approaches one. The results help to understand the limitation of the
boolean model and quantify the advantage of collaboration among sensors under the data fusion model.

A. Full Coverage using Fixed Fusion Range

According to Eqs. (6) and (11), when coverage is equal to one, the sensor densities under both the fusion and boolean models approach to infinity. However, the speed that network density increases with coverage is different under two models. In this section, we study the ratio of sensor densities required to achieve full coverage under the boolean model and the data fusion with a fixed fusion range. We have the following theorem.

Theorem 3: Suppose $\rho$ and $\rho_0$ denote the sensor densities required to achieve the same coverage $c$ under the data fusion and boolean models, respectively. For a given fusion range $R$, the ratio of $\rho$ to $\rho_0$ satisfies $\lim_{c \to 1} \frac{\rho}{\rho_0} = \frac{2}{\pi R}$.

Proof: In order to provide full coverage, $\rho$ approaches to infinity. Hence, we can adopt the normal approximation of $c$ under the fusion model:

$$c = 1 - \Phi \left(\frac{\Gamma(R) - \rho \pi R^2}{\sqrt{\rho \pi R^2}}\right)$$

$$= Q \left(\frac{\Gamma(R) - \frac{1}{\sqrt{\rho}}}{\sqrt{\frac{1}{\sqrt{\rho}}}}\right)$$

(15)

When $\rho \to +\infty$, the second item in (15), $-\sqrt{\pi R \cdot \rho}$, dominates, because

$$\lim_{\rho \to +\infty} \frac{\Gamma(R)}{\sqrt{\rho \pi R \cdot \rho}} = 0$$

Accordingly,

$$c = Q \left(-\sqrt{\pi R \cdot \rho}\right), \quad \rho \to +\infty \quad (16)$$

We now compare the sensor densities required by the fusion model and the boolean model, when the coverage approaches 1. Define $x = Q^{-1}(c)$, according to (16), we have:

$$\rho = \frac{1}{\pi R^2} x^2$$

And according to (11), we have:

$$\rho_0 = \frac{1}{\pi} \ln \frac{1}{1-c} = \frac{1}{\pi} \ln \frac{1}{\Phi(x)}$$

Note that $\Phi'(x) = \phi(x)$ and $\phi'(x) = -x\phi(x)$. We have:

$$\lim_{x \to +\infty} \frac{\rho}{\rho_0} = \lim_{x \to +\infty} \frac{\frac{1}{\pi R^2} x^2}{\frac{1}{\pi} \ln \frac{1}{\Phi(x)}}$$

$$= \frac{r^2}{\pi^2} \lim_{x \to +\infty} \frac{2x}{\sqrt{x^2 + 1}} \quad (\ast)$$

$$= \frac{2r^2}{\pi^2} \lim_{x \to +\infty} \frac{x\phi(x)}{\phi(x)}$$

$$= \frac{2r^2}{\pi^2} \lim_{x \to -\infty} \frac{\Phi(x) + x\phi(x)}{-x\phi(x)} \quad (\ast)$$

$$= \frac{2r^2}{\pi^2} \left[\lim_{x \to -\infty} \frac{\phi(x)}{x - 2\phi(x)} + 1\right] \quad (\ast)$$

$$= \frac{2r^2}{\pi^2} \left[\lim_{x \to -\infty} \frac{1}{1 - x^2} + 1\right]$$

$$= \frac{2}{\pi R}$$

where the derivation steps marked by (\ast) follow from the L' Hospital’s Rule\(^4\). Note that $\lim_{x \to -\infty} x\phi(x) = 0$ and $\lim_{x \to +\infty} x\phi(x) = 0$.

Theorem 3 shows that $\rho < \rho_0$ if $R > \sqrt{2}r$. From Theorem 2, we know that the optimal fusion range increases with network density $\rho$. As $\rho$ approaches to infinity when full coverage is required, the optimal fusion range also approaches to infinity accordingly. On the other hand, according to Eq. (14), sensing range $r$ is a constant independent of network density. Therefore, the condition $R > \sqrt{2}r$ is easily satisfied under the full coverage. For instance, the acoustic sensor on Mica2 motes has a sensing range of 3~5 meters if a high performance (e.g., $\alpha = 5\%$ and $\beta = 95\%$) is required [10]. On the other hand, the fusion range can be set to be much larger. Figure 4 plots $R_{opt}$ (numerically computed from (27)) and $r$ versus network density.

B. Full Coverage using Optimal Fusion Range

As shown in Theorem 2, the optimal fusion range increases with network density. Data fusion based on the optimal fusion range allows the maximum number of informative sensors to contribute to the final detection decision. An interesting question is “what is the minimum network density to achieve full coverage when the optimal fusion range is always used?”.

However, the answer to this question is challenging because the closed-form expression of optimal fusion range is difficult to derive. The following theorem shows that $\rho$ reduces to $O(\sqrt{\rho_0})$ as long as the fusion range is order-optimal (i.e., in the same order of the optimal fusion range).

Theorem 4: When fusion range $R$ is order-optimal, $\rho = O(\sqrt{\rho_0})$ when coverage approaches 1.

Proof: We define the following notations:

$$h_1(\rho, R) = \frac{\Gamma(R)}{\sqrt{\pi R}} \frac{1}{\sqrt{\rho}}$$

$$h_2(\rho, R) = \sqrt{\pi R} \cdot \sqrt{\rho}$$

According to (15), the coverage $c = Q(h_1(\rho, R) - h_2(\rho, R))$. As both $h_1(\rho, R)$ and $h_2(\rho, R)$ are positive, the following inequality is a sufficient condition to ensure $c \to 1$:

$$\lim_{\rho \to \infty, R \to \infty} \frac{h_2(\rho, R)}{h_1(\rho, R)} = \lim_{\rho \to \infty, R \to \infty} \frac{\rho \pi R^2}{\Gamma(R)} > 1 \quad (17)$$

We choose $R$ to satisfy:

$$\frac{\Delta_3^2}{\pi^2} \ln^2(1 + R^2) = \rho \quad (18)$$

where $\Delta_3$ is a constant which satisfies $\Delta_3 > \Delta_1$ such that

$$\lim_{\rho \to \infty, R \to \infty} \frac{\rho \pi R^2}{\Gamma(R)} = \lim_{R \to \infty} \frac{\Delta_3^2 R^4}{\ln^2(1 + R^2)} > 1$$

\(^4\)L’Hôpital’s Rule says that $\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}$, if $\lim f(x)$ and $\lim g(x)$ are both 0 or $\pm\infty$ and $\lim \frac{f'(x)}{g'(x)}$ has a finite value.
Therefore, Eq. (18) guarantees inequality (17). The order of the fusion range $R$ in (18) is the same as the optimal fusion range $R_{opt}$ given by Theorem 2. We denote $x = \frac{\beta^2}{\ln^2(1 + R^2)}$ in the following derivations. Let the coverage achieved under the data fusion and boolean models be the same:

\[
1 - e^{-\rho_b \pi r^2} = Q\left(\frac{\Gamma(R)}{\sqrt{\pi R}} \cdot \frac{1}{\sqrt{2 \Delta_3^2}} - \frac{\sqrt{\pi R} \cdot \sqrt{2 \Delta_3^2}}{\pi}\right) = 1 - \Phi\left(\frac{\Gamma(R)}{x \Delta_3} - x \Delta_3\right)
\]

Solving $\rho_b$ from (19):

\[
\rho_b = -\frac{1}{\pi^2} \cdot \ln \Phi\left(\frac{\Gamma(R)}{x \Delta_3} - x \Delta_3\right)
\]

Accordingly,

\[
\lim_{c \to 1} \frac{\rho_b}{\rho} = \lim_{R \to +\infty} \frac{\frac{x^2 \Delta_3^2}{\pi^2} - \frac{x^2 \Delta_3^2}{x \Delta_3}}{\frac{1}{\pi^2} \cdot \ln \Phi\left(\frac{\Gamma(R)}{x \Delta_3} - x \Delta_3\right)} \leq \lim_{R \to +\infty} \frac{\frac{x^2 \Delta_3^2}{\pi^2} - \frac{x^2 \Delta_3^2}{x \Delta_3}}{\frac{1}{\pi^2} \cdot \ln \Phi\left(\frac{\Gamma(R)}{x \Delta_3} - x \Delta_3\right)} = \frac{r^2 \Delta_3^2}{\pi} \cdot \ln \Phi(kx)
\]

Therefore, if the optimal fusion range is chosen, we have

\[
\lim_{c \to 1} \rho_b \leq \frac{2r^2 \Delta_3^2}{k^2 \pi} = \Theta\left(\frac{1}{\delta^2}\right)
\]

Therefore, if $\delta < (\alpha' - \beta') (\frac{\Delta_3^2}{3} + 1)$, $\rho < \rho_b$ when $c \to 1$; otherwise, $\rho > \rho_b$. This observation shows, for a fixed $R$, the fusion model is more effective than the boolean model when SNR is low (e.g., in a noisy environment). On the other hand, if SNR is high, it suffices to employ the boolean model. In practice, the SNRs in many application scenarios are low. For instance, the SNR is about 50 in the vehicle detection experiments based on Mica2 motes [10]. Under such a setting, $\rho < \rho_b$ holds if $R > 6m$.

We now analyze the case of optimal fusion range. Recall $\Delta_1 = \frac{\alpha'}{\beta'} - \frac{\beta'}{\alpha'}$, $\Delta_3$ and $k^2$ in (20) with their orders with respect to $\delta$, we have

\[
\rho = \mathcal{O}\left(\sqrt{\frac{1}{\delta^2}}\right)
\]

Therefore, if the optimal fusion range is chosen, we have

\[
\lim_{c \to 1} \rho = \mathcal{O}\left(\sqrt{\frac{\rho_b}{\delta}}\right)
\]

VII. COVERAGE USING MOBILE SENSORS

It has been shown that random network deployments can lead to undesirable over-provision of sensing coverage [30]. Under the boolean model, many areas in a fully covered region have far more than one sensor covering them. Recent work [30], [29], [4] has exploited mobility to reduce the network density for achieving full coverage under the boolean model. In such a scheme, randomly distributed mobile sensors can relocate themselves to fill coverage holes in the initial network deployment. In this section, we extend an existing mobile relocation algorithm [30] to the data fusion model. Our analysis shows that the mobile sensor density required to achieve full coverage is no greater than that under the boolean model, i.e., $\rho \leq \rho_b$. Moreover, the ratio between $\rho_b$ and $\rho$ increases with SNR, i.e., $\rho_b/\rho = \Theta(\delta)$. At the same time, the maximum mobile moving distance under the fusion model remains the same order as under the boolean model.
A. Mobile Relocation Algorithm

We now briefly describe a mobile relocation algorithm for achieving \( k \)-coverage proposed in [30] and our extension to it. A region is \( k \)-covered if every point in the region is within the sensing range of at least \( k \) sensors. As the notion of \( k \)-coverage is not applicable to our problem, we assume \( k = 1 \) in the following discussion. Suppose a mobile sensor network is randomly deployed in a region of area \( L \) divided by grids of size \( d_b \). After the initial deployment, the mobile sensors relocate themselves such that each grid point has exactly one mobile on it. It is shown in [30] that full coverage can be achieved by setting the grid length to \( d_b = \sqrt{2}r \). The maximum moving distance of mobile is bounded as as \( O(\sqrt{2}r \log^\frac{3}{2} \frac{L}{\sqrt{2}r}) \).

We extend the above relocation algorithm to the data fusion model as follows. The new algorithm divides the region into grids of width \( d \). Similar to the original algorithm, the mobile sensors relocate themselves such that each grid point has exactly one mobile. We choose grid width \( d \) such that at least \( M \) sensors fall into the fusion range of any point in the region. This is different from the original algorithm that ensures any point in the region to fall in the sensing range of at least one sensor. Fig. 5 illustrates the two relocation algorithms.

B. Maximum Mobile Moving Distance

In this section, we first derive the grid width \( d \) in the relocation algorithm. We then show that the algorithm achieves the same order of maximum mobile moving distance as the original algorithm. The following lemma from stochastic geometry theory [15] is used in our analysis.

**Lemma 3 ([15]):** Suppose \( W(R) \) grid points of side width \( d \) are covered by a disc of radius \( R \) centered at an arbitrary point. Then \( W(R) \geq \frac{\pi((\sqrt{2}R-d)^2)}{2d^2} \).

We now derive the condition for the relocation algorithm to achieve full coverage.

**Theorem 5:** The full coverage is achieved after the relocation of mobile sensors, if the grid width satisfies \( d \leq \frac{\sqrt{2}R}{\sqrt{\Delta_2 (2 + \pi) + 1}} \), where \( \Delta_2 \) is defined in (9).

**Proof:** After the relocation, the network has a regular rather than a random deployment. Suppose \( N \) grid points are covered by a fusion disc centered at an arbitrary point \( p \). In order to meet the upper bound of the false alarm rate, we set the detection threshold \( T = N\mu + \sqrt{N}\sigma\alpha' \). Accordingly, the detection probability

\[
P_D = Pr(Y > T|\theta_1) = Q \left( \frac{T - \sum_{i=1}^{N} s_i - N\mu}{\sqrt{N}\sigma} \right)
\]

\[
= Q \left( \frac{N\mu + \sqrt{N}\sigma\alpha' - N\bar{s} - N\mu}{\sqrt{N}\sigma} \right) = Q \left( \alpha' - \sqrt{N}\bar{s}/\sigma \right)
\]

where \( \bar{s} \) is the average signal energy measured by all sensors on the grid points in the fusion disc. Note that \( \bar{s} \) depends on the position of point \( p \). To meet the performance requirement, i.e., \( P_D \geq \beta \), we have

\[
N \geq \left( \frac{\sigma(\alpha' - \beta')}{\bar{s}} \right)^2 = \frac{a^2\Delta_2^2}{\bar{s}^2}
\]

As \( \bar{s} \) depends on the position of \( p \), we denote \( \bar{s}_{\min} = \min\{\bar{s}\} \) as the minimum \( \bar{s} \) subject to all positions in a grid (including grid sides). Let \( M \) be

\[
M = \frac{a^2\Delta_2^2}{\bar{s}_{\min}^2}
\]

If the number of sensors within any fusion disc is greater than \( M \), the network provides full coverage after relocation. According to Lemma 3, the number of sensors (i.e., grid points) within any disc of radius \( R \) is lower bounded. Therefore, the following inequality should be satisfied:

\[
\frac{\pi(R - \frac{1}{\sqrt{2}}d)^2}{d^2} \geq M
\]

Accordingly, from (21) and (22), \( d \) must satisfy:

\[
d \leq \frac{R}{\sqrt{\bar{s}_{\min}^2 + \frac{1}{\sqrt{2}}}}
\]

In the worst case, all sensors within the fusion disc are on the perimeter, \( \bar{s} \) equals \( \frac{a^2\Delta_2^2}{\sqrt{2}} \). Accordingly, \( \bar{s}_{\min} \geq \frac{a^2\Delta_2^2}{\sqrt{2}} \).

Therefore, after replacing \( \bar{s}_{\min} \) by its lower bound in (23), the sufficient condition to guarantee full coverage can be obtained as \( d \leq \frac{\sqrt{2}R}{\sqrt{\Delta_2 (2 + \pi) + 1}} \).

It is shown that the relocation algorithm based on the boolean model bounds the maximum moving distance of mobile nodes as \( O(\sqrt{2}r \log^\frac{3}{2} \frac{L}{\sqrt{2}r}) \). We now show that the maximum mobile moving distance under the fusion model remains the same order with respect to the network size \( L \) as under the boolean model. The following discussion is similar to the analysis in [30]. The main difference is that our relocation algorithm adopts a different grid width given by Theorem 5. The analysis is based on the minimax grid matching theorem from [17]. When \( L = l^2 \) points are randomly and independently scattered in the region according to a uniform distribution, with high probability, there exists a perfect match between the \( L \) points and the \( L \) grid points in which the maximum distance between any matched pairs is \( O(\log^\frac{3}{2} L) \).

The total number of grid points is \( L/l^2 \) in our network instead of \( L \) and grid width is \( d \) instead of 1. Therefore, the maximum moving distance of mobiles is \( O(d \log^\frac{3}{2} L) \), which has the same order \( O(\log^\frac{3}{2} L) \) with respect to the network size \( L \) as under the boolean model.

C. Mobile Sensor Density for Achieving Full Coverage

In this section, we compare the sensor densities required by the two sensing models to guarantee full coverage after mobile sensor relocation.

**Lemma 4:** Suppose \( \rho_b \) and \( \rho \) are the minimum sensor densities required by the relocation algorithms under the boolean and data fusion models to achieve full coverage, respectively. There exists a fusion range \( R \) such that \( \rho \leq \rho_b \).

**Proof:** In [30], the grid side width \( d_b \) is chosen to be \( \sqrt{2}r \) to guarantee full coverage under the boolean model. By replacing \( r \) with \( \sqrt{1/\Delta_2 - 1} \) according to (14), the network density required is

\[
\rho_b = \frac{1}{d_b^2} = \frac{1}{2 (\Delta_2 - 1)}
\]
As shown in [30], when the grid side width is set to be $\sqrt{2}R/\sqrt{M}$, the number of sensors that are within distance $R$ from any point is no smaller than $M$. Accordingly, the network density under the fusion model is

$$\rho = \frac{1}{a^2} = \frac{M}{2R^2} = \frac{2a^2\Delta_2^2}{2R^2\delta_{\min}}$$

As $\delta_{\min} \geq \frac{a}{1 + R^2}$ always holds, the network density ratio satisfies:

$$\frac{\rho_0}{\rho} = \frac{R^2\delta_{\min}}{a^2} \geq \frac{1}{\Delta_2(1 - \Delta_2)} \geq \frac{2a^2\Delta_2^2}{2R^2\delta_{\min}}$$

The following inequality guarantees that $\rho \leq \rho_0$ provided $\frac{R}{1 + R^2} \geq \sqrt{\Delta_2(1 - \Delta_2)}$.

$$\rho \leq \rho_0$$

Assume $\Delta_2 \in (0, 1)$, we have $\sqrt{\Delta_2(1 - \Delta_2)} \in (0, \frac{1}{2}]$. Moreover, since $\frac{R}{1 + R^2} \in (0, \frac{1}{4}]$, there must exist a $R$ which satisfies (25) for any $\Delta_2$. Consequently, by choosing proper fusion range $R$, $\rho \leq \rho_0$ always holds.

Theorem 4 shows that $\rho_0/\rho \geq 1$ when the fusion range $R$ is chosen appropriately. This result is derived under the condition that $\Delta_2 = \frac{\alpha' - \beta'}{\alpha' - \beta}$ can be an arbitrary number within $(0, 1)$. Note that $\delta > \alpha' - \beta'$ is necessary for any system to be able to detect a target. Once fusion range $R$ is fixed, the density ratio depends on $\frac{1}{\Delta_2(1 - \Delta_2)}$. When $\delta = 2(\alpha' - \beta')$, i.e., $\Delta_2 = \frac{1}{2}$, the density ratio $\frac{\rho_0}{\rho}$ has the minimum value. If $\delta > 2(\alpha' - \beta')$, the density ratio increases with $\delta$; otherwise, it decreases with $\delta$. In practice, $\delta$ is much larger than $\alpha' - \beta'$. For instance, $2(\alpha' - \beta') \approx 6.6$ if $\alpha = 0.05$ and $\beta = 0.95$ while $\delta$ is about 50 in several vehicle detection experiments based on Mica2 motes [10]. We now study the impact of SNR on the density of mobile sensors for achieving full coverage.

**Theorem 6:** If $\delta > 2(\alpha' - \beta')$, the density ratio satisfies $\frac{\rho_0}{\rho} = \Omega(\delta)$ when full coverage is required after relocation of mobile sensors.

**Proof:** By replace $\Delta_2$ with $\alpha' - \beta'$ in (24), we have

$$\frac{\rho_0}{\rho} \geq \left(\frac{R}{1 + R^2}\right)^2 \frac{1}{\alpha' - \beta'} \frac{\delta}{\delta'}$$

As $\Delta_2 \in (0, 1)$, we have $\sqrt{\Delta_2(1 - \Delta_2)} \in (0, \frac{1}{2}]$. Moreover, since $\frac{R}{1 + R^2} \in (0, \frac{1}{4}]$, there must exist a $R$ which satisfies (25) for any $\Delta_2$. Consequently, by choosing proper fusion range $R$, $\rho \leq \rho_0$ always holds.

**Theorem 6:** If $\delta > 2(\alpha' - \beta')$, the density ratio satisfies $\frac{\rho_0}{\rho} = \Omega(\delta)$ when full coverage is required after relocation of mobile sensors.

**Proof:** By replace $\Delta_2$ with $\alpha' - \beta'$ in (24), we have

$$\frac{\rho_0}{\rho} \geq \left(\frac{R}{1 + R^2}\right)^2 \frac{1}{\alpha' - \beta'} \frac{\delta}{\delta'}$$

Accordingly, $\frac{\rho_0}{\rho} = \Omega(\delta)$ holds.

**VIII. Simulations**

In this section, we carry out several sets of simulations to validate our analytical results.

**A. Simulation Settings**

In each set of simulations, the mean of Gaussian noise, $\mu$, is set to be 1, and the variance, $\sigma^2$, is also set to be 1. The detection performance requirements, $\alpha$, $\beta$, are set to be 0.05 and 0.95, respectively. The energy of target, $a$, is set to be 4. The fusion range $R$ is set to be 5m except in Figure 7.

**B. Coverage of Static Networks**

Due to the stationarity of the Poisson point process, it suffices to consider an arbitrary point in the region for evaluating the area coverage of the whole network. Without loss of generality, we assume that the target is located at the origin of the region. To eliminate the edge effect, we deploy square networks with size of $4R \times 4R$ centered at the origin.

Our simulations evaluate the impact of several practical issues. For instance, the location of target may not be accurately localized in practice. As a result, the set of sensors that actually participate in data fusion may be different from the set of sensors within the fusion range from target location.
However, this issue has little impact on our asymptotic results that are derived under the condition of full coverage. In such a case, both the fusion range and the number of nodes within the fusion range approach to infinity and hence a constant localization error does not change our results.

We simulate location errors as follows. Every time the target appears, we randomly pick an angle $\theta$ that follows an uniform distribution, i.e., $\theta \sim \text{Unif}(0, 2\pi)$. The localized target position is at point $p' (\epsilon \cos \theta, \epsilon \sin \theta)$. That is, the sensors within $R$ from $p'$ participate in the fusion. Each sensor samples 1000 times and the measurements of all sensors are fused to make a final detection decision. The detection threshold is set according to the upper bound of false alarm rate $\alpha$ as discussed in Section V. The number of network deployments is 200. For each deployment, the detection probability is computed as the ratio of the number of successful detections to the total number of samples. The coverage is computed as the ratio of the number of network deployments whose detection probability exceeds the requirement $\beta$ to the total number of networks.

Figure 6 plots the analytic and measured coverage versus network density, "sim($\epsilon$)" stands for the coverage measured in the simulations under given location error $\epsilon$. The maximum location error is $2m$, which is 40% of the fusion range in our setting. We can see that the simulation result matches well the analytic result, which validates our derivations. A network density of 0.8 is enough to provide high coverage under the fusion model.

Figure 6 shows that the coverage decreases when there is a location error. However, the location error has little impact on the speed of convergence to full coverage. Figure 7 plots the analytic and measured coverage versus the fusion range. The network density is set to be 1. In this figure, the solid curve represents the analytic coverage predicted by Theorem 1. The fluctuation of the curve is due to the property of CDF of the Poisson distribution. We can see that the derived optimal fusion range matches the measured value.

Figure 8 plots the ratio of sensor densities under different sensing models that are required to achieve the same coverage. The SNR is set to be 4. We can see that the location error has little impact on the density ratio. The ratio increases to 4 when the detection probability is required to be above 90%. The ratio is larger than 10 when near 100% detection probability is required. Figure 9 plots the density ratio when SNR is 50 and 100. In such a case, the sensing range under the boolean model is $3.77m$ and $5.42m$, respectively. We can see that the density ratio increases to 4 when the required detection probability approaches to 1. However, when the required detection probability is small, the network densities under the boolean and fusion models are similar. This is due to the conservative setting of the fusion range (about 5m). When the fusion range is large, the density ratio yields a similar behavior as the case of small SNRs (as shown in Figure 8).

C. Coverage of Mobile Networks

In this set of simulation, we measure the required mobile sensor densities to achieve full coverage. We deploy a grid network in which a sensor is located at each grid point. This scenario simulates the network after mobile sensors relocation to grid points. We choose the grid side width $d$ under the fusion model according to (23). The target appears at a point $p$ randomly chosen in a particular grid. The sensors with $R$ from $p$ sample for 1000 times. Point $p$ is covered if the detection probability exceeds $\beta$. We check 900 points uniformly scattered in the grid. The coverage is computed as the ratio of the number of covered points to 900.

Figure 10 plots the mobile density ratio versus the fusion range under different SNRs. In the settings, $\delta < 2(\alpha' - \beta')$. We can see that the density ratio $\frac{\rho}{p}$ increases when SNR becomes lower, which suggests that the data fusion model improves in presence of low SNR. This is consistent with our analysis in Section VII-C. When the value of SNR is higher, Figure 11 shows that the density ratio $\frac{\rho}{p}$ increases with SNR, which is consistent with our result $\frac{\rho}{p} = O(\frac{1}{\delta})$.

IX. Conclusion

In this paper, we explore the fundamental limits of coverage in large-scale sensor networks based on a data fusion model. We derive several network density bounds for achieving full coverage for both static and mobile sensor networks and compare against the existing results based on the boolean sensing model. Our results show that the data fusion model can achieve the same level of coverage using significantly fewer sensors by exploiting the collaboration among sensors. Our results also provide insights into the design of large-scale data fusion based sensor networks by quantifying the relation between various system parameters including coverage, network density, false alarm rate, detection probability and signal to noise ratio. In the future, we will study the data fusion based coverage in other network scenarios such grid and uniform network deployments. This work is conducted based on a particular data fusion scheme. In the future, we plan to extend our results to different data fusion schemes.

REFERENCES

APPENDIX

A. Proof of Lemma 2

Proof: As $\beta > 0.5$ (according to the definition of coverage in Section IV-C), $\tilde{\beta} < 0$. Accordingly,

\[
\Gamma(R) = \left(\frac{\alpha^2 R^2 - \beta^2 \sqrt{R^4} \ln(1 + R^2) + \frac{1}{\alpha} R^4}{\ln(1 + R^2)}\right)^2
\]

< \left(\frac{\alpha^2 R^2 - \beta^2 \sqrt{R^4} \ln(1 + R^2) + \frac{1}{\alpha} R^4}{\ln(1 + R^2)}\right)^2

= \left(\frac{\alpha^2 - \beta^2 \sqrt{R^4} \ln(1 + R^2) + \frac{1}{\alpha} R^4}{\ln(1 + R^2)}\right)^2

= \left(\frac{\alpha^2 - \beta^2 \sqrt{R^4} \ln(1 + R^2) + \frac{1}{\alpha} R^4}{\ln(1 + R^2)}\right)^2

= \Delta_i^2 \frac{R^4}{\ln^2(1 + R^2)}

Moreover, according to (3),

\[
\Gamma(R) = \left(\frac{\alpha^2 - \sqrt{\sigma_2^2 + \sigma_2^2 \beta^2}}{\mu_s}\right)^2
\]

\[
= \left(\frac{\alpha^2 - \beta^2 \sqrt{R^4} \ln(1 + R^2) + \frac{1}{\alpha} R^4}{\ln(1 + R^2)}\right)^2
\]

\[
= \left(\frac{\alpha^2 - \beta^2 \sqrt{R^4} \ln(1 + R^2) + \frac{1}{\alpha} R^4}{\ln(1 + R^2)}\right)^2
\]

\[
= \Delta_i^2 \frac{R^4}{\ln^2(1 + R^2)}
\]

where $\mu_s$ is given by (7). Hence, the tight bound of $\Gamma(R)$ is given by (10).

B. Proof of Theorem 2

Proof: Since we are only interested in the asymptotic order relation between $c$ and $R$, the Poisson representation of $c$ in Eq. (6) can be safely approximated by:

\[
c = 1 - \Phi\left(\frac{\Gamma(R) - \rho_m R^2}{\sqrt{\rho_m R^2}}\right) = Q(g(R))
\]

where $g(R) = \frac{R^4}{\sqrt{\rho_m R^2}}$. As $Q(x)$ is a monotonically decreasing function and $R_{\text{opt}}$ maximizes of $c$, $R_{\text{opt}}$ must minimize $g(R)$. According to Lemma 2, both the following two functions are asymptotic tight bounds of $g(R)$:

\[
g_1(R) = \Delta_i^2 \frac{R^3}{\ln^2(1 + R^2)} - \sqrt{\rho_m} \cdot R, \quad i = 1, 2
\]

The first and second order derivative of $g_1(R)$ are given by

\[
dg_1(R) / dR = \Delta_i^2 \frac{3R^2 \ln(1 + R^2) - 4R^4}{\ln^3(1 + R^2)} - \sqrt{\rho_m}
\]

\[
d^2g_1(R) / dR^2 = \Delta_i^2 \frac{6R^2 \ln^2(1 + R^2) - 28R^4 + 20R^4 \ln(1 + R^2) + 24R^6}{\ln^4(1 + R^2)}
\]

As $d^2g_1(R) / dR^2 = \Theta\left(\frac{R}{\ln R}\right)$, $d^2g_1(R) / dR^2 > 0$ if $R$ is large enough. Therefore, the root of $d^2g_1(R) / dR^2 = 0$, $R_i^1$, minimize $g_1(R)$. By letting Eq. (26) be zero, we have

\[
\Delta_i^2 = \frac{3R^2 \ln(1 + R^2) - 4R^4}{\ln^3(1 + R^2)} = \Theta\left(\frac{R}{\ln R}\right)
\]

As both $g_1(R)$ and $g_2(R)$ are tight bounds of $g(R)$, we know the optimal fusion range $R_{\text{opt}}$ satisfies $R_{\text{opt}} / \ln R_{\text{opt}} = \Theta(\sqrt{\rho_m})$. ■