Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Freshness constraints of an age of information based event-triggered Kalman consensus filter algorithm over a wireless sensor network*

Rui WANG[1], Yahui LI[1], Hui SUN[†‡1], Youmin ZHANG[2]

[1]*College of Information Engineering and Automation, Civil Aviation University of China, Tianjin 300300, China*

[2]*Department of Mechanical, Industrial and Aerospace Engineering, Concordia University, Montreal, Quebec H3G 1M8, Canada*

[†]E-mail: h-sun@cauc.edu.cn

Received Apr. 30, 2020; Revision accepted Aug. 30, 2020; Crosschecked Dec. 11, 2020

Abstract: This paper presents the design of a new event-triggered Kalman consensus filter (ET-KCF) algorithm for use over a wireless sensor network (WSN). This algorithm is based on information freshness, which is calculated as the age of information (AoI) of the sampled data. The proposed algorithm integrates the traditional event-triggered mechanism, information freshness calculation method, and Kalman consensus filter (KCF) algorithm to estimate the concentrations of pollutants in the aircraft more efficiently. The proposed method also considers the influence of data packet loss and the aircraft's loss of communication path over the WSN, and presents an AoI-freshness-based threshold selection method for the ET-KCF algorithm, which compares the packet AoI to the minimum average AoI of the system. This method can obviously reduce the energy consumption because the transmission of expired information is reduced. Finally, the convergence of the algorithm is proved using the Lyapunov stability theory and matrix theory. Simulation results show that this algorithm has better fault tolerance compared to the existing KCF and lower power consumption than other ET-KCFs.

Key words: Distributed Kalman consensus filter (KCF); Event-triggered mechanism; Age of information (AoI); Stability analysis; Energy optimization

https://doi.org/10.1631/FITEE.2000206 **CLC number:** TP391; TP393

1 Introduction

As the civil aviation industry is rapidly developing, safety and traveling comfort have become a focus for aircraft environment design, and a hot research topic is the safe and economic use of commercial aircraft. Once the narrow and closed cabin or cargo hold encounters mechanical failures, pipe rupture, or seal failures, the cabin environment may become contaminated, resulting in passenger and crew dizziness, headaches, and other symptoms, even leading to serious neurological diseases (Li F et al., 2016; Farag, 2018), which will seriously threaten flight safety. Therefore, it is quite important to detect urgent cabin environment issues as quickly as possible to ensure the life and health of the passengers and crew. With the development of wireless sensor technology, the wireless sensor network (WSN) has demonstrated great value and broad commercial prospects in environmental, military, health, agriculture, space exploration, and other fields (Saha and Majumdar, 2017; Hudhajanto et al., 2018). When monitoring pollutant concentrations over a WSN, it is often difficult for the actual operating system to detect pollutant concentrations accurately due to the environmental interference such as dusts and other communication factors such as data packet loss and

‡ Corresponding author

* Project supported by the Civil Aviation Science and Technology Project (No. MHRD20150220), the Fundamental Research Funds for the Central Universities, China (No. 3122017003), and the Natural Sciences and Engineering Research Council of Canada

ORCID: Rui WANG, https://orcid.org/0000-0003-4452-6795; Hui SUN, https://orcid.org/0000-0001-7199-9059

network bandwidth limitation. Therefore, it is necessary to develop an effective concentration estimation algorithm to quickly and accurately monitor sudden environmental problems in the aircraft to reduce the false alarm rate.

Distributed networks have been applied in wide areas such as WSN applications and fault-tolerant cooperative controls for unmanned aerial vehicle systems (Yu et al., 2019a, 2019b, 2020). Specifically, distributed estimation algorithms in WSNs are focused on network systems with interference. The distributed Kalman consensus filter (KCF) algorithm is widely used because of its fast convergence, high fusion precision, and strong robustness. Olfati-Saber et al. studied the problem of distributed consensus estimation systematically (Olfati-Saber and Shamma, 2005; Olfati-Saber, 2007, 2009). In Olfati-Saber and Shamma (2005), a distributed filter based on average consensus was introduced; it plays a crucial role in solving the data fusion problem of distributed Kalman filtering in sensor networks. Convergence analysis was provided. Olfati-Saber (2007) proposed three new distributed Kalman filter (DKF) algorithms for sensor networks, and the high estimation accuracy of these algorithms was verified by simulation results. To effectively track and estimate dynamic networks, a DKF algorithm for sensor networks with a variable topological structure was presented in Olfati-Saber (2009). In practical applications, packet loss and path loss are inevitable in the process of data transmission due to the network bandwidth limitation, node energy limitation, adverse environments, and other influencing factors. Furthermore, packet loss and path loss affect mainly the accuracy of the filter algorithm and the stability of the system for practical applications. Liu YG et al. (2013) designed a Kalman filter algorithm that can tolerate continuous packet losses, and the simulation results proved that the algorithm has good robustness. Fan et al. (2017) obtained a state prediction consensus mechanism for linear discrete systems with packet loss, and provided a sufficient condition for the convergence of the error estimation system. Sun SL et al. (2016) designed a predictor to compensate for packet or path loss, and proposed an optimal linear estimation based on the packet arrival rate in the sense of linear minimum variance. Liu XD et al. (2017), Paul et al. (2018), and Shi et al. (2018) also discussed the stochastic stability of an extended Kalman filter under intermittent observations.

When considering the real communication limitation, event-triggered mechanisms in KCF are introduced and studied by researchers. Amini et al. (2018) proposed a distributed guaranteed-performance event-triggered average consensus (GP-ETAC) algorithm for WSNs (i.e., the agent selectively restricts or transmits state updates to the locally adjacent region), and the algorithm stability theorem was given using the Lyapunov method. To reduce data transmission in the sensor network, Li WL et al. (2016) designed a KCF with an event-triggered communication protocol, which allows each sensor to deliver its local estimates to its neighbors only when the difference between the latest transmitted estimate and the current estimate exceeds the tolerable threshold, and verified the effectiveness of the filter algorithm. Wang et al. (2019) analyzed the influence of the triggering threshold on the energy consumption of the algorithm and proposed that the triggering threshold can be adjusted according to practical needs to control the power consumption and the accuracy of the algorithm. Zhang and Zhang (2018) focused on the Kalman filter algorithm based on the activation strategy of the fully distributed event-triggered sensor, and proposed a node-detection activation strategy based on the event-triggered mechanism, to save energy on the premise that each node has a small tracking error. It is known that timely delivery of sampling information over the network is quite crucial for the aircraft cabin pollutant monitoring system. However, in reality, because of system delay or transmission policy, some information arrives so late that it no longer correctly estimates the real information at that moment—the information is not "fresh." Sometimes, however, even if the information is delayed, it is still effective and can be used. The concept of age of information (AoI) was first proposed by Kaul S et al. (2011) to quantify the freshness of system state information. Because the average AoI for all packets in the network affects the real-time performance of the system, and because of the freshness limitation, the transmission of expired information is reduced. Considerable studies have concentrated on measuring information freshness. Among them, Kaul SK et al. (2012) employed a time-averaged age metric (i.e., AoI) to characterize the performance of such state updating systems and evaluate the freshness of information carried by packets. The relationship between the average AoI of the system and the real-time performance of the

system has been studied. Talak et al. (2018) presented the minimum value calculation of the network average AoI under double-interference constraints for time- varying channels and pointed out that the network average AoI is most efficient in providing information updates when it is the smallest. In an embedded system with energy constraints, Zhou et al. (2020) obtained the minimum average AoI of the whole network by determining whether to transmit, discard, or save the real-time data update sequences. With regard to designing a sampler, Sun Y and Cyr (2019) developed a method for extracting samples from data sources to remotely improve the freshness of received data samples; i.e., the AoI of sampled data is small. Tang et al. (2019) measured the freshness of data from the central controller perspective and developed a scheduling algorithm that decouples multi-user scheduling into a single-user constrained Markov decision process to realize minimum AoI performance. In Wang et al. (2019), each sensor transmitted its local estimates to its neighbors only if the difference between the most recent transmitted estimate and the current estimate exceeded a tolerable threshold. This approach can reduce the network energy consumption, and the accuracy and energy consumption of the system are related to this threshold. In addition, in Song et al. (2018), each sensor node transmitted its latest measurement update to the corresponding remote estimator based on its measurement value as an event-triggered condition. This method can significantly reduce the average communication rate. Although these event-triggered mechanisms can reduce the energy consumption to a certain extent, no exact threshold has been developed in the existing research to minimize the energy consumption while ensuring system accuracy. In this study, an event-triggered mechanism is presented based on judging the freshness of packet information, and an algorithm is proposed to determine the accurate triggering threshold. This algorithm requires additional calculation of the AoI of each packet, which reduces the convergence speed of the algorithm to some extent, but has no influence on the stability of the algorithm.

In this study we develop a new event-triggered KCF (ET-KCF) algorithm that relies on information freshness constraints to effectively estimate pollutant concentrations in an aircraft, and the proposed freshness constraints of the AoI-based ET-KCF algorithm (FCET-KCF) reduce the system energy consumption. Network energy consumption, system usage extension, and the correctness of the detection and stability of the system in the case of packet loss and path loss are considered to further reduce the false alarms.

2 Preliminaries: integrated monitoring network and Kalman consensus filter

2.1 Integrated sensor structures

In the authors' previous work (Wang et al., 2017), three integrated sensor structures were introduced and compared, and it was concluded that feedback correction can reflect the real dynamic process of system errors more accurately. However, when considering aviation system safety requirements, the output correction method is the best choice. Therefore, in this study we adopt the output correction structure for each integrated sensor node; the corresponding scheme is shown in Fig. 1. It can be seen that the measurement result of the primary sensor is z_{pi}, and the measurement result of the secondary sensor is z_{si}. The difference between two measurements, z_i, is fused with the estimated values of the neighbor nodes by KCF. The optimal state estimate of the sensor error, $\hat{\boldsymbol{x}}_i$, is obtained. Finally, the error estimate is used to correct the measured value of the primary sensor to obtain an optimal estimate of the concentration, z_{oi}.

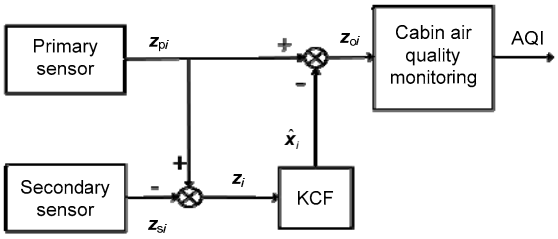

Fig. 1 Schematic of the output correction structure for an integrated sensor node (Wang et al., 2017)

2.2 Kalman consensus filter algorithm

The topology diagram of the WSN is defined as $G=(V, E, A)$, where $V=\{v_1, v_2, \ldots, v_n\}$ is the set of sensor nodes and $E=V\times V$ is the set of edges between nodes. N_i represents the set of neighbor nodes for node i, i.e., $N_i=\{v_j|v_j\in V, (v_i, v_j)\in E\}$. Define the adjacency matrix $\boldsymbol{A}=[a_{ij}]$ to indicate whether any two

sensors are connected. If node i can receive information from node j, a_{ij}=1 ($i\neq j$); otherwise, a_{ij}=0.

We consider a WSN G with m cluster-head nodes and n member nodes in the cluster, where the cluster-head nodes in the network use the output correction structure mentioned above. The state model of the target system and the observation model of the sensor are shown as

$$\begin{cases} \boldsymbol{x}_{i,k+1} = \boldsymbol{A}\boldsymbol{x}_{i,k} + \boldsymbol{B}\boldsymbol{w}_{i,k}, \\ \boldsymbol{z}_{i,k+1} = \boldsymbol{H}_{i,k}\boldsymbol{x}_{i,k} + \boldsymbol{F}_{i,k}\boldsymbol{v}_{i,k}, \end{cases} \tag{1}$$

where $\boldsymbol{x}_{i,k}$ and $\boldsymbol{w}_{i,k}$ represent the error state vector and process noise vector of the integrated error system, respectively. $\boldsymbol{z}_{i,k}$ is the observation vector and $\boldsymbol{v}_{i,k}$ is the observed noise vector of the primary sensor. $\boldsymbol{A}$ and $\boldsymbol{B}$ are the system matrices with appropriate dimensions. $\boldsymbol{H}_{i,k}$ and $\boldsymbol{F}_{i,k}$ are the measurement matrix and the fault matrix respectively, both assumed to be invertible. It is assumed that $\boldsymbol{w}_{i,k}$ and $\boldsymbol{v}_{i,k}$ are independent zero-mean Gaussian white noises, and their covariance matrices satisfy

$$E[\boldsymbol{w}_k \boldsymbol{w}_l^{\mathrm{T}}] = \boldsymbol{Q}_k \delta_{kl},\ E[\boldsymbol{v}_k^i (\boldsymbol{v}_l^j)^{\mathrm{T}}] = \boldsymbol{R}_k^{ij} \delta_{kl},$$

where δ_{kl} is an impulse response function. When k=l, δ_{kl} takes a value of 1, otherwise 0. $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k^{ij}$ are the corresponding noise covariance matrices. When i=j, $\boldsymbol{R}_k^{ij}$ is denoted as $\boldsymbol{R}_k^i$.

3 Information freshness constraints of the AoI-based event-triggered Kalman consensus filter algorithm

In this section the average AoI is minimized to optimize the real-time performance of systems and regarded as the threshold for the FCET-KCF algorithm, to reduce the measurement error and system energy consumption. The assumptions for the algorithm over the WSN are as follows:

Assumption 1 For each node in the network, all packet sizes obtained by sampling are the same.

Assumption 2 In the absence of the external interference and the failure of sensors deployed on each node in the WSN, each packet is of the same size. The average AoI of all packets during the sample interval $(0, T_1)$ is the same as that during the left interval (T_1, T) for a whole period $(0, T)$, where $0<T_1<T$.

3.1 Measurement method for information freshness

In a multi-cluster-head WSN, once all the cluster heads confirm the head identity, they can receive data transmitted from the member nodes in a cluster and information connected to adjacent cluster heads. When the sampled information is transmitted over the network, each sensor node packs the sampling information as one packet and transmits it to the cluster-head node. The information in the packet arriving at the cluster head contains the current environment information and information freshness. The average AoI for all packets in the network affects the real-time performance of the measurement system; the smaller the average AoI is, the stronger the real-time performance of the system will be (Kaul SK et al., 2012).

Consider a cluster containing n source nodes and one destination node for a WSN, where n is the number of member nodes in a cluster, and the cluster-head node corresponds to the destination node (Fig. 2). The source node observes the cabin environment parameters and extracts samples. The communication link transmits the samples collected from the source nodes to the destination node for Kalman filter consensus calculation. At the transmit terminal of the source node, there is a buffer zone which stores the samples that contain measurements and timestamp $U(t)$ as a packet. Packets are sent along the source-destination communication link, and each packet arriving at the destination provides a status update for the cluster-head node. The packet transmission model in the cluster with energy harvesting satisfies the $M/M/1$ queue mode.

Sensor (i.e., source) node in a cluster: Each sensor node with energy harvesting transmits the environment information to the cluster head, and $\{u_i, U_i\}$ represents the packet formed on the buffer in cluster i. Because there are many different sensor nodes in the cluster and the position of each node changes continuously, it can be assumed that the process of sending packets from each source node to the destination node follows a Poisson distribution with parameter λ (λ represents the packet generation rate, $1/\lambda$ is the time interval for packet generation, and λ_i^n is the packet generation rate on source node n for cluster i). Because the time-varying availability of

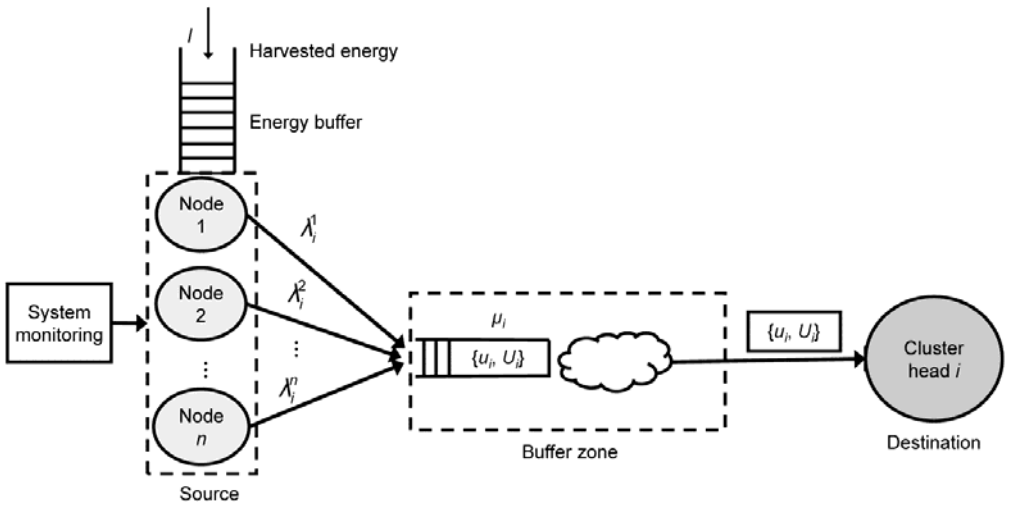

Fig. 2 Packet transfer model with energy harvesting in a cluster

energy harvesting at the transmitter limits the sampling rate at the source, it can be assumed that the process of energy harvesting energy supply to the source node obeys the Poisson distribution with parameter l (Zhou et al., 2020). Here l represents the energy harvested rate, also called the energy supply rate of the source node, T represents the total time of energy supply, and N is the total number of node samples. As T increases, $l=N/T$.

Buffer zone: This zone stores sampling information from the source nodes, including state measurements and their associated timestamps. Because the transmission time between the source node and the destination node depends mainly on the size of the packet and the channel state, it can be assumed that the packet service process from the source node to the destination node follows an exponential distribution with an exponent of μ, and $1/\mu$ is the average service time (Farazi et al., 2018).

Cluster-head (i.e., destination) node: The cluster-head node receives sample data from the source nodes and records the arrival time of the packet.

It is assumed that the AoI of the packet for the transmission model mentioned above is represented by $1/\lambda_i^q+1/\mu_i$, where $i=1, 2, \ldots, m$ and $q=1, 2, \ldots, n$.

Definition 1 (Zhou et al., 2020) The AoI is the time elapsed since the generation of the last sample packet. At sampling time t, if the latest packet generation has a timestamp $U(t)$, the AoI is $\Delta(t)=t-U(t)$.

Based on Definition 1, the average AoI of the real-time data during interval $(0, T)$ is obtained as

$$\Delta_T=\frac{1}{T}\int_0^T \Delta(t)\mathrm{d}t. \tag{2}$$

Fig. 3 shows the evolution of the AoI of each packet in a cluster. Without loss of generality, assume that at $t=0$, the queue is empty and the AoI of each source node is $\Delta(0)=0$. It is assumed that in a time order, the packets formed by any source node in a cluster after the w^{th} sampling are generated at moment t_w and received by the destination node at moment T_w. In time interval $(0, T)$, sampling number w satisfies $w\in\{1, 2, \ldots, N(T)\}$, where $N(T)=\max\{w|T_w\le T\}$. In addition, assume that $t_w=0\ \forall w\in\{1, 2, \ldots, N(T)\}$. Between T_{w-1} and T_w, if the destination node does not receive the packet, the AoI increases linearly with time. Otherwise, if it receives the packet, the AoI is reset to the time period that the packet has experienced in the communication link represented by the area during this period, marked as Q_w. Therefore, the evolution of the AoI in a cluster with continuous samplings is shown as a sawtooth in Fig. 3. Set X_w as the packet consuming time in the link, where $X_w=T_w-t_w$. Set Y_w as the time interval when two consecutive packets arrive at the destination node, where $Y_w=T_w-T_{w-1}$. Q_w represents the AoI of the system at time interval Y_w.

It is known that the calculation of the integral of a function is determined by the area enclosed by the function in an interval. For example, for calculation of $\int_{T_2}^{T_3}\Delta(t)\mathrm{d}t$, the area enclosed below the line $\Delta(t)$ in

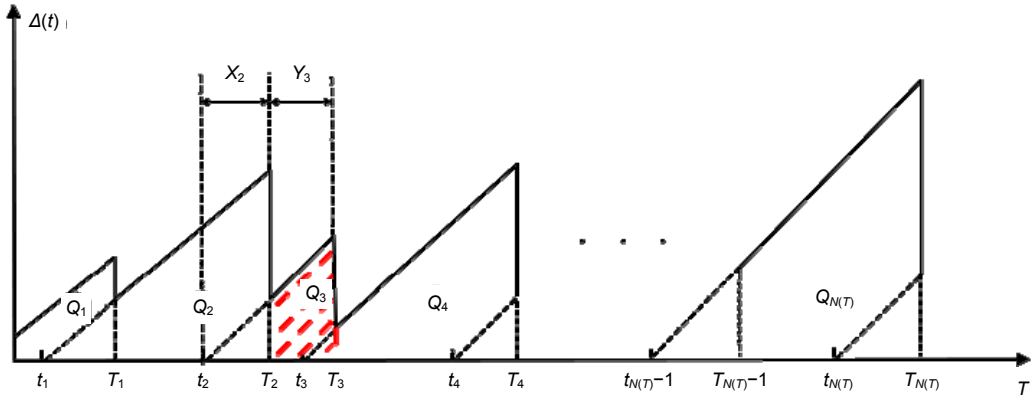

Fig. 3 The evolution of a packet for AoI

interval $[T_2, T_3]$ is calculated, that is, the area of a trapezoid Q_3. The total AoI is then calculated as the sum of the areas of these disjoint geometric parts, as shown in Fig. 3. For the convenience of interpretation, time interval $(0, T)$ is selected. During this period, the area of the disjoint geometric region can be considered as the sum of the areas Q_w ($1 \le w \le N(T)$). Let S be the sum of the area of the regions represented by

$$S = \sum_{w=1}^{N(T)} Q_w. \tag{3}$$

The region Q_w ($w \ge 2$) is a trapezoid, and its area can be calculated by the following formula:

$$Q_w = \frac{1}{2}(X_{w-1} + Y_w)^2 - \frac{1}{2}X_{w-1}^2 = \frac{1}{2}Y_w^2 + X_{w-1}Y_w. \tag{4}$$

Because the average AoI Δ_T is defined as

$$\Delta_T = \frac{S}{T} = \frac{1}{T}\left(Q_1 + \sum_{w=2}^{N(T)} Q_w\right), \tag{5}$$

as T increases, $Q_1/T \triangleright\!\triangleleft 0$, $N(T)/T \triangleright\!\triangleleft l$. Then

$$\Delta_T = \lim_{T\to+\infty}\left(\frac{Q_1}{T} + \frac{N(T)}{T}\frac{1}{N(T)}\sum_{w=2}^{N(T)} Q_w\right) = lE[Q_w]. \tag{6}$$

Because X_{w-1} and Y_w are independent, the following formula can be derived:

$$E[Q_w] = E[XY] + \frac{1}{2}E[Y^2]. \tag{7}$$

Therefore, the average AoI of packets generated by all nodes in a cluster is

$$\Delta = l\left(E[XY] + \frac{1}{2}E[Y^2]\right). \tag{8}$$

3.2 Threshold selection for ET-KCF by minimizing the average AoI

In this study we specifically present a threshold selection method for the ET-KCF algorithm. We design an effective method for calculating the minimum value of the average AoI. To minimize the average AoI (Δ) in Eq. (8), it is necessary to obtain X and Y first. Denote the time elapsed from packet generation to packet arrival at the destination node for the packet generated by the w^{th} sampling as X_w. We have

$$X_w = S_w + W_w, \tag{9}$$

where S_w represents the service time of the packet and W_w is the time of the packet waiting in the system.

The interval between two consecutive packets arriving at the destination node is Y_w and $Y_w = T_w - T_{w-1}$.

To describe the calculation of the minimum average AoI, Theorem 1 is presented as follows:

Theorem 1 Suppose that the packet generated by the w^{th} sampling of the source node is sent to the destination node immediately after it is generated at any time. Then the delay of the packet in the system is 0, i.e., $X_w = S_w$.

Proof Suppose that there is a delay when the packet travels to the destination node. The time traveling to the destination node, T_w, has been changed into T'_w ($T'_w > T_w$) (Fig. 4). Compared to Q_w, Q'_w contains two

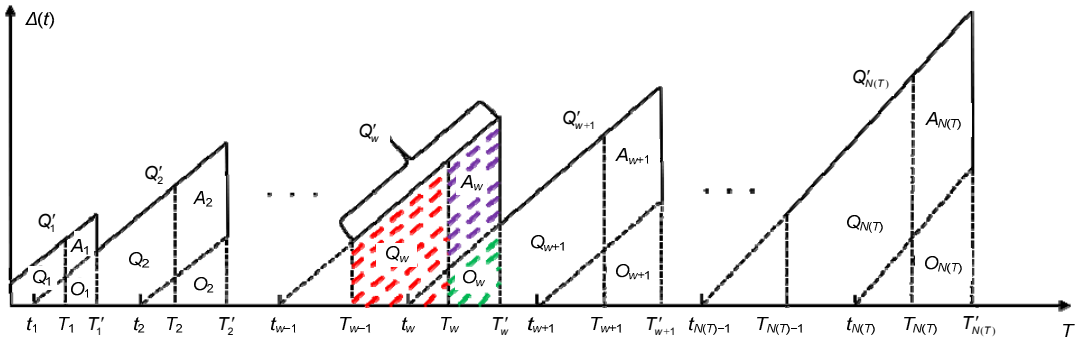

Fig. 4 The evolution process of packet AoI with transmission delay

additional portions, O_w and A_w. O_w is the portion that overlaps with Q_{w+1}, and A_w is the other additional portion that is added due to the packet delay. Therefore, Q'_w increases by A_w compared to Q_w, and Q'_{w+1} reduces by O_w compared to Q_{w+1}. The variation of the average AoI, Δ', in the whole cluster can be calculated by the following equation:

$$\begin{aligned}\Delta' &= \frac{1}{T}[(Q'_1-Q_1)+(Q'_2-Q_2)+\cdots\\ &\quad +(Q'_w-Q_w)+\cdots+(Q'_{N(T)}-Q_{N(T)})]\\ &= \frac{1}{T}[(A_1+O_1)+(A_2+O_2-O_1)+(A_3+O_3-O_2)+\cdots\\ &\quad +(A_w+O_w-O_{w-1})+\cdots+(A_{N(T)}+O_{N(T)}-O_{N(T)-1})]\\ &= \frac{1}{T}\left(\sum_{w=1}^{N(T)}A_w+O_{N(T)}\right).\end{aligned}$$

If delay exists so that $A_w>0$ and $O_{N(T)}>0$ as shown in Fig. 4, Δ' will continue to be accumulated as the number of samplings increases. Thus, the average AoI increases continuously. This is in contradiction to the goal of minimizing the average AoI. Therefore, the packets generated through sampling are sent immediately after generation, i.e., $X_w=S_w$.

According to Theorem 1, Eq. (8) is changed into

$$\Delta = l\left(\frac{1}{\mu}E[Y]+\frac{1}{2}E[Y^2]\right). \tag{10}$$

To minimize the average AoI, the sum of Y_w is equal to the total energy supply time T (Zhou et al., 2020). Because $T=N/l$, $E[Y]=1/l$. Thus, the average AoI can be obtained by minimizing $\sum_{w=2}^{N(T)}Y_w^2$ according to the Cauchy inequality (Steele, 2004).

According to the Cauchy inequality, when Y_w is close to the mean value of $\sum_{w=2}^{N(T)}Y_w^2$ such that $E[Y]\approx 1/l$, the minimum average AoI value Δ^* is

$$\Delta^*=l\left(\frac{1}{\mu}E[Y]+\frac{1}{2}E[Y^2]\right)\approx\frac{1}{\mu}+\frac{1}{2l}, \tag{11}$$

where $E[Y]$ is close to $1/l$, and $E[Y^2]$ approaches $1/l^2$.

The parameters μ and l of different clusters in the system are not the same, so the minimum value of the average AoI in different clusters can be represented by $\Delta^*=\frac{1}{\mu_i}+\frac{1}{2l_i}$, $i=1,2,...,m$.

3.3 Event-triggered Kalman consensus filter algorithm based on information freshness constraints of the AoI

In this subsection, the FCET-KCF algorithm is presented. The event-triggered function is given first:

$$\xi_{i,k}=\begin{cases}1, & 0\le\frac{1}{\lambda_i^q}+\frac{1}{\mu_i}\le\Delta^*,\\ 0, & \frac{1}{\lambda_i^q}+\frac{1}{\mu_i}>\Delta^*,\end{cases} \tag{12}$$

where $i=1, 2, \ldots, m$ and $q=1, 2, \ldots, n$, Δ^* is the minimum value of the average AoI in cluster i, and $1/\lambda_i^q+1/\mu_i$ represents the AoI of the packet generated at a sampling on source node q.

Suppose that within the same cluster, the average service time of each packet is the same, and equal to $1/\mu$. The difference value of AoI between each packet and the event-triggered threshold can be expressed as $(1/\lambda_i^q+1/\mu_i)-\Delta^*$.

According to Eq. (11), the AoI difference is $1/\lambda_i^q - 1/(2l)$. Because $1/\lambda_i^q$ represents the time interval of the packet generated in node q within cluster i, l represents the energy harvested rate, $l=N/T$, in a sampling time T. As the value of $1/\lambda_i^q - 1/(2l)$ can be positive, negative, or zero during communication, the proposed event-triggered method can avoid the Zeno behavior.

The event-triggered principle: When a member node in a cluster sends a packet to the cluster head, the cluster-head node determines the AoI value of this packet and the minimum average AoI (Δ^*). If it is smaller than Δ^*, then $\xi_{i,k}$=1. This means that the measured environmental information contained in the packet is fresh and can reflect the real-time pollutant concentrations. At this time, the event is triggered, and the cluster head transmits the updated estimate to other connected neighbor nodes. Otherwise, it indicates that the packet information is not fresh, the cluster head does not transmit the update, and the consensus algorithm continues to use the previous estimate for calculation.

In each cluster of the network, the member nodes of the cluster periodically collect target information and send the information $m_i=\{u_i, U_i\}$ to the cluster-head node. Next, each cluster-head node performs KCF to estimate the target state depending on the information of other connected neighbor cluster-head nodes. Here, $L=\{\boldsymbol{L}_1, \boldsymbol{L}_2, \ldots, \boldsymbol{L}_m\}$ represents the set of network cluster heads, and m is the total number of cluster-head nodes.

In addition, in this study we consider the packet drop phenomenon, which often occurs due to random factors such as network delays, network blockage, sensor failures, and path loss of wireless signal transmission caused by the dimensions of different materials. The existence of packet loss reduces the accuracy of algorithm estimation and even leads to divergence. Packet loss can be divided into two types: loss of measurement data for each node in the network and loss of communication information between nodes in the network. The Bernoulli process is used here for packet receiving. Binary variables $\alpha_{i,k}$ and $\beta_{i,k}$ are defined to describe the packet arrival process on cluster-head node i at time k. $\alpha_{i,k}$=1 indicates the observed packet successfully received; $\beta_{i,k}$=1 indicates that the communication packet is received successfully. Furthermore, $P\{\alpha_{i,k}=1\}=\varpi_1$, $P\{\beta_{i,k}=1\}=\varpi_2$. When considering wireless signal transmission path loss, assume that θ_{ij} is the path loss rate between cluster-head nodes i and j. From the practical point of view, we will design an ET-KCF algorithm based on the information freshness constraints of the AoI under packet and path loss scenarios.

The procedures for the FCET-KCF algorithm (Algorithm 1) are as follows:

Step 1: initialization, $P_i=P_0$, $\hat{\boldsymbol{x}}_i=\boldsymbol{x}_0$.

Step 2: All cluster member sensor nodes measure the target state information $\boldsymbol{z}_{i,k}$.

Step 3: Each cluster member node sends message $m_i=\{u_i, U_i\}$ and the corresponding AoI to the cluster-head node.

Step 4: Each cluster-head node fuses the information sent by the member nodes and the adjacent nodes:

$$\begin{cases}\boldsymbol{u}_{i,k}=\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{R}_{i,k}^{-1}\alpha_{i,k}\boldsymbol{z}_{i,k},\\ \quad \forall j\in N_i,\ \boldsymbol{y}_{i,k}=\sum_{j\in N_i}\beta_{i,k}\boldsymbol{u}_{j,k}+\boldsymbol{u}_{j,k},\\ \boldsymbol{U}_{i,k}=\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{R}_{i,k}^{-1}\alpha_{i,k}\boldsymbol{H}_{i,k},\\ \quad \forall j\in N_i,\ \boldsymbol{S}_{i,k}=\sum_{j\in N_i}\beta_{i,k}\boldsymbol{U}_{j,k}+\boldsymbol{U}_{j,k}.\end{cases}$$

Step 5: Check the packet information freshness of $\hat{\boldsymbol{x}}_{i,k}$:

$$\hat{\boldsymbol{x}}_{i,k}=\begin{cases}\hat{\boldsymbol{x}}_{i,k}, & \xi_{i,k}=1,\\ \hat{\boldsymbol{x}}_{i,k-1}, & \xi_{i,k}=0.\end{cases}$$

The latest estimate is transmitted when $\xi_{i,k}$=1; otherwise, for consensus calculation the previous estimate will be used to reduce the network energy consumption.

Step 6: According to the classical KCF algorithm shown in Olfati-Saber (2007), the estimated values based on FCET-KCF are (the detailed introduction to the consensus strategy is given in Appendix A)

$$\begin{aligned}\hat{\boldsymbol{x}}_{i,k+1}=&\boldsymbol{A}\hat{\boldsymbol{x}}_{i,k}+\varpi_1\boldsymbol{A}\boldsymbol{K}_k^i\xi_k^i(\boldsymbol{Z}_{i,k}-\boldsymbol{H}_{i,k}\hat{\boldsymbol{x}}_{i,k})\\&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{j\in N_i}[(1-\theta_{ij})\hat{\boldsymbol{x}}_{j,k}^o-\hat{\boldsymbol{x}}_{i,k})],\end{aligned}\tag{13}$$

where $\hat{\boldsymbol{x}}_{j,k}^o$ is the broadcast estimate at time k.

Step 7: Each cluster-head node updates matrices $\boldsymbol{M}_{i,k}$ and $\boldsymbol{P}_{i,k}$:

$$\boldsymbol{M}_{i,k}=(\boldsymbol{P}_{i,k}^{-1}+\boldsymbol{S}_{i,k})^{-1},$$
$$\boldsymbol{P}_{i,k+1}=\boldsymbol{A}\boldsymbol{M}_{i,k}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_{i,k}\boldsymbol{B}^{\mathrm{T}}.$$

Step 8: Return to step 2.

The aim of the optimal event-triggered KCF is to minimize the following mean squared error: $E\{(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{i,k})(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{i,k})^{\mathrm{T}}\}$.

For further analysis, we define the estimation error of node i at each time instant k as

$$\begin{cases}\boldsymbol{e}_{i,k}=\hat{\boldsymbol{x}}_{i,k}-\boldsymbol{x}_k,\\ \boldsymbol{e}_{i,k}^{o}=\hat{\boldsymbol{x}}_{i,k}^{o}-\boldsymbol{x}_k,\\ \tilde{\boldsymbol{e}}_{i,k}=\boldsymbol{e}_{i,k}^{o}-\boldsymbol{e}_{i,k},\end{cases}\tag{14}$$

and the corresponding estimation error covariance matrices are

$$\boldsymbol{P}_{ij,k}=E\left\{\boldsymbol{e}_{i,k}\boldsymbol{e}_{j,k}^{\mathrm{T}}\right\},\ \boldsymbol{P}_{ij,k}^{o}=E\left\{\boldsymbol{e}_{i,k}^{o}\boldsymbol{e}_{j,k}^{\mathrm{T}}\right\},$$
$$\tilde{\boldsymbol{P}}_{ij,k}=E\left\{\tilde{\boldsymbol{e}}_{i,k}\tilde{\boldsymbol{e}}_{j,k}^{\mathrm{T}}\right\},\ \hat{\boldsymbol{P}}_{ij,k}=E\left\{\hat{\boldsymbol{x}}_{i,k}\hat{\boldsymbol{x}}_{j,k}^{\mathrm{T}}\right\},$$
$$\hat{\boldsymbol{P}}_{io,k}'=E\left\{\hat{\boldsymbol{x}}_{i,k}\boldsymbol{x}_k^{\mathrm{T}}\right\},\ \hat{\boldsymbol{P}}_{io,k}^{o}=E\left\{\hat{\boldsymbol{x}}_{i,k}^{o}\boldsymbol{x}_k^{\mathrm{T}}\right\},$$
$$\hat{\boldsymbol{P}}_{ij,k}^{o}=E\left\{\hat{\boldsymbol{x}}_{i,k}^{o}\hat{\boldsymbol{x}}_{j,k}^{\mathrm{T}}\right\}.$$

The optimal gain matrix, $\boldsymbol{K}_{i,k}$, can be obtained by solving

$$\frac{\partial\mathrm{tr}(\boldsymbol{P}_{i,k+1})}{\partial\boldsymbol{K}_{i,k}}=0,$$

where the estimation error of node i is

$$\mathrm{tr}(\boldsymbol{P}_{i,k+1})=E\left\{(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{i,k})(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{i,k})^{\mathrm{T}}\right\}.$$

Then the optimal gain can be derived and the detailed proofs are presented in Appendix B.

3.4 Stability analysis of the FCET-KCF algorithm

Without considering noise, the estimated error of cluster-head node i in the network is defined as $\boldsymbol{\eta}_{i,k}=\hat{\boldsymbol{x}}_{i,k}-\boldsymbol{x}_k$, $i\in\{1,\ 2,\ \cdots,\ m\}$. Combined with step 6 in Algorithm 1, the dynamic error can be obtained as

$$\boldsymbol{\eta}_{i,k+1}=\boldsymbol{A}\boldsymbol{\eta}_{i,k}-\varpi_1\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k}\boldsymbol{\eta}_{i,k}+\varpi_2\boldsymbol{C}_{i,k}\xi_{i,k}(1-\theta_{ij})\boldsymbol{u}_{i,k},\tag{15}$$

where $\boldsymbol{u}_{i,k}=\boldsymbol{A}\sum_{j\in N_i}\left(\boldsymbol{\eta}_{j,k}-\boldsymbol{\eta}_{i,k}/(1-\theta_{ij})\right)$. Select the Lyapunov function as

$$V(\boldsymbol{\eta}_k)=\sum_{i=1}^{m}\boldsymbol{\eta}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k}^{-1}\boldsymbol{\eta}_{i,k}.\tag{16}$$

From the forward difference formula $\Delta V(\boldsymbol{\eta}_k)=V(\boldsymbol{\eta}_{k+1})-V(\boldsymbol{\eta}_k)$ of the Lyapunov function for discrete systems, one can obtain

$$\begin{aligned}&\Delta V(\boldsymbol{\eta}_k)\\&=\textstyle\sum_i\boldsymbol{\eta}_{i,k+1}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{\eta}_{i,k+1}-\boldsymbol{\eta}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k}^{-1}\boldsymbol{\eta}_{i,k}\\&=\textstyle\sum_i\boldsymbol{\eta}_{i,k}^{\mathrm{T}}[(\boldsymbol{A}-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\\&\quad\cdot(A-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})-\boldsymbol{M}_{i,k}^{-1}]\boldsymbol{\eta}_{i,k}\\&\quad+\textstyle\sum_i\boldsymbol{\eta}_{i,k}^{\mathrm{T}}[(\boldsymbol{A}-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\\&\quad\cdot\varpi_2\xi_{i,k}(1-\theta_{ij})\boldsymbol{C}_{i,k}]\boldsymbol{u}_{i,k}\\&\quad+\textstyle\sum_i\boldsymbol{\eta}_{i,k}^{\mathrm{T}}[\varpi_2\xi_{i,k}(1-\theta_{ij})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\\&\quad\cdot(\boldsymbol{A}-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})]\boldsymbol{u}_{i,k}\\&\quad+\textstyle\sum_i\boldsymbol{u}_{i,k}^{\mathrm{T}}[\varpi_2\xi_{i,k}(1-\theta_{ij})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{C}_{i,k}]\boldsymbol{u}_{i,k}.\end{aligned}\tag{17}$$

To prove the stability of the proposed algorithm, the discrete Kalman filter equation and two lemmas are given below (Olfati-Saber, 2009):

$$\begin{cases}\hat{\boldsymbol{x}}_k=\bar{\boldsymbol{x}}_k+\boldsymbol{K}_k(\boldsymbol{z}_k-\boldsymbol{H}\bar{\boldsymbol{x}}),\\ \boldsymbol{K}_k=\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{R}_k+\boldsymbol{H}\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}})^{-1},\\ \boldsymbol{M}_k=\boldsymbol{P}_k-\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{R}_k+\boldsymbol{H}\boldsymbol{P}_k\boldsymbol{H}^{\mathrm{T}})^{-1}\boldsymbol{H}\boldsymbol{P}_k,\\ \boldsymbol{P}_{k+1}=\boldsymbol{A}\boldsymbol{M}_k\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_k\boldsymbol{B}^{\mathrm{T}},\\ \bar{\boldsymbol{x}}_{k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_k.\end{cases}\tag{18}$$

Lemma 1 For the discrete Kalman filter (18), define $\boldsymbol{F}_k=\boldsymbol{I}-\boldsymbol{K}_k\boldsymbol{H}$ and information matrix $\boldsymbol{S}_k=\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}_k^{-1}\boldsymbol{H}$. The following conclusions can be drawn:

(1) $\boldsymbol{F}_k=\boldsymbol{I}-\boldsymbol{M}_k\boldsymbol{S}_k=\boldsymbol{M}_k\boldsymbol{P}_k^{-1}$,

(2) $\boldsymbol{M}_{k+1}=\boldsymbol{F}_k\boldsymbol{G}_k\boldsymbol{F}_k^{\mathrm{T}}$,

where $\boldsymbol{G}_k=\boldsymbol{A}\boldsymbol{M}_k\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_k\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{P}_k\boldsymbol{S}_k\boldsymbol{P}_k^{\mathrm{T}}$. If the information matrix $\boldsymbol{S}$ is positive definite, then $\boldsymbol{G}$ is also positive definite, and vice versa.

Lemma 2 For the discrete Kalman filter (18), $\boldsymbol{\eta}_k=\hat{\boldsymbol{x}}_k-\boldsymbol{x}_k$ and $\bar{\boldsymbol{\eta}}_k=\bar{\boldsymbol{x}}_k-\boldsymbol{x}_k$ represent the estimation error and prediction error, respectively. If the information matrix $\boldsymbol{S}$ is positive definite at any time, choose the Lyapunov function as $V(\boldsymbol{\eta}_k)=\boldsymbol{\eta}_k^{\mathrm{T}}\boldsymbol{M}_k^{-1}\boldsymbol{\eta}_k$, and then $\Delta V(\boldsymbol{\eta}_k)=-\boldsymbol{\eta}_k^{\mathrm{T}}\boldsymbol{\Phi}_k\boldsymbol{\eta}_k<0, \forall\boldsymbol{\eta}_k\neq\boldsymbol{0}$; that is, the error of the filter is globally asymptotically stable. Among them,

$$\boldsymbol{\Phi}_k=\boldsymbol{M}_k^{-1}(\boldsymbol{M}_k^{-1}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}_k^{-1}\boldsymbol{A})^{-1}\boldsymbol{M}_k^{-1},$$
$$\boldsymbol{W}_k=\boldsymbol{B}\boldsymbol{Q}_k\boldsymbol{B}+\boldsymbol{P}_k\boldsymbol{S}_k\boldsymbol{P}_k>0.$$

Theorem 2 Consider an undirected connected network with the number of cluster-head nodes m. Each cluster-head node estimates the state of noise system (1). This estimation employs Algorithm 1, given the Kalman filter gain $\boldsymbol{K}_{i,k}=\boldsymbol{A}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}(\boldsymbol{H}_{i,k}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}}+\boldsymbol{R}_{i,k})^{-1}$ and the consensus gain $\boldsymbol{C}_{i,k}=\gamma\boldsymbol{A}^{-1}\boldsymbol{F}_{i,k}\boldsymbol{G}_{i,k}$, where $\boldsymbol{F}_{i,k}=\boldsymbol{I}-\boldsymbol{M}_{i,k}\boldsymbol{S}_{i,k}$, $\boldsymbol{G}_{i,k}=\boldsymbol{A}\boldsymbol{M}_{i,k}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_{i,k}\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{P}_{i,k}\boldsymbol{S}_{i,k}\boldsymbol{P}_{i,k}^{\mathrm{T}}$. When the number of iteration steps $k\geq0$, select a sufficiently small number γ ($\gamma>0$). If

$$\boldsymbol{\Lambda}_k-\gamma^2\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{N}_k\hat{\boldsymbol{L}}\hat{\boldsymbol{A}}>0 \tag{19}$$

is also satisfied, where $\boldsymbol{\Lambda}_k=\mathrm{diag}(\Lambda_k^1,\Lambda_k^2,\cdots,\Lambda_k^m)$ and $\boldsymbol{N}_k=\mathrm{diag}(\boldsymbol{N}_k^1,\boldsymbol{N}_k^2,\cdots,\boldsymbol{N}_k^m)$ are partition diagonal matrices, the correlation expression is

$$\begin{aligned}-\boldsymbol{\Lambda}_k=&(\boldsymbol{A}-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\\&\cdot(\boldsymbol{A}-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})-\boldsymbol{M}_{i,k}^{-1},\end{aligned} \tag{20}$$

$$\boldsymbol{N}_{i,k}=\varpi_2\xi_{i,k}(1-\theta_{ij})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{C}_{i,k}. \tag{21}$$

Therefore, the error of FCET-KCF is globally asymptotically stable, and the estimated values of all cluster-head nodes are approximately equal, that is, $\hat{\boldsymbol{x}}_1=\hat{\boldsymbol{x}}_2=\cdots=\hat{\boldsymbol{x}}_m=\boldsymbol{x}$. The expression $\xi_{i,k}$ in Eqs. (20) and (21) is as shown in Eq. (12).

Proof According to Lemma 1, if $\boldsymbol{F}_{i,k}=\boldsymbol{I}-\boldsymbol{M}_{i,k}\boldsymbol{S}_{i,k}=\boldsymbol{M}_{i,k}\boldsymbol{P}_{i,k}^{-1}<0$, then $(\boldsymbol{A}-\varpi_1\xi_{i,k}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\varpi_2\xi_{i,k}\cdot(1-\theta_{ij})\boldsymbol{C}_{i,k}\leq0$, and the second and third terms in Eq. (17) are negative semi-definite. So, it is necessary to make sure that $\boldsymbol{\Lambda}_{i,k}-\gamma^2\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{L}}N_{i,k}\hat{\boldsymbol{L}}\hat{\boldsymbol{A}}>0$. The filter error dynamics of FCET-KCF is globally asymptotically stable. At the same time, note that the value of $\boldsymbol{\Lambda}_{i,k}-\gamma^2\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{N}_{i,k}\hat{\boldsymbol{L}}\hat{\boldsymbol{A}}$ is related to the value of the event-triggered function, $\xi_{i,k}$, packet arrival rates ϖ_1 and ϖ_2, path loss rate θ_{ij}, and consensus gain $\boldsymbol{C}_i$. When no clusters meet the information freshness requirement, that is, $\xi_{i,k}=0$ ($i=1, 2, \ldots, m$), we have $\hat{\boldsymbol{x}}_{i,k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_{i,k}$. This means that there is no sufficient information exchange between cluster heads, and Algorithm 1 will not converge.

When $\xi_{i,k}=1$ ($i=1, 2, \ldots, m$), $\boldsymbol{N}_k^*$ denotes $\boldsymbol{N}_{i,k}$ and $\boldsymbol{\Lambda}_k^*$ denotes $\boldsymbol{\Lambda}_{i,k}$; therefore, Eq. (17) can be further simplified as

$$\begin{aligned}\Delta V(\boldsymbol{\eta}_k)=&\sum\nolimits_i\boldsymbol{\eta}_{i,k}^{\mathrm{T}}[\varpi_1\boldsymbol{A}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{F}_{i,k}\boldsymbol{A}-\boldsymbol{M}_{i,k}^{-1}]\boldsymbol{\eta}_{i,k}\\&+2\sum\nolimits_i\boldsymbol{\eta}_{i,k}^{\mathrm{T}}\varpi_1\boldsymbol{A}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\varpi_2\boldsymbol{C}_{i,k}(1-\theta_{ij})\boldsymbol{u}_{i,k}\\&+\sum\nolimits_i\varpi_2(1-\theta_{ij})\boldsymbol{u}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{C}_{i,k}u_{i,k},\end{aligned} \tag{22}$$

where $\boldsymbol{F}_{i,k}=\boldsymbol{I}-\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k}=\boldsymbol{I}-\boldsymbol{M}_{i,k}\boldsymbol{S}_{i,k}$.

From the conclusion in Lemma 2, the following can be derived:

$$\begin{aligned}&\boldsymbol{A}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{F}_{i,k}\boldsymbol{A}-\boldsymbol{M}_{i,k}^{-1}\\&=-\boldsymbol{M}_{i,k}^{-1}[\boldsymbol{M}_{i,k}^{-1}+(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}_{i,k}^{-1}\boldsymbol{A})^{-1}\boldsymbol{M}_{i,k}^{-1}]<0,\end{aligned} \tag{23}$$

where $\boldsymbol{W}_{i,k}=\boldsymbol{B}\boldsymbol{Q}_{i,k}\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{P}_{i,k}\boldsymbol{S}_{i,k}\boldsymbol{P}_{i,k}^{\mathrm{T}}>0$.

According to the consensus theory (Godsil and Royle, 2001), any undirected network G satisfies

$$\sum_{i=1}^{m}\sum_{j\in N_i}\boldsymbol{\zeta}_i(\boldsymbol{\zeta}_j-\boldsymbol{\zeta}_i)=-\frac{1}{2}\sum_{(i,j)\in E}\left\|\boldsymbol{\zeta}_j-\boldsymbol{\zeta}_i\right\|^2=-\boldsymbol{\zeta}^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{\zeta}, \tag{24}$$

where $\boldsymbol{\zeta}=\mathrm{col}(\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,\ldots,\boldsymbol{\zeta}_n)$, $\hat{\boldsymbol{L}}=\boldsymbol{L}_m\otimes\boldsymbol{L}$.

Thus, the following can be obtained from Eqs. (22) and (24) when $\boldsymbol{C}_{i,k}=\gamma(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1})^{-1}$:

$$\begin{aligned}&\sum_i\boldsymbol{\eta}_{i,k}^{\mathrm{T}}\varpi_1\boldsymbol{A}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\varpi_2\boldsymbol{C}_{i,k}(1-\theta_{ij})\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{u}_{i,k}\\&=\gamma\varpi_1\varpi_2(1-\theta_{ij})\sum_{i=1}^{m}\sum_{j\in N_i}(\boldsymbol{\eta}_k^i)^{\mathrm{T}}[\boldsymbol{\eta}_{i,k}^{\mathrm{T}}-\boldsymbol{\eta}_{i,k}/(1-\theta_{ij})]\\&=-\frac{\gamma}{2}\varpi_1\varpi_2(1-\theta_{ij})\sum_{(i,j)\in E_G}\left\|\boldsymbol{\eta}_{j,k}-\boldsymbol{\eta}_{i,k}/(1-\theta_{ij})\right\|^2\\&=-\gamma\varpi_1\varpi_2(1-\theta_{ij})\boldsymbol{\eta}_k^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{\eta}_k.\end{aligned} \tag{25}$$

Similarly, $\varpi_2(1-\theta_{ij})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{C}_{i,k}$ of the third term in Eq. (22) can be derived as

$$\begin{aligned}&\varpi_2(1-\theta_{ij})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}\boldsymbol{C}_{i,k}=\boldsymbol{N}_k^*\\&=\varpi_2(1-\theta_{ij})(\gamma\boldsymbol{A}^{-1}\boldsymbol{K}_{i,k}\boldsymbol{G}_{i,k})^{\mathrm{T}}\boldsymbol{M}_{i,k+1}^{-1}(\gamma\boldsymbol{A}^{-1}\boldsymbol{K}_{i,k}\boldsymbol{G}_{i,k})\\&=\varpi_2(1-\theta_{ij})\gamma^2(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}(\boldsymbol{A}\boldsymbol{M}_{i,k}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_{i,k}\boldsymbol{B}^{\mathrm{T}}\\&\quad+\boldsymbol{P}_{i,k}\boldsymbol{S}_{i,k}\boldsymbol{P}_{i,k}^{\mathrm{T}}).\end{aligned}\tag{26}$$

As concluded from Lemma 1, if the information matrix $\boldsymbol{S}$ is positive definite, then $\boldsymbol{G}=\boldsymbol{A}\boldsymbol{M}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^{\mathrm{T}}+\boldsymbol{P}\boldsymbol{S}\boldsymbol{P}$ is also positive definite, and thus the last term in Eq. (22) is positive definite. Combining the first and third terms in Eq. (22), $\Delta V(\boldsymbol{\eta}_k)$ can be obtained as

$$\begin{aligned}\Delta V(\boldsymbol{\eta}_k)=&-\boldsymbol{\eta}_k^{\mathrm{T}}[\boldsymbol{\Lambda}_k^*-\gamma^2\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{N}_k^*\hat{\boldsymbol{L}}\hat{\boldsymbol{A}}]\boldsymbol{\eta}_k\\&-\gamma\varpi_1\varpi_2(1-\theta_{ij})\boldsymbol{\eta}_k^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{\eta}_k.\end{aligned}\tag{27}$$

So, for all $\boldsymbol{\eta}_k\neq 0$, as long as the coefficient γ is small enough to make $\boldsymbol{\Lambda}_k^*-\gamma^2\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{N}_k^*\hat{\boldsymbol{L}}\hat{\boldsymbol{A}}>0$, there is $\Delta V(\boldsymbol{\eta})<0$. Meanwhile, it is necessary to prevent γ from being too small to affect the convergence of $V(\boldsymbol{\eta}_k)$. The upper bound of γ is shown as

$$\gamma\le\gamma^*=\left(\frac{\min_i\lambda_{\min}(\boldsymbol{\Lambda}_i^*)}{\max_i\lambda_{\max}(\boldsymbol{N}_i^*)}\right)^{1/2}\frac{1}{\sigma_{\max}(\boldsymbol{A})\lambda_n(\boldsymbol{L})},\tag{28}$$

where $\lambda_n(\boldsymbol{L})$ is the maximum eigenvalue of Laplacian matrix $\boldsymbol{L}$ for graph G.

Similarly, as long as the data packets generated at any sampling time meet the freshness requirements and the system guarantees to meet

$$\boldsymbol{\Lambda}_k^*-\gamma^2\hat{\boldsymbol{A}}^{\mathrm{T}}\hat{\boldsymbol{L}}\boldsymbol{N}_k^*\hat{\boldsymbol{L}}\hat{\boldsymbol{A}}>0,\tag{29}$$

the filter error dynamics of FCET-KCF is globally asymptotically stable.

4 Energy consumption analysis of a WSN based on FCET-KCF

In this study, the energy consumption model is adopted from Heinzelman et al. (2002). When the sensor node sends an l-bit packet to the neighbor node at a distance of d, the energy consumption of the transmitter and receiver and the total energy consumption of radio frequency are

$$\begin{aligned}E_{\mathrm{Tx}}(l,d)&=E_{\mathrm{Tx\text{-}elec}}(l)+E_{\mathrm{Tx\text{-}amp}}(l,d)\\&=\begin{cases}E_{\mathrm{elec}}l+\varsigma_{\mathrm{mp}}ld^2, & d<d_0,\\E_{\mathrm{elec}}l+\varsigma_{\mathrm{mp}}ld^4, & d\ge d_0,\end{cases}\\E_{\mathrm{Rx}}(l)&=E_{\mathrm{Rx\text{-}elec}}(l)=E_{\mathrm{elec}}l,\\E_{\mathrm{total}}&=l(2E_{\mathrm{elec}}+\varsigma_{\mathrm{mp}}d^2)+E_{\mathrm{DA}},\end{aligned}$$

where E_{elec} represents the energy consumption for a 1-bit packet sent by the transmitter or received by the receiver, $\varsigma_{\mathrm{mp}}d^2$ or $\varsigma_{\mathrm{mp}}d^4$ represents the energy consumption of the amplifier, ς_{mp} represents the power consumption when the transmitting amplifier transmits packets for 1 m^2/bit, d_0 is the critical distance, and E_{DA} represents the energy required for data fusion. In this study it is assumed that the process of packet transmission from the cluster member node to the cluster-head node does not consume energy, and the energy consumed by the cluster-head node in calculating the AoI of packets is ignored.

When the event is not triggered, the sensor node uses the previous estimation value for consensus calculation, the current estimation value is not transmitted, and the energy consumption is only $E_{\mathrm{elec}}l+E_{\mathrm{DA}}$. Therefore, the energy consumption of cluster-head node i at time k can be expressed as

$$E=\xi_{i,k}l(E_{\mathrm{elec}}+\varsigma_{\mathrm{mp}}d^2)+E_{\mathrm{elec}}l+E_{\mathrm{DA}}.\tag{30}$$

The value of $\xi_{i,k}$ is as shown in Eq. (12). According to Eq. (30), it can be seen that the consumption of network energy is affected by the packet freshness constraint. Therefore, the monitoring system can greatly reduce the network energy consumption with the FCET-KCF algorithm on the premise of ensuring the stability of the system.

5 Performance analysis of FCET-KCF

Fig. 5 describes the node configuration of a distributed integrated sensor network for monitoring pollutant concentrations for aircraft, and the following simulations are performed based on this configuration.

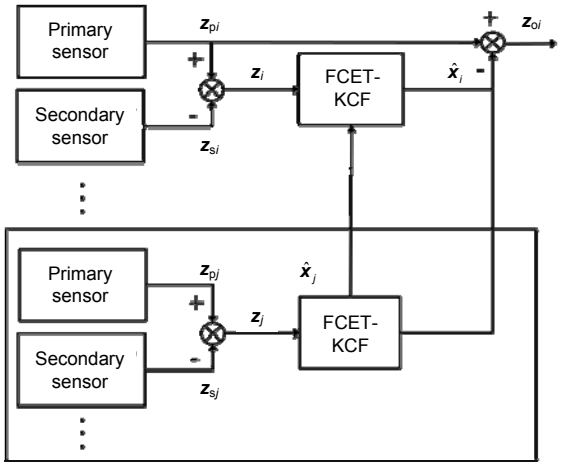

Fig. 5 Simulation configuration of the distributed monitoring sensor network

As shown in Fig. 5, z_{pi} and z_{si} are the measured values of pollutant concentrations in the aircraft measured by primary and secondary sensors, respectively. Correspondingly, the measurement error is z_i. The output value $\hat{\boldsymbol{x}}_i$ of the distributed filter modifies the measurement value of the primary sensor through the FCET-KCF algorithm (Algorithm 1) to obtain the final measurement values. The Monte Carlo method is used in the simulation process to carry out a large number of independently repeated experiments. Using the statistical mean value of each time, we analyze the error response of the monitoring network, by adopting the following performance indexes given in Wang et al. (2017).

Mean estimation error (MEE):

$$\mathrm{MEE}_k=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(\boldsymbol{e}_{i,k}^{\mathrm{T}}\boldsymbol{e}_{i,k})},\ \boldsymbol{e}_{i,k}=\hat{\boldsymbol{x}}_{i,k}-\boldsymbol{x}_{i,k}.$$

Mean consistency error (MCE):

$$\mathrm{MCE}_k=\sqrt{\frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\delta}_{i,k}^{\mathrm{T}}\boldsymbol{\delta}_{i,k}},\ \boldsymbol{\delta}_{i,k}=\hat{\boldsymbol{x}}_{i,k}-\frac{1}{m}\sum_{i=1}^{m}\hat{\boldsymbol{x}}_{i,k}.$$

Here k is the instantaneous time, and m is the number of cluster-head nodes.

For comparison with the existing methods, the ET-KCF algorithm in Wang et al. (2019) and KCF, we select the same parameters for our proposed FCET-KCF algorithm. The error dynamic equation is selected as

$$\boldsymbol{x}_{i,k+1}=\begin{bmatrix}0.9995 & -0.003+0.0001\cdot\mathrm{rand}\\ 0.003+0.0001\cdot\mathrm{rand} & 0.9995\end{bmatrix}\boldsymbol{x}_{i,k}+\begin{bmatrix}0.015 & 0\\ 0 & 0.015\end{bmatrix}\boldsymbol{w}_{i,k},$$

$$\boldsymbol{z}_{i,k+1}=\left(\begin{bmatrix}1+0.005\sin j & -0.003\cos j\\ 0.003\cos j & 1+0.005\sin j\end{bmatrix}+0.0001\cdot\mathrm{rand}\cdot\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\right)\boldsymbol{x}_{i,k}+\left(\begin{bmatrix}1+0.5\sin j & -0.03\cos j\\ 0.03\cos j & 1+0.5\sin j\end{bmatrix}+0.001\cdot\mathrm{rand}\cdot\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\right)\boldsymbol{v}_{i,k}.$$

The initial measurement error is set as $\boldsymbol{x}_0=(15, 10)^{\mathrm{T}}$, and the initial prediction error matrix is $\boldsymbol{P}_0=10\boldsymbol{I}_2$. The process noise and observed noise are independent Gaussian white noises with covariances of $10i$ and $100i$, respectively, where i is the node index. E_{elec}= 50 nJ/bit, E_{DA}=50 nJ/(bit·signal), ς_{mp}=100 pJ/(bit·m^2). The communication radius of the cluster-head node is $r=\sqrt{2}$ m. The number of bits k=5000 and the initial energy of each node is 4 J. The path loss rate θ_{ij} between nodes i and j is 0.2, and the observed packet loss rate ϖ_1 and communication packet loss rate ϖ_2 are 0.6. μ_i=0.5, l_i=0.8 (i=1, 2, …, m). λ is a random number in (0, 1).

As shown in Fig. 6, when packet loss and path loss exist, KCF will diverge, while the ET-KCF algorithm proposed in Wang et al. (2019) and our proposed FCET-KCF algorithm can still converge stably.

According to Fig. 7, the average consensus error performance of FCET-KCF is nearly the same as those of ET-KCF and KCF. In addition, in the case of packet loss and path loss, the average consensus error of the system is much lower than that of the local Kalman filter.

Figs. 6 and 7 show that ET-KCF and FCET-KCF converge and have smaller estimation errors and consensus errors in practical systems with packet loss and path loss, so the pollutant concentrations in the aircraft can be monitored more effectively.

Fig. 8 compares the energy consumption of FCET-KCF and ET-KCF. FCET-KCF consumes node

energy after about 350 sampling steps, while ET-KCF consumes energy after about 100 steps. Simulation results demonstrate that FCET-KCF can reduce the bandwidth pressure and save system energy while ensuring the stability of the estimation algorithm.

Fig. 9 presents the relationship between the system energy loss and the number of sampling steps for ET-KCF and FCET-KCF. The simulation results show that FCET-KCF has less energy loss and better energy saving performance when the system has the same number of sampling steps.

Figs. 8 and 9 demonstrate that FCET-KCF can effectively reduce the network energy consumption and extend the service life of the system compared to ET-KCF when monitoring aircraft pollutant concentrations.

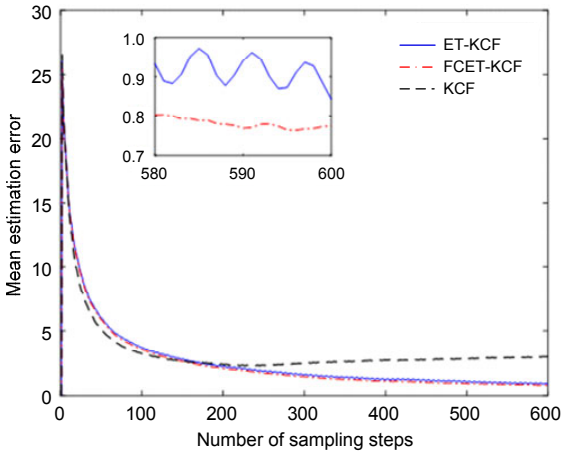

Fig. 6 Comparison of the average estimation error for different filters under packet loss and path loss

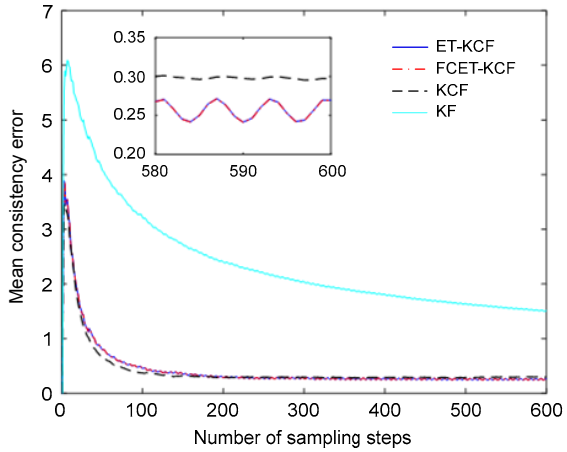

Fig. 7 Comparison of the average consensus error for different filters under packet loss and path loss

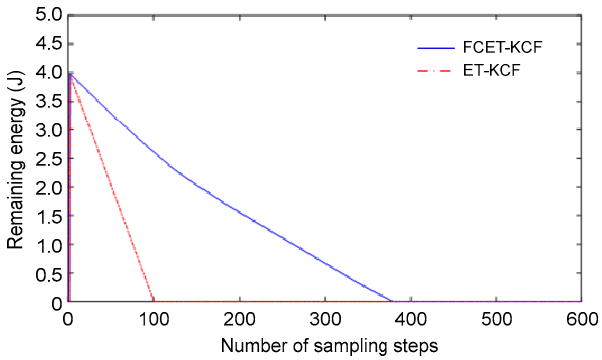

Fig. 8 Comparison of residual energy for the two filter algorithms

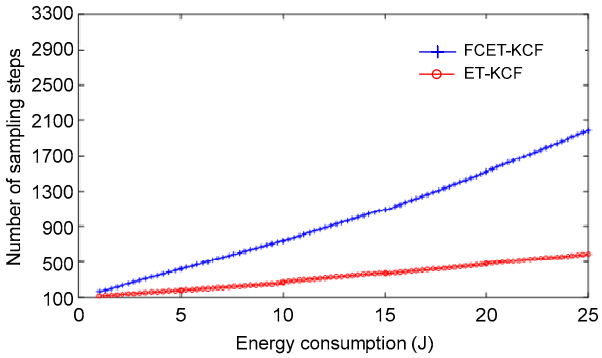

Fig. 9 Relationship between system energy loss and the number of sampling steps for the two filter algorithms

6 Conclusions

In this paper, we have proposed a new event-triggered Kalman consensus filter based on freshness constraints of the AoI algorithm to monitor pollutant concentrations in aircraft cabins. The algorithm comprehensively considers the observed packet loss, communication packet loss, and path loss in wireless sensor networks. Compared to the traditional KCF, this algorithm has better convergence in real detection systems. To enhance the real-time performance of the estimation system and reduce the system energy consumption, we have introduced a method to optimize the freshness of the sampled information for an event-triggered mechanism, so that a suitable threshold can be selected. The FCET-KCF algorithm also reduces the transmission of expired packets and saves energy. Finally, the stability of the algorithm is proved. Simulation results showed that the algorithm can accurately estimate pollutant concentrations in aircraft cabins in the presence of packet loss and path loss while greatly reducing network energy loss.

Contributors

Rui WANG and Yahui LI designed the research. Hui SUN and Youmin ZHANG guided the research. Rui WANG and Yahui LI drafted the manuscript. All authors contributed to the interpretation of the results and the revision of the paper.

Compliance with ethics guidelines

Rui WANG, Yahui LI, Hui SUN, and Youmin ZHANG declare that they have no conflict of interest.

References

Amini A, Asif A, Mohammadi A, 2018. An event-triggered average consensus algorithm with performance guarantees for distributed sensor networks. Proc IEEE Int Conf on Acoustics, Speech and Signal Processing, p.3409-3413. https://doi.org/10.1109/ICASSP.2018.8462664

Fan S, Yan HC, Zhang H, et al., 2017. Prediction consensus-based distributed Kalman filtering with packet loss. Proc 36th Chinese Control Conf, p.7950-7955. https://doi.org/10.23919/ChiCC.2017.8028613

Farag AM, 2018. Thermal comfort investigation for commercial aircraft cabin by using CFD. Proc IEEE Aerospace Conf, p.1-10. https://doi.org/10.1109/AERO.2018.8396465

Farazi S, Klein AG, Brown DR, 2018. Age of information in energy harvesting status update systems: when to preempt in service? Proc IEEE Int Symp on Information Theory, p.2436-2440. https://doi.org/10.1109/ISIT.2018.8437904

Godsil CD, Royle G, 2001. Algebraic Graph Theory. Springer, New York, USA.

Heinzelman WB, Chandrakasan AP, Balakrishnan H, 2002. An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans Wirel Commun*, 1(4):660-670. https://doi.org/10.1109/TWC.2002.804190

Hudhajanto RP, Fahmi N, Prayitno E, et al., 2018. Real-time monitoring for environmental through wireless sensor network technology. Proc Int Conf on Applied Engineering, p.1-5. https://doi.org/10.1109/INCAE.2018.8579377

Kaul S, Gruteser M, Rai V, et al., 2011. Minimizing age of information in vehicular networks. Proc 8th Annual IEEE Communications Society Conf on Sensor, Mesh and ad hoc Communications and Networks, p.350-358. https://doi.org/10.1109/SAHCN.2011.5984917

Kaul SK, Yates RD, Gruteser M, 2012. Status updates through queues. Proc 46th Annual Conf on Information Sciences and Systems, p.1-6. https://doi.org/10.1109/CISS.2012.6310931

Li F, Liu JJ, Ren JL, et al., 2016. Numerical investigation of airborne contaminant transport under different vortex structures in the aircraft cabin. *Int J Heat Mass Transf*, 96:287-295. https://doi.org/10.1016/j.ijheatmasstransfer.2016.01.004

Li WL, Jia YM, Du JP, 2016. Event-triggered Kalman consensus filter over sensor networks. *IET Contr Theory Appl*, 10(1):103-110. https://doi.org/10.1049/iet-cta.2015.0508

Liu XD, Li LY, Li Z, et al., 2017. Stochastic stability condition for the extended Kalman filter with intermittent observations. *IEEE Trans Circ Syst II Expr Briefs*, 64(3):334-338. https://doi.org/10.1109/TCSII.2016.2578956

Liu YG, Xu BG, Shi BH, 2013. Kalman filtering for stochastic systems with consecutive packet losses and measurement time delays. *Contr Theory Appl*, 30(7):898-908. https://doi.org/10.7641/CTA.2013.12052

Olfati-Saber R, 2007. Distributed Kalman filtering for sensor networks. Proc 46th IEEE Conf on Decision and Control, p.5492-5498. https://doi.org/10.1109/CDC.2007.4434303

Olfati-Saber R, 2009. Kalman-consensus filter: optimality, stability, and performance. Proc 48th IEEE Conf on Decision and Control held jointly with 28th Chinese Control Conf, p.7036-7042. https://doi.org/10.1109/CDC.2009.5399678

Olfati-Saber R, Shamma JS, 2005. Consensus filters for sensor networks and distributed sensor fusion. Proc 44th IEEE Conf on Decision and Control, p.6698-6703. https://doi.org/10.1109/CDC.2005.1583238

Paul A, Kamwa I, Jóos G, 2018. Centralized dynamic state estimation using a federation of extended Kalman filters with intermittent PMU data from generator terminals. *IEEE Trans Power Syst*, 33(6):6109-6119. https://doi.org/10.1109/TPWRS.2018.2834365

Saha S, Majumdar A, 2017. Data centre temperature monitoring with ESP8266 based wireless sensor network and cloud based dashboard with real time alert system. Proc Devices for Integrated Circuit, p.307-310. https://doi.org/10.1109/DEVIC.2017.8073958

Shi J, Qi GQ, Li YY, et al., 2018. Stochastic convergence analysis of cubature Kalman filter with intermittent observations. *J Syst Eng Electron*, 29(4):823-833. https://doi.org/10.21629/JSEE.2018.04.17

Song WH, Wang JN, Wang CY, et al., 2018. Event-triggered cooperative unscented Kalman filtering. Proc 15th Int Conf on Control, Automation, Robotics and Vision, p.1004-1009. https://doi.org/10.1109/ICARCV.2018.8581328

Steele JM, 2004. The Cauchy-Schwarz Master Class: an Introduction to the Art of Mathematical Inequalities. Cambridge University Press, New York, USA.

Sun SL, Tian T, Lin HL, 2016. Optimal linear estimators for systems with finite-step correlated noises and packet dropout compensations. *IEEE Trans Signal Process*, 64(21):5672-5681. https://doi.org/10.1109/TSP.2016.2576420

Sun Y, Cyr B, 2019. Sampling for data freshness optimization: non-linear age functions. *J Commun Netw*, 21(3):204-219. https://doi.org/10.1109/JCN.2019.000035

Talak R, Karaman S, Modiano E, 2018. Distributed scheduling algorithms for optimizing information freshness in wireless networks. Proc IEEE 19th Int Workshop on Signal Processing Advances in Wireless Communications, p.1-5. https://doi.org/10.1109/SPAWC.2018.8445979

Tang HY, Wang JT, Song LQ, et al., 2019. Scheduling to minimize age of information in multi-state time-varying networks with power constraints. Proc 57th Annual Allerton Conf on Communication, Control, and Computing, p.1198-1205. https://doi.org/10.1109/ALLERTON.2019.8919867

Wang R, Li YX, Sun H, et al., 2017. Analyses of integrated aircraft cabin contaminant monitoring network based on Kalman consensus filter. *ISA Trans*, 71:112-120. https://doi.org/10.1016/j.isatra.2017.06.027

Wang R, Wang XY, Sun H, et al., 2019. Analysis of estimator and energy consumption with multiple faults over the distributed integrated WSN. *Int J Model Ident Contr*, 32(2):154-168. https://doi.org/10.1504/IJMIC.2019.102365

Yu ZQ, Zhang YM, Liu ZX, et al., 2019a. Distributed adaptive fractional-order fault-tolerant cooperative control of networked unmanned aerial vehicles via fuzzy neural networks. *IET Contr Theory Appl*, 13(17):2917-2929. https://doi.org/10.1049/iet-cta.2018.6262

Yu ZQ, Qu YH, Zhang YM, 2019b. Distributed fault-tolerant cooperative control for multi-UAVs under actuator fault and input saturation. *IEEE Trans Contr Syst Technol*, 27(6):2417-2429. https://doi.org/10.1109/TCST.2018.2868038

Yu ZQ, Liu ZX, Zhang YM, et al., 2020. Distributed finite-time fault-tolerant containment control for multiple unmanned aerial vehicles. *IEEE Trans Neur Netw Learn Syst*, 31(6):2077-2091. https://doi.org/10.1109/TNNLS.2019.2927887

Zhang LL, Zhang Y, 2018. A distributed energy-efficient target tracking algorithm based on event-triggered strategy for sensor networks. Proc 33rd Youth Academic Annual Conf of Chinese Association of Automation, p.12-17. https://doi.org/10.1109/YAC.2018.8405798

Zhou ZM, Fu CC, Xue CJ, et al., 2020. Energy-constrained data freshness optimization in self-powered networked embedded systems. *IEEE Trans Comput-Aided Des Integr Circ Syst*, 30(10):2293-2306. https://doi.org/10.1109/TCAD.2019.2948905

Appendix A: KCF-based consensus algorithm

The consensus strategy is a key tenet in the Kalman filter algorithm, and can make the estimation value of each sensor converge to the true value consistently. Olfati-Saber (2007) proposed the classical Kalman consensus filter algorithm:

$$\begin{cases}\hat{\boldsymbol{x}}_{i,k}=\bar{\boldsymbol{x}}_{i,k}+\boldsymbol{K}_{i,k}(\boldsymbol{Z}_{i,k}-\boldsymbol{H}_{i,k}\bar{\boldsymbol{x}}_{i,k})+\boldsymbol{C}_{i,k}\sum\limits_{j\in N_i}(\bar{\boldsymbol{x}}_{j,k}-\bar{\boldsymbol{x}}_{i,k}),\\ \boldsymbol{K}_{i,k}=\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}}(\boldsymbol{R}_{i,k}+\boldsymbol{H}_{i,k}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}})^{-1},\\ \boldsymbol{M}_{i,k}=(\boldsymbol{I}-\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})\boldsymbol{P}_{i,k}(\boldsymbol{I}-\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}+\boldsymbol{K}_{i,k}\boldsymbol{R}_{i,k}\boldsymbol{K}_{i,k}^{\mathrm{T}},\\ \boldsymbol{P}_{i,k+1}=\boldsymbol{A}\boldsymbol{M}_{i,k}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_{i,k}\boldsymbol{B}^{\mathrm{T}},\\ \bar{\boldsymbol{x}}_{i,k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_{i,k}.\end{cases} \tag{A1}$$

The sensor nodes in the algorithm do not need to exchange information with all other nodes. The node measured value can achieve a global consensus estimation by exchanging information with only a few neighbor sensor nodes. In the classical Kalman consensus filter algorithm (A1), substituting the intermediate variable $\boldsymbol{M}_{i,k}$ into $\boldsymbol{P}_{i,k+1}$ and substituting $\hat{\boldsymbol{x}}_{i,k}$ into $\bar{\boldsymbol{x}}_{i,k+1}$, the compact form of the filter algorithm can be obtained:

$$\begin{cases}\hat{\boldsymbol{x}}_{i,k+1}=\boldsymbol{A}\hat{\boldsymbol{x}}_{i,k}+\boldsymbol{A}\boldsymbol{K}_{i,k}(\boldsymbol{Z}_{i,k}-\boldsymbol{H}_{i,k}\hat{\boldsymbol{x}}_{i,k})\\ \qquad\qquad +\boldsymbol{A}\boldsymbol{C}_{i,k}\sum\limits_{j\in N_i}(\hat{\boldsymbol{x}}_{j,k}-\hat{\boldsymbol{x}}_{i,k}),\\ \boldsymbol{K}_{i,k}=\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}}(\boldsymbol{R}_{i,k}+\boldsymbol{H}_{i,k}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}})^{-1},\\ \boldsymbol{P}_{i,k+1}=(\boldsymbol{A}-\boldsymbol{A}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})\boldsymbol{P}_{i,k}(\boldsymbol{A}-\boldsymbol{A}\boldsymbol{K}_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\ \qquad\qquad +\boldsymbol{A}\boldsymbol{K}_{i,k}\boldsymbol{R}_{i,k}\boldsymbol{K}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{Q}_{i,k}\boldsymbol{B}^{\mathrm{T}}.\end{cases}$$

Appendix B: Derivation of optimal gain $K_{i,k}$

For simplicity, the notation $\boldsymbol{P}_{ij,k}=\boldsymbol{P}_{i,k}$ when $i=j$. Substituting Eq. (13) into Eq. (14) yields:

$$\begin{aligned}\boldsymbol{e}_{i,k+1}&=\hat{\boldsymbol{x}}_{i,k+1}-\boldsymbol{x}_{i,k+1}\\ &=\boldsymbol{A}\hat{\boldsymbol{x}}_{i,k}+\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}(\boldsymbol{Z}_{i,k}-\boldsymbol{H}_{i,k}\hat{\boldsymbol{x}}_{i,k})\\ &\quad+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{j\in N_i}[(1-\theta_{ij})\hat{\boldsymbol{x}}_{j,k}^{o}-\hat{\boldsymbol{x}}_{i,k}]-\boldsymbol{A}\boldsymbol{x}_{i,k}-\boldsymbol{B}\boldsymbol{w}_{i,k}\\ &=\boldsymbol{A}\boldsymbol{e}_{i,k}+\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}(\boldsymbol{H}_{i,k}\boldsymbol{x}_{i,k}-\boldsymbol{H}_{i,k}\hat{\boldsymbol{x}}_{i,k}+\boldsymbol{F}_{i,k}\boldsymbol{v}_{i,k})\\ &\quad+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{j\in N_i}[(1-\theta_{ij})\hat{\boldsymbol{x}}_{j,k}^{o}-(1-\theta_{ij})\boldsymbol{x}_{i,k}\\ &\quad+(1-\theta_{ij})\boldsymbol{x}_{i,k}-\hat{\boldsymbol{x}}_{i,k}]-\boldsymbol{B}\boldsymbol{w}_{i,k}\\ &=(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\boldsymbol{e}_{i,k}+\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{F}_{i,k}\boldsymbol{v}_{i,k}\\ &\quad+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{j\in N_i}[(1-\theta_{ij})\hat{\boldsymbol{x}}_{j,k}^{o}-(1-\theta_{ij})\boldsymbol{x}_{i,k}\end{aligned}$$

$$
\begin{aligned}
&+(1-\theta_{ij})\boldsymbol{x}_{i,k}-\hat{\boldsymbol{x}}_{i,k}]-\boldsymbol{B}\boldsymbol{w}_{i,k}\\
&=(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\boldsymbol{e}_{i,k}+\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{F}_{i,k}\boldsymbol{v}_{i,k}\\
&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum\nolimits_{j\in N_i}(1-\theta_{ij})\tilde{\boldsymbol{e}}_{j,k}\\
&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum\nolimits_{j\in N_i}(1-\theta_{ij})(\boldsymbol{e}_{j,k}-\boldsymbol{e}_{i,k})\\
&-\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum\nolimits_{j\in N_i}\theta_{ij}\hat{\boldsymbol{x}}_{i,k}-\boldsymbol{B}\boldsymbol{w}_{i,k}.
\end{aligned}
$$

Then, we can obtain the covariance matrix $\boldsymbol{P}_{ij,k+1}$:

$$
\begin{aligned}
&\boldsymbol{P}_{ij,k+1}=E\{\boldsymbol{e}_{i,k+1}\boldsymbol{e}_{j,k+1}^{\mathrm{T}}\}\\
&=(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\boldsymbol{P}_{ij,k}(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&+\varpi_1^2\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{F}_{i,k}\boldsymbol{v}_{i,k}\boldsymbol{v}_{i,k}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\boldsymbol{K}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}(1-\theta_{ir})(1-\theta_{js})\tilde{\boldsymbol{P}}_{rs,k}\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}(1-\theta_{ir})(1-\theta_{js})\\
&\cdot(\boldsymbol{P}_{rs,k}-\boldsymbol{P}_{rj,k}-\boldsymbol{P}_{is,k}+\boldsymbol{P}_{ij,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}\theta_{ir}\theta_{js}\hat{\boldsymbol{P}}_{ij,k}\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}\boldsymbol{w}_{i,k}\boldsymbol{w}_{i,k}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\\
&+\varpi_2(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\sum_{r\in N_j}(1-\theta_{jr})((\boldsymbol{P}_{ri,k}^{o})^{\mathrm{T}}-\boldsymbol{P}_{ij,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\sum_{r\in N_j}(1-\theta_{jr})(\boldsymbol{P}_{ir,k}-\boldsymbol{P}_{i,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_2(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\sum_{r\in N_j}\theta_{jr}(\hat{\boldsymbol{P}}_{ij,k}-(\hat{\boldsymbol{P}}'_{oj,k})^{\mathrm{T}})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}(1-\theta_{ir})(\boldsymbol{P}_{rj,k}^{o}-\boldsymbol{P}_{rj,k})(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&+\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}(1-\theta_{ir})(1-\theta_{js})\\
&\cdot(\boldsymbol{P}_{rs,k}^{o}-\boldsymbol{P}_{rj,k}^{o}-\boldsymbol{P}_{rs,k}+\boldsymbol{P}_{rj,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}(1-\theta_{ir})\theta_{js}(\hat{\boldsymbol{P}}_{rj,k}^{o}-\hat{\boldsymbol{P}}_{rj,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}(1-\theta_{ir})(\boldsymbol{P}_{rj,k}-\boldsymbol{P}_{ij,k})(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&+\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}(1-\theta_{ir})(1-\theta_{js})\\
&\cdot((\boldsymbol{P}_{rs,k}^{o})^{\mathrm{T}}-\boldsymbol{P}_{rs,k}-\boldsymbol{P}_{si,k}^{o}+\boldsymbol{P}_{is,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}(1-\theta_{ir})\theta_{js}(\hat{\boldsymbol{P}}_{rj,k}^{o}-\hat{\boldsymbol{P}}_{ij,k}^{o})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\theta_{ir}(\hat{\boldsymbol{P}}_{ij,k}-\hat{\boldsymbol{P}}'_{io,k})(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&-\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}\theta_{ir}(1-\theta_{js})((\hat{\boldsymbol{P}}_{si,k}^{o})^{\mathrm{T}}-\hat{\boldsymbol{P}}_{is,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_2^2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\sum_{s\in N_j}\theta_{ir}(1-\theta_{js})(\hat{\boldsymbol{P}}_{is,k}-\hat{\boldsymbol{P}}_{ij,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}.
\end{aligned}
$$

Therefore, the optimal gain matrix $\boldsymbol{K}_{i,k}$ can be obtained by solving the following equation, where the estimation error of node i is $\mathrm{tr}(\boldsymbol{P}_{i,k+1})=E\{(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{i,k})(\boldsymbol{x}_k-\hat{\boldsymbol{x}}_{i,k})^{\mathrm{T}}\}$:

$$
\begin{aligned}
&\frac{\partial\mathrm{tr}(\boldsymbol{P}_{i,k+1})}{\partial\boldsymbol{K}_{i,k}}=0\Rightarrow\\
&\partial\mathrm{tr}\{(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\boldsymbol{P}_{i,k}(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&+\varpi_1^2\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{F}_{i,k}\boldsymbol{v}_{i,k}\boldsymbol{v}_{i,k}^{\mathrm{T}}\boldsymbol{F}_{i,k}^{\mathrm{T}}\boldsymbol{K}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\sum_{r\in N_i}(1-\theta_{ir})((\boldsymbol{P}_{ri,k}^{o})^{\mathrm{T}}-\boldsymbol{P}_{i,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\sum_{r\in N_i}(1-\theta_{ir})(\boldsymbol{P}_{ir,k}-\boldsymbol{P}_{i,k})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_2(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})\sum_{r\in N_i}\theta_{ir}(\hat{\boldsymbol{P}}_{i,k}-(\hat{\boldsymbol{P}}'_{oi,k})^{\mathrm{T}})\boldsymbol{C}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}(1-\theta_{ir})(\boldsymbol{P}_{ri,k}^{o}-\boldsymbol{P}_{ri,k})(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&+\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}(1-\theta_{ir})(\boldsymbol{P}_{ri,k}-\boldsymbol{P}_{i,k})(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\\
&-\varpi_2\boldsymbol{A}\boldsymbol{C}_{i,k}\xi_{i,k}\sum_{r\in N_i}\theta_{ir}(\hat{\boldsymbol{P}}_{i,k}-\hat{P}'_{io,k})\\
&\cdot(\boldsymbol{A}-\varpi_1\boldsymbol{A}\boldsymbol{K}_{i,k}\xi_{i,k}\boldsymbol{H}_{i,k})^{\mathrm{T}}\}/\partial\boldsymbol{K}_{i,k}=0.
\end{aligned}
$$

As the matrix calculus theory illustrates, for any two matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, the following two properties hold:

$$
\frac{\partial\mathrm{tr}(\boldsymbol{X}\boldsymbol{Y})}{\partial\boldsymbol{Y}}=\boldsymbol{X}^{\mathrm{T}},\quad\frac{\partial\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{X})}{\partial\boldsymbol{X}}=(\boldsymbol{Y}+\boldsymbol{Y}^{\mathrm{T}})\boldsymbol{X}.
$$

Therefore, we have

$$
\begin{aligned}
&-2\varpi_1\xi_{i,k}\boldsymbol{A}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}+2\varpi_1^2\xi_{i,k}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}\boldsymbol{A}\boldsymbol{K}_{i,k}\\
&+2\varpi_1^2\xi_{i,k}\boldsymbol{A}\boldsymbol{R}_{i,k}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{K}_{i,k}\\
&-\varpi_1\varpi_2\sum_{r\in N_i}(1-\theta_{ir})\boldsymbol{A}((\boldsymbol{P}_{ri,k}^{o})^{\mathrm{T}}-\boldsymbol{P}_{i,k})^{\mathrm{T}}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_1\varpi_2\sum_{r\in N_i}(1-\theta_{ir})\boldsymbol{A}(\boldsymbol{P}_{ri,k}^{o}-\boldsymbol{P}_{i,k})^{\mathrm{T}}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_1\varpi_2\sum_{r\in N_i}\theta_{ir}\boldsymbol{A}(\hat{\boldsymbol{P}}_{i,k}-(\boldsymbol{P}'_{oi,k})^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_1\varpi_2\sum_{r\in N_i}(1-\theta_{ir})\boldsymbol{A}(\boldsymbol{P}_{ri,k}^{o}-\boldsymbol{P}_{i,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&-\varpi_1\varpi_2\sum_{r\in N_i}(1-\theta_{ir})\boldsymbol{A}(\boldsymbol{P}_{ri,k}-\boldsymbol{P}_{i,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&+\varpi_1\varpi_2\sum_{r\in N_i}\theta_{ir}\boldsymbol{A}(\hat{\boldsymbol{P}}_{i,k}-\hat{\boldsymbol{P}}'_{io,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}=0.
\end{aligned}
$$

The following equation can be obtained by reorganization:

$$
\begin{aligned}
&2\varpi_1\xi_{i,k}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}\boldsymbol{A}+\boldsymbol{A}\boldsymbol{R}_{i,k}\boldsymbol{A}^{\mathrm{T}})\boldsymbol{K}_{i,k}\\
&=2\boldsymbol{A}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}+\varpi_2\xi_{i,k}\sum_{r\in N_i}(1-\theta_{ir})\boldsymbol{A}\\
&\quad\cdot(2(\boldsymbol{P}_{ri,k}^{o})^{\mathrm{T}}+\boldsymbol{P}_{ri,k}^{o}-2\boldsymbol{P}_{i,k}-\boldsymbol{P}_{ri,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&\quad-\varpi_2\xi_{i,k}\sum_{r\in N_i}\theta_{ir}\boldsymbol{A}(2\hat{\boldsymbol{P}}_{i,k}-2\hat{\boldsymbol{P}}'_{oi,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\boldsymbol{K}_{i,k}=&[2\varpi_1\xi_{i,k}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}\boldsymbol{A}+\boldsymbol{A}\boldsymbol{R}_{i,k}\boldsymbol{A}^{\mathrm{T}})]^{-1}\\
&\cdot\Bigg[2\boldsymbol{A}\boldsymbol{P}_{i,k}\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}+\varpi_2\xi_{i,k}\sum_{r\in N_i}(1-\theta_{ir})\boldsymbol{A}\\
&\quad\cdot(2(\boldsymbol{P}_{ri,k}^{o})^{\mathrm{T}}+\boldsymbol{P}_{ri,k}^{o}-2\boldsymbol{P}_{i,k}-\boldsymbol{P}_{ri,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\\
&\quad-\varpi_2\xi_{i,k}\sum_{r\in N_i}\theta_{ir}\boldsymbol{A}(2\hat{\boldsymbol{P}}_{i,k}-2\hat{\boldsymbol{P}}'_{oi,k})\boldsymbol{H}_{i,k}^{\mathrm{T}}\boldsymbol{C}_{i,k}\boldsymbol{A}^{\mathrm{T}}\Bigg].
\end{aligned}
$$