

Routing Optimization and Secure Target Tracking in Distributed Wireless Sensor Networks

Majdi Mansouri¹, Lyes Khoukhi², Hichem Snoussi¹ and Cédric Richard³

¹ ICD/LM2S, University of Technology of Troyes, FRANCE

² ICD/ERA, University of Technology of Troyes, FRANCE

³ Laboratoire FIZEAU UMR CNRS 6525 Université de Nice Sophia-Antipolis, France.

Abstract

Due to the limited energy supplies of nodes in wireless sensor networks (WSN), optimizing their design under energy constraints, reducing their communication costs, and securing their aggregated data are of paramount importance. To this goal and in order to efficiently solve the problem of target tracking in WSN with quantized measurements, we propose to jointly estimate the target position, the relay location, select the secure sensor nodes and the best communication path. Firstly, we select the appropriate group in order to balance the energy dissipation and to provide the required data of the target in the WSN. This selection is also based on the transmission power between a single sensor and a cluster head (CH). Secondly, we detect the malicious sensor nodes based on the information relevance of their measurements. Thirdly, we select the best communication path between the candidate sensor and the CH. Then, we estimate jointly the target position and the relay location using Quantized Variational Filtering (QVF) algorithm. The selection of candidate sensors is based on multi-criteria function, which is computed by using the predicted target position provided by the QVF algorithm, the malicious sensor nodes detection is based on Kullback-Leibler distance between the current target position distribution and the predicted sensor observation, while the best communication path is selected as well as the highest signal-to-noise ratio (SNR) at the CH. The efficiency of the proposed method is validated by extensive simulations in target tracking for wireless sensor networks.

Index Terms

Wireless sensor networks, target tracking, quantized variational filtering, relay localization, best

communication path, malicious sensor.

I. INTRODUCTION

The WSN is defined as highly distributed networks of small and lightweight wireless nodes, which are deployed in large numbers to monitor the environment or system by measuring physical parameters such as temperature, pressure, or relative humidity [1]. Due to small dimension, sensor nodes are typically operated by lightweight batteries, which are difficult to be replaced or recharged. So WSN is a typical energy-restricted network. As observed in almost applications, the communication power consumption accounts for about 70% of the total power in WSNs [2], i.e. most energy of sensor nodes is spent in the exchange of routing information and data. A direct consequence is that sensor nodes will use up their energy quickly, which is uneconomical. Recently, researchers have proposed a special type of node called a "relay node" which is responsible for routing data packets. By using relay nodes, each sensor node would be able to send only its own data without a need to relay traffic from others [3].

To this goal, minimizing the communication costs between sensor nodes is critical to extend the lifetime of sensor networks. Another important metric of sensor networks is the accuracy of the sensing result of the target in that several sensors in the same cluster can present redundant data. Because of physical characteristics of sensor networks, such as distance, modality, or noise model of individual sensors, data from different sensors can have various qualities. Hence, the accuracy depends on the selection achieved by the cluster head on sensors and communication links.

In this chapter, we address the problem of secure target tracking, relay localization and sensors selection using a WSN based on quantized proximity sensors. Target tracking using quantized observations is a nonlinear estimation problem that can be solved using estimation solutions such as unscented Kalman filter (KF) [4], particle filters (PF) [5] or variational filtering (VF) [6]. Recently, a VF has been proposed as efficient solution for solving the target tracking problem since: (i) it respects the communication constraints of sensor, (ii) the online update of the filtering distribution and its compression are simultaneously performed, and (iii) it has the nice property to be model-free, ensuring the robustness of data processing.

There has already been a certain amount of research in the area of relay node placement estimation, sensors selection, and secure target tracking for wireless sensor networks. Li et al. [7] have proposed a Voronoi-based relay node placement scheme to balance energy consumption of each sensor node spent in communication. Bari et al. [8] have studied the relay node placement with a mobile data collector. Iranly et al. [9] have addressed the joint problem of energy provisioning and relay node placement. Hou et al. [10] have formulated the issues discussed above as a mixed-integer nonlinear programming problem. In [11], Ergen and Varaiya have proposed an approximation algorithm where the location of each relay node is restricted to a square lattice. An energy-constrained relay node placement problem has been explored by Wang et al. [12] aiming at minimizing the network cost with constraint on the network lifetime. The work in [13] has explored a topic similar to tracking security by proposing a protocol to verify securely the time of encounters in multi-hop networks. However, [13] did not address the security issue in the context of Bayesian tracking. Sensor selection based on expected information gain was introduced for decentralized sensing systems in [14]. The mutual information between the predicted sensor observation and the current target location distribution was proposed to evaluate the expected information gain about the target position attributable to a sensor in [15]–[17]. On the other hand, without using information theory, Yao et al. [18] found that the overall localization accuracy depends not only on the accuracy of individual sensors but also on the sensor locations relative to the target position during the development of localization algorithms. In addition, sensor selection can be discussed within the framework of optimal estimation, where the estimation error is normally used as the cost function of the optimization. In previous works [19]–[21], the problem of sensors selection is formulated as a constrained 0-1 integer programming problem, where the estimation error is minimized.

Most of these previous works do not consider a tradeoff between the quality of sensed data, the transmit power and the power stored in candidate nodes to select the candidate sensors. Moreover, they have simplified or ignored the security issue in the context of Bayesian tracking and the communication costs through a sensors-cluster head path.

Our contribution lies in the following aspects: 1) we improve the use of variational filtering algorithm (VF), which perfectly fits the highly non-linear conditions and eliminates the transmission error; 2) we jointly estimate the target position and relay location by using QVF algorithm; 3) we jointly select the best group of sensors that participates in data collection and detect the

malicious sensor nodes; 4) we explore the impact of the communication path selection on the QVF algorithm performances and introduce an adaptive quantization algorithm.

The remainder of the chapter is organized as follows. Section II presents both the observation model and the general state evolution models. Section III is devoted to the technique aimed at adaptively estimating the target position and the relay node location. The appropriate group of sensors selection is presented in Section IV. The malicious sensor nodes detection method is presented in Section V. Then, Section VI presents the best communication path selection scheme. Section VII gives some numerical results. Finally, Section VIII concludes the chapter.

II. PROBLEM FORMULATION

The variational filtering (VF) algorithm for target tracking inherits many desirable properties from the Bayesian Inference framework. An important step in the Bayesian target tracking is the recursive estimation of the predictive distribution described as follows,

$$p(\mathbf{X}_t|\mathbf{Z}_{1:t-1}) = \int p(\mathbf{X}_t|\mathbf{X}_{t-1})p(\mathbf{X}_{t-1}|\mathbf{Z}_{1:t-1})d\mathbf{X}_{t-1}, \quad (1)$$

Where, $p(\mathbf{X}_t|\mathbf{X}_{t-1})$ is the distribution used to model the prior time evolution of the target state. By incorporating the observation model $p(\mathbf{Z}_t|\mathbf{X}_t)$, the new estimate of the targets state \mathbf{X}_t is updated using the following predictive distribution $p(\mathbf{X}_t|\mathbf{Z}_{1:t-1})$:

$$p(\mathbf{X}_t|\mathbf{Z}_{1:t}) = \frac{p(\mathbf{Z}_t|\mathbf{X}_t)p(\mathbf{X}_t|\mathbf{Z}_{1:t-1})}{p(\mathbf{Z}_t|\mathbf{Z}_{1:t-1})}, \quad (2)$$

$$\text{where } p(\mathbf{Z}_t|\mathbf{Z}_{1:t-1}) = \int p(\mathbf{Z}_t|\mathbf{X}_t)p(\mathbf{X}_t|\mathbf{Z}_{1:t-1})d\mathbf{X}_t.$$

The observation model $p(\mathbf{Z}_t|\mathbf{X}_t)$ depends on the sensing mode employed by the sensors, while the state evolution model $p(\mathbf{X}_t|\mathbf{X}_{t-1})$ is always described by a parametric model.

A. General State Evolution Model

In this chapter, we use a General State Evolution Model (GSEM) described in [22]–[25], which is more adaptive to practical situations and has no restriction on velocity or moving direction of the target. As defined above, at instant t , the joint hidden state $\mathbf{X}_t = \{\mathbf{x}_t, \mathbf{R}_t\}$ to be estimated contains the target position \mathbf{x}_t , a set of activated relays locations $\mathbf{R}_t = \{\mathbf{r}_t^i\}_{i=1}^{m_t}$, where m_t denotes the number of activated relays. Take relay i for example, \mathbf{r}_t^i is assumed to be a Gaussian variable, whose expectation is its latest estimate value $\hat{\mathbf{r}}^i$, and the precision matrix

ν^i indicates its position offset due to deployment error and other spatial factors. The target \mathbf{x}_t is assumed to follow an extended Gaussian model, where the expectation $\boldsymbol{\mu}_t$ and the precision matrix $\boldsymbol{\lambda}_t$ are both random, with a Gaussian distribution and a Wishart distribution respectively:

$$\begin{cases} \mathbf{r}_t^i & \sim \mathcal{N}(\hat{\mathbf{r}}^i, \nu^i) \\ \mathbf{x}_t & \sim \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\lambda}_t) \\ \boldsymbol{\mu}_t & \sim \mathcal{N}(\boldsymbol{\mu}_{t-1}, \bar{\boldsymbol{\lambda}}) \\ \boldsymbol{\lambda}_t & \sim \mathcal{W}_n(\bar{\mathbf{V}}, \bar{n}) \end{cases}, \quad (3)$$

where $\bar{\boldsymbol{\lambda}}$ is the initial precision matrix to reflect the uncertainty of the target location estimation at instant t with respect to the previous one. The target state precision matrix $\boldsymbol{\lambda}_t$ is modeled by a d dimensional Wishart distribution ($n = 2$ in this work), with $\bar{\mathbf{V}}$ and \bar{n} denote the precision matrix and the degrees of freedom, respectively. Notice that $\bar{\cdot}$ denotes the initial fixed parameter. Assuming random mean and covariance for the state \mathbf{x}_t leads to a probability distribution covering a wide range of tail behaviors, which allows discrete jumps in the target trajectory. In fact, the marginal state distribution is obtained by integrating over the mean and precision matrix (Eq. (4)):

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = \int \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\lambda}_t) p(\boldsymbol{\mu}_t, \boldsymbol{\lambda}_t | \mathbf{x}_{t-1}) d\boldsymbol{\mu}_t d\boldsymbol{\lambda}_t, \quad (4)$$

In (4), the integration with respect to the precision matrix leads to the known class of scale mixture distributions given by Barndorff-Nielsen [26]. The low values of the degrees of freedom \bar{n} reflects the heavy tails of the marginal distribution $p(\mathbf{x}_t | \mathbf{x}_{t-1})$.

B. Quantized Proximity Observation Model QPOM

Consider a wireless sensor network, in which the sensor locations, for an activated sensor i are given by $\mathbf{s}^i = (s_1^i, s_2^i)$, $i = 1, 2, \dots, N_s$. Taking the activated sensor i (the activation procedure is explained in section IV), their observations are modeled by:

$$\gamma_t^{i,\mathbf{x}} = C \|\mathbf{x}_t - \mathbf{s}^i\|^\eta + \epsilon_t, \quad (5)$$

Similarly,

$$\gamma_t^{i,\mathbf{r}} = C \|\mathbf{r}_t - \mathbf{s}^i\|^\eta + \epsilon_t, \quad (6)$$

where ϵ_t , is a Gaussian noise with zero mean and variance σ_ϵ^2 and η and C are a known constants. The observation for target tracking problem is quantized, before being transmitted, by partitioning

the observation space into N_t^i intervals $\mathcal{R}_k = [\tau_k(t), \tau_{k+1}(t)]$, where $j \in \{1, \dots, N_t^i\}$. Similarly, where the sensor collects information to estimate the relay location, it partitions the observation space into M_t^i regions $\mathcal{R}_j = [\tau_j(t), \tau_{j+1}(t)]$, where $j \in \{1, \dots, M_t^i\}$.

The equations (7) and (7) show the quantized signals,

$$y_t^{i,\mathbf{x}} = d_{k,\mathbf{x}} \quad \text{if} \quad \gamma_t^{i,\mathbf{x}} \in [\tau_k(t), \tau_{k+1}(t)], \quad (7)$$

Similarly,

$$y_t^{i,\mathbf{r}} = d_{j,\mathbf{r}} \quad \text{if} \quad \gamma_t^{i,\mathbf{r}} \in [\tau_j(t), \tau_{j+1}(t)], \quad (8)$$

where $d_{k,\mathbf{x}}$ (resp. $d_{j,\mathbf{r}}$) is the centroid of k -th cell during the target tracking problem (resp. the centroid of j -th cell during the relay localization). Then, the signals received by the cluster head from the sensor i at the sampling instant t are written as,

$$z_t^{i,\mathbf{x}} = \beta_i y_t^{i,\mathbf{x}} + n_t; \quad (9)$$

Similarly,

$$z_t^{i,\mathbf{r}} = \alpha_i y_t^{i,\mathbf{r}} + n_t; \quad (10)$$

where n_t is a random Gaussian noise sensor with a zero mean and a variance σ_n^2 , and β_i is the i -th sensor channel attenuation coefficient.

1) *Formulation of the Observation Model $p(\mathbf{z}_t | \mathbf{X}_t)$:*

As illustrated above, the Bayesian filtering involves the construction of the observation model $p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{R}_t)$. To track the target \mathbf{x} , the available observations at the activated CH are denoted by $\mathbf{z}_t^{\mathbf{s},\mathbf{x}} = \{z_t^{i,\mathbf{x}}\}_{i=1}^{m_t'}$, where m_t' is the number of sensors in the activated cluster. Assuming that the noise samples ϵ_t are independently distributed, we have,

$$p(\mathbf{z}_t^{\mathbf{s},\mathbf{x}} | \mathbf{x}_t) = \prod_{i=1}^{m_t'} \sum_{j=0}^{N_t^i-1} p(\tau_j(t) < \gamma_t^{\mathbf{s},\mathbf{x}} < \tau_{j+1}(t)) \mathcal{N}(\beta_i d_j, \sigma_\epsilon^2) \quad (11)$$

where

$$p(\tau_j(t) < \gamma_t^{\mathbf{s},\mathbf{x}} < \tau_{j+1}(t)) = \int_{\tau_j(t)}^{\tau_{j+1}(t)} \mathcal{N}(\rho_{\gamma_t^{\mathbf{s},\mathbf{x}}}(\mathbf{x}_t), \sigma_n^2) d\gamma_t^{\mathbf{s},\mathbf{x}} \quad (12)$$

is computed according to the quantization rule defined in (7), in which

$$\rho_{\gamma_t^{\mathbf{s},\mathbf{x}}}(\mathbf{x}_t) = K \|\mathbf{x}_t - \mathbf{s}\|^\eta \quad (13)$$

Similarly,

$$p(z_t^{\mathbf{s},\mathbf{r}}|\mathbf{r}_t) = \prod_{i=1}^{m_t} \sum_{k=0}^{n_t-1} p(\tau_k(t) < \gamma_t^{\mathbf{s},\mathbf{r}} < \tau_{k+1}(t)) \mathcal{N}(\alpha_i d_k, \sigma_\epsilon^2) \quad (14)$$

Explicitly, $z^{i,\mathbf{r}}$ denotes the observations collected during the relay localization phase. At sampling instant t , the observation of the target $z_t^{i,\mathbf{x}}$ is incorporated together with $z^{i,\mathbf{r}}$ to help refining the localization of the relay i .

C. Overview of the proposed algorithm

Generally, in order to ensure a precise target tracking, the relay locations need to be known a priori. Initially, we suppose that relays are placed in roughly known positions. The true position \mathbf{r}^i of the relay i is Gaussian distributed around its initial setting value $\bar{\mathbf{r}}^i$ with precision ν^i . This initial assignment results in a node layout that resembles an unfolded and scaled version of the actual deployment, roughly preserving the topological ordering of nodes. After the deployment, sensors proceed to exchange information with neighboring relays in their communication range. The observations between them can thus be detected, and stored in the corresponding sensors. Then, a localization phase is launched locally by incorporating only the observations between sensors and relays, in the aim to improve *a priori* information on the relay locations. Consider the relay i for example,

$$p(\hat{\mathbf{r}}^i|z^{i,\mathbf{r}}) \propto \mathcal{N}(\bar{\mathbf{r}}^i, \nu^i) p(z^{i,\mathbf{r}}|\mathbf{r}^i). \quad (15)$$

Thus, the estimation of the relay position is refined by incorporating the observation $z^{i,\mathbf{r}}$ according to the prior distribution $\mathbf{r}^i \sim \mathcal{N}(\bar{\mathbf{r}}^i, \nu^i)$. After having localized all the relays, much more precise information on their locations is provided, which is incorporated as the prior information $\hat{\mathbf{r}}^i$ for the adaptive scheme.

Once an intrusion in the WSN is identified, a cluster of sensors around the phenomenon of interest is activated. Note that, as mentioned above, the adaptive scheme procedure is distributively executed on a cluster head in order to minimize energy and bandwidth consumption. The sensors that have detected the presence of the target in their sensing ranges broadcast their residual energy level. The one with the maximum residual energy is elected as the cluster head (CH) to take charge of signal processing. The other clustered detecting sensors transfer to the CH their timely observation of the target, and the observations between their neighboring relays which are stored during the pre-localization phase. Note that, the information concerning the detection

of relay locations and the estimations of the target are simultaneously updated in the CH by the adaptive scheme algorithm based on these observations. According to the Bayesian framework, the adaptive scheme algorithm works in a recursive way, where the filtering distribution is transferred to the next CH for further use.

III. ADAPTIVE SCHEME ALGORITHM

The classical Bayesian framework is adopted by the adaptive scheme algorithm adopts to estimate the unknown state over time by using incoming observations. As mentioned in section II, two distinct phases compose the Bayesian filtering framework: Prediction and Update. The prediction phase uses the state estimated from the previous sampling instant to produce an estimate of the state at the current instant according to Eq. (1). In the update phase of Eq. (2), measurement information at the current instant is used to refine this prediction to arrive at a new and hopefully more accurate state estimate. In a distributed context, the estimations of the detecting sensors are updated and stored locally, whereas the filtering distribution of the target needs to be transferred for future use. Concerning energy and bandwidth efficiency, the variational approach compresses the filtering distribution of the target to a single Gaussian distribution between successive clusters in a consistent manner [6]. Thus distributed signal processing is achieved effectively. The details of the adaptive scheme algorithm are illustrated in what follows.

A. Variational Filtering Approximation

The variational approach is employed here to approximate the posterior probability $p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t})$ by a separable distribution $q(\boldsymbol{\alpha}_t)$, which minimizes the Kullback-Leibler (KL) divergence error:

$$D_{\text{KL}}(q||p) = \int q(\boldsymbol{\alpha}_t) \log \frac{q(\boldsymbol{\alpha}_t)}{p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t})} (d\boldsymbol{\alpha}_t), \quad (16)$$

where $q(\boldsymbol{\alpha}_t) = \prod_i q(\boldsymbol{\alpha}_t^i) = q(\mathbf{x}_t)q(\boldsymbol{\mu}_t)q(\boldsymbol{\lambda}_t)q(\mathbf{R}_t)$,

and $q(\mathbf{R}_t) = \prod_{i=1}^{m_t} q(\mathbf{r}_t^i)$.

With variational computation, the following approximate distribution yields [6],

$$q(\boldsymbol{\alpha}_t^i) \propto \exp\langle \log p(\mathbf{Z}_{1:t}, \boldsymbol{\alpha}_t) \rangle_{\prod_{j \neq i} q(\boldsymbol{\alpha}_t^j)}, \quad (17)$$

where $\langle \cdot \rangle_q$ denotes the expectation operator relative to the distribution q . Taking into consideration the separable approximate distribution at time $t-1$, that is, $\hat{p}(\boldsymbol{\alpha}_{t-1} | \mathbf{Z}_{1:t-1}) = q(\boldsymbol{\alpha}_{t-1})$, the filtering distribution at time t is deduced,

$$\begin{aligned} \hat{p}(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t}) &= \frac{p(\mathbf{z}_t | \boldsymbol{\alpha}_t) \int p(\boldsymbol{\alpha}_t | \boldsymbol{\alpha}_{t-1}) q(\boldsymbol{\alpha}_{t-1}) d\boldsymbol{\alpha}_{t-1}}{p(\mathbf{z}_t | \mathbf{Z}_{1:t-1})} \\ &\propto p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{R}_t) p(\mathbf{x}_t | \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t) p(\boldsymbol{\lambda}_t) p(\mathbf{R}_t) q_p(\boldsymbol{\mu}_t), \\ \text{with } q_p(\boldsymbol{\mu}_t) &= \int p(\boldsymbol{\mu}_t | \boldsymbol{\mu}_{t-1}) q(\boldsymbol{\mu}_{t-1}) d\boldsymbol{\mu}_{t-1}. \end{aligned} \quad (18)$$

Therefore, the filtering distribution $p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t})$ can be sequentially updated, through a simple integral with respect to $\boldsymbol{\mu}_{t-1}$. Considering the general state evolution model proposed in (3), the evolution of $\boldsymbol{\mu}_{t-1}$ is Gaussian, namely $p(\boldsymbol{\mu}_t | \boldsymbol{\mu}_{t-1}) = \mathcal{N}(\boldsymbol{\mu}_t | \boldsymbol{\mu}_{t-1}, \bar{\boldsymbol{\lambda}})$. Defining $q(\boldsymbol{\mu}_{t-1}) = \mathcal{N}(\boldsymbol{\mu}_{t-1}^*, \boldsymbol{\lambda}_{t-1}^*)$, $q_p(\boldsymbol{\mu}_t)$ is also Gaussian [27], with the following parameters,

$$\begin{aligned} q_p(\boldsymbol{\mu}_t) &= \mathcal{N}(\boldsymbol{\mu}_t^p, \boldsymbol{\lambda}_t^p), \\ \text{where } \boldsymbol{\mu}_t^p &= \boldsymbol{\mu}_{t-1}^* \quad \text{and} \quad \boldsymbol{\lambda}_t^p = (\boldsymbol{\lambda}_{t-1}^{*-1} + \bar{\boldsymbol{\lambda}}^{-1})^{-1}. \end{aligned} \quad (19)$$

Hence, the temporal dependence is hence reduced to the incorporation of only one Gaussian component approximation $q(\boldsymbol{\mu}_{t-1})$. The update and the approximation of the filtering distribution $p(\boldsymbol{\alpha}_t | \mathbf{z}_{1:t})$ are jointly performed, yielding a natural and adaptive compression [6], [28]. According to the equation (17) and taking into account (18) and (19), variational calculus leads to closed-form expressions of $q(\boldsymbol{\mu}_t)$ and $q(\boldsymbol{\lambda}_t)$:

$$\begin{cases} q(\boldsymbol{\mu}_t) = \mathcal{N}(\boldsymbol{\mu}_t^*, \boldsymbol{\lambda}_t^*) \\ q(\boldsymbol{\lambda}_t) = \mathcal{W}_2(\mathbf{V}_t^*, n^*) \end{cases}, \quad (20)$$

where the parameters are iteratively updated until convergence, according to the following scheme:

$$\begin{cases} \boldsymbol{\mu}_t^* &= \boldsymbol{\lambda}_t^{*-1} (\langle \boldsymbol{\lambda}_t \rangle \langle \mathbf{x}_t \rangle + \boldsymbol{\lambda}_t^p \boldsymbol{\mu}_t^p) \\ \boldsymbol{\lambda}_t^* &= \langle \boldsymbol{\lambda}_t \rangle + \boldsymbol{\lambda}_t^p \\ n^* &= \bar{n} + 1 \\ \mathbf{V}_t^* &= (\langle \mathbf{x}_t \mathbf{x}_t^T \rangle - \langle \mathbf{x}_t \rangle \langle \boldsymbol{\mu}_t \rangle^T - \langle \boldsymbol{\mu}_t \rangle \langle \mathbf{x}_t \rangle^T + \langle \boldsymbol{\mu}_t \boldsymbol{\mu}_t^T \rangle + \bar{\mathbf{V}}^{-1})^{-1} \end{cases}. \quad (21)$$

Notice that $\langle \cdot \rangle$ designates the expectation relative to the distribution $q(\cdot)$. The mean state and the precision matrix distributions represented respectively by $q(\boldsymbol{\mu}_t)$ and $q(\boldsymbol{\lambda}_t)$ have closed forms,

such that their expectations are easily derived:

$$\begin{cases} \langle \boldsymbol{\mu}_t \rangle &= \boldsymbol{\mu}_t^* \\ \langle \boldsymbol{\mu}_t \boldsymbol{\mu}_t^T \rangle &= \boldsymbol{\lambda}_t^{*-1} + \boldsymbol{\mu}_t^* \boldsymbol{\mu}_t^{*T} \\ \langle \boldsymbol{\lambda}_t \rangle &= n^* \mathbf{V}_t^* \end{cases} . \quad (22)$$

Nevertheless, the target state \mathbf{x}_t and the activated relay positions \mathbf{R}_t do not have closed forms. Combining the equations (17) and (18), $q(\mathbf{x}_t)$ and $q(\mathbf{r}_t^i)$ have the following expressions:

$$\begin{cases} q(\mathbf{x}_t) \propto \mathcal{N}(\langle \boldsymbol{\mu}_t \rangle, \langle \boldsymbol{\lambda}_t \rangle) \prod_{i=1}^{m_t'} p(z_t^{i,\mathbf{x}} | \mathbf{x}_t), \\ q(\mathbf{r}_t^i) \propto \mathcal{N}(\hat{\mathbf{r}}^i, \nu^i) \prod_{i=1}^{m_t} p(z_t^{i,\mathbf{r}} | \mathbf{r}_t), \end{cases} \quad (23)$$

Hence, the general state evolution model (3) and the observation model (5) and (6) are naturally incorporated to update $q(\mathbf{x}_t)$ and $q(\mathbf{r}_t^i)$. Their distribution forms immediately suggest an Importance Sampling (IS) procedure, where samples are drawn from Gaussian distributions $\mathcal{N}(\langle \boldsymbol{\mu}_t \rangle, \langle \boldsymbol{\lambda}_t \rangle)$ and $\mathcal{N}(\hat{\mathbf{r}}^i, \nu^i)$ respectively, and are weighted according to their likelihoods (taking into account the observation model):

$$\begin{cases} \mathbf{x}_t^{(k)} \sim \mathcal{N}(\langle \boldsymbol{\mu}_t \rangle, \langle \boldsymbol{\lambda}_t \rangle), w_t^{(k)} \propto \prod_{i=1}^{m_t'} p(z_t^{i,\mathbf{x}} | \mathbf{x}_t^{(k)}), \\ \mathbf{r}_t^{i,(k)} \sim \mathcal{N}(\hat{\mathbf{r}}^i, \nu^i), w_t^{i,(k)} \propto \prod_{i=1}^{m_t} p(z_t^{i,\mathbf{r}} | \mathbf{r}_t^{i,(k)}), \end{cases} \quad (24)$$

Therefore, $q(\mathbf{x}_t)$ and $q(\mathbf{r}_t^i)$ can be approximated by the Monte Carlo method:

$$\begin{cases} \langle \mathbf{x}_t \rangle = \sum_{k=1}^N w_t^{(k)} \mathbf{x}_t^{(k)} \\ \langle \mathbf{r}_t^i \rangle = \sum_{k=1}^N w_t^{i,(k)} \mathbf{r}_t^{i,(k)} \end{cases} . \quad (25)$$

As mentioned above, the standard adaptive scheme solution includes both prediction and update steps. Besides the update of the filtering distribution $p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t})$, the predictive distribution $p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t-1})$ can also be efficiently calculated by the variational approach. In fact, by incorporating the separable approximate distribution $q(\boldsymbol{\alpha}_{t-1})$ in the place of $p(\boldsymbol{\alpha}_{t-1} | \mathbf{Z}_{1:t-1})$, the recursive adaptive scheme algorithm calculates the predictive distribution $p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t-1})$ in the following form:

$$\hat{p}(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t-1}) \propto \int p(\boldsymbol{\alpha}_t | \boldsymbol{\alpha}_{t-1}) q(\boldsymbol{\alpha}_{t-1}) d\boldsymbol{\alpha}_{t-1} \propto p(\mathbf{x}_t | \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t) p(\boldsymbol{\lambda}_t) p(\mathbf{R}_t) q_p(\boldsymbol{\mu}_t). \quad (26)$$

The exponential form solution, which minimizes the Kullback-Leibler divergence between the predictive distribution $p(\boldsymbol{\alpha}_t | \mathbf{Z}_{1:t-1})$ and the separable approximate distribution $q_{t|t-1}(\boldsymbol{\alpha}_t)$, yields

Gaussian distributions for the predicted expectations, and a Wishart distribution for the target precision matrix:

$$\begin{cases} q_{t|t-1}(\mathbf{x}_t) & \propto \mathcal{N}(\langle \boldsymbol{\mu}_t \rangle_{q_{t|t-1}}, \langle \boldsymbol{\lambda}_t \rangle_{q_{t|t-1}}), \\ q_{t|t-1}(\boldsymbol{\mu}_t) & \propto \mathcal{N}(\boldsymbol{\mu}_{t|t-1}^*, \boldsymbol{\lambda}_{t|t-1}^*), \\ q_{t|t-1}(\boldsymbol{\lambda}_t) & \propto \mathcal{W}_2(V_{t|t-1}^*, n_{t|t-1}^*), \\ q_{t|t-1}(\mathbf{R}_t) & \propto \mathcal{N}(\hat{\mathbf{R}}, \boldsymbol{\nu}), \end{cases} \quad (27)$$

where the parameters are updated according to the same iterative schemes (21) and (22). The target state and the activated relays are now evaluated by the following expressions:

$$\begin{aligned} \langle \mathbf{x}_t \rangle_{q_{t|t-1}} &= \langle \boldsymbol{\mu}_t \rangle_{q_{t|t-1}}, \langle \mathbf{x}_t \mathbf{x}_t^T \rangle_{q_{t|t-1}} = \langle \boldsymbol{\lambda}_t \rangle_{q_{t|t-1}}^{-1} + \langle \boldsymbol{\mu}_t \rangle_{q_{t|t-1}} \langle \boldsymbol{\mu}_t \rangle_{q_{t|t-1}}^T, \\ \langle \mathbf{R}_t \rangle_{q_{t|t-1}} &= \langle \hat{\mathbf{R}} \rangle. \end{aligned} \quad (28)$$

In the next section, we present a technique aiming at adaptively (and jointly) selecting the appropriate group of candidate sensors that participate in data collection for tracking the target.

IV. OPTIMAL GROUP OF CANDIDATE SENSORS SELECTION BASED ON MULTI-CRITERIA FUNCTION (MCF)

Since the predicted target position $\mathbf{x}_p(t) = \langle \mathbf{x}_t \rangle_{q_{t+1|t}}$ at time t is available, it could be used to select sensors. Sensors inside the disk centered at $\mathbf{x}_p(t)$ with radius R_{max} are pre-selected. After the pre-selection of sensors within R_{max} range, the CH divides the pre-selected sensors into $M = \sum_{j=4}^{N_s} C_{N_s}^j$ groups \mathcal{G}_t (at least four sensors are needed to sense the target within their range [13]). Then, it computes for each group of sensors the MCF (detailed in IV-A) and activates the appropriate group which has the highest MCF value to participate in data aggregation.

In the next subsection, we detail the MCF function with the used parameters permitting to select the appropriate group of sensors that participate in data collection for tracking the target.

A. Multi-criteria function

The multi-criteria function for sensors selection aims is to define the main parameters that may influence the relevance of the participation in cooperation, which are: 1) $(MI(\mathbf{x}_t, z^{\mathcal{G}_t}))$: the information that can be transferred from the group of candidate sensors, 2) $\mathcal{G}_t(\mathbf{D}(i))$: the transmitting distance between the sensor i and the CH in group, and 3) $(E(i))$: the energy

stored in candidate sensor. The problem is how to formulate the criteria for the CH to select the appropriate group of sensors that provide relevant data and balances the energy level among all sensors. We define a combinative measurement for the group of candidate nodes \mathcal{G}_t , denoted as \mathbf{F} , which is given by:

$$\mathbf{F}(\mathcal{G}_t) = n_1 MI(\mathbf{x}_t, z^{\mathcal{G}_t}) - n_2 \sum_{i=1}^M \mathbf{R}_t^2(i). \quad (29)$$

Where $\mathbf{R}(i) = \frac{E(i)}{D(i)}$, n_1 and n_2 are the importance factor for the Mutual Information (MI) function and the ratio between the energy stored in the sensor (E) and the transmit distance (D). In order to select the best group of candidate sensors based on equation (29), the objective is to choose the appropriate group of sensors \mathcal{G}_t so that,

$$\widehat{\mathcal{G}}_t = \arg \max_{\mathcal{G}_t \subset \mathcal{C}} \mathbf{F}(\mathcal{G}_t) \quad (30)$$

The efficiency measurement of given information is often achieved by the mutual information function. This later is a quantity measuring the amount of information that the observable variable $z^{\mathcal{G}_t}$ carries about the unknown parameter \mathbf{x}_t . The mutual information between the observations $z^{\mathcal{G}_t}$ and the source \mathbf{x}_t is proportional to:

$$MI(\mathbf{x}_t, z^{\mathcal{G}_t}) \propto p(z^{\mathcal{G}_t} | \mathbf{x}_t) \log(p(z^{\mathcal{G}_t} | \mathbf{x}_t)) \quad (31)$$

The likelihood function (L) is expressed as,

$$L(\mathbf{s}^{\mathcal{G}_t}) = p(z^{\mathcal{G}_t} | \mathbf{x}_t) = \prod_{i=1}^M \sum_{j=0}^{N_t^i - 1} p(\tau_j(t) < \gamma^i < \tau_{j+1}(t)) \mathcal{N}(d_j, \sigma_\epsilon^2) \quad (32)$$

where,

$$\begin{aligned} p(\tau_j(t) < \gamma^i < \tau_{j+1}(t)) &= \int_{\tau_j(t)}^{\tau_{j+1}(t)} \mathcal{N}(\rho_{\gamma_t^i}(\mathbf{s}^i), \sigma_n^2) d\gamma_t \\ &= \frac{1}{\sqrt{\pi}} \left[\operatorname{erfc} \left(\frac{\tau_j(t) - \rho_{\gamma_t^i}(\mathbf{x}_t)}{\sqrt{2\sigma_n^2}} \right) - \operatorname{erfc} \left(\frac{\tau_{j+1}(t) - \rho_{\gamma_t^i}(\mathbf{x}_t)}{\sqrt{2\sigma_n^2}} \right) \right] \end{aligned} \quad (33)$$

using the quantization rule defined in (7), in which

$$\rho_{\gamma_t^i}(\mathbf{s}^i) = K \|\mathbf{x}_t - \mathbf{s}^i\|^\eta, \quad (34)$$

It is worth noting that the expression of the MI given in (31) depends on the target position \mathbf{x}_t at the sampling instant t and on the group of candidate sensors \mathcal{G}_t . However, as the target

position is unknown, the MI is replaced by its expectation according to the predictive distribution $p(\mathbf{x}_t | z_{1:t-1}^{\mathcal{G}_t})$ of the target position:

$$\langle MI(\mathbf{s}^{\mathcal{G}_t}) \rangle = E_{p(\mathbf{x}_t | z_{1:t-1}^{\mathcal{G}_t})} [MI(\mathbf{s}^{\mathcal{G}_t})] \quad (35)$$

The computation of the above expectation is analytically untractable. However, as the VF algorithm yields a Gaussian predictive distribution $\mathcal{N}(\mathbf{x}_t; \mathbf{x}_{t/t-1}, \lambda_{t/t-1})$, the expectation (35) can be efficiency approximated by a Monte Carlo scheme:

$$\begin{aligned} \langle MI(\mathbf{s}^{\mathcal{G}_t}) \rangle &\simeq \frac{1}{J} \sum_{j=1}^J MI(\tilde{\mathbf{x}}_t^j, \mathbf{s}^{\mathcal{G}_t}), \\ \tilde{\mathbf{x}}_t^j &\sim \mathcal{N}(\mathbf{x}_p(t); \mathbf{x}_{t/t-1}, \lambda_{t/t-1}) \end{aligned} \quad (36)$$

where $\tilde{\mathbf{x}}_t^j$ is the j -th drawn sample at instant t , and J is the total number of drawn vectors $\tilde{\mathbf{x}}_t$. The next section presents the malicious sensor nodes detection based on the Kullback-Leibler distance (KLD) between the current target position distribution and the predicted sensor observation.

V. MALICIOUS SENSOR NODE DETECTION BASED ON KULLBACK-LEIBLER DISTANCE

(KLD)

A. Problem definition

The following are the assumptions made to resolve the malicious sensor detection problem:

- 1) We assume that malicious nodes can successfully authenticate with the sensor network, and their data can be collected with other nodes in the network. Otherwise, the CHs are assumed to well-behave and not malicious.
- 2) We assume a centralized scenario, in which a CH processing unit collects tracking reports from sensor nodes, determines which of them are malicious and removes them from sensor networks.
- 3) The purpose of the fictitious path is to allow the enemy to avoid surveillance. However, the fictitious path does not go beyond the sensing range of the nodes.
- 4) We assume that sensor nodes have unlimited communication bandwidth among each other. Meanwhile, an unknown number of the nodes are malicious and they are injecting false tracking reports into the network. The problem is how to detect those malicious nodes, and to provide a correct target trajectory.

B. The computing of the Kullback Leibler distance (KLD)

The measurement of the distance between two statistical models is needed to resolve certain problems. For example, this distance can be used in evaluating the training algorithm or classifying the estimated models [29]. The Kullback-Leibler distance or the relative entropy arises in many contexts as an appropriate measurement of the distance between two distributions. The KLD between the two probability density functions p and \hat{p} is defined as [30]:

$$KLD(p||\hat{p}) = \int p \log \frac{p}{\hat{p}} \quad (37)$$

The computation of the distribution function is very complex for hidden Markov models, and practically it can be only computed via a recursive procedure; the "forward/backward" or "upward/downward" algorithms [31], [32]. Thus there is no simple closed form expression for the KLD for these models. Commonly, the Monte-Carlo method is used to numerically approximate the integral in (37) as

$$KLD(p||\hat{p}) = \mathbf{E}_p(\log(p) - \log(\hat{p})) \quad (38)$$

Hereafter, we detail the two proposed schemes for the detection of malicious sensors.

C. Case 1: Reactive cluster formation

At every sampling instant, the cluster (CH and candidate secure sensors) is dynamically formed. The candidate sensors are activated in a manner explained in Section IV. The cluster head CH_t is chosen to be the nearest sensor to $\langle \mathbf{x}_t \rangle_{q_{t|t-1}}$ i.e :

$$CH_t = \arg \min_{i=1, \dots, |\mathcal{B}_t|} \{ \|\langle \mathbf{x}_t \rangle_{q_{t|t-1}} - \mathbf{s}_t^m\| \quad m \in \mathcal{B}_t \} \quad (39)$$

where $|\cdot|$ denotes the cardinality, and \mathcal{B}_t is the set of activated sensors.

Let's assume that sensors i , j and k are activated and the sensors j and k can overhear i 's transmission. At time step t , the sensors j and k compute their predictive target distributions $p(\mathbf{x}_t|z_{t+1|t}^j)$ and $p(\mathbf{x}_t|z_{t+1|t}^k)$ by executing the QVF algorithm. At time step $(t+1)$, they overhear the i 's transmitted value, and compare the i -th sensor predictive distribution $p(\mathbf{x}_t|z_{t+1|t}^i)$ to the two predictive distributions $p(\mathbf{x}_t|z_{t+1|t}^j)$ and $p(\mathbf{x}_t|z_{t+1|t}^k)$ by computing the Kullback Leibler distances (detailed in V-B) $K_{i,j}$ and $K_{i,k}$. If the $K_{i,j}$ and $K_{i,k}$ are larger than K_0 (a predefined threshold), as shown in Fig. 1, then the sensors j and k estimate that the sensor i is malicious and send

a notification to the CH. This later compares the $p(\mathbf{x}_t|z_{t+1}^i)$ to the $p(\mathbf{x}_t|z_{t+1}^i)$ distribution by computing the distance $K_{i,i}$, if this latter is also large than K_0 , then, it deletes the i -th fictitious path from the sensor network.

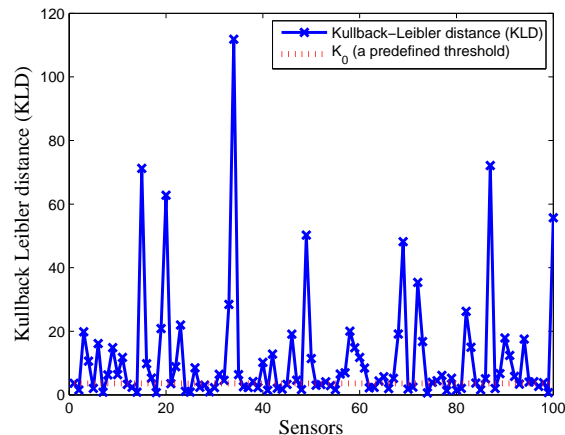


Fig. 1. A simple example for KLD computing

The major advantages of this algorithm can be summarized as follows. i) The accuracy of tracking is guaranteed by choosing the most potential sensors to dynamically form a cluster. ii) Sensors Kullback Leibler distance (KLD) Kullback?Leibler distance (KLD) (a predefined threshold). ii) This scheme is much more robust to external attacks. iii) As the lifetime of WSN is defined as the time elapsed until the first sensor depletes its energy [33], it is essential to evenly distribute the energy consumption over the whole WSN. By dynamically forming the clusters, CHs performing series of energy-intensive functions are changing frequently in order to balance the energy expenditure. However, since the signal processing task is assigned to all the slave sensors, all these advantages cited above are at the expense of homogeneous high hardware configuration.

The scheme discussed previously will be referred to as Re-QVF; its procedure is described in **Algorithm 1**.

Algorithm 1:

— Initialization:

- 1) Determine the CH by using the equation (39).
- 2) Select the appropriate group of sensors according to Section IV
- 3) Quantize the sensors' measurements according to Sub.II-B.
- 4) Execute the QVF algorithm.

— Iterations:

- 1) Compute the multi-criteria function by using the equation (29).
 - 2) Select the appropriate group of sensors according according to Section IV
 - 3) Determine the CH by using the equation (39).
 - 4) Detect the malicious sensors according to Sub.V-C.
 - 5) Delete the fictitious path according to Sub.V-C.
 - 6) Quantize the sensors' measurements according to Sub.II-B.
 - 7) Execute the QVF algorithm.
-

D. Case 2: Proactive cluster formation

For this second case, we suppose that CHs are statically selected at the time of sensors deployment. At sampling instant t , the QVF provides at the sampling instant t , the predicted target position $\mathbf{x}_{t/t-1} = \langle \mathbf{x}_t \rangle_{q_{t|t-1}}$. As shown in Fig.2, based on this predicted information, the cluster head CH_{t-1} at sampling instant $t-1$ selects the next cluster head CH_t . If the predicted target position $\langle \mathbf{x}_t \rangle_{q_{t|t-1}}$ remains in the vicinity of CH_{t-1} , which means that at least four of its slave sensors can detect the target, then $CH_t = CH_{t-1}$. Otherwise, if $\langle \mathbf{x}_t \rangle_{q_{t|t-1}}$ is going beyond the sensing range of the current cluster, then a new CH_t is activated based on the target position prediction $\langle \mathbf{x}_t \rangle_{q_{t|t-1}}$ and its future tendency.

$$CH_t = \arg \max_{k=1, \dots, K} \left\{ \frac{\cos \theta_t^k}{d_t^k} \right\} \quad (40)$$

where $d_t^k = \|\langle \mathbf{x}_t \rangle_{q_{t|t-1}} - L_{CH_t^k}\|$

and $\theta_t^k = \text{angle}(\overrightarrow{\langle \mathbf{x}_{t-1} \rangle \langle \mathbf{x}_t \rangle_{q_{t|t-1}}}, \overrightarrow{\langle \mathbf{x}_{t-1} \rangle L_{CH_t^k}})$

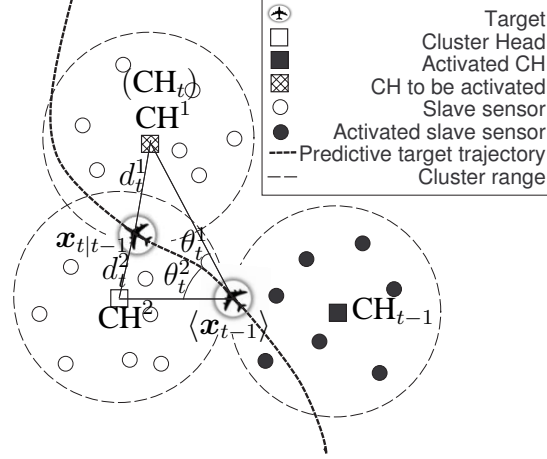


Fig. 2. Prediction-based CH_t activation.

where K is the number of CHs in the neighborhood of CH_{t-1} and $L_{CH_t^k}$ is the location of the k -th neighboring CH_t .

The initial distribution of the target position; $p(\mathbf{x}_0)$, at the instant $t = 0$, is assumed to be known and stored in CH unit. Then, at time step t , the activated sensor i sends its quantized observation y_t^i to the CH_t . When, the CH_t receives z_t^i , which is corrupted by an additive white Gaussian noise n_t , it executes the QVF algorithm, which provides in addition to the estimated target distribution $\hat{p}(z_t|\mathbf{x}_t)$, the predictive target distribution for the i -th sensor $p(\mathbf{x}_t|z_{t+1|t}^i)$. The CH_t sends this predicted information to the $CH_{(t+1)}$, which receives also a measurement from the i -th sensor. Based on this information, the $CH_{(t+1)}$ computes for i -th sensor, the Kullback Leiber distance $K_{i,i}$ between the predictive distribution $p(\mathbf{x}_t|z_{t+1|t}^i)$ received from CH_t and the distribution $p(\mathbf{x}_t|z_{t+1}^i)$. In the similar way, the $CH_{(t+1)}$ calculates similarly, $K_{i,j}$ and $K_{i,k}$ and compares them with the constraint K_0 , if the differences are greater than 0, then it estimates that the i -th sensor node is malicious. Finally, the CH deletes the fictitious path corresponding to the malicious sensor from the sensor networks.

In summary, our proposed scheme includes many several advantages. Firstly, the consumed energy and the required bandwidth in communication are considerably reduced. The tracking process is performed only by the activated CH, while the slave sensors are unable to take over

this task. The candidate slave sensors are required to collect and transmit their measurements over short distances to the CH. Secondly, the cost of hardware configuration drops sharply owing to the low-cost of slave sensors. Furthermore, avoiding CH competition puts an end to unnecessary resource and energy consumption. Only when the hand-off operation occurs does the active CH need to communicate the temporal dependence information to the subsequent CH. However, the occurrence of hand-off operations is reduced by the non-myopic selective CH activation rule. In addition, the QVF algorithm reduces the dependence of temporal informations to only statistic parameters of a Gaussian distribution. We refer to the approach described in this section by Pro-QVF summarized by the following steps (**Algorithm 2**).

Algorithm 2:

— Initialization:

- 1) Determine the cluster head by using the equation (40).
- 2) Select the best group of candidate sensors according to Section.IV.
- 3) Quantize the sensors' measurements according to Sub.II-B.
- 4) Execute the QVF algorithm.

— Iterations:

- 1) Compute the predictive target distribution.
 - 2) Compute the multi-criteria function by using the equation (29).
 - 3) Select the appropriate group of sensors according to Section.IV.
 - 4) Determine the CH by using the equation (40).
 - 5) Detect the malicious sensors according to Sub.V-C.
 - 6) Delete the fictitious path according to Sub.V-C.
 - 7) Quantize the sensors' measurements according to Sub.II-B.
 - 8) Execute the QVF algorithm.
-

The next section is devoted to the method aimed at adaptively selecting the best communication path between the candidate sensor and the cluster head.

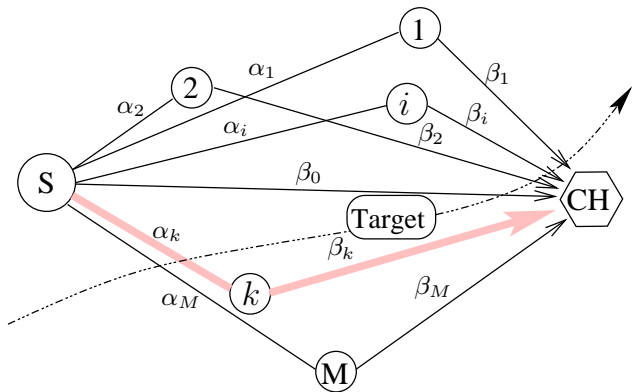


Fig. 3. Illustration of the diversity network with the best communication path selection scheme.

VI. BEST COMMUNICATION PATH SELECTION METHOD

This section aims at selecting the best communication path as well as the highest SNR at the CH. In what follows, we assume that the relay placement is already estimated.

A. System model

Fig. 3 illustrates the idea of the proposed model. The communication is established between a sensor node and the cluster head through the selected relay including direct and indirect links. The selected relay is achieved according to the best communication procedure. The cluster head can get the best copy of the source signal transmitted by the source sensor S , via the best communication path selection scheme allows. The first one is from the source sensor (direct link), while the second one is from the best path as shown in Fig. 3. The parameter α_i shown in Fig. 3 is the channel coefficient between the source sensor S and the i -th sensor. α_i and α_j , β_i and β_j are flat Rayleigh fading coefficients and mutually independent and non identical for all i and j .

The signal is simply amplified, at the relay sensor i , the signal is simply amplified using the gain $g = 1/\sqrt{E_s\alpha_i^2 + N_\epsilon}$, where E_s is the transmitted signal energy of the source. It is easy to prove that the source to cluster head SNR of the indirect path, $S \rightarrow i \rightarrow CH$ can be written as

$$\gamma_{S \rightarrow i \rightarrow CH} = \frac{\gamma_{\alpha_i} \gamma_{\beta_i}}{\gamma_{\alpha_i} + \gamma_{\beta_i} + 1} \quad (41)$$

where $\gamma_{\alpha_i} = \alpha_i^2 \frac{E_s}{N_c}$ is the instantaneous SNR of the source signal at the sensor i , $\gamma_{\beta_i} = \beta_i^2 \frac{E_i}{E_0}$ is the instantaneous SNR of the sensor signal (sensor i) measured at the cluster head, and E_i is the signal transmitted energy of the relay i .

The best communication path will be selected as the one that achieves the highest source-to-end SNR of indirect and direct paths. The SNR for the best path is given by,

$$\gamma_b = \max(\gamma_{\beta_0}, \gamma_{S \rightarrow i \rightarrow CH})_{i=1 \dots M} \quad (42)$$

where $\gamma_{\beta_0} = \beta_0^2 \frac{E_s}{N_n}$ is the instantaneous SNR between the source and the cluster head and M is the total number of links inside cluster. The upper-bound of $\gamma_{S \rightarrow i \rightarrow CH}$ is given by [34]. Computing the above expression is analytically untractable. However, (42) can be efficiency approximated [35] by

$$\gamma_{S \rightarrow i \rightarrow CH} \leq \gamma_i = \min(\gamma_{\alpha_i} \gamma_{\beta_i}), \quad (43)$$

this approximation simplifies the derivation of the SNR statistics: Cumulative Distribution Function (CDF), Probability Distribution Function (PDF) and Moment Generating Function (MGF).

B. Error performance analysis

The error probability p_e of the best path is given as

$$p_e(\gamma_{\beta_0}, \gamma_b) = A \operatorname{erfc} \left(\sqrt{B \max(\gamma_{\beta_0}, \gamma_b)} \right) \quad (44)$$

Where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2) dt$, and A and B are dependent on the modulation type. Over the random variables representing the SNR values of the best communication path the average error probability is given by

$$P_E = \int_0^\infty p_e(\gamma_b) f_{\gamma_b} d\gamma_b \quad (45)$$

using the alternative definition of the $\operatorname{erfc}(x)$ function as [36]

$$\operatorname{erfc}(x) = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \exp \left(-\frac{x^2}{\sin^2 \theta} \right) d\theta \quad (46)$$

and by substituting (46) into (45) we obtain

$$P_E = \int_0^\infty \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \exp \left(-\frac{B \gamma_b}{\sin^2 \theta} \right) f_{\gamma_b}(\gamma_b) d\gamma_b = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} M_{\gamma_b} \left(\frac{B}{\sin^2 \theta} \right) d\theta \quad (47)$$

where, $M_{\gamma_b} = \int_0^\infty f_{\gamma_b}(\gamma_b) \exp(-s\gamma_b) d\gamma_b$ is the MGF of γ_b and f_{γ_b} is PDF of γ_b .

In order to find the P_E , it is necessary to find the PDF and MGF of γ_b . We can write the CDF of γ_b as follows,

$$F_{\gamma_b}(\gamma) = P(\gamma_b \leq \gamma) = P(\gamma_i \leq \gamma)P(\gamma_{\beta_0} \leq \gamma)_{i=1\dots M} = \left[\prod_{i=1}^M (1 - e^{-\gamma/\gamma_i}) \right] (1 - e^{-\gamma/\gamma_{\beta_0}}) \quad (48)$$

In the aim of giving a coherent analytic form, we assume $\gamma_{M+1} = \gamma_{\beta_0}$, the PDF can be found by taking the derivative of (48), with respect to γ , and after doing some manipulations, $f_{\gamma_b}(\gamma)$ can be written as

$$f_{\gamma_b}(\gamma) = \sum_{n=1}^{M+1} (-1)^{(n+1)} \sum_{k_1=1}^{M-n+2} \sum_{k_2=k_1+1}^{M-n+3} \sum_{k_n=k_{n-1}+1}^{M+1} \prod_{j=1}^n (e^{-\gamma/\bar{\gamma}_{k_j}}) \sum_{j=1}^n \left(\frac{1}{\bar{\gamma}_{k_j}} \right) \quad (49)$$

by using the PDF in (49), the MGF can be written as

$$M_{\gamma_b}(s) = \int_0^{\infty} e^{-s\gamma} \sum_{n=1}^{M+1} (-1)^{(n+1)} \sum_{k_1=1}^{M-n+2} \sum_{k_2=k_1+1}^{M-n+3} \sum_{k_n=k_{n-1}+1}^{M+1} \prod_{j=1}^n (e^{-\gamma/\bar{\gamma}_{k_j}}) \sum_{j=1}^n \left(\frac{1}{\bar{\gamma}_{k_j}} \right) d\gamma \quad (50)$$

and the integral can be evaluated in a closed form as

$$M_{\gamma_b}(s) = \sum_{n=1}^{M+1} (-1)^{(n+1)} \sum_{k_1=1}^{M-n+2} \sum_{k_2=k_1+1}^{M-n+3} \sum_{k_n=k_{n-1}+1}^{M+1} \frac{\Psi_n}{s + \Psi_n} \quad (51)$$

where $\Psi_n = \sum_{j=1}^n (1/\bar{\gamma}_{k_j})$. Substituting (51) in (47) and finalizing the integration using [37], we can write PE in a closed form as follows,

$$P_E = A \sum_{n=1}^{M+1} (-1)^{(n+1)} \sum_{k_1=1}^{M-n+2} \sum_{k_2=k_1+1}^{M-n+3} \sum_{k_n=k_{n-1}+1}^{M+1} \left(1 - \sqrt{\frac{B/\Psi_n}{1 + 1/\Psi_n}} \right) \quad (52)$$

VII. RESULTS

In this section, we evaluate the proposed algorithm based on a synthetic example, which involves the mobile target tracking, the relay localization, the secure sensor node detection and the optimal communication path selection. The purpose of the synthetic example is to establish a baseline performance comparison on a relatively difficult problem.

In what follows, we proceed to compare the tracking accuracy of the adaptive quantized variational filtering (AQVF) algorithm, with the binary variational filtering (BVF) algorithm [28], and the centralized quantized particle filter (QPF) algorithm [38]. In the simulation, we have considered the following parameters: $\eta = 2$ for free space environment, the constant characterizing the sensor range is fixed for simplicity to $C = 1$, the cluster head noise power is

$\sigma_n^2 = 10^{-3}$, the sensor noise power is $\sigma_\epsilon^2 = 10^{-4}$, the maximum sensing range R_{max} (resp. the minimum sensing range R_{min}) is fixed to 10 m (resp. 0 m) and 200 particles were used in QVF, BVF and QPF algorithms. All sensors have equal initial battery energies of $E_i = 1$ Joule. All the simulations shown in this chapter are implemented with Matlab version 7.1, using an Intel Pentium CPU 3.4 GHz, 1.0 G of RAM PC.

The quantized proximity observation model, formulated in equation (7), was adopted for the QPF algorithm.

Fig. 4 and Fig. 5 show the performances comparison of AQVF and BVF, we can observe from Fig. 4.a) and Fig. 5.a) that, even with sudden changes in the target trajectory, the desired quality is achieved by AQVF algorithm and outperforms the BVF algorithm. Their tracking accuracies are compared in Fig. 4.b) and Fig. 5.b), in terms of Mean Square Error (MSE) where the sensors number varying in $\{200, \dots, 600\}$.

$$\text{MSE} = E((\mathbf{x} - \hat{\mathbf{x}})^2). \quad (53)$$

Where \mathbf{x} (resp. $\hat{\mathbf{x}}$) is the true trajectory (resp. the estimated trajectory).

The proposed method shows performances that demonstrate the effectiveness of the adaptive quantization and the impact of neglecting the information relevance of sensor measurements, when we used binary proximity sensors. As can be expected, with the increase in the amount of particles, QPF algorithm demonstrates much more accurate tracking at the cost of a higher computation complexity. Particularly, the computation time grows proportionally to the increment of the number of particles. The tracking accuracy of QPF algorithm is compared to that of the proposed method in Fig. 6 and Fig. 7. The simulation results obtained by the comparison of the smaller MSE of the proposed method to that of QPF method confirms the effectiveness of the the proposed method in terms of tracking accuracy and efficiency in a non-Gaussian context.

As can be expected, the MSE decreases for all the algorithms when the nodes density increases, as shown in Fig. 4, Fig. 5, Fig. 6 and Fig. 7). These figures show clearly that the proposed AQVF technique outperforms all other techniques when varying the nodes density

Concerning the precision of relay localization, Fig. 8.a) compares MSE of initial relay deployment and that of relay localization in the proposed scheme. One can notice that the proposed algorithm accomplishes a significant improvement in the precision of relay localization. Owing to including quantized measurements between activated relays, the proposed algorithm has much

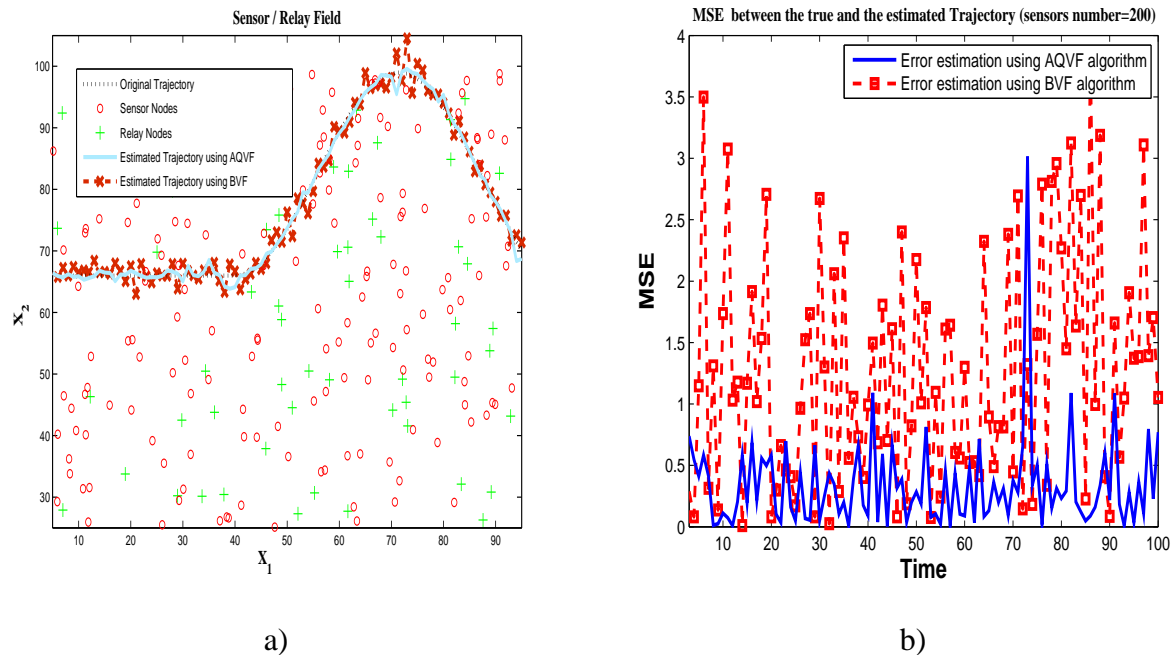


Fig. 4. a) Tracking accuracy between BVF and AQVF algorithms. b) Mean Square Error comparison between QPF and AQVF algorithms where the sensors number=200.

Comparison	Mean of MSE (MMSE)	Execution time
AQVF algorithm	0.2511 m	1.5938 s
BVF algorithm	2.69 m	1.5163 s
QPF algorithm	2.832 m	1.1273 s

TABLE I
COMPARISON OF TARGET TRACKING ALGORITHMS

more information for localization at every sampling instant. As shown in Tab. I, the proposed algorithm outperforms the classical algorithms by interdependently and continuously improving the estimation of both relays and target where the number of sensors is fixed to 400. The tracking accuracy and the localization precision are evaluated by their Mean of Mean Square Error (MMSE). The results illustrate that BVF and QPF outperform the proposed algorithm, with respect to the execution time, since our algorithm jointly detects the secure sensor node and terminates the relay re-localization phase during the target.

Fig. 9 shows the bit error rate for BPSK with different numbers of paths sensors-CH (M). As

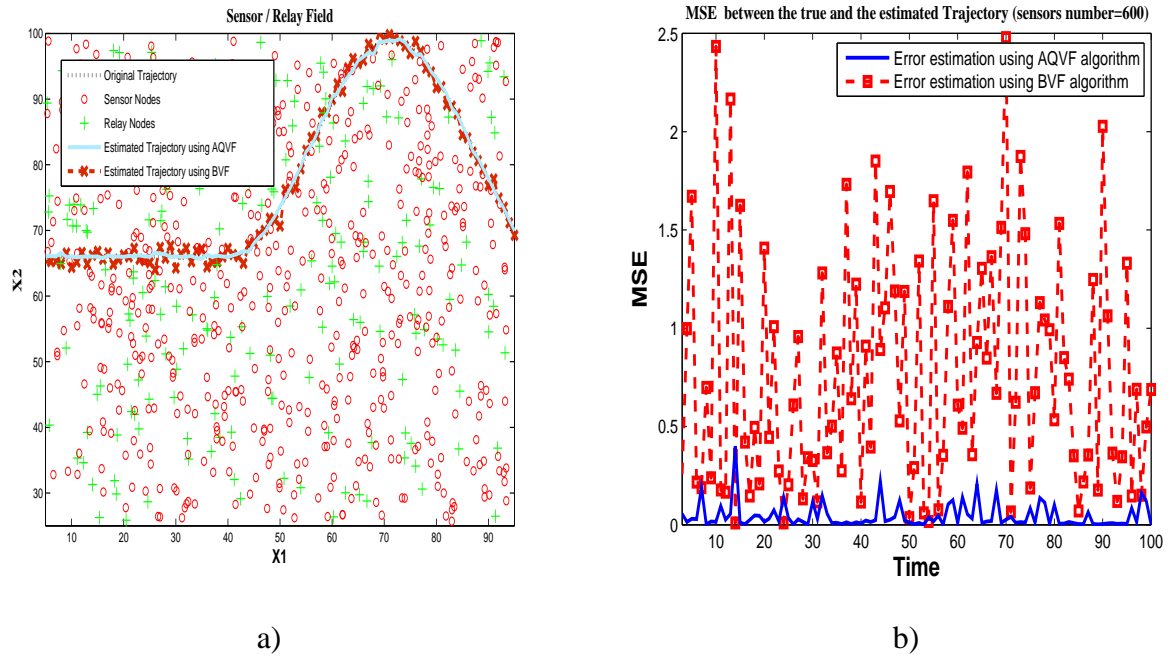


Fig. 5. a) Tracking accuracy between BVF and AQVF algorithms. b) Mean Square Error comparison between QPF and AQVF algorithms where the sensors number=300.

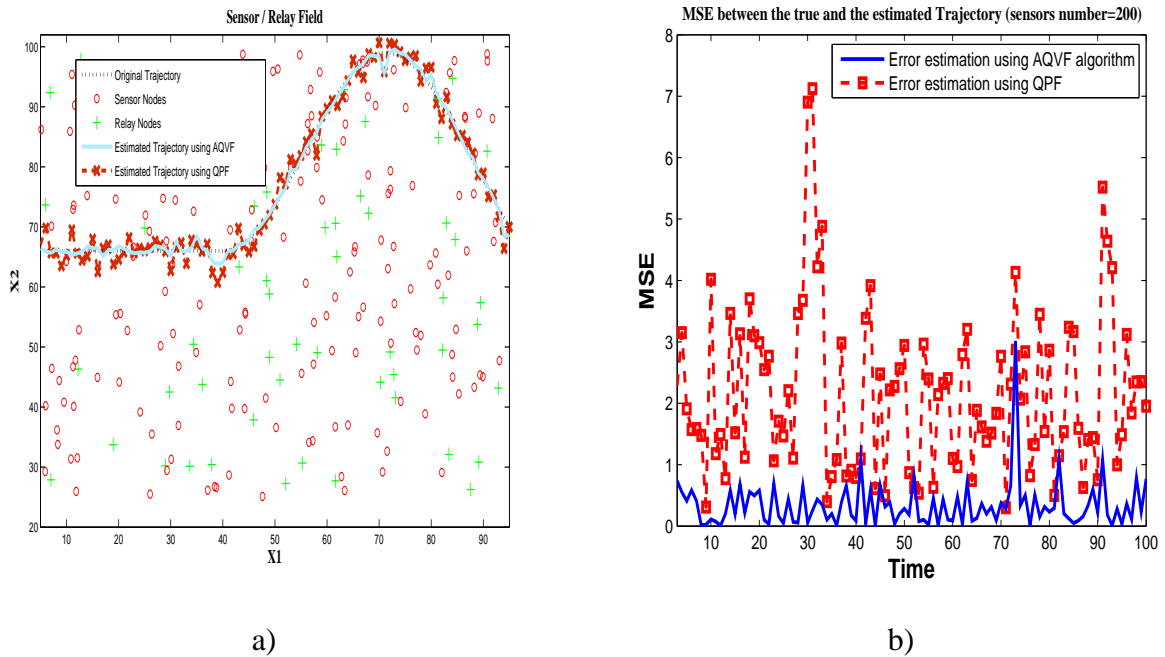


Fig. 6. a) Tracking accuracy between QPF and AQVF algorithms. b) Mean Square Error comparison between QPF and AQVF algorithms where the sensors number=200.

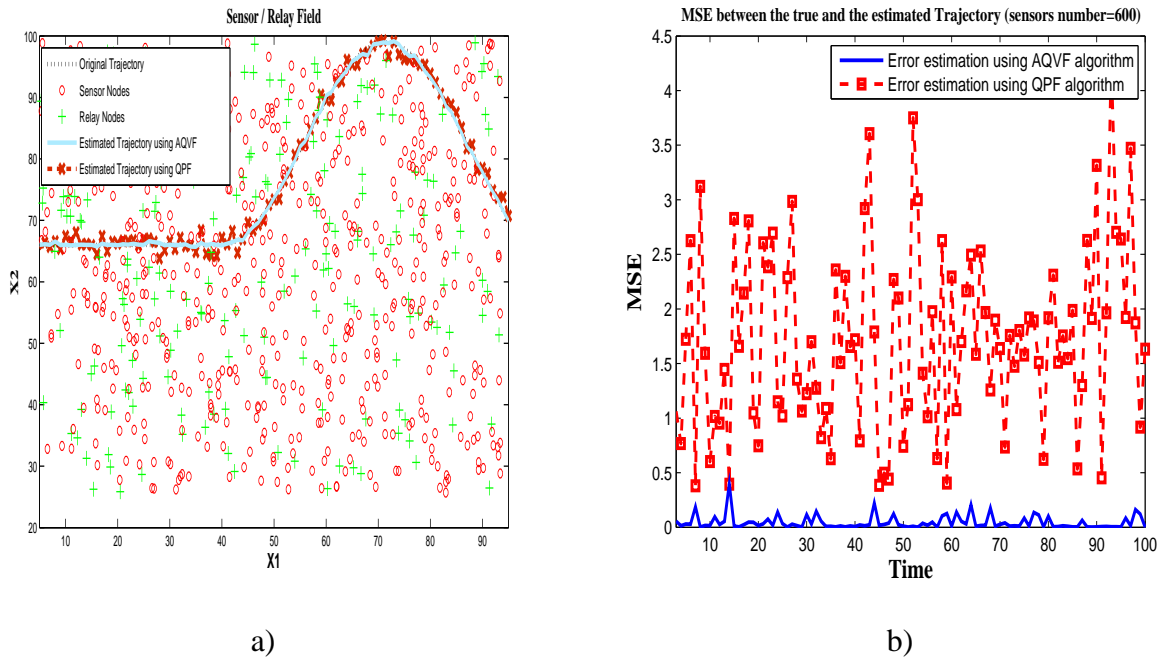


Fig. 7. a) Tracking accuracy between QPF and AQVF algorithms. b) Mean Square Error comparison between QPF and AQVF algorithms where the sensors number=600.

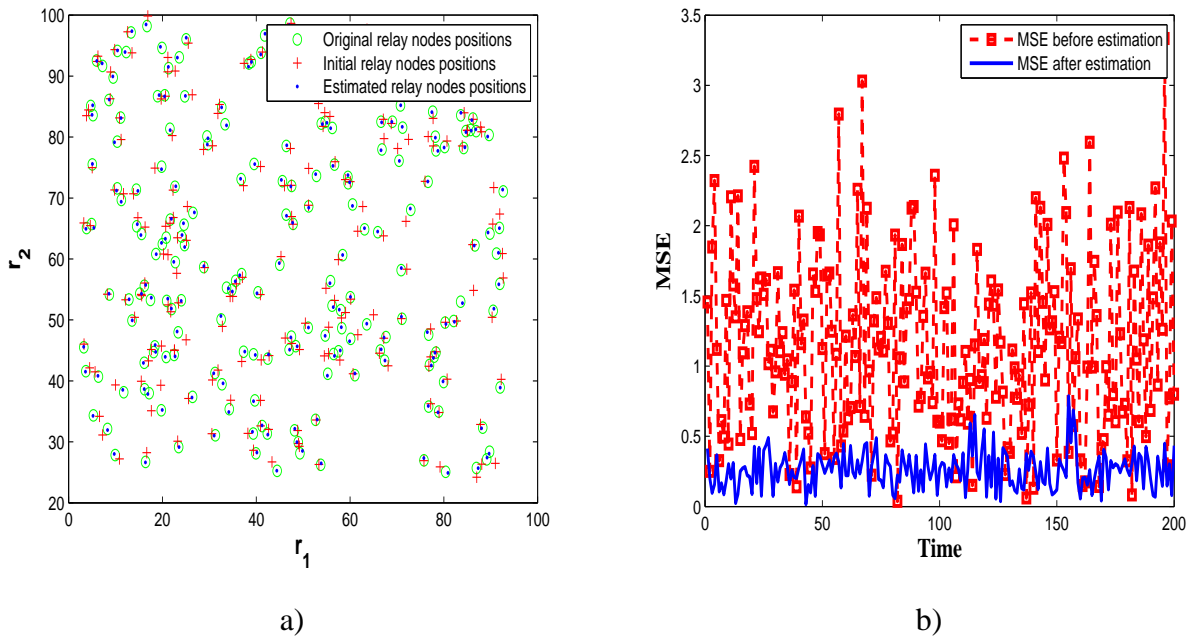


Fig. 8. a) Relay nodes estimation. b) MSE before and after estimation.

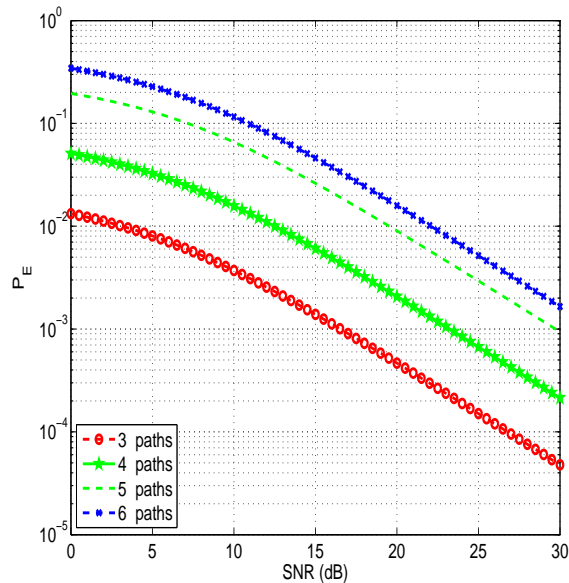


Fig. 9. Error performance for the best path selection scheme over Rayleigh fading channels.

can clearly be observed in a high SNR regime, the improvement of bit error rate is proportional to the number of sensors-CH links (M).

The evaluation of the energy consumption is done following the model proposed in [39]. We can observe from Fig. 10 that our model successfully balances the trade-off between the energy consumption even with several abrupt changes in the trajectory where the bits quantization number is fixed to 3.

Hereafter, we evaluate the performance of the malicious sensor node detection schemes presented in Section V in terms of ROC curves. The evaluation of a malicious sensor detection technique for WSNs depends on whether it can satisfy the mining accuracy requirements while maintaining the resource consumptions of WSNs to a minimum [40]. While keeping the false alarm rate low, the techniques of malicious sensor nodes detection are required to maintain a high detection rate. The detection rate represents the percentage of vulnerable sensor that are correctly considered as malicious, and the false positive rate, represents the percentage of normal sensor that are incorrectly considered as malicious. ROC curves [41] is frequently used to represent the trade-off between the detection rate and the false positive rate. The larger the area under the ROC curve, the better the performance of the corresponding techniques. An example of ROC

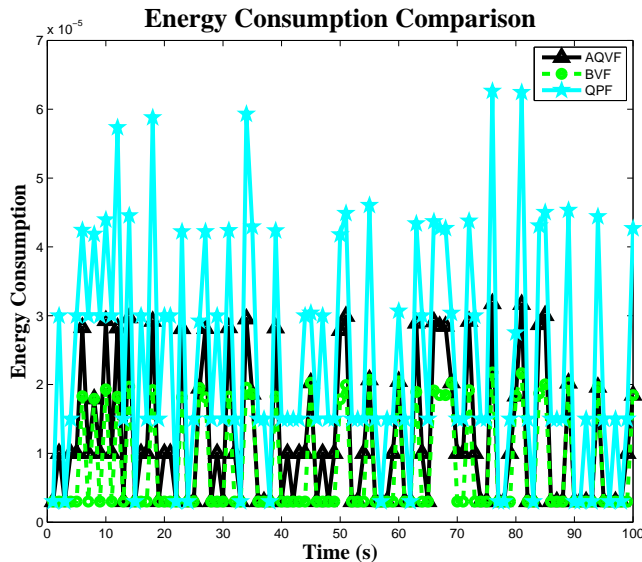


Fig. 10. Energy consumption comparison where the quantization level is fixed to 3.

curves is illustrated in Fig. 11.

VIII. CONCLUSION AND PERSPECTIVES

In this chapter, a distributed variational filtering solution to simultaneously localize relays, track mobile target, select the secure candidate sensor and select the best communication path is proposed in the context of WSN. Without any a priori information on the target motion, the proposed adaptive scheme algorithm aims at continuously updating and improving the estimation of the activated relay locations and the target trajectory, and optimizing the routing by selecting the best communication path. As the target can travel arbitrarily and the location information of the activated relays is rather coarse, a general state evolution model is proposed in this chapter to describe the hidden state, which is more adaptable to the non-linear / non-gaussian situation than other kinematic parameter models. The adaptive scheme algorithm is executed on a fully distributed cluster scheme In addition, which permits to minimize the resource consumption in WSN. The variational method allows an implicit compression of the exchanged statistics between clusters. This method permits not only to reduce inter-cluster communication, but also it terminates the error propagation problem, which is always unavoidable in other approximation methods. Furthermore, by incorporating the quantized proximity observation model, the

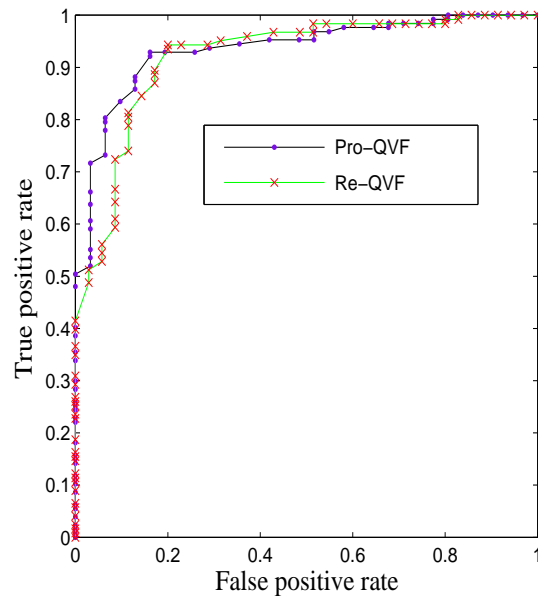


Fig. 11. ROC curves for the proposed malicious sensors detection techniques.

energy and bandwidth consumed by intra-cluster communication are dramatically reduced. In conclusion, as the target moves freely in WSN, a large number of quantized measurements are generated, which facilitates both the activated relays localization and the target tracking. The promising results obtained by simulations and presented in this chapter have shown clearly that the estimation of relay locations and that of the target are interdependently and continuously improved on-line.

REFERENCES

- [1] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," *IEEE Communications Letters*, vol. 9, no. 11, pp. 976–978, 2005.
- [2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, 2002.
- [3] Q. Wang, K. Xu, G. Takahara, and H. Hassanein, "Locally optimal relay node placement in heterogeneous wireless sensor networks."
- [4] S. Julier and J. Uhlmann, "Unscented filtering and non-linear estimation," in *Proceedings of the IEEE*, vol. 92, Mar. 2004, pp. 401–422.
- [5] P. Djuric, J. Z. J. Kotecha, Y. Huang, T. Ghirmai, M. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, pp. 19–38, Sep. 2003.

- [6] H. Snoussi and C. Richard, "Ensemble learning online filtering in wireless sensor networks," in *IEEE ICCS International Conference on Communications Systems*, 2006.
- [7] J. Li, H. Kao, and J. Ke, "Voronoi-based relay placement scheme for wireless sensor networks," 2009.
- [8] A. Bari, D. Teng, and A. Jaekel, "Optimal Relay Node Placement in Hierarchical Sensor Networks with Mobile Data Collector," in *Proceedings of the 2009 Proceedings of 18th International Conference on Computer Communications and Networks-Volume 00*. IEEE Computer Society, 2009, pp. 1–6.
- [9] A. Iranli, M. Maleki, and M. Pedram, "Energy efficient strategies for deployment of a two-level wireless sensor network," in *Proceedings of the 2005 international symposium on Low power electronics and design*. ACM, 2005, p. 238.
- [10] Y. Hou, Y. Shi, H. Sherali, and S. Midkiff, "On energy provisioning and relay node placement for wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, p. 2579, 2005.
- [11] S. Ergen and P. Varaiya, "Optimal placement of relay nodes for energy efficiency in sensor networks," in *Proceedings of the IEEE International Conference on Communications*, vol. 8, 2006, pp. 3473–3479.
- [12] Y. Wang, "Topology control for wireless sensor networks," *Wireless Sensor Networks and Applications*, pp. 113–147, 2008.
- [13] W. Chen, J. Hou, and L. Sha, "Dynamic clustering for acoustic target tracking in wireless sensor networks," *IEEE Transactions on Mobile Computing*, pp. 258–271, 2004.
- [14] J. Manyika and H. Durrant-Whyte, *Data fusion and sensor management: a decentralized information-theoretic approach*. Prentice Hall PTR Upper Saddle River, NJ, USA, 1995.
- [15] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for target tracking," *EURASIP Journal on Applied Signal Processing*, pp. 378–391, 2003.
- [16] E. Ertin, J. Fisher, and L. Potter, "Maximum mutual information principle for dynamic sensor query problems," in *Information Processing in Sensor Networks*. Springer, 2003, pp. 558–558.
- [17] X. Luo and G. Giannakis, "Energy-constrained optimal quantization for wireless sensor networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, pp. 1–12, 2008.
- [18] K. Yao, R. Hudson, C. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1555–1567, 1998.
- [19] D. Uciński and M. Patan, "D-optimal design of a monitoring network for parameter estimation of distributed systems," *Journal of Global Optimization*, vol. 39, no. 2, pp. 291–322, 2007.
- [20] L. Kaplan, "Global node selection for localization in a distributed sensor network," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 42, no. 1, pp. 113–135, 2006.
- [21] G. Han, L. Shu, J. Ma, J. Park, and J. Ni, "Power-Aware and Reliable Sensor Selection Based on Trust for Wireless Sensor Networks," *Journal of Communications*, vol. 5, no. 1, p. 23, 2010.
- [22] H. Snoussi and C. Richard, "Ensemble learning online filtering in wireless sensor networks," in *10th IEEE Singapore International Conference on Communication systems, 2006. ICCS 2006*, 2006, pp. 1–5.
- [23] J. Vermaak, N. Lawrence, and P. Perez, "Variational inference for visual tracking," in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*, vol. 1, 2003.
- [24] J. Teng, H. Snoussi, and C. Richard, "Prediction-Based Proactive Cluster Target Tracking Protocol for Binary Sensor Networks," in *2007 IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 234–239.
- [25] M. Mansouri, I. Ouchani, H. Snoussi, and C. Richard, "Cramer-Rao Bound-based adaptive quantization for target tracking in wireless sensor networks," in *IEEE/SP Workshop on Statistical Signal Processing, 2009. SSP'09*, 2009.

- [26] O. Barndorff-Nielsen, "Exponentially decreasing distributions for the logarithm of particle size," in *Proc. Roy. Soc.*, vol. 353, London, 1977, pp. 401–419.
- [27] J. Vermaak, N. Lawrence, and P. Perez, "Variational inference for visual tracking," in *Conf. Computer Vision and Pattern Recog, CVPR'03*, Jun. 2003.
- [28] J. Teng, H. Snoussi, and C. Richard, "Binary variational filtering for target tracking in sensor networks," in *IEEE/SP 14th Workshop on Statistical Signal Processing, 2007. SSP'07, 2007*, pp. 685–689.
- [29] B. Juang and L. Rabiner, "A probabilistic distance measure for hidden Markov models," *AT&T Bell Laboratories technical journal*, vol. 64, no. 2, pp. 391–408, 1985.
- [30] T. Cover and J. Thomas, *Elements of information theory*. Wiley-Interscience, 2006.
- [31] O. Ronen, J. Rohlicek, and M. Ostendorf, "Parameter estimation of dependence tree models using the EM algorithm," *IEEE Signal Processing Letters*, vol. 2, no. 8, pp. 157–159, 1995.
- [32] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [33] I. Misra, S. Dolui, and A. Das, "Enhanced-Efficient Adaptive Clustering Protocol for distributed sensor networks," *ICON*, 2005.
- [34] S. Ikki and M. Ahmed, "Performance analysis of cooperative diversity wireless networks over Nakagami-m fading channel," *IEEE Communications Letters*, vol. 11, no. 4, p. 334, 2007.
- [35] R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, 1987.
- [36] M. Simon and M. Alouini, *Digital communication over fading channels*. IEEE, 2005.
- [37] A. Papoulis and S. Pillai, *Probability, random variables and stochastic processes*. McGraw-Hill Education (India) Pvt Ltd, 2002.
- [38] L. Zuo, R. Niu, and P. Varshney, "A sensor selection approach for target tracking in sensor networks with quantized measurements," in *Proceedings of the 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing, II*, pp. 2521 – 2524.
- [39] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences*, vol. 8. Citeseer, 2000, p. 8020.
- [40] J. Gama and M. Gaber, *Learning from data streams: processing techniques in sensor networks*. Springer-Verlag New York Inc, 2007.
- [41] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection," in *Proceedings of the Third SIAM International Conference on Data Mining*, 2003, pp. 25–36.