SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2016, Vol. 59 022316:1–022316:15
doi: 10.1007/s11432-015-5415-6

Distributed state estimation and data fusion in wireless sensor networks using multi-level quantized innovation

Zhi ZHANG[1,2*], Jianxun LI[1,2] & Liu LIU[3]

[1]*Department of Automation, Shanghai Jiao Tong University, Shanghai* 200240*, China;*
[2]*Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai* 200240*, China;*
[3]*College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing* 210000*, China*

Received April 17, 2015; accepted May 27, 2015; published online October 29, 2015

Abstract Low energy consumption and limited power supply are significant factors for wireless sensor networks (WSNs); thus, distributed state estimation and data fusion with quantized innovation are explored. The universal features of practical WSNs are investigated, and a dynamic transmission strategy is introduced. Furthermore, quantization state estimation based on Bayesian theory is derived. Unlike previous algorithms suitable for processing scalar measurement, the proposed distributed data fusion algorithm is applicable to general vector measurement. Furthermore, the efficiency of the proposed dynamic transmission strategy is analyzed. It is concluded that the proposed algorithm is more efficient than previous methods, and its estimation accuracy comparable to that of the standard Kalman filtering, which is based on analog-amplitude vector measurement.

Keywords data fusion, distributed state estimation, target tracking, Kalman filtering, quantization, wireless sensor networks

Citation Zhang Z, Li J X, Liu L. Distributed state estimation and data fusion in wireless sensor networks using multi-level quantized innovation. Sci China Inf Sci, 2016, 59(2): 022316, doi: 10.1007/s11432-015-5415-6

1 Introduction

Distribution state estimation and data fusion are essential considerations in the study of wireless sensor networks (WSN), which have many applications, such as surveillance and target tracking [1–3]. In general, a WSN comprises a fusion center (FC) and a large number of sensor nodes with limited computation capability and energy. Each sensor node acquires measurements and transmits them to the FC. Then, the FC processes the measurements to obtain the state estimation with minimum mean square error (MMSE). The sensor nodes have limited energy; thus, it is necessary to minimize communication bandwidth, which can be realized by quantizing measurements before transmission. Thus, the focus of the subjects is transformed to state estimation and data fusion based on quantized noisy measurements.

Several data fusion methods have been proposed to conserve energy in WSNs [4–6]. One approach utilizes data fusion to set a node as active in the WSN, thereby resulting in reduced network resource

* Corresponding author (email: jasmine@sjtu.edu.cn)

consumption [4]. Another study proposed a delay-aware network structure for WSNs with in-network data fusion [5]. Furthermore, a quality-based data fusion approach to increase WSN performance has been proposed [6]. However, the main drawback to these approaches is that the resource and communication bandwidth limitations of sensor nodes were not considered.

On the other hand, many approaches with quantized innovations have been investigated [7–11]. For example, an SOI-KF approach for distributed state estimation based on one-bit innovation quantization and transmission was proposed [4]. Through quantization at multiple levels, another type of KF to quantize innovation was developed [5–6]. An efficient quantization scheme that exploited the fact that wireless data packets may not be transmitted to the FC in practical WSNs has also been proposed [7]. In another method, an efficient quantization scheme and state estimation quantization at multiple levels were combined [8], greatly improving performance and reducing bandwidth consumption.

However, these methods cannot be used for data fusion in WSNs because they are only suitable for vector state-scalar measurement, which means that the algorithms only process a single measurement scalar for each time step. To conserve energy and bandwidth, other algorithms [4–8] focus on the design of quantization; however, the processes of transmitting and encoding are ignored, which can also be exploited to conserve energy. Therefore, we propose a dynamic transmission strategy. The proposed dynamic transmission strategy, which combines dynamic encoding with the data packet structure, outperforms other previously reported methods [5–8]. Furthermore, even if the structure of the data packet is not considered, the proposed dynamic transmission strategy is more efficient than previous methods [5–8].

Finally, the algorithms in previous methods [4–8] are only suitable for processing measurement of a scalar transmitted from a single activated sensor per time step. However, measurements transmitted from activated sensors to the FC are vector valued at each time step. Thus, we introduce distributed data fusion for vector state-vector measurement. The proposed approach adopts Bayesian theory and broadens the scope of the previous algorithms [4–8].

The main objective of this paper is to propose a distributed estimation data fusion approach for WSNs. In the proposed method, an efficient quantization scheme that carefully considers different communication modes is introduced. By considering dynamic encoding to conserve energy and communication bandwidth, a wireless transmission strategy ignored in previous research is proposed, and a data fusion algorithm for vector state-vector measurement is derived.

The remainder of this paper is organized as follows. Model assumptions and preliminary information are introduced in Section 2. The proposed distributed data fusion approach with quantized innovation is presented in Section 3. A performance analysis is given in Section 4. Section 5 gives details about an experimental WSN test-bed, and several experiments are discussed in Section 6. Finally, Section 7 concludes the paper.

Notation: The probability density function of a random variable x conditioned on D is represented as $p(x|D)$, and the probability mass function is denoted by $\Pr\{x\}$. The Gaussian probability density function with mean $\mathrm{E}\{x\} = \mu$ and covariance matrix $\mathrm{cov}\{x\} = P$ is presented as $p(x) = \mathcal{N}[x; \mu, P]$, and $\Phi(\cdot)$ is the standard normal distribution function, namely $\Phi(z) = \int_{-\infty}^{z} \mathcal{N}[x; 0, 1]\mathrm{d}x$.

2 Model assumptions and preliminaries

Primarily, this paper considers a discrete-time linear Gaussian-Markov stochastic system in a WSN composed of N sensor nodes and a FC with M activated sensor nodes at time step k. This is expressed as follows:

$$\begin{cases} X_k = F_{k-1}X_{k-1} + w_{k-1}, \\ Y_k = H_k X_k + v_k, \end{cases} \tag{1}$$

where $X_k \in \mathbb{R}^n$ is the state vector, $Y_k = [y_k^1, \ldots, y_k^i, \ldots, y_k^M]^{\mathrm{T}}$ $(1 \leqslant i \leqslant M)$, $Y_k \in \mathbb{R}^{M\times 1}$ is the measurement vector, $y_k^i \in \mathbb{R}$ is the measurement scalar of the ith activated sensor, $F_{k-1} \in \mathbb{R}^{n\times n}$ is the dynamic model, $H_k = [h_k^1, \ldots, h_k^i, \ldots, h_k^M]^{\mathrm{T}}$ $(1 \leqslant i \leqslant M)$, $H_k \in \mathbb{R}^{M\times n}$ is the measurement model of the activated

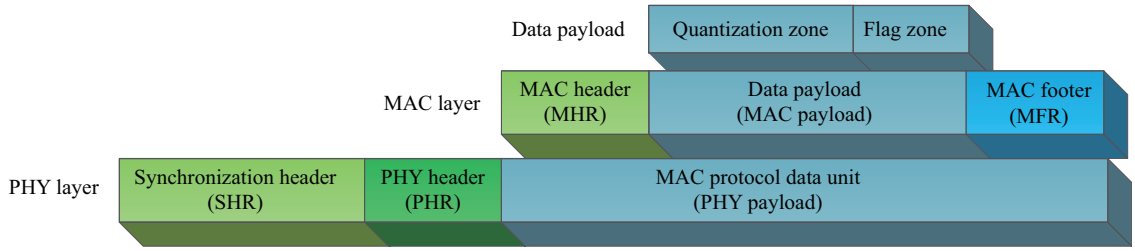

Figure 1 (Color online) IEEE 802.15.4 data packet structure.

sensors, and $h_k^i \in \mathbb{R}^{1\times n}$ is the measurement model of the ith activated sensor. Both $w_{k-1} \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^M$ are assumed to be zero mean, white, and Gaussian noises with covariance matrixes $Q_{k-1} \in \mathbb{R}^{n\times n}$ and $R_k \in \mathbb{R}^{M\times M}$, respectively. It is assumed that the initial state X_0 is Gaussian distributed with mean $\hat{X}_{0|0}$ and covariance matrix $P_{0|0}$. The two noise sequences and the initial state are assumed to be mutually independent. Furthermore, the measurement noises for different sensors are uncorrelated. In this paper, it is assumed that there is no communication loss in the wireless transmission process.

Rather than quantizing measurement y_k^i directly, the proposed method quantizes the normalized innovation ψ_k^i in the activated sensor. At time step k, the one-step predicted measurement $\hat{y}_{k|k-1}^i$ and the square root of the innovation's covariance matrix Δ_k^i are transmitted from the FC to the ith $(1 \leqslant i \leqslant M)$ activated sensor one at a time. For the ith activated sensor, $\hat{y}_{k|k-1}^i$ and Δ_k^i are given as follows:

$$\begin{cases} \hat{y}_{k|k-1}^i = h_k^i \hat{X}_{k|k-1}, \\ \Delta_k^i = \sqrt{h_k^i P_{k|k-1} {h_k^i}^{\mathrm{T}} + R_k^i}, \end{cases} \tag{2}$$

where $\hat{X}_{k|k-1}$ is the mean of the one-step predicted state at time step k and $P_{k|k-1}$ is the covariance matrix.

The normalized innovation ψ_k^i can be computed locally as follows:

$$\psi_k^i = (y_k^i - \hat{y}_{k|k-1}^i)/\Delta_k^i. \tag{3}$$

Then, ψ_k^i is quantized and expressed as follows:

$$d_k^i = q[\psi_k^i], \tag{4}$$

where d_k^i is the quantized normalized innovation of the activated sensor at time step k and $q_i[\cdot]$ is the quantization scheme. Note that ψ_k^i can be quantized and encoded into binary data.

The format of the data packet defined by this protocol is shown in Figure 1. The details of the quantization zone and flag zone are introduced in Section 3. Several remarks about the protocol are provided to facilitate the introduction of the data fusion algorithm in Section 3.

Remark 1: Data packets can be transmitted from the activated sensor nodes to the FC in two different modes, i.e., broadcast and peer-to-peer communication. These modes can be selected by changing a parameter in the MAC header (MHR). The same energy management mechanism is used in the PHY header (PHR) [9]; therefore, they consume equal energy.

Remark 2: The length of the MAC payload is a variable, the length of the PHY payload is contained in the PHR, and the lengths of the MHR and MAC footer (MFR) are constants according to the protocol. Thus, the length of the MAC payload in each packet can be derived.

3 Data fusion using quantized innovation vector

Here, a quantization scheme is introduced for normalized innovation. Furthermore, an innovative dynamic transmission strategy is explored by considering dynamic encoding and the structure of the data packet. Then, an innovative quantization KF is derived for vector state-vector measurement by adopting Bayesian theory. Finally, the data fusion algorithm is proposed based on this quantization scheme and transmission strategy.

3.1 Quantization scheme

The quantization scheme is as follows:

$$d_k^i = q[\psi_k^i] = \begin{cases} \lambda_l^b, & \text{if } \psi_k^i \in (\tau_1, \tau_2]; \\ \lambda_l^p, & \text{if } \psi_k^i \in (\tau_2, \tau_3]; \\ \vdots & \vdots \\ \lambda_1^b, & \text{if } \psi_k^i \in (\tau_{2l-1}, \tau_{2l}]; \\ \lambda_1^p, & \text{if } \psi_k^i \in (\tau_{2l}, \tau_{2l+1}]; \\ 0, & \text{if } \psi_k^i \in (\tau_{2l+1}, \tau_{2l+2}]; \\ -\lambda_1^p, & \text{if } \psi_k^i \in (\tau_{2l+2}, \tau_{2l+3}]; \\ -\lambda_1^b, & \text{if } \psi_k^i \in (\tau_{2l+3}, \tau_{2l+4}]; \\ \vdots & \vdots \\ -\lambda_l^p, & \text{if } \psi_k^i \in (\tau_{4l}, \tau_{4l+1}]; \\ -\lambda_l^b, & \text{if } \psi_k^i \in (\tau_{4l+1}, \tau_{4l+2}), \end{cases} \tag{5}$$

where $\tau_1 = -\infty$, $\tau_{4l+2} = +\infty$, and $\{\tau_1, \tau_2, \ldots, \tau_{4l+1}, \tau_{4l+2}\}$ represent the thresholds of the scheme. Here, L is the number of quantization levels, i.e., $L = 4l+1$. In the proposed method, the Lloyd-Max quantization scheme [12,13] is employed to obtain the optimal thresholds.

In the case of $d_k^i = 0$, there will be no transmission of the data packet to the FC. If $d_k^i = \pm\lambda_m^b$ $(1 \leqslant m \leqslant l)$, the data packet will be transmitted to the FC in the broadcast manner. By contrast, if $d_k^i = \pm\lambda_m^p$ $(1 \leqslant m \leqslant l)$, the peer to peer manner will be chosen. After receiving the packet, the FC can confirm the level of ψ_k^i by reading the transmission mode parameter in MHR and the data payload.

For instance, if $d_k^i \in \{\lambda_1^p, -\lambda_1^p\}$, it can be transmitted by 1 bit through peer to peer communication mode, and when $d_k^i \in \{\lambda_1^b, -\lambda_1^b\}$, d_k^i will be transmitted by 1 bit through broadcast communication mode. Therefore, when $d_k^i \in \{0, \pm\lambda_1^b, \pm\lambda_1^p\}$, it can be transmitted using binary data no more than 1 bit. Hence, if the normalized innovation ψ_k^i quantized by 1 bit in the ith activated sensor, this quantization scheme has five quantization levels. There are three levels in [7,8] and two in [5,6].

On the basis of the quantization scheme in [5,6] which has taken the feature that the data packet may not be transmitted to FC in practical WSNs, this scheme broadens the scope of the quantizer in [5,6] by taking into account the communication modes in the practical WSNs which are based on IEEE 802.15.4 standard.

3.2 Transmission strategy

For the ith activated sensor, a dynamic transmission strategy is put forward for data payload and it is different from those used in [4–8].

In terms of the data packet format, a quantization zone (i.e. quantized normalized innovation d_k^i) and a binary flag zone are defined in data payload. The structure is presented in Figure 1. According to Remark 2, the length of data payload (MAC payload) is a known number, and the length of binary flag zone is stored in PHR where there are several reserved zones which are long enough [9]. Thus, the length of quantization zone can be derived. The transmission strategy is summarized in Table 1. In Tables 1 and 2, $\{\pm\lambda_t^{b,p}\} = \{\pm\lambda_t^b, \pm\lambda_t^p\}$ $(1 \leqslant t \leqslant l)$.

In Table 1, a_i and b_i are the lengths of flag zone and quantization zone respectively; c_i is the total length of data payload $c_i = a_i + b_i$. According to Table 1, if $d_k^i \in \{\pm\lambda_1^{b,p}\}$, d_k^i can be transmitted with $c_i = 1$ bit (See the example in Subsection 3.1); if $d_k^i \in \{\pm\lambda_2^{b,p}, \ldots, \pm\lambda_5^{b,p}\}$, d_k^i can be transmitted with $c_i = 2$ bits. Furthermore, if $d_k^i \in \{\pm\lambda_6^{b,p}, \ldots, \pm\lambda_{17}^{b,p}\}$, d_k^i can be transmitted with $c_i = 3$ bits.

Therefore, when $d_k^i \in \{0\} \cup \{\pm\lambda_1^{b,p}\}$, it can be transmitted using data payload no more than 1 bit ($c_i = 0$ bit or $c_i = 1$ bit). So, the number of quantization levels L and the length of data payload c_i are summarized in Table 2.

Table 1 Dynamic transmission strategy

Flag zone	a_i (bit)	Quantization zone	b_i (bit)	c_i (bit)
	0	$\pm\lambda_1^{b,p}$	1	1
	0	$\pm\lambda_2^{b,p}, \pm\lambda_3^{b,p}$	2	2
0	1	$\pm\lambda_4^{b,p}$	1	2
1	1	$\pm\lambda_5^{b,p}$	1	2
	0	$\pm\lambda_6^{b,p}, \pm\lambda_7^{b,p}, \pm\lambda_8^{b,p}, \pm\lambda_9^{b,p}$	3	3
0	1	$\pm\lambda_{10}^{b,p}, \pm\lambda_{11}^{b,p}$	2	3
1	1	$\pm\lambda_{12}^{b,p}, \pm\lambda_{13}^{b,p}$	2	3
0 0	2	$\pm\lambda_{14}^{b,p}$	1	3
0 1	2	$\pm\lambda_{15}^{b,p}$	1	3
1 0	2	$\pm\lambda_{16}^{b,p}$	1	3
1 1	2	$\pm\lambda_{17}^{b,p}$	1	3

Table 2 Transmission scheme

d_k^i	L	c_i (bit)
If $d_k^i \in \{0, \pm\lambda_1^{b,p}\}$	$L = 5$	$c_i \leqslant 1$
If $d_k^i \in \{0, \pm\lambda_1^{b,p}, \ldots, \pm\lambda_5^{b,p}\}$	$L = 21$	$c_i \leqslant 2$
If $d_k^i \in \{0, \pm\lambda_1^{b,p}, \ldots, \pm\lambda_{17}^{b,p}\}$	$L = 69$	$c_i \leqslant 3$

Table 3 Dynamic transmission strategy without flag zone

Quantization zone	b_i (bit)	c_i (bit)
$\pm\lambda_1^{b,p}$	1	1
$\pm\lambda_2^{b,p}, \pm\lambda_3^{b,p}$	2	2
$\pm\lambda_4^{b,p}, \pm\lambda_5^{b,p}, \pm\lambda_6^{b,p}, \pm\lambda_7^{b,p}$	3	3

According to Table 2 and [4–8], with the same number of quantization levels L, the quantized innovation d_k^i can be transmitted with less bits using transmission strategy proposed in this paper than using those in [4–8]. The details of the comparison are discussed in Subsection 4.2. In other words, using the same number of transmission bits c_i, the state estimation algorithm based on this strategy is more accurate than those of [4–8]. Therefore, this transmission strategy can efficiently reduce the bandwidth consumption.

According to Table 7 in Subsection 4.2, when $c_i \leqslant 2$ bits, the coefficient of the error covariance matrix correction $\rho_k = 0.9943$ and the length of flag zone $a_i \leqslant 1$ bit. This means that the algorithm put forward in this paper is very close to that of the standard KF which is based on analogy-amplitude measurements. According to the network protocol, in PHR and MHR there are several reservation zones [9] which are long enough to store a_i (such as $a_i \leqslant 1$ bits and $\rho_k = 0.9943$), and the data packet contains the lengths of data payload and flag zone. After the FC's successful reception of the data packet, the FC can obtain all parameters about the transmission.

Therefore, the FC can confirm the level of ψ_k^i by reading the lengths of a_i, b_i, c_i, the flag and quantization zones in data payload, and the transmission mode parameter in the MHR. As a result, this transmission strategy can be applied to the majority of the practical low power consumption WSNs which are based on IEEE 802.15.4 protocol.

Furthermore, without considering the structure of data packet, the dynamic transmission strategy is only concentrated on the quantized normalized innovation d_k^i. And the dynamic transmission strategy is proposed in Table 3.

In Table 3, the quantization zone is the data payload, and $b_i = c_i$. Similarly, the number of quantization levels L and the length of data payload c_i are summarized in Table 4. And the details of the performance analysis are discussed in Subsection 4.2.

After quantization and transmission, the quantized normalized innovation vector D_k can be obtained

Table 4 Transmission scheme without flag zone

d_k^i	L	c_i (bit)
If $d_k^i \in \{0, \pm\lambda_1^{b,p}\}$	$L = 5$	$c_i \leqslant 1$
If $d_k^i \in \{0, \pm\lambda_1^{b,p}, \ldots, \pm\lambda_3^{b,p}\}$	$L = 13$	$c_i \leqslant 2$
If $d_k^i \in \{0, \pm\lambda_1^{b,p}, \ldots, \pm\lambda_7^{b,p}\}$	$L = 29$	$c_i \leqslant 3$

by FC at time step k. This can be given by

$$D_k = [d_k^1, \ldots, d_k^i, \ldots, d_k^M], \qquad 1 \leqslant i \leqslant M, \tag{6}$$

and

$$D_1^k = [D_1, \ldots, D_k], \qquad 1 \leqslant k. \tag{7}$$

3.3 State estimation and data fusion with quantized innovation

Here, we derive the distributed state estimation fusion, which can process vector measurement. This is expressed as follows:

$$\hat{X}_{k|k} = E[X_k|b_1^k] = E[E[X_k|b_1^{k-1}, y_k]|b_k], \tag{8}$$

where b_k is the quantized innovation scalar at time step k in [5,6]; y_k is the analogy-amplitude measurement scalar in [5,6]; b_1^k and b_1^{k-1} are the quantized innovation vectors which comprise the entire quantized innovation history to the current time k and $k-1$ respectively in [5,6].

Different from this method, this paper studies the MMSE of state estimation and data fusion for vector state-vector measurement case in WSNs.

The data fusion with quantized innovation vector is demonstrated as follows:

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + \sum_{i=1}^{M} [(P_{k|k}^i)^{-1} - (P_{k|k-1}^i)^{-1}], \tag{9}$$

where

$$P_{k|k}^i = P_{k|k-1}^i - \eta_k^i \frac{P_{k|k-1}^i (h_k^i)^{\mathrm{T}} h_k^i P_{k|k-1}^i}{h_k^i P_{k|k-1}^i (h_k^i)^{\mathrm{T}} + R_k^i}, \tag{10}$$

$$\eta_k^i = \frac{1}{2\pi} \times \left[\frac{(\mathrm{e}^{-\tau_j^2/2} - \mathrm{e}^{-\tau_{j+1}^2/2})}{\Phi(\tau_{j+1}) - \Phi(\tau_j)} \right]^2 - \frac{1}{\sqrt{2\pi}} \times \frac{\tau_j \mathrm{e}^{-\tau_j^2/2} - \tau_{j+1} \mathrm{e}^{-\tau_{j+1}^2/2}}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{11}$$

And

$$P_{k|k}^{-1} \hat{X}_{k|k} = P_{k|k-1}^{-1} \hat{X}_{k|k-1} + \sum_{i=1}^{M} [(P_{k|k}^i)^{-1} \hat{X}_{k|k}^i - (P_{k|k-1}^i)^{-1} \hat{X}_{k|k-1}^i], \tag{12}$$

where

$$\hat{X}_{k|k}^i = \hat{X}_{k|k-1}^i + \sigma_k^i \frac{P_{k|k-1}^i (h_k^i)^{\mathrm{T}}}{\sqrt{h_k^i P_{k|k-1}^i (h_k^i)^{\mathrm{T}}}}, \tag{13}$$

$$\sigma_k^i = \frac{1}{\sqrt{2\pi}} \times \frac{(\mathrm{e}^{-\tau_j^2/2} - \mathrm{e}^{-\tau_{j+1}^2/2})}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{14}$$

Proof. See Appendix A.

In (9) and (12), $P_{k|k-1} = P_{k|k-1}^i$ and $\hat{X}_{k|k-1} = \hat{X}_{k|k-1}^i$ respectively. When the number of the activated sensors $M = 1$ at time step k, it is obvious that the data fusion can become the decentralized Kalman filtering based on quantized innovation scalar.

Table 5 Transmission bandwidth with different quantization levels

	Zhang's bandwidth c_i (bit)	You's bandwidth c_i (bit)	Msechu's bandwidth c_i (bit)
$L = 5$	1	2	3
$L = 21$	2	5	5

Table 6 Transmission bandwidth without flag zone

	Zhang's bandwidth c_i (bit)	You's bandwidth c_i (bit)	Msechu's bandwidth c_i (bit)
$L = 5$	1	2	3
$L = 13$	2	4	4

3.4 Algorithm

Based on the quantization scheme and dynamic transmission strategy, a new data fusion algorithm using quantized innovation vector is proposed (Algorithm 1).

Algorithm 1 Data fusion algorithm using quantized innovation vector

(1) Initialization:
$\hat{X}_{0|0}$, $P_{0|0}$, L, and $\{\tau_1, \ldots, \tau_{4l+2}\}$ are set in the FC.
for $i = 1$; $i \leqslant N$; $i++$ **do**
  L and $\{\tau_1, \ldots, \tau_{4l+2}\}$ are set in the ith sensor node.
end for
(2) Estimation:
while $k \geqslant 1$ **do**
  According to (A3) and (A4) in Appendix A, the one-step state prediction $\hat{X}_{k|k-1}$ and its covariance matrix $P_{k|k-1}$ can be calculated in the FC.
  for $j = 1$; $j \leqslant M$; $j++$ **do**
    $\hat{y}^j_{k|k-1}$ and Δ^j_k are computed in the FC which transmits them to the jth activated sensor node. Then in the sensor nodes ψ^j_k and d^j_k can be calculated by (3) and (5) respectively. When $d^j_k = 0$, there will be no transmission of the data packet to the FC. By contrast, when $d^j_k \neq 0$, the data packet will be transmitted from the sensor node to the FC by applying the transmission strategy in this paper;
  end for
  Finally, the state estimation $P_{k|k}$ and $\hat{X}_{k|k}$ can be obtained by (9) and (12).
end while

4 Performance analysis

This section mainly discusses the performance of the proposed data fusion algorithm.

4.1 Performance analysis of the transmission strategy

For the same transmission bandwidth bytes c_i in [5,7], the number of the quantization levels L which can be transmitted is computed by $L = 2^{c_i}$ and $L = 2^{c_i} + 1$ respectively.

On the other hand, when the numbers of the quantization levels L are the same, the maximal transmission bytes c_i in Msechu [5], You [7], and the proposed algorithm are summarized in Table 5.

This table indicates that the transmission strategy proposed in this paper is much better than those in [5,7] since it reduces more bandwidth consumption. And when the transmission bandwidth bytes c_i are the same, normalized innovation can be quantized into more levels by this dynamic transmission strategy than in [5,7].

Similarly, if do not considering the structure of data packet, the dynamic transmission strategy is only concentrated on the quantized normalized innovation d^i_k. And the maximal transmission bytes c_i in this dynamic transmission strategy, You [7], and Msechu [5] are proposed in Table 6.

Table 7 The coefficient of the error covariance matrix correction ρ_k with different algorithms

	Zhang's ρ_k	You's ρ_k	Msechu's ρ_k
$c_i \leqslant 1$ bit	0.9201	0.8098	0.6366
$c_i \leqslant 2$ bits	0.9943	0.9201	0.8825

4.2 Performance analysis of the distributed state estimation

The error covariance matrix correction can be obtained by

$$\Delta P_k = P_{k|k} - P_{k|k-1} = \rho_k \frac{P_{k|k-1} h_k^{\mathrm{T}} h_k P_{k|k-1}}{h_k P_{k|k-1} h_k^{\mathrm{T}} + R_k}, \tag{15}$$

where ρ_k is the coefficient of the error covariance matrix correction.

Here, different quantization state estimation algorithms can be evaluated by the coefficient of the error covariance matrix correction ρ_k. Obviously, for standard KF $\rho_k = 1$; for estimators using quantized innovation, ρ_k is calculated by [5],

$$\rho_k = \frac{1}{2\pi} \sum_{j=1}^{L} \frac{\left[\exp(-\tau_j^2/2) - \exp(-\tau_{j+1}^2/2)\right]^2}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{16}$$

When the amounts of transmission bandwidth c_i are the same for different algorithms, the values of ρ_k are presented in Table 7.

According to Table 7, the data fusion algorithm put forward in this paper is more accurate than those in [5–8].

5 The testbed of wireless ultra sound sensor network

This section presents the design, architecture, and framework of a WSN testbed for target tracking.

5.1 Testbed architecture

This testbed comprises three main components: (a) a networked wireless ultrasound sensor localization system; (b) control, monitoring, and fusion centers; and (c) single or multiple mobile targets. The architecture of the testbed is shown in Figure 2(a). Typically, a moving target can be performed by two different modes, i.e., active mode and passive mode. For the active mode, the target is equipped with an active signal transmitter that can broadcast a signal to sensor nodes. In contrast, in the passive approach, the sensor nodes can transmit the signal to detect the mobile target. The passive mode is adopted in this testbed.

The targets and the FC are shown in Figure 2(b). The left shows two mobile targets, and the others are the monitor screen of the FC.

5.2 Testbed subsystem and software architecture

The subsystem and software architecture of the testbed are shown in Figure 3. Initiation of the testbed is performed as follows. First, a beacon node is activated and communicates with the FC via RS-232. Second, the sensor network is developed using the Zigbee networks protocol, which is based on IEEE 802.15.4. Third, the activated sensor node acquires measurements and transmits them to the beacon node via wireless communication. Finally, the beacon node transmits measurements to the FC. A sensor node has three main parts, i.e., a wireless RF module, an ultrasound scheduling module, and an ultrasonic module. The wireless RF module and the scheduling module are connected by RS-232, and the scheduling module can control the three ultrasound sensors through an inter IC bus interface. In consideration of the low energy consumption requirement, the Zigbee wireless communication protocol and CC2430 chip are adopted. The scheduling module is connected to the wireless module via RS-232, which can transmit

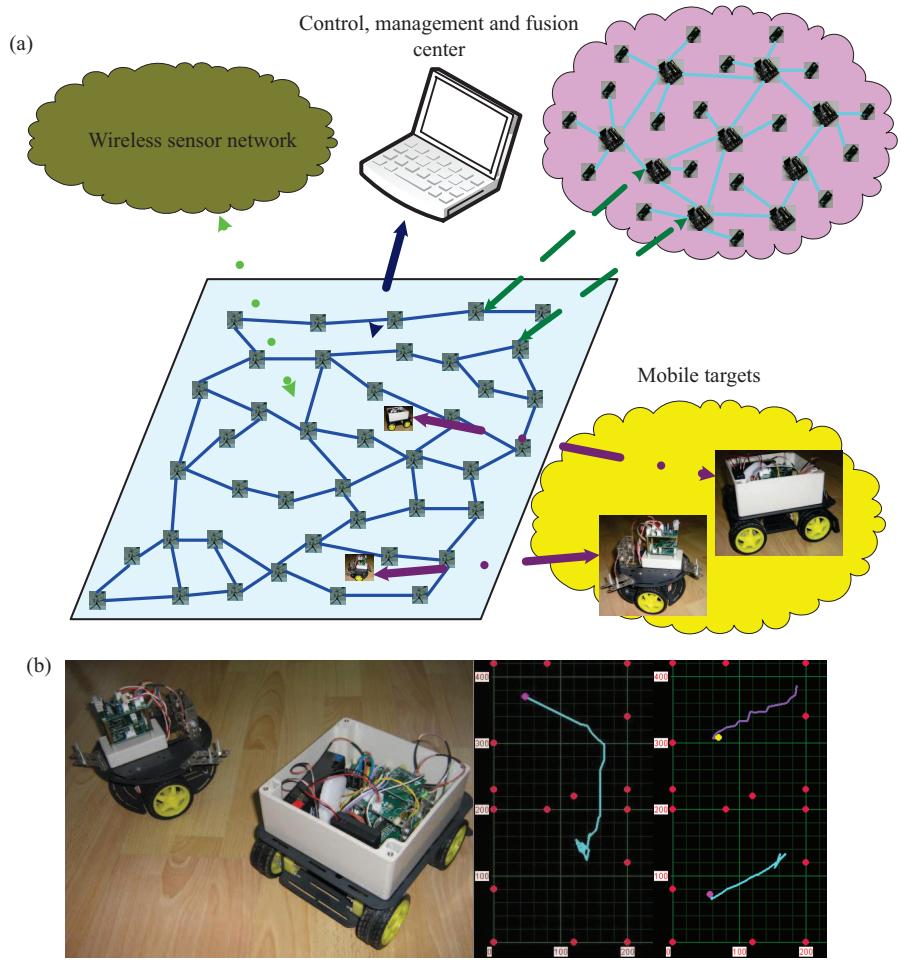

Figure 2 (Color online) Target tracking testbed. (a) Testbed components; (b) targets and monitor screen of the FC.

instructions to the scheduling module. Furthermore, the scheduling module can communicate with the ultrasound modules by inter IC bus. Depending on the instruction, the scheduling module can control one, two, or three ultrasonic modules simultaneously. An ATmega8 microcontroller is used for the scheduling and ultrasonic modules.

6 Results of simulation experiments

Primarily, a WSN is simulated for target tracking experiments. It consists of a FC and $N(N = 50)$ sensors. The sensors are randomly scattered and their coordinates are known. The single target in this region is regarded as a point object. The target state at time step k is defined by $X_k = [x_k, \dot{x}_k, \ddot{x}_k, y_k, \dot{y}_k, \ddot{y}_k]^{\mathrm{T}}$. In this formula, x_k and y_k are the positions at the X and Y coordinates respectively; $\dot{x}_k$ and $\dot{y}_k$ are the velocities at the two axises; $\ddot{x}_k$ and $\ddot{y}_k$ are the acceleration of the target at the X and Y coordinates respectively.

The time step is assumed to be constant, namely $T = 1$ s. During this period of time, the three sensors which are nearest of the moving target will always be activated. Then, each activated sensor measures the distance between the moving target and itself. In this WSN, the state estimation algorithms are run Monte-Carlo 1000 times in the experiments.

This paper assumes that the target moves in the 2D Cartesian coordinate system, and the dynamic model is assumed as the constant velocity (CV) model for the state vector. For the ith activated sensor,

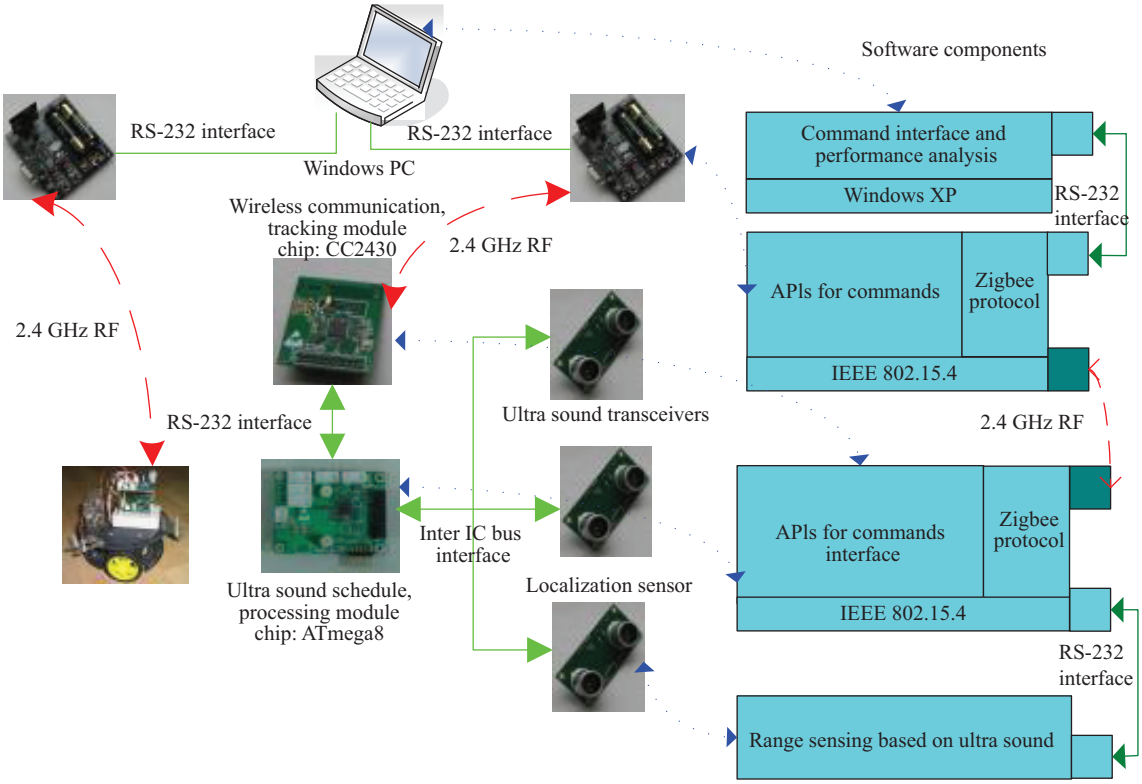

Figure 3 (Color online) Subsystem of the test-bed and software architecture.

its coordinate is (x_i, y_i), and the measurements are modeled as

$$Z_k^i = h_k^i(X_k) + v_k^i = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} + v_k^i, \tag{17}$$

where the white Gaussian noise sequence v_k^i and the process noise sequence w_{k-1} in (1) are independent from each other and their variances are $\mathrm{var}(w_{k-1}) = 0.01^2$ and $\mathrm{var}(v_k) = 10^2$. The measurement equation is nonlinear and is linearized in the simulation experiments. A comparison is made among the MQI-KF in this paper, the other algorithms in [5,7], and the standard KF (EKF) based on analogy-amplitude measurements.

6.1 Simulation without the transmission strategy

To compare the performance of data fusion without the transmission strategy to that of previous methods [5,7], the quantized innovation was obtained with transmission bandwidth of 1 bit.

Figure 4 shows the root mean square error (RMSE) of the positions at the X and Y coordinates. According to Figure 4, when the transmission bandwidth is 1 bit, the performance of the proposed algorithm is superior to that of previous methods [5,7] because the proposed quantization scheme can quantize the normalized innovation into more quantization levels using only 1 bit.

6.2 Simulation with the transmission strategy

In the second experiment, to evaluate the efficiency of the different transmission strategies, a comparison was made based on equal transmission bandwidth.

Figure 5 shows that by adopting equal transmission bandwidth (2 bits), the proposed data fusion algorithm is obviously more accurate than previous methods [5,7], and its estimation is very close to the standard KF (EKF), which is based on analog-amplitude measurements. The reason for this is that the algorithm is based on an efficient transmission strategy. In other words, based on the equal transmission

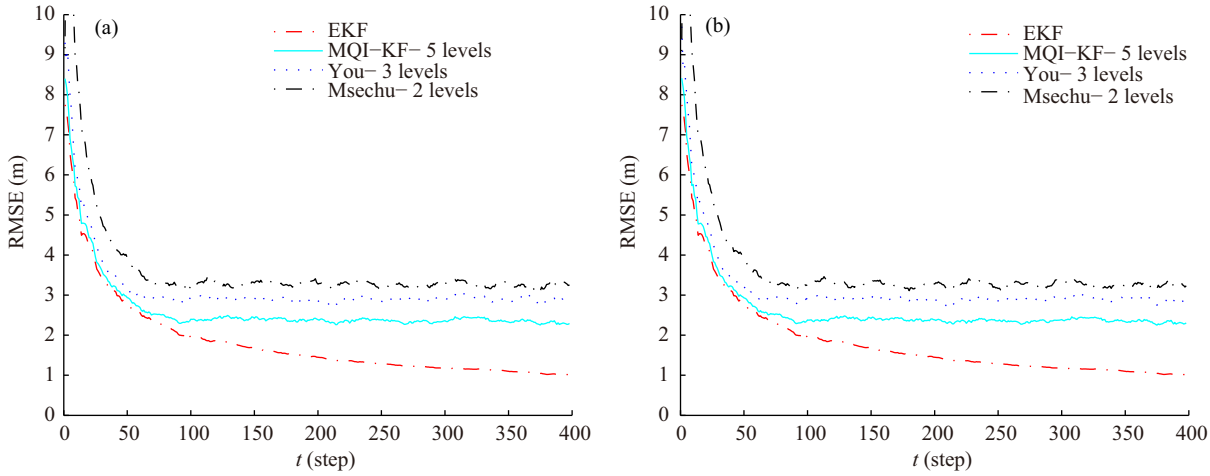

Figure 4 (Color online) RMSE of the positions for the algorithms using 1 bit. (a) X-axis; (b) Y-axis.

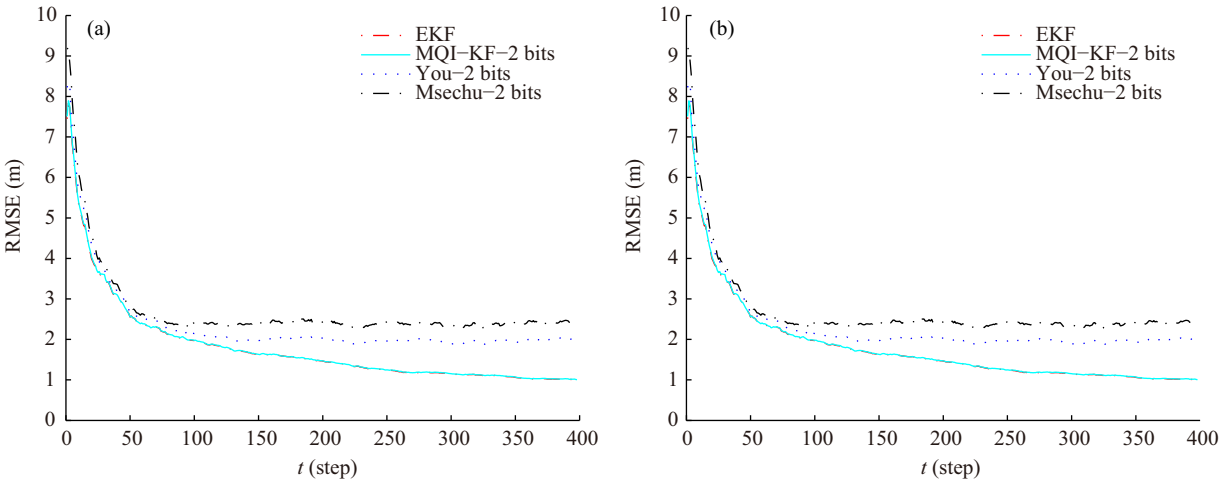

Figure 5 (Color online) RMSE of the positions for the algorithms using 2 bits. (a) X-axis; (b) Y-axis.

bandwidth, the normalized innovation can be quantized and transmitted into more levels than previous methods [5,7].

7 Conclusion

A transmission strategy for WSNs has been proposed. A distributed data fusion algorithm based on quantized innovation has also been proposed for general vector measurement. This algorithm is derived by adopting Bayesian theory. The proposed algorithm is based on the presented quantization scheme and transmission strategy and considers the features of WSNs. According to a performance analysis and simulation experiments, the proposed data fusion algorithm is more efficient than previously proposed methods [5–8].

Acknowledgements This work was jointly supported by National Natural Science Foundation of China (Grant No. 61175008), State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (Grant No. CEMEE2014K0301A), National Natural Science Foundation of Jiangsu Province (Grant No. BK20140896).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Masazade E, Niu R, Varshney P K. Dynamic bit allocation for object tracking in wireless sensor networks. IEEE Trans Signal Process, 2012, 60: 5048–5063
2 Leng M, Tay W, Quek T, et al. Distributed local linear parameter estimation using Gaussian SPAWN. IEEE Trans Signal Process, 2015, 63: 244–257
3 Braca P, Willett P, LePage K, et al. Bayesian tracking in underwater wireless sensor networks with port-starboard ambiguity. IEEE Trans Signal Process, 2014, 62: 1864–1878
4 Soltani M, Hempel M, Sharif H. Data fusion utilization for optimizing large-scale wireless sensor networks. In: Proceedings of the IEEE International Conference on Communications, Sydney, 2014. 367–372
5 Cheng C, Leung H, Maupin P. A delay-aware network structure for wireless sensor networks with in-network data fusion. IEEE Sens J, 2013, 13: 1622–1631
6 Kreibich O, Neuzil J, Smid R. Quality-based multiple-sensor fusion in an industrial wireless sensor network for MCM. IEEE Trans Ind Electron, 2014, 61: 4903–4911
7 Riberio A, Giannaki G B, Rounmeliotis S I. SOI-KF: distributed Kalman filtering with low-cost communications using the sign of innovations. IEEE Trans Signal Process, 2006, 54: 4782–4795
8 Msechu E J, Roumeliotis S I, Ribeiro A, et al. Decentralized quantized Kalman filtering with scalable communication cost. IEEE Trans Signal Process, 2008, 56: 3727–3741
9 Msechu E J, Ribeiro A, Roumeliotis S I, et al. Distributed Kalman filtering based on quantized innovation. In: IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, 2008. 3293–3296
10 You K, Xie L, Sun S, et al. Multiple-level quantized innovation Kalman filtering. In: Proceedings of the 17th IFAC World Congress, COEX, 2008. 1420–1425
11 You K, Xie L, Sun S, et al. Quantized filtering of linear stochastic system. Trans Inst Meas Contr, 2011, 33: 683–689
12 Ben-Israel A, Greville T. Generalized Inverses: Theory and Applications. 2nd ed. New York: Springer, 2003
13 Bar-Shalom Y, Li X, Kirubarajan T. Estimation with Applications to Tracking and Navigation. New York: Wiley, 2001

Appendix A

For the $p(X_k|D_1^k)$, it can be obtained by two steps.

(1) Prediction step

For the $p(X_k|D_1^k)$, it can be assumed as a Gaussian conditional density

$$p(X_{k-1}|D_1^{k-1}) = [(2\pi)^{\frac{n}{2}}|P_{k-1|k-1}|^{\frac{1}{2}}]^{-1}\exp\left\{-\frac{(X_{k-1}-\hat{X}_{k-1|k-1})^{\mathrm{T}}P_{k-1|k-1}^{-1}(X_{k-1}-\hat{X}_{k-1|k-1})}{2}\right\}, \tag{A1}$$

where $\hat{X}_{k-1|k-1}$ is the mean and $P_{k-1|k-1}$ is its covariance.

According to (1), the X_k is a linear combination of X_{k-1} and w_{k-1}, and $p(w_{k-1})$ is also Gaussian. Hence the density $p(X_k|D_1^{k-1})$ is Gaussian.

Thus, assumption that X_k and D_1^{k-1} are independent, the mean of the $p(X_k|D_1^{k-1})$ can be obtained by

$$\hat{X}_{k|k-1} = E[X_k|D_1^{k-1}] = F_{k-1}\hat{X}_{k-1|k-1}, \tag{A2}$$

and its covariance $P_{k|k-1}$ is

$$\begin{aligned} P_{k|k-1} &= E[(X_k-\hat{X}_{k|k-1})(X_k-\hat{X}_{k|k-1})^{\mathrm{T}}|D_1^{k-1}] \\ &= F_{k-1}P_{k-1|k-1}F_{k-1}^{\mathrm{T}}+Q_{k-1}. \end{aligned} \tag{A3}$$

So the $p(X_k|D_1^{k-1})$ can be presented by

$$p(X_k|D_1^{k-1}) = [(2\pi)^{\frac{n}{2}}|P_{k|k-1}|^{\frac{1}{2}}]^{-1}\exp\left\{-\frac{(X_k-\hat{X}_{k|k-1})^{\mathrm{T}}P_{k|k-1}^{-1}(X_k-\hat{X}_{k|k-1})}{2}\right\}. \tag{A4}$$

(2) Correction step

At time step k, $D_k = \{d_k^1, \ldots, d_k^i, \ldots, d_k^M\}$ $(1 \leqslant i \leqslant M)$.

Firstly, we assume $M = 1$ (i.e. $D_k = d_k^1$) and $\psi_k^1 \in (\tau_j, \tau_{j+1}]$ $(1 \leqslant j \leqslant L)$. To be brief, let $d = d_k^1$, $\psi = \psi_k^1$, $h = h_k^1$, $\hat{y}_{k|k-1} = \hat{y}_{k|k-1}^1$, $\Delta = \Delta_k^1$, $v = v_k^1$, $R = R_k^1$, $\tilde{X} = X_k - \hat{X}_{k|k-1}$, and $P = P_{k|k-1}$ at this time step.

The posterior density $p(X_k|D_1^{k-1}, D_k)$ can be derived by following the Bayes' rule:

$$p(X_k|D_1^{k-1}, D_k) = p(X_k|D_1^{k-1}, d) = \frac{p(d|X_k, D_1^{k-1})}{p(d|D_1^{k-1})} \times p(X_k|D_1^{k-1}). \tag{A5}$$

Since the quantized normalized innovation d is a discrete value, by referring to (5) and $\psi \in (\tau_j, \tau_{j+1}]$, (A5) can be rewritten as

$$p(X_k|D_1^{k-1}, D_k) = \frac{\Pr\{\psi \in (\tau_j, \tau_{j+1}]|X_k, D_1^{k-1}\}}{\Pr\{\psi \in (\tau_j, \tau_{j+1}]|D_1^{k-1}\}} \times p(X_k|D_1^{k-1}). \tag{A6}$$

Because the $p(X_k|D_1^{k-1})$ has already been determined at the prediction step, the aim is to describe the posterior conditional density $p(X_k|D_1^{k-1}, D_k)$ explicitly. For the (A6), we have that

$$p(X_k|D_1^{k-1}, D_k) = \frac{\int_{\tau_j}^{\tau_{j+1}} p(\psi|X_k, D_1^{k-1})\mathrm{d}\psi}{\int_{\tau_j}^{\tau_{j+1}} p(\psi|D_1^{k-1})\mathrm{d}\psi} \times p(X_k|D_1^{k-1}). \tag{A7}$$

By referring to (1) and (3), there will be

$$\psi = (hX_k + v - \hat{y}_{k|k-1})/\Delta, \tag{A8}$$

where $\Delta^2 = hPh^{\mathrm{T}} + R$ in (2).

For the denominator $p(\psi|D_1^{k-1})$ in (A7), ψ in (A8) is a linear combination of X_k and v, and $p(v)$ in (1) are Gaussian densities. Hence the $p(\psi|D_1^{k-1})$ is a Gaussian density, its mean $\hat{\psi} = 0$ and the conditional covariance $P(\hat{\psi}) = 1$.

So $p(\psi|D_1^{k-1}) = \mathcal{N}(\psi; 0, 1)$, and there will be

$$p(\psi|D_1^{k-1}) = (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{\psi^{\mathrm{T}}\psi}{2}\right). \tag{A9}$$

For the numerator $p(\psi|X_k, D_1^{k-1})$ in (A7), the ψ conditioning on X_k means that X_k is regarded as a known vector in the conditional density at this time step. On the other hand, Δ, h and $\hat{y}_{k|k-1}$ in (A8) are constants. Because ψ is a linear combination of a known vector in (A8) and a Gaussian variable v, the density $p(\psi|X_k, D_1^{k-1})$ is Gaussian, its mean $\hat{\psi}^X = h\tilde{X}/\Delta$ and the covariance $P(\hat{\psi}^X) = R/\Delta^2$.

Hence $p(\psi|X_k, D_1^{k-1}) = \mathcal{N}[\psi; \hat{\psi}^X, P(\hat{\psi}^X)]$, there is

$$p(\psi|X_k, D_1^{k-1}) = [(2\pi)^{\frac{1}{2}}|P(\hat{\psi}^X)|^{\frac{1}{2}}]^{-1} \exp\left\{-\frac{(\psi - \hat{\psi}^X)^{\mathrm{T}} P(\hat{\psi}^X)^{-1}(\psi - \hat{\psi}^X)}{2}\right\}. \tag{A10}$$

Substituting (A4), (A9), and (A10) into (A7), we can obtain

$$p(X_k|D_1^{k-1}, D_k) = [(2\pi)^{\frac{n+1}{2}}|PR/\Delta^2|^{\frac{1}{2}}]^{-1} \frac{\gamma}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}, \tag{A11}$$

where

$$\gamma = \int_{\tau_j}^{\tau_{j+1}} \exp\left[-\frac{(\psi - \hat{\psi}^X)^{\mathrm{T}} P(\hat{\psi}^X)^{-1}(\psi - \hat{\psi}^X)}{2}\right] \mathrm{d}\psi \times \exp\left(-\frac{\tilde{X}^{\mathrm{T}} P^{-1} \tilde{X}}{2}\right). \tag{A12}$$

Furthermore,

$$E[X_k|D_1^{k-1}, D_k] = \hat{X}_{k|k-1} + E[\tilde{X}|D_1^{k-1}, D_k]. \tag{A13}$$

And

$$E[\tilde{X}|D_1^{k-1}, D_k] = \frac{[(2\pi)^{\frac{n+1}{2}}|PR/\Delta^2|^{\frac{1}{2}}]^{-1}}{\Phi(\tau_{j+1}) - \Phi(\tau_j)} \times \gamma, \tag{A14}$$

where

$$\gamma = \int_{\mathbb{R}^n} \tilde{X} \exp\left(-\frac{\tilde{X}^{\mathrm{T}} P^{-1} \tilde{X}}{2}\right) \int_{\tau_j}^{\tau_{j+1}} \exp\left[-\frac{1}{2}\left(\psi - \frac{h\tilde{X}}{\Delta}\right)^{\mathrm{T}} \left(\frac{R}{\Delta^2}\right)^{-1} \left(\psi - \frac{h\tilde{X}}{\Delta}\right)\right] \mathrm{d}\psi \mathrm{d}\tilde{X}. \tag{A15}$$

Define $\beta = \psi - h\tilde{X}/\Delta$, γ can be rewritten as

$$\begin{aligned}\gamma &= \int_{\mathbb{R}^n} \tilde{X} \exp\left(-\frac{\tilde{X}^{\mathrm{T}} P^{-1} \tilde{X}}{2}\right) \int_{\tau_j - h\tilde{X}/\Delta}^{\tau_{j+1} - h\tilde{X}/\Delta} \exp\left[-\frac{\beta^{\mathrm{T}}(R/\Delta^2)^{-1}\beta}{2}\right] \mathrm{d}\beta \mathrm{d}\tilde{X} \\ &= -P \int_{\mathbb{R}^n} \frac{\partial}{\partial \tilde{X}} \left[\exp\left(-\frac{\tilde{X}^{\mathrm{T}} P^{-1} \tilde{X}}{2}\right)\right] \int_{\tau_j - h\tilde{X}/\Delta}^{\tau_{j+1} - h\tilde{X}/\Delta} \exp\left(-\frac{\beta^{\mathrm{T}}(R/\Delta^2)^{-1}\beta}{2}\right) \mathrm{d}\beta \mathrm{d}\tilde{X}.\end{aligned} \tag{A16}$$

Defining $\mathrm{I}_n \in \mathbb{R}^n$, $\mathrm{I}_n = [1, 0, \ldots, 0]^{\mathrm{T}}$, $\alpha \in \mathbb{R}^n$, $\alpha = \beta \times \mathrm{I}_n$, $l_j = (\tau_j - h\tilde{X}/\Delta) \times \mathrm{I}_n$, $l_{j+1} = (\tau_{j+1} - h\tilde{X}/\Delta) \times \mathrm{I}_n$, $A \in \mathbb{R}^{n \times n}$, $A = R/\Delta^2 \times \mathrm{I}_{n \times n}$, $\mathrm{I}_{n \times n}$ is an N-order identity matrix, β, R and Δ are scalar.

Then, (A16) will be this

$$\begin{aligned}\gamma &= -P \int_{\mathbb{R}^n} \frac{\partial}{\partial \tilde{X}} \left[\exp\left(-\frac{\tilde{X}^{\mathrm{T}} P^{-1} \tilde{X}}{2}\right)\right] \int_{l_j}^{l_{j+1}} \exp\left(-\frac{\alpha^{\mathrm{T}} A^{-1} \alpha}{2}\right) \mathrm{d}\alpha \mathrm{d}\tilde{X} \\ &= -P \int_{\mathbb{R}^n} \exp\left(-\frac{\alpha^{\mathrm{T}} A^{-1} \alpha}{2}\right) \exp\left(-\frac{\tilde{X}^{\mathrm{T}} P^{-1} \tilde{X}}{2}\right) \Bigg|_{h^+(\tau_j - \beta)\Delta}^{h^+(\tau_{j+1} - \beta)\Delta} \mathrm{d}\alpha \\ &= -P \int_{\mathbb{R}^n} \exp\left(-\frac{\sigma_1}{2}\right) \mathrm{d}\alpha + P \int_{\mathbb{R}^n} \exp\left(-\frac{\sigma_2}{2}\right) \mathrm{d}\alpha,\end{aligned} \tag{A17}$$

where h^+ is a Moore-Penrose pseudoinverse matrix [11], and

$$\sigma_1 = \alpha^{\mathrm{T}} A^{-1}\alpha + [h^+(\tau_{j+1}-\beta)\Delta]^{\mathrm{T}} P^{-1}[h^+(\tau_{j+1}-\beta)\Delta], \tag{A18}$$

$$\sigma_2 = \alpha^{\mathrm{T}} A^{-1}\alpha + [h^+(\tau_{j}-\beta)\Delta]^{\mathrm{T}} P^{-1}[h^+(\tau_{j}-\beta)\Delta]. \tag{A19}$$

Furthermore, we can obtain

$$\sigma_1 = \mathrm{b}^{\mathrm{T}}\mathrm{B}^{-1}\mathrm{b} + \tau_{j+1}^2, \tag{A20}$$

where $\mathrm{b} = \alpha - \tau_{j+1}R(hPh^{\mathrm{T}}+R)^{-1}\mathrm{I}_n$ and $\mathrm{B} = hPh^{\mathrm{T}}R(hPh^{\mathrm{T}}+R)^{-2}\mathrm{I}_{n\times n}$.

And

$$\sigma_2 = \mathrm{c}^{\mathrm{T}}\mathrm{B}^{-1}\mathrm{c} + \tau_{j}^2, \tag{A21}$$

where $\mathrm{c} = \alpha - \tau_{j}R(hPh^{\mathrm{T}}+R)^{-1}\mathrm{I}_n$. Substituting (A20) and (A21) into (A17), we have that

$$\gamma = [(2\pi)^{n/2}|\mathrm{B}|^{1/2}]P[\exp(-\tau_j^2/2) - \exp(-\tau_{j+1}^2/2)]. \tag{A22}$$

So (A14) will be

$$E[\tilde{X}|D_1^{k-1}, D_k] = (2\pi)^{-1/2}\frac{Ph^{\mathrm{T}}}{\sqrt{hPh^{\mathrm{T}}+R}} \times \frac{\exp(-\tau_j^2/2) - \exp(-\tau_{j+1}^2/2)}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{A23}$$

According to (A13)

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \sigma_k \frac{Ph^{\mathrm{T}}}{\sqrt{hPh^{\mathrm{T}}+R}}, \tag{A24}$$

where

$$\sigma_k = (2\pi)^{-1/2}\frac{\exp(-\tau_j^2/2) - \exp(-\tau_{j+1}^2/2)}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{A25}$$

According to [8], the covariance matrix $P_{k|k}$ is

$$P_{k|k} = P - \xi_k \frac{Ph^{\mathrm{T}}hP}{hPh^{\mathrm{T}}+R},$$

where

$$\eta_k = \sigma_k^2 - (2\pi)^{-1/2}\frac{\tau_j\exp(-\tau_j^2/2) - \tau_{j+1}\exp(-\tau_{j+1}^2/2)}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{A26}$$

Thus, in the case of one activated sensor (i.e. $M = 1$) at time step k, the multi-level quantized innovation KF is demonstrated as follows:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \sigma_k^1 \frac{P_{k|k-1}(h_k^1)^{\mathrm{T}}}{\sqrt{h_k^1 P_{k|k-1} h_k^{\mathrm{T}} + R_k^1}}, \tag{A27}$$

where

$$\sigma_k^1 = (2\pi)^{-1/2}\frac{\exp(-\tau_j^2/2) - \exp(-\tau_{j+1}^2/2)}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}, \qquad 1 \leqslant j \leqslant L. \tag{A28}$$

And

$$P_{k|k} = P_{k|k-1} - \eta_k^1 \frac{P_{k|k-1}(h_k^1)^{\mathrm{T}} h_k^1 P_{k|k-1}}{h_k^1 P_{k|k-1}(h_k^1)^{\mathrm{T}} + R_k^1}, \tag{A29}$$

where

$$\eta_k^1 = (\sigma_k^1)^2 - (2\pi)^{-1/2}\frac{\tau_j\exp(-\tau_j^2/2) - \tau_{j+1}\exp(-\tau_{j+1}^2/2)}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}, \qquad 1 \leqslant j \leqslant L. \tag{A30}$$

When the number of the activated sensors $M \geqslant 1$ at time step k, $D_k = \{d_k^1, \ldots, d_k^M\}$ $(M \geqslant 1)$. Since the measurements noise and measurement models for different sensors are independent, using the distributed multi-sensor data fusion [1)], the data fusion with quantized innovation vector is given by

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + \sum_{i=1}^{M}[(P_{k|k}^i)^{-1} - (P_{k|k-1}^i)^{-1}], \tag{A31}$$

where

$$P_{k|k}^i = P_{k|k-1}^i - \eta_k^i \frac{P_{k|k-1}^i (h_k^i)^{\mathrm{T}} h_k^i P_{k|k-1}^i}{h_k^i P_{k|k-1}^i (h_k^i)^{\mathrm{T}} + R_k^i}, \tag{A32}$$

1) Palattella M, Accettura N, Vilajosana X, et al. Standardized protocol stack for the internet of (important) things. IEEE Commun Surv Tut, 2013, 15: 1389–1406

$$\eta_k^i = \frac{1}{2\pi} \times \left[\frac{(\mathrm{e}^{-\tau_j^2/2} - \mathrm{e}^{-\tau_{j+1}^2/2})}{\Phi(\tau_{j+1}) - \Phi(\tau_j)} \right]^2 - \frac{1}{\sqrt{2\pi}} \times \frac{\tau_j \mathrm{e}^{-\tau_j^2/2} - \tau_{j+1}\mathrm{e}^{-\tau_{j+1}^2/2}}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{A33}$$

And

$$P_{k|k}^{-1}\hat{X}_{k|k} = P_{k|k-1}^{-1}\hat{X}_{k|k-1} + \sum_{i=1}^{M} [(P_{k|k}^i)^{-1}\hat{X}_{k|k}^i - (P_{k|k-1}^i)^{-1}\hat{X}_{k|k-1}^i], \tag{A34}$$

where

$$\hat{X}_{k|k}^i = \hat{X}_{k|k-1}^i + \sigma_k^i \frac{P_{k|k-1}^i (h_k^i)^{\mathrm{T}}}{\sqrt{h_k^i P_{k|k-1}^i (h_k^i)^{\mathrm{T}}}}, \tag{A35}$$

$$\sigma_k^i = \frac{1}{\sqrt{2\pi}} \times \frac{(\mathrm{e}^{-\tau_j^2/2} - \mathrm{e}^{-\tau_{j+1}^2/2})}{\Phi(\tau_{j+1}) - \Phi(\tau_j)}. \tag{A36}$$