# Attack-Resilient Map-based Localization

Yulin Yang and Guoquan Huang

## I. INTRODUCTION

Due to increasing proliferation of autonomous vehicles, securing robot navigation against malicious attacks becomes a matter of urgent societal interest, because attackers can fool these vehicles by manipulating their sensors, exposing us to unprecedented vulnerabilities and ever-increasing possibilities for malicious attacks. To address this issue, we seek to secure state estimation for *stochastic nonlinear* systems with the particular application to map-based localization. In particular, based on the MCC-KF [1] [2], we first perform in-depth analysis of the maximum correntropy criterion (MCC)-based **E**KF. Then, we analytically derive the weighted MCC-EKF (WMCC-EKF) that shows to improve accuracy and robustness to unbounded attacks as compared to the state-of-the-art methods. Different with [3], the proposed WMCC-EKF is derived for nonlinear measurement model and the weights are determined partially according to the known noise level. Furthermore, as a conservative solution, we generalize the secure estimation algorithm [4] to nonlinear systems and develop the Secure Estimation (SE)-EKF that integrates the attack detection based on $\ell_0(\ell_1)$-optimization within a sliding-window filtering framework. The proposed secure EKFs are validated through both Monte-Carlo simulations and experiments on real datasets.

## II. PROBLEM STATEMENT

Consider a nonlinear system with measurements possibly attacked by adversaries:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{w}_k) \tag{1}$$
$$\mathbf{y}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{n}_{k+1} + \mathbf{a}_{k+1} \tag{2}$$
$$\mathbf{z}_{k+1} = \mathbf{y}_{k+1} - \mathbf{a}_{k+1} = \mathbf{h}(\mathbf{x}_{k+1}) + \mathbf{n}_{k+1} \tag{3}$$

where $\mathbf{x}_k \in \mathbb{R}^{m\times 1}$ represents the system states at the time step $k$, $\mathbf{f}$ represents the system dynamic model and $\mathbf{w}$ is the input white Gaussian noise with covariance $\mathbf{Q}$. $\mathbf{y} \in \mathbb{R}^{p\times 1}$ denotes the measurements from $p$ sensors, $\mathbf{h}$ represents the nonlinear measurement model function. $\mathbf{a} \in \mathbb{R}^{p\times 1}$ denotes the attack signals and is assumed to be sparse vector that at least one sensor cannot be attacked. We also define $\mathbf{z} \in \mathbb{R}^{p\times 1}$ as the un-attacked output. $\mathbf{n} \in \mathbb{R}^{p\times 1}$ represents zero-mean Gaussian white noises with covariance $\mathbf{R} = \mathbf{diag}\{\sigma_1^2 \dots \sigma_i^2 \dots \sigma_p^2\}$, where $\sigma_i, i = 1 \dots p$ represents the $i$-th sensor's noise variance and $\mathbf{diag}\{\cdot\}$ is the diagonal matrix form. If the $\mathbf{R}$ is a full (not diagonal or block diagonal) matrix, a noise pre-whitening operation (see [5]) can be performed to transform $\mathbf{R}$ into diagonal form. The corresponding linearized system can be computed as follows:

$$\tilde{\mathbf{x}}_{k+1} \simeq \mathbf{F}_k \tilde{\mathbf{x}}_k + \mathbf{G}_k \mathbf{w}_k \tag{4}$$
$$\tilde{\mathbf{y}}_{k+1} \simeq \mathbf{H}_{k+1} \tilde{\mathbf{x}}_{k+1} + \mathbf{n}_{k+1} + \mathbf{a}_{k+1} \tag{5}$$
$$\tilde{\mathbf{z}}_{k+1} \simeq \mathbf{H}_{k+1} \tilde{\mathbf{x}}_{k+1} + \mathbf{n}_{k+1} \tag{6}$$

where $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$ denotes the error states, the $\mathbf{F}_k$ and $\mathbf{G}_k$ represent the Jacobians regarding to the state $\mathbf{x}_k$ and the noise $\mathbf{w}_k$ respectively. $\tilde{\mathbf{y}}$ denotes the measurement residual, while $\tilde{\mathbf{z}}$ describes the un-attacked measurement residual. $\mathbf{H}_{k+1}$ represents the measurement Jacobian with respect to the state $\mathbf{x}_{k+1}$.

It is important to note that, instead of assuming a fixed set of attacked sensors [6], [7], we consider that the attacker can attack different sensors randomly at different time steps. Note also that as compared to [4], instead of assuming that less than a half of the sensors can be attacked, we only assume that at least one bearing or range sensor is not attacked.

## III. WEIGHTED MAXIMUM CORRENTROPY CRITERION (WMCC)-BASED FILTERS

Based on maximum Correntropy criterion [1], [2], given the initial state in the form of Gaussian distribution, $\mathcal{N}(\hat{\mathbf{x}}_{0|0}, \mathbf{P}_0)$, the propagation of WMCC-EKF filter can be derived as:

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{f}(\hat{\mathbf{x}}_{k|k}, \mathbf{0}) \tag{7}$$
$$\mathbf{P}_{k+1|k} = \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^\top + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^\top \tag{8}$$

Then, EKF-like update can be written as:

$$\hat{\mathbf{y}}_{k+1|k} = \hat{\mathbf{z}}_{k+1|k} = \mathbf{h}(\hat{\mathbf{x}}_{k+1|k}) \tag{9}$$
$$d_{i,k+1} = \frac{\mathbf{G}_{\hat{\sigma}_{i,k+1}}\left(\left\|\mathbf{y}_{i,k+1} - \mathbf{h}_{i,k+1}(\mathbf{x}_{k+1})\right\|\right)}{\mathbf{G}_{\hat{\sigma}_{0,k+1}}\left(\left\|\mathbf{x}_{k+1} - \mathbf{f}(\hat{\mathbf{x}}_{k|k}, \mathbf{0})\right\|_{\mathbf{P}_{k+1|k}^{-1}}\right)} \tag{10}$$
$$\mathbf{D}_{k+1} = \mathbf{diag}\{d_{1,k+1}, \dots, d_{i,k+1}, \dots, d_{p,k+1}\} \tag{11}$$
$$\mathbf{K}_{k+1|k} = \left[\mathbf{H}_{k+1}^\top \mathbf{D}_{k+1} \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1} + \mathbf{P}_{k+1|k}^{-1}\right]^{-1} \mathbf{H}_{k+1}^\top \mathbf{D}_{k+1} \mathbf{R}_{k+1}^{-1} \tag{12}$$
$$= \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^\top \left(\mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^\top + \mathbf{R}_{k+1} \mathbf{D}_{k+1}^{-1}\right)^{-1} \tag{13}$$
$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1|k}\left(\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1|k}\right) \tag{14}$$
$$\mathbf{P}_{k+1|k+1} = \left[\mathbf{H}_{k+1}^\top \mathbf{D}_{k+1} \mathbf{R}_{k+1}^{-1} \mathbf{H}_{k+1} + \mathbf{P}_{k+1|k}^{-1}\right]^{-1} \tag{15}$$

Note that $\mathbf{D}_{k+1}$ is a diagonal matrix computed based on the current measurements, which serves as a weight matrix for the measurement information. Hence, we will inspect WMCC-EKF from an information perspective. Compared to the MCC-EKF [2], the information matrix for the WMCC-

The authors are with the Department of Mechanical Engineering, University of Delaware, Newark, DE 19716, USA. Email: {yuyang, ghuang}@udel.edu

EKF can be written as:

$$\mathbf{P}_{k+1|k+1}^{-1} = \mathbf{P}_{k+1|k}^{-1} + \mathbf{H}_{k+1}^{\top}(\mathbf{D}_{k+1}\mathbf{R}_{k+1}^{-1})\mathbf{H}_{k+1} \tag{16}$$

$$= \underbrace{\mathbf{P}_{k+1|k}^{-1}}_{\Sigma_{w1}} + \underbrace{\sum_{i=1}^{p} d_{i,k+1}\frac{\mathbf{H}_{i,k+1}^{\top}\mathbf{H}_{i,k+1}}{\sigma_{i,k+1}^{2}}}_{\Sigma_{w2}} \tag{17}$$

where $\Sigma_{w1}$ and $\Sigma_{w2}$ denote the information from motion model (1) and the measurement model (2), respectively. Note that $d_{i,k+1}\frac{\mathbf{H}_{i,k+1}^{\top}\mathbf{H}_{i,k+1}}{\sigma_{i,k+1}^{2}}$ represents the information contribution from the $i$-th sensor's measurement, and thus, $\Sigma_{w2}$ in (17) can be seen as the sum of single information matrix from all the $p$ sensors. If the $i$-th sensor is attacked, $d_{i,k+1}$ will decrease exponentially and the corresponding information contribution $d_{i,k+1}\frac{\mathbf{H}_{i,k+1}^{\top}\mathbf{H}_{i,k+1}}{\sigma_{i,k+1}^{2}}$ will be dramatically reduced. However, this process will not affect the information contribution from other sensors. Therefore, the WMCC-EKF is able to utilize the information from un-attacked sensor measurements.

## IV. Secure Estimation (SE)-EKF

Ideally, we would like to identify the attacked measurements so that we can ensure estimation security by excluding them from the EKF update. To this end, we introduce the Secure-estimation (SE)-EKF by generalizing the SE-KF [4] to the nonlinear system under consideration. We define the state vector with window size $N$ at time step $k$ as:

$$\mathbf{x}_{c_k} = \begin{bmatrix}\mathbf{x}_k^{\top} & \mathbf{x}_{k-1}^{\top} & \cdots & \mathbf{x}_{k-N+1}^{\top} & \mathbf{x}_{k-N}^{\top}\end{bmatrix}^{\top} \tag{18}$$

where $\mathbf{x}_k$ represents the current robot state, $\mathbf{x}_{k-i}$ represents the cloned robot state at time step $k-i, i \in \{1 \ldots N\}$. Thus, $\mathbf{x}_{k-N}$ is the oldest cloned state. Let $\mathbf{F}_{k-1,k-N}$ represent the state transition matrix from cloned state $\tilde{\mathbf{x}}_{k-N}$ to the current robot state $\tilde{\mathbf{x}}_k$. Thus, stacked the observation can be written regarding to the first state in the window as:

$$\underbrace{\begin{bmatrix}\tilde{\mathbf{z}}_k \\ \tilde{\mathbf{z}}_{k-1} \\ \vdots \\ \tilde{\mathbf{z}}_{k-N}\end{bmatrix}}_{\tilde{\mathbf{Z}}} \simeq \underbrace{\begin{bmatrix}\mathbf{H}_0 \\ \mathbf{H}_{k-1} \\ \vdots \\ \mathbf{H}_{k-N}\end{bmatrix}\begin{bmatrix}\mathbf{F}_{k,k-N} \\ \mathbf{F}_{k-1,k-N} \\ \vdots \\ \mathbf{I}\end{bmatrix}}_{\Phi}\tilde{\mathbf{x}}_{k-N} + \underbrace{\begin{bmatrix}\mathbf{n}_0 \\ \mathbf{n}_{k-1} \\ \vdots \\ \mathbf{n}_{k-N}\end{bmatrix} + \begin{bmatrix}\mathbf{a}_0 \\ \mathbf{a}_{k-1} \\ \vdots \\ \mathbf{a}_{k-N}\end{bmatrix}}_{\mathbf{E}} \tag{19}$$

where $\tilde{\mathbf{Z}}$ represents the stacked measurement residuals, and $\mathbf{E}$ denotes the sum of stacked noise and attack vectors, $\Phi$ denotes the stacked state transition matrix from $\tilde{\mathbf{x}}_{k-N}$ to each state in the window. Similar to [4] we apply left null space operation to $\Phi$ to simplify (19). Let $\mathbf{U}_n$ be the left null space of $\Phi$, that is $\mathbf{U}_n^{\top}\Phi = \mathbf{0}$, then we can have:

$$\mathbf{Z}_o = \mathbf{U}_n^{\top}\mathbf{Z} = \mathbf{U}_n^{\top}\mathbf{E} \tag{20}$$

where $\mathbf{U}_n$ can be computed from the $QR$ decomposition of $\Phi$. Given the strong sparse attack assumption that less than a half of the all the sensors can be attacked, $\mathbf{E}$ can be solved by formulating the following optimization problem with $\ell_1$ norm regularization as:

$$\hat{\mathbf{E}} = \arg\min_{\mathbf{E}}\left[\left\|\mathbf{Z}_o - \mathbf{U}_n^{\top}\mathbf{E}\right\|_2^2 + \lambda\|\mathbf{E}\|_{\ell_1}\right] \tag{21}$$

where $\lambda$ is the regularization parameter.

With the attack identification, the SE-EKF algorithm will be able to remove the attacked measurements and perform the state update only with un-attacked measurements [5].

## V. Experimental Results

We have extensively validated the proposed proposed WMCC-EKF and SE-EKF in map-based localization through both 50 Monte-Carlo simulations and real experiments using the Victoria Park dataset.

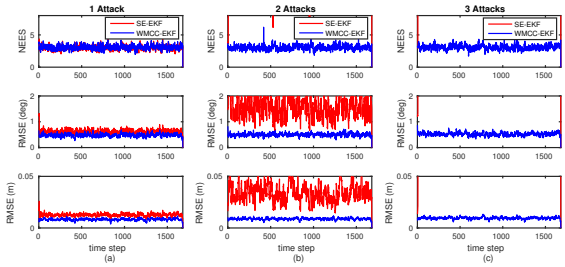


Fig. 1. Monte Carlo results of the proposed WMCC-EKF and SE-EKF.

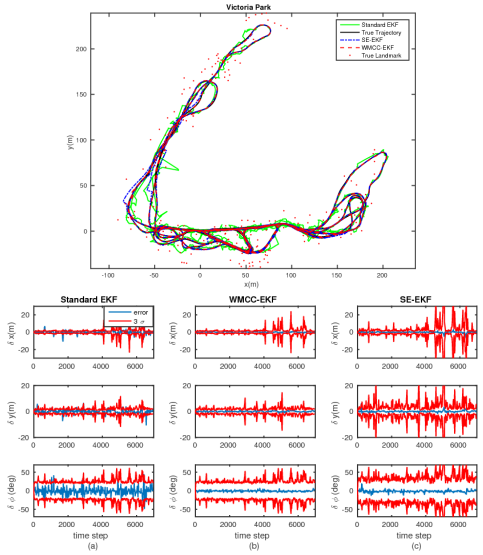


Fig. 2. Estimated trajectories of the WMCC-EKF, SE-EKF and the Standard EKF with synthetic attacks on the Victoria Park dataset.

## References


[1] R. Izanloo, S. A. Fakoorian, H. S. Yazdi, and D. Simon, "Kalman filtering based on the maximum correntropy criterion in the presence of non-gaussian noise," in *Conference on Information Science and Systems (CISS)*, March 2016, pp. 500–505.

[2] M. Kulikova, "Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-gaussian noise," *arXiv preprint arXiv:1610.00257*, 2016.

[3] X. Liu, H. Qu, J. Zhao, and B. Chen, "Extended kalman filter under maximum correntropy criterion," in *Inter. Joint Conf. on Neural Networks*, July 2016, pp. 1733–1737.

[4] Q. Hu, Y. H. Chang, and C. J. Tomlin, "Secure estimation for unmanned aerial vehicles against adversarial cyber attacks," *arXiv preprint arXiv:1606.04176*, 2016.

[5] Y. Yang and G. Huang, "Map-based localization under adversarial attacks," in *Proc. of the International Symposium on Robotics Research*, Puerto Varas, Chile, Dec.11-14, 2017.

[6] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, Jun. 2014.

[7] M. Pajic, P. Tabuada, I. Lee, and G. J. Pappas, "Attack-resilient state estimation in the presence of noise," in *Conference on Decision and Control*. IEEE, 2015, pp. 5827–5832.