SCIENCE CHINA
Information Sciences

• LETTER •

November 2018, Vol. 61 119201:1–119201:3
https://doi.org/10.1007/s11432-017-9350-y

A distributed consensus filter for sensor networks with heavy-tailed measurement noise

Peng DONG*, Zhongliang JING*, Kai SHEN & Minzhe LI

School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, *China*

Received 25 October 2017/Revised 1 January 2018/Accepted 16 January 2018/Published online 2 August 2018

Citation Dong P, Jing Z L, Shen K, et al. A distributed consensus filter for sensor networks with heavy-tailed measurement noise. Sci China Inf Sci, 2018, 61(11): 119201, https://doi.org/10.1007/s11432-017-9350-y

Dear editor,

Distributed state estimation is very important in distributed sensor networks (DSNs) [1]. The consensus estimation can make the sensor networks achieve global consistency according to the data of all nodes [2]. It is very useful for the state estimation of DSNs. The fusion center and full connection between network nodes are not required. The information only exchanges between the neighboring nodes, which eliminates the need of local observability, and the stability of the state estimation can be guaranteed by the global observability. These characteristics lead to a simplified network topology, enhanced flexibility and robustness of the network structure. An effective approach to consensus estimation is consensus on information (CI) proposed in [3], and the stability is also proved. In addition, there are other distributed state estimation methods such as [4, 5]. However, measurement outliers with the heavy-tailed feature occur relatively often in practice and they may cause the divergence of estimates of states. The consideration of this problem is absent in the consensus approaches. Recently, some robust filters using the Student-t distribution and variational Bayesian (VB) method are proposed to deal with the heavy-tailed measurement noise [6–8].

We propose a novel distributed consensus filter to deal with the heavy-tailed measurement noise for sensor networks. We model the measurement noise of each sensor node as the multivariate Student-t distribution rather than the Gaussian distribution. The Student-t distribution can be constructed by a Gaussian variable and a Gamma variable with an auxiliary parameter. Since the sufficient statistic of the parameter and the state of each sensor node are coupled, an iterative solution of VB formula is applied to approximate the joint posterior density. Each iteration resembles the distributed CI filter. Simulation results demonstrate that the proposed method outperforms the conventional distributed CI filter.

Principles of the proposed method. A two-tuple $(\mathcal{N}, \mathcal{A})$ can be used to represent the DSN, where $\mathcal{N}$ means the set of sensor nodes, and $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of connections between nodes. For each node $i \in \mathcal{N}$, $\mathcal{N}^i$ is the set of its neighbors. That is $\mathcal{N}^i \triangleq \{j : (j, i) \in \mathcal{A}\}$. A linear dynamic system is considered here.

$$x_{k+1} = F_k x_k + w_k, \tag{1}$$

$$z_k^i = H_k^i x_k + r_k^i, \quad i \in \mathcal{N}, \tag{2}$$

where x_k is the n-dimension state vector, F_k is the state matrix, $w_k \sim N(0, Q_k)$ is the Gaussian process noise with zero mean and covariance Q_k, z_k^i is the d-dimension measurement vector of sensor node i, H_k^i is the measurement matrix, and r_k^i is the measurement noise. Let the information matrix $\Omega_k \triangleq P_k^{-1}$ and the information vector $q_k \triangleq P_k^{-1}\hat{x}_k$, where $\hat{x}$ denotes the estimation of the state, and P denotes the covariance of the

* Corresponding author (email: dongpengkty@sjtu.edu.cn, zljing@sjtu.edu.cn)

estimation error. In addition, the process noise information matrix is defined by $W_k \triangleq Q_k^{-1}$.

Unlike [7] that can only manage the estimate of a single sensor with heavy-tailed measurement noise, our derivation starts from the multi-sensor scenario and further extends to the distributed consensus filter. The proposed algorithm is derived step by step as follows.

Suppose that $Z_k^i = \{z_1^i, z_2^i, \ldots, z_k^i\}$ is the set of measurements of the node i till time k and $\mathcal{Z}_k = \{Z_k^i, i \in \mathcal{N}\}$ is the set of measurements of all sensor nodes till time k. We model the measurement noise as a Student-t distribution $z_k^i|x_k \sim \mathrm{St}(H_k^i x_k, (\Lambda_k^i)^{-1}, \nu^i)$, where Λ_k^i denotes the precision matrix and ν^i is degrees of freedom (DOF).

By introducing an auxiliary random variable λ_k^i, the distribution of the measurement can be determined by the following marginal density $p(z_k^i|x_k) = \int p(z_k^i|x_k, \lambda_k^i)\mathrm{d}\lambda_k^i$, where $z_k^i|x_k, \lambda_k^i \sim \mathrm{N}(H_k^i x_k, (\lambda_k^i \Lambda_k^i)^{-1})$ is a Gaussian distribution and $\lambda_k^i \sim G(\nu^i/2, \nu^i/2)$ is a Gamma distribution. The aim of Bayesian filter is to estimate joint posterior $p(x_k, \lambda_k^i|\mathcal{Z}_k)$, that is

$$p(x_k, \lambda_k^i|\mathcal{Z}_k) \propto (\lambda_k^i)^{\frac{d}{2}} \mathrm{e}^{-\frac{\lambda_k^i}{2}(z_k^i - H_k^i x_k)^{\mathrm{T}} \Lambda_k^i (z_k^i - H_k^i x_k)} \times \mathrm{e}^{(x_k - \hat{x}_{k|k-1})^{\mathrm{T}} \Omega_{k|k-1}(x_k - \hat{x}_{k|k-1})} (\lambda_k^i)^{\frac{d}{2}-1} \mathrm{e}^{-\frac{\nu^i \lambda_k^i}{2}}. \quad (3)$$

Then the state posterior can be recovered by marginalizing λ_k^i out of the joint posterior.

Because the density of x_k and λ_k^i are coupled, the VB approximation approach is utilized to make the computation of the joint posterior density tractable. Let $\lambda_k = \{\lambda_k^i, i \in \mathcal{N}\}$ and assume that x_k and λ_k are independent, then the joint posterior distribution can be approximated via the free-form VB approximation by $p(x_k, \lambda_k|\mathcal{Z}_k) \approx q(x_k)q(\lambda_k)$, where $q(x_k)$ and $q(\lambda_k)$ are approximating densities. The log marginal probability of $p(\mathcal{Z}_k)$ can be decomposed as

$$\ln p(\mathcal{Z}_k) = \mathrm{L}(q(x_k)q(\lambda_k)) + \mathrm{KL}[q(x_k)q(\lambda_k) \parallel p(x_k, \lambda_k|\mathcal{Z}_k)], \quad (4)$$

where $\mathrm{L}(\cdot)$ is a lower bound on the log likelihood function $\ln p(\mathcal{Z}_k)$ and the $\mathrm{KL}(\cdot)$ means Kullback-Leibler divergence (KLD). The definitions of the lower bound and the KLD can be seen in [6]. By maximizing lower bound, we can obtain the minimized KLD, which can result in an optimal solution. Suppose all the elements $\lambda_k^i (i \in \mathcal{N})$ in λ_k are independent and disjoint, thus we have $q(\lambda_k) = \prod_{i \in \mathcal{N}} q_i(\lambda_k^i)$. Then the joint posterior distribution of x_k and λ_k can be rewritten as

$$p(x_k, \lambda_k|\mathcal{Z}_k) \approx q(x_k) \prod_{i \in \mathcal{N}} q_i(\lambda_k^i). \quad (5)$$

Maximization will be reached by

$$\ln q(x_k) = \int \ln p(x_k, \lambda_k, \mathcal{Z}_k) \prod_{i \in \mathcal{N}} q_i(\lambda_k^i)\mathrm{d}\lambda_k + \mathrm{C} = E_\lambda[\ln p(x_k, \lambda_k, \mathcal{Z}_k)] + \mathrm{C}, \quad (6)$$

$$\ln q_i(\lambda_k^i) = \int \ln p(x_k, \lambda_k, \mathcal{Z}_k) q(x_k) \times \prod_{j \in \mathcal{N}, j \neq i} q_j(\lambda_k^j)\mathrm{d}x_k \mathrm{d}\lambda_k^j + \mathrm{C} = E_{x,\bar{\lambda}}[\ln p(x_k, \lambda_k, \mathcal{Z}_k)] + \mathrm{C}, \quad (7)$$

where C is a constant. Since the state and the parameter are inter-dependent, the iteration method should be applied to solve these equations. The solutions can be obtained via alternating between them until settling at a fixed point.

Given the expected $q(\lambda_k^i)$, $\ln q(x_k)$ can be given according to (6):

$$\ln q(x_k) = -\frac{1}{2}\sum_{i \in \mathcal{N}} \mathrm{tr}(\mathrm{E}[\lambda_k^i]\Lambda_k^i (z_k^i - H_k^i x_k)(z_k^i - H_k^i x_k)^{\mathrm{T}}) - \frac{1}{2}\mathrm{tr}(\Omega_{k|k-1}(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}) + \mathrm{C}. \quad (8)$$

Similarly, $\ln q(\lambda_k^i)$ can be determined according to (7):

$$\ln q(\lambda_k^i) = -\frac{1}{2}\lambda_k^i \bar{\gamma}_k^i + \left(\frac{\nu^i + d}{2} - 1\right)\ln \lambda_k^i - \frac{\nu^i \lambda_k^i}{2} + \mathrm{C}, \quad (9)$$

where

$$\bar{\gamma}_k^i = E_x[(z_k^i - H_k^i x_k)^{\mathrm{T}} \Lambda_k^i (z_k^i - H_k^i x_k)] = \mathrm{tr}(D_k^i \Lambda_k^i), \quad (10)$$

$$D_k^i = E_x[(z_k^i - H_k^i x_k)(z_k^i - H_k^i x_k)^{\mathrm{T}}] = H_k^i P_k^i (H_k^i)^{\mathrm{T}} + (z_k^i - \hat{z}_k^i)(z_k^i - \hat{z}_k^i)^{\mathrm{T}}. \quad (11)$$

The posterior distribution of λ_k^i is still a Gamma, thus $q(\lambda_k^i)$ is the density of the distribution $\mathrm{G}(a_k^i, b_k^i)$ with expectation $\bar{\lambda}_k^i = \mathrm{E}[\lambda_k^i] = \int \lambda_k^i q(\lambda_k^i)\mathrm{d}\lambda_k^i = \frac{a_k^i}{b_k^i}$, where $a_k^i = \frac{\nu^i + d}{2}$ and $b_k^i = \frac{\nu^i + \bar{\gamma}_k^i}{2}$.

We derive the above formulations of the state and parameter update in a centralized way for node i and it can be extended to distributed estimation in a consensus way directly. Since the posterior distribution of the state x_k^i is a Gaussian, $q(x_k^i)$ can be approximated by

$$\Omega_{k,0}^i = \Omega_{k|k-1}^i + \left(H_k^i\right)^{\mathrm{T}} \bar{L}_k^i H_k^i, \quad q_{k,0}^i = q_{k|k-1}^i + \left(H_k^i\right)^{\mathrm{T}} \bar{L}_k^i z_k^i, \quad (12)$$

$$q^i_{k,l} = \sum_{j\in\mathcal{N}^i} \pi^{i,j} q^j_{k,l-1}, \Omega^i_{k,l} = \sum_{j\in\mathcal{N}^i} \pi^{i,j} \Omega^j_{k,l-1}, \quad (13)$$

where $\bar{L}^i_k = \bar{\lambda}^i_k \Lambda^i_k$, $l = 0, 1, \ldots, L-1$, L is the consensus step, $\pi^{i,j}$ is the consensus weight with $\pi^{i,j} \geqslant 0$ and $\sum_{j\in\mathcal{N}^i} \pi^{i,j} = 1$. The prediction distribution can be approximated with the Gaussian density according to the prediction of the CI filter.

$$\Omega^i_{k+1|k} = W_k - W_k F_k \left(\Omega^i_k + F^{\mathrm{T}}_k W_k F_k\right)^{-1} F^{\mathrm{T}}_k W_k, \quad (14)$$

$$\hat{x}^i_{k+1|k} = F_k \hat{x}^i_k, q^i_{k+1|k} = \Omega^i_{k+1|k} \hat{x}^i_{k+1|k}. \quad (15)$$

The proposed distributed consensus filter named distributed variational Bayesian Student-t CI (DVBSCI) filter for DSNs with heavy-tailed measurement noise is summarized in Table A1.

Simulations. A two-dimension tracking scenario of 20 sensors (see Figure A1) is considered here. Configurations of this scenario are given in Appendix A. We compare the proposed DVBSCI filter with the distributed CI filter in [3] with the nominal R (DCI). Figure A2 shows the position errors of x-axis in a typical run for sensor node 1. It can be seen that the DCI filter is subject to outliers obviously. However, the proposed method is hardly affected by outliers. The averaged root mean square error (RMSE) of the estimated position and velocity obtained by 100 Monte Carlo runs are shown in Figures A3 and A4. It can be seen that the convergence speed of the DVBSCI filter is faster than that of the DCI filter and the RMSE of the DVBSCI filter is significantly smaller than that of the DCI filter. Therefore, with merely the nominal R at hand, the proposed DVBSCI filter performs significantly better than the DCI filter. The RMSE of position and velocity with different p_o are given in Table A2. Observe that when $p_o = 0$, which means that the measurement noise is the standard Gaussian with none heavy tails, performance of the DVBSCI filter and the DCI filter are almost the same. When $p_o > 0$, the DVBSCI filter always outperforms the DCI filter. It means the proposed method has a certain adaptive ability to adjust different heavy-tailed noises.

Conclusion. This article presents a novel consensus filter for distributed state estimation with the heavy-tailed measurement noise in DSNs. The Student-t distribution of the measurement noise is used to deal with the heavy-tailed noise. The variational Bayesian is applied to iteratively estimate the joint density of the state and the parameter under the framework of the consensus filter. The consistency on information can be obtained in the presence of the heavy-tailed measurement noise. Simulation results verify that the proposed method can perform better than the traditional DCI filter. The merits of this method are that the computational burden is small, and there is a certain adaptability to the noise. However, it cannot deal with situations where there are outliers on both the state and measurement noises. In the future, we will extend this framework to combine other robust Student-t based estimation methods with different consensus strategies.

Acknowledgements This work was jointly supported by National Natural Science Foundation of China (Grant Nos. 61673262, 61175028), Major Program of National Natural Science Foundation of China (Grant Nos. 61690210, 61690212), and Shanghai Key Project of Basic Research (Grant No. 16JC1401100).

Supporting information Appendix A providing configurations of the simulation scenario, Tables A1, A2 and Figures A1–A4. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Wang Y H, Lin P, Hong Y G. Distributed regression estimation with incomplete data in multi-agent networks. Sci China Inf Sci, 2018, 61: 092202

2 Shen K, Jing Z, Dong P. A consensus nonlinear filter with measurement uncertainty in distributed sensor networks. IEEE Signal Process Lett, 2017, 24: 1631–1635

3 Battistelli G, Chisci L. Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. Automatica, 2014, 50: 707–718

4 Liu Q, Wang Z, He X, et al. Event-based recursive distributed filtering over wireless sensor networks. IEEE Trans Automat Contr, 2015, 60: 2470–2475

5 Liu Q, Wang Z, He X, et al. A resilient approach to distributed filter design for time-varying systems under stochastic nonlinearities and sensor degradation. IEEE Trans Signal Process, 2017, 65: 1300–1309

6 Agamennoni G, Nieto J I, Nebot E M. Approximate inference in state-space models with heavy-tailed noise. IEEE Trans Signal Process, 2012, 60: 5024–5037

7 Piché R, Särkkä S, Hartikainen J. Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate Student-t distribution. In: Prcoeedings of IEEE International Workshop on Machine Learning for Signal Processing, Santander, 2012. 1–6

8 Nurminen H, Ardeshiri T, Piche R, et al. Robust inference for state-space models with skewed measurement noise. IEEE Signal Process Lett, 2015, 22: 1898–1902