2016

Resilient dynamic state estimation in the presence of false information injection attacks

Jingyang Lu

RESILIENT DYNAMIC STATE ESTIMATION IN THE PRESENCE OF FALSE INFORMATION INJECTION ATTACKS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

JINGYANG LU

Bachelor of Science, Harbin Institute of Technology - August 2008 to July 2012

Advisor: Ruixin Niu,

Assistant Professor, Department of Electrical and Computer Engineering

Virginia Commonwealth University

Richmond, Virginia

December, 2016

Acknowledgments

First, I would like to express my sincere gratitude to my advisor Dr. Ruixin Niu for his continuous support of my Ph.D. study and related research, and for his patience, motivation, and immense knowledge. His guidance helped me in all the time of my Ph.D. research and writing of this dissertation. I cannot imagine having a better advisor and mentor for my Ph.D. study.

I would also like to thank my committee members, Dr. Alen Docef, Dr. Carl R Elks, Dr. Vojislav Kecman, and Dr. Hong-sheng Zhou for serving as my committee members. I also want to thank them for letting my defense be an enjoyable moment, and for their brilliant comments and suggestions.

I would like to thank my parents. Even though I have not seen them for four and half years, the daily calls make me feel they are by my side supporting me all the time. They are always the first people I will turn to when I am faced with difficulties. They are always encouraging me and telling me not to lose heart and to be brave. I really appreciate everything they have done for me.

I would also like to thank my lab mates, Armond Conte, Mengqi Ren, and Puxiao Han. I still remember the scenes we discussed difficult problems for hours and hours and came up with excellent ideas in the end.

I would also like to thank VCU for providing the excellent environment for me to conduct the research. I can always find a good studying spot in the library. The facilities in the gym on the Cary Street are awesome. The gym is my favorite, and I can always relax myself whenever I feel stressful.

In the end, I would express my best love to my fiancée Hebing Liu. She has been supporting me and taking care of me since we met each other in 2014. We are

both studying in Ph.D. programs in different departments at VCU. Even though her research topic is totally different from mine, she tries her best to help me modify the dissertation, go through countless rounds of dissertation defense rehearsals, and prepare refreshments for the final defense. I do really appreciate her love and kindness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

RESILIENT DYNAMIC STATE ESTIMATION IN THE PRESENCE OF FALSE INFORMATION INJECTION ATTACKS

By Jingyang Lu

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2016.

Advisor: Ruixin Niu,

Assistant Professor, Department of Electrical and Computer Engineering

In this dissertation, the problem of resilient dynamic system state estimation in the presence of false information injection attacks is investigated. First, it is assumed that the system is unaware of the existence of false information and the adversary tries to maximize the negative effect of the false information on Kalman filter's estimation performance under a power constraint. The false information attack under different conditions is mathematically characterized. For the adversary, many closed-form results for the optimal attack strategies that maximize Kalman filter's estimation error are theoretically derived. It is shown that by choosing the optimal correlation coefficients among the false information and allocating power optimally among sensors, the adversary could significantly increase Kalman filter's estimation errors.

In order to detect the false information injected by an adversary, we investigate the strategies for the Bayesian estimator to detect the false information and defend itself from such attacks. We assume that the adversary attacks the system with certain

probability, and that he/she adopts the worst possible strategy that maximizes the mean squared error (MSE) if the attack is undetected. An optimal Bayesian detector is designed which minimizes the average system estimation error instead of minimizing the probability of detection error, as a conventional Bayesian detector typically does.

The case where the adversary attacks the system continuously is also studied. In this case, sparse attack strategies in multi-sensor dynamic systems are investigated from the adversary's point of view. It is assumed that the defender can perfectly detect and remove the sensors once they are corrupted by false information injected by an adversary. The adversary's goal is to maximize the covariance matrix of the system state estimate by the end of the attack period under the constraint that the adversary can only attack the system a few times over the sensors and over the time, which leads to an integer programming problem. In order to overcome the prohibitive complexity of the exhaustive search, polynomial-time algorithms, such as greedy search and dynamic programming, are proposed to find the suboptimal attack strategies. As for greedy search, it starts with an empty set and one sensor is added at each iteration, whose elimination will lead to the maximum system estimation error. The process terminates when the cardinality of the active set reaches the sparsity constraint. Greedy search based approaches such as sequential forward selection (SFS), sequential backward selection (SBS), and simplex improved sequential forward selection (SFS-SS) are discussed and corresponding attack strategies are provided. Dynamic programming is also used in obtaining a sub-optimal attack strategy. The validity of dynamic programming lies on a straightforward but important nature of dynamic state estimation systems: the credibility of the state estimate at current step is in accordance with that at previous step.

The problem of false information attack on and Kalman filter's defense of state estimation in dynamic multi-sensor systems is also investigated from a game theo-

retic perspective. The relationship between Kalman filter and the adversary can be regarded as a two-person zero-sum game. The condition under which both sides of the game will reach a Nash equilibrium is investigated.

CHAPTER 1

INTRODUCTION

1.1 Motivation and Background

System state estimation aims at monitoring the system state and providing accurate information for the controller to make reliable actuation decisions for the system. For example, the control center of a power system conducts system state estimation to distribute the power to different regions properly [1]. Self-driving vehicles have drawn a lot of attentions which involve the vehicle state estimation [2]. Medical diagnosis concerns the determination of the true physiological state of the patient by gathering the test measurements. The system becomes more vulnerable to attacks as it gets more complicated and the adversary finds more ways to access it. For example, according to an inspector general's report sent to the Federal Aviation Administration (FAA) in 2009, hackers have broken into air traffic control mission-support systems several times in recent years [3]. Some hackers were also able to hack wireless medical devices implanted in human bodies [4].

As for the electric power system, it consists of apparatus, generators, electrical transformers, and lines that can be damaged or destroyed as a result of short circuits, thermal overload, weather, and even physical attacks. For example, in 2013, Pacific Gas and Electric Company's Metcalf Transmission Substation in San Jose, California was broken by gunmen who fired on 17 electrical transformers resulting in over $15 million worth of damage. Certain detection mechanisms and corresponding defending strategies are urgently needed in order to detect these types of abnormal conditions and attacks and protect the system to ensure the safe and reliable operation of the whole electric power system. A protection system must be dependable and secure in all its operations. The protection devices should properly respond when an abnormal or dangerous condition is indicated.

An increasing demand for reliable energy has motivated the development of smart electric grid. The U.S. Department of Energy (DOE) has identified seven properties required for the smart grid to meet future demands including attack resistance, self-healing, consumer motivation, power quality, generation and storage accommodation, enabling markets, and asset optimization [5, 6]. The smart grid is applied for sensor data reading and system control in two-way communications. The development of a trustworthy smart grid system depends on a deeper understanding of the potential effects of false information. A comprehensive approach to understand the security of the system is to appropriately quantify the effect of the false information injection attack. Studying the relation between the false information attack and the physical system effect and designing the countermeasures to mitigate risks from the attack will help increase the robustness of the smart grid system.

System state estimation of the power system is a key function in building real-time models of electricity networks in the energy management centers (EMC). A real-time model usually utilizes the data every few seconds from energy control center to conduct the system state estimation. It is not practical and economical to measure all the possible states in the network. System state estimation is a useful tool for estimating the system state by using a limit set of measurements. Two kinds of measurement information - analog and digital data are usually used in system state estimation [7]. The control center of the system can take use of the measurement to estimate the system state and make certain control decisions. Anomaly detection turns to be essential when abnormal conditions like topology error or false information injection by malicious attacks occur. Without knowing the existence of the false information, the system state estimation will mislead the control center in making decisions. The traditional detector such as Chi-square detector works by comparing the residue between the measurement and its prediction with a threshold. However, this detector cannot detect the false information when the adversary has knowledge of the system configuration and launch a carefully designed attack [8]. Therefore it is very important to design the detection mechanisms and defending strategies to avoid the case where false information is injected into the system incurring large system state estimation error.

System state estimation is of importance in the area of driverless cars development. Autonomous navigation is one of the most important technologies for driverless cars. Accurate system state estimation is generally the basis of any other functions such as path planning and environment perception. An accurate system state estimate ensures the safety of a driverless car. The control center of a driverless car conducts system state estimation based on the Global Positioning System (GPS). An enhanced differential GPS receiver with phase carrier signal measurements may run in operating modes of real time kinematics, which has the highest absolute position accuracy. In addition to the driverless cars, system state estimation and anomaly detection also play a key role in the development of Unmanned Aerial Vehicles (UAVs). The theoretical models sometimes may not work because the natural environment is very complicated, and many abnormal conditions may happen. The control center of the system has to be able to detect anomalies and improve the performance of system state estimation. Different types of attacks may be launched by an adversary to the UAV. Hardware attacks can happen when the adversary has direct access to the UAV's autopilot components. An adversary can corrupt the data stored on the board or add extra data to mislead the whole control system. Wireless attacks can also happen when the adversary has access to the communication channel so that they can change the data stored on-board in real time. In order to overcome this, a more accurate and resilient state estimation system needs to be designed.

For the radar system, the false information is usually caused by jammers, which can apply various techniques of misleading a radar system. It could be either mechanical or electronic [9]. An electronic jammer misleads the radar system by injecting jamming signals through the communication channels. The distributed MIMO radar system consists of multiple of transmitters, receivers, and a fusion center where the final system state estimate is made [10]. Even though it has been shown that the distributed MIMO radar system can provide better performance than the traditional radar system, it increases the vulnerability of the system itself as well. If the distributed MIMO radar system is built under the nominal condition that there are no false information attacks, system state estimation is significantly affected even under a low-level attack.

A lot of techniques have been utilized in developing system state estimation. Reinforcement learning, a machine learning approach, is concerned with how the control center makes corresponding actions by optimizing the cumulative reward. It has been heavily used in advertising, robot design, deriving complex hierarchical schemes, and learning non-ambiguous models. Reinforcement learning can be applied to cases where a model of the environment is known, but an analytic solution is not available. The way to get the information about the environment is by interacting with it. Reinforcement learning uses samples to optimize the performance and uses function approximation to deal with large environments. The reinforcement problems are specified by a Markov Decision Process, which in some cases can be shown to be equivalent to a shortest-path problem.

Another popular state estimation approach is Kalman filter, which uses series of sensor measurements overtime to conduct system state estimation in the presence of random noise. At each recursion, the algorithm works in two steps: prediction and update. In the prediction step, Kalman filter makes a prediction of the current state. In the update step, the current state estimate is updated by the residue between the measurement and its prediction. Kalman filter estimates the system state by recursively conducting Bayesian estimation. Kalman filter has been applied in wide and diverse areas. Kalman filter is widely used in robotic motion planning and control. It also works for characterizing the human's central nervous system's control of movement, and supports the realistic model by making system state estimation and issuing the updated commands [11].

System state estimation in the presence of an adversary that injects false information into sensor readings has attracted much attention in wide application areas, such as target tracking with compromised sensors, secure monitoring of dynamic electric power systems, and radar tracking and detection in the presence of jammers. This topic has been studied in [8, 12, 13, 14, 15, 16, 17, 18, 19, 20]. In [8], the problem of taking advantage of the power system configuration to introduce arbitrary bias to the system without being detected was investigated and inspired many researchers to further study false information injection along this direction. In [12] the impact of malicious attacks on real-time electricity market concerning the locational marginal price was investigated and how the attackers can make profit by manipulating certain values of the measurements was shown. Some strategies are also provided to find the optimal single attack vector. The relationship between the attackers and the control center was discussed in [13], where both the adversary's attack strategies and the control center's detection algorithms have been proposed. Readers are referred to [14] and [15] for more about false information attacks on the electricity market. Inspired by [8], in [16] it was shown that the data frame attack can be formulated as a quadratically constrained quadratic program (QCQP) problem, in which deleting the comprised sensors which the defender system detects will make the system unobservable. In [17], the relation between a target and an MIMO radar was characterized as a two-person zero-sum game. However, in the aforementioned publications, only the problem of *static* system state estimation has been considered.

For a linear *dynamic* system, the impact of the injected false information on Kalman filter's state estimation performance over time has not got much attention in the literature. In many problems with multiple target information variables [21], one is interested in the mean squared error (MSE) matrix of the state estimate. As the defender, the object is minimizing the system state estimation MSE matrix, i.e. to achieve the smallest system estimation error. Here we introduce several measures of the system state estimation MSE matrix and describe their physical meanings:

- The trace of the state estimation MSE matrix, that is the summation of the diagonal entries, is mostly used to evaluate the performance for numerous estimation tasks. The trace captures the total expected squared error in estimation problems.

- The determinant of the state estimation MSE matrix, is also used in the system state estimation problems, which captures the volume of the error ellipsoid around the true state value. It also measures the mutual information between the unknown state and observations in estimation problems.

- The MSE matrix itself can also be used in formulating the objective function. From an adversary's point of view, the optimal attack strategies under certain constraints such as power constraints would be the optimal solutions leading the state estimation MSE matrix to be the largest positive semidefinite matrix. This guarantees the optimality in terms of trace or determinant of the MSE matrix.

Sensor management plays an important role in increasing the resilience of the system state estimation. Some related publications exist on sensor management [22, 23], where the problem of arranging the sensors to minimize the covariance of the state estimation error so that a more accurate state estimate can be obtained, was investigated. In [24], the problem of sensor bias estimation and compensation for target tracking has been addressed. Interested readers are referred to [24] and the references therein for details.

The impact of the injected biases on a Kalman filter's estimation performance was presented in [18], showing that if the false information is injected at a single time, its impact converges to zero as time goes on; if the false information is injected into the system continuously, the estimation error tends to reach a steady state.

1.2 New Contributions

Based on [18], we have obtained some results regarding optimal false information attacks. In [25], we have found that the best strategies for the adversary to attack Kalman filter system from the perspective of the trace of the MSE matrix, and obtained some closed-form results. In [26], a closed-form optimal attack strategy was found for the adversary, which maximizes the impact of the false information injection on Kalman filter's state estimation from the determinant perspective. By adopting the objective function as the determinant of the MSE matrix, we change the problem significantly. The optimal attack strategy that maximizes the determinant of the MSE matrix is a function of Kalman filter's state estimation covariance and hence "adaptive" to Kalman filter; whereas the optimal solution that maximizes the trace of the MSE matrix is not a function of Kalman filter's state estimation covariance.

In this dissertation, we also investigate the detection of false information injection attacks [27]. More particularly, our goal is to design the optimal Bayesian detector minimizing the average system estimation error. For a Bayesian estimator whose sensors could be attacked by false information injected by an adversary, we investigate the strategies for the Bayesian estimator to detect the false information and defend itself from such attacks. We assume that the adversary attacks the system with certain probability, and that he/she adopts the worst possible strategy which maximizes the MSE if the attack is undetected. The defender's goal is to minimize the average system estimation MSE instead of minimizing the probability of error, as a conventional Bayesian detector typically does. The cost functions are based on the traces of the MSE matrices of the estimation error. Numerical results show that the new detection-estimation structure outperforms that based on the traditional detectors such as the conventional Bayesian detector and the chi-square detector significantly in terms of the average MSE. One proposed detection-estimation strategy, discarding sensor data when the presence of attack is declared, is very robust even when the attacker uses an attack strategy significantly different from the one assumed by the defender.

There are still a lot of problems left to be solved. In [28, 29], the optimal attack strategies are studied when the adversary aims to maximize the state estimation MSE matrix of the system state estimate by the end of the attack period under the constraint that the adversary can only attack the system a few times over time and over sensors and the defender has the perfect detection mechanism, which leads to an integer programming problem. The exhaustive-search is intractable even when the size of problem increases moderately. Greedy search based approaches such as sequential forward selection (SFS), sequential backward selection (SBS), and simplex improved sequential forward selection(SFS-SS) have been discussed in the dissertation and corresponding attack strategies are provided. Considering the credibility of the current estimate is in accordance with that of the previous estimate in dynamic state estimation systems, dynamic programming (DP) is also used, which helps reduce the time complexity by memorizing the internal results during the process to obtain a suboptimal attack strategy.

As for the case where the defender knows the existence of the false information, game theory is utilized to find the Nash Equilibrium between the defender and adversary [30]. The relationship between Kalman filter and the adversary can be regarded as a two-person zero-sum game. Under which condition both sides of the game will reach a Nash equilibrium is investigated. The multi-sensor Kalman filter system and the adversary are supposed to be rational players. Kalman filter and the adversary have to choose their respective subsets of sensors to perform system state estimation and false information injection. It is shown how both sides pick their strategies in order to gain more and lose less.

1.3 Dissertation Outline

The rest of dissertation is organized as follows. In Chapter 2, the discrete-time linear dynamic system and Kalman filter system are introduced. The optimal attack strategies which the adversary can adopt under a power constraint are investigated and studied. An optimal Bayesian detector which minimizes the average system estimation MSE is designed in Chapter 3. In Chapter 4, the sparse attack strategies are analyzed under the assumption that Kalman filter has the perfect detection of such attacks. The adversary aims to maximize the covariance matrix of the system state estimate by the end of the attack period with the sparsity constraint. The relation between the defender and the adversary is characterized and studied using game theory in Chapter 5. Conclusion is drawn in Chapter 6.

CHAPTER 2

ATTACK STRATEGY ANALYSIS

In this chapter, Kalman filter system is presented and the impact of false information injection is investigated for linear dynamic systems with multiple sensors. It is assumed that the system is unsuspecting the existence of false information and the adversary is trying to maximize the negative effect of the false information on Kalman filter's estimation performance. The false information attack under different conditions is mathematically characterized. For the adversary, many closed-form results for the optimal attack strategies that maximize Kalman filter's estimation error are theoretically derived. It is shown that by choosing the optimal correlation coefficients among the bias noises and allocating power optimally among sensors, the adversary could significantly increase Kalman filter's estimation errors.

2.1 Kalman Filter System

2.1.1 Linear Dynamic State Estimation

The discrete-time linear dynamic system [31] can be described as below,

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{u}_k + \mathbf{v}_k \tag{2.1}$$

where $\mathbf{F}_k$ is the system state transition matrix, $\mathbf{x}_k$ is the system state vector at time k, $\mathbf{u}_k$ is a known input vector, $\mathbf{G}_k$ is the input gain matrix, and $\mathbf{v}_k$ is a zero-mean white Gaussian process noise with covariance matrix $E[\mathbf{v}_k\mathbf{v}_k^T] = \mathbf{Q}_k$. The measurement equation is

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k \tag{2.2}$$

where $\mathbf{w}_k$ is zero-mean white Gaussian measurement noise, and

$$E[\mathbf{w}_k \mathbf{w}_k^T] = \mathbf{R}_k \tag{2.3}$$

The matrices $\mathbf{F}_k$, $\mathbf{G}_k$, $\mathbf{H}_k$, $\mathbf{Q}_k$, and $\mathbf{R}_k$ are assumed to be known with proper dimensions and possibly time varying. The initial state $\mathbf{x}_0$ in general is unknown and modeled as Gaussian distributed with known mean and covariance. The two noise sequences and the initial state are mutually independent. Sometimes, $\mathbf{v}_k$ is taken as $\mathbf{\Gamma}_k \mathbf{v}_k$ with $\mathbf{v}_k$ being an n_v-dimensional vector and $\mathbf{\Gamma}_k$ a known $n_x \times n_v$ matrix. Then the covariance matrix of the noise in the state equation can be written as

$$E\left[(\mathbf{\Gamma}_k \mathbf{v}_k)(\mathbf{\Gamma}_k \mathbf{v}_k)^T\right] = \mathbf{\Gamma}_k \mathbf{Q}_k \mathbf{\Gamma}_k^T \tag{2.4}$$

The linearity of (2.1) and (2.2) ensures the preservation of the Gaussian property of the state and measurements. The estimate of the system state $\mathbf{x}_i$ based on the observations up to time k can be written as,

$$\hat{\mathbf{x}}_{i|k} = E\left[\mathbf{x}_i | \mathbf{Z}^k\right] \tag{2.5}$$

where

$$\mathbf{Z}^k = \{\mathbf{z}_j : j = 1, \cdots, k\} \tag{2.6}$$

If $i = k$, the conditional mean is called the estimate of the system; if $i < k$, the conditional mean is called the smoothed value of the state; if $i > k$, the conditional mean is called predicted value of the state. The estimation error is defined as

$$\tilde{\mathbf{x}}_{i|k} = \mathbf{x}_i - \hat{\mathbf{x}}_{i|k} \tag{2.7}$$

The conditional covariance matrix of $\mathbf{x}_i$ given the data $\mathbf{Z}^k$ or the covariance associated with the estimate is

$$\mathbf{P}_{i|k} = E\left[\left(\mathbf{x}_i - \hat{\mathbf{x}}_{i|k}\right)\left(\mathbf{x}_i - \hat{\mathbf{x}}_{i|k}\right)^T |\mathbf{Z}^k\right] \tag{2.8}$$

2.1.2 The Recursive Estimation Algorithm

In terms of a linear and Gaussian observation $\mathbf{z}$ according to the minimum mean squared error (MMSE) criterion, the estimate of $\mathbf{x}$ with prior information $\mathbf{x} \sim N(\bar{\mathbf{x}}, \mathbf{P}_{xx})$ is

$$\hat{\mathbf{x}} = E\left[\mathbf{x}|\mathbf{z}\right] = \bar{\mathbf{x}} + \mathbf{P}_{xz}\mathbf{P}_{zz}^{-1}(\mathbf{z} - \bar{\mathbf{z}}) \tag{2.9}$$

and the corresponding MSE is

$$\mathbf{P}_{xx|z} = E\left[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T\right] = \mathbf{P}_{xx} - \mathbf{P}_{xz}\mathbf{P}_{zz}^{-1}\mathbf{P}_{zx} \tag{2.10}$$

Given the initial estimate $\hat{\mathbf{x}}_{0|0}$ of $\mathbf{x}_0$ and the associated initial covariance $\mathbf{P}_{0|0}$, the cycle of the dynamic estimation will consider mapping the estimate

$$\hat{\mathbf{x}}_{k|k} = E\left[\mathbf{x}_k|\mathbf{Z}^k\right] \tag{2.11}$$

which is the conditional mean of the state at the time k, and the covariance matrix

$$\mathbf{P}_{k|k} = E\left[[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}][\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}]^T|\mathbf{Z}^k\right] \tag{2.12}$$

into the corresponding variables at the next stage, that is to say, $\hat{\mathbf{x}}_{k+1|k+1}$ and $\mathbf{P}_{k+1|k+1}$. Since the process noise is white and Gaussian, the predicted state $\hat{\mathbf{x}}_{k+1|k}$ is

$$\begin{aligned}\hat{\mathbf{x}}_{k+1|k} &= E\left[\mathbf{x}_{k+1}|\mathbf{Z}^k\right] = E\left[\mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{u}_k + \mathbf{v}_k|\mathbf{Z}^k\right] \\ &= \mathbf{F}_k\hat{\mathbf{x}}_{k|k} + \mathbf{G}_k\mathbf{u}_k\end{aligned} \tag{2.13}$$

The state prediction error, namely the difference between the system state and its prediction is

$$\tilde{\mathbf{x}}_{k+1|k} = \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k \tilde{\mathbf{x}}_{k|k} + \mathbf{v}_k \tag{2.14}$$

Using the equation above, we can get the state prediction covariance as

$$\begin{aligned}
\mathbf{P}_{k+1|k} &= E\left[\tilde{\mathbf{x}}_{k+1|k}\tilde{\mathbf{x}}_{k+1|k}^T|\mathbf{Z}^k\right] \qquad (2.15)\\
&= \mathbf{F}_k E\left[\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T|\mathbf{Z}^k\right]\mathbf{F}_k^T + E\left[\mathbf{v}_k\mathbf{v}_k^T\right]\\
&= \mathbf{F}_k\mathbf{P}_{k|k}\mathbf{F}_k^T + \mathbf{Q}_k
\end{aligned}$$

The predicted measurement is the expectation of the measurement conditioned on $\mathbf{Z}^k$,

$$\begin{aligned}
\mathbf{z}_{k+1|k} &= E\left[\mathbf{z}_{k+1}|\mathbf{Z}^k\right] \qquad (2.16)\\
&= E\left[\mathbf{H}_{k+1}\mathbf{x}_{k+1} + \mathbf{w}_{k+1}|\mathbf{Z}^k\right]\\
&= \mathbf{H}_{k+1}\hat{\mathbf{x}}_{k+1|k}
\end{aligned}$$

The measurement prediction error is

$$\tilde{\mathbf{z}}_{k+1|k} = \mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k} = \mathbf{H}_{k+1}\tilde{\mathbf{x}}_{k+1|k} + \mathbf{w}_{k+1} \tag{2.17}$$

Thus the measurement prediction covariance, which is defined as $\mathbf{S}_{k+1}$, is

$$\mathbf{S}_{k+1} = \mathbf{H}_{k+1}\mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T + \mathbf{R}_{k+1} \tag{2.18}$$

The covariance between the state and measurement is

$$\begin{aligned}
E\left[\tilde{\mathbf{x}}_{k+1|k}\tilde{\mathbf{z}}_{k+1|k}^T|\mathbf{Z}^k\right] &= E\left[\tilde{\mathbf{x}}_{k+1|k}\left[\mathbf{H}_{k+1}\tilde{\mathbf{x}}_{k+1|k} + \mathbf{w}_{k+1}\right]^T|\mathbf{Z}^k\right] \qquad (2.19)\\
&= \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T
\end{aligned}$$

The filter gain can be calculated as

$$\mathbf{W}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T\mathbf{S}_{k+1}^{-1} \tag{2.20}$$

Thus the updated state estimate can be written as

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{W}_{k+1}\tau_{k+1} \tag{2.21}$$

where

$$\tau_{k+1} = \mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k} = \tilde{\mathbf{z}}_{k+1|k} \tag{2.22}$$

which is called innovation or measurement residual. Finally, the updated covariance of the state at time $k+1$ is,

$$\begin{aligned}\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{P}_{k+1|k}\mathbf{H}_{k+1}^T\mathbf{S}_{k+1}^{-1}\mathbf{H}_{k+1}\mathbf{P}_{k+1|k} \\ = \mathbf{P}_{k+1|k} - \mathbf{W}_{k+1}\mathbf{S}_{k+1}\mathbf{W}_{k+1}^T\end{aligned} \tag{2.23}$$

An alternative form for the covariance update can be provided as

$$\mathbf{P}_{k+1|k+1}^{-1} = \mathbf{P}_{k+1|k}^{-1} + \mathbf{H}_{k+1}^T R_{k+1}^{-1}\mathbf{H}_{k+1} \tag{2.24}$$

2.1.3 Statistical Test for Filter Consistency

Under the linear-Gaussian assumption, the conditional probability density function of the state $\mathbf{x}_k$ at the time k is

$$p(\mathbf{x}_k|\mathbf{Z}^k) = \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_{k|k}) \tag{2.25}$$

Based on (2.25), we can get the first two moments,

$$E\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right] = E\left[\tilde{\mathbf{x}}_{k|k}\right] = 0 \tag{2.26}$$
$$E\left[\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right]\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right]^T\right] = E\left[\tilde{\mathbf{x}}_{k|k}\tilde{\mathbf{x}}_{k|k}^T\right] = \mathbf{P}_{k|k}$$

Define the normalized estimation error squared as

$$\epsilon_k = \tilde{\mathbf{x}}_{k|k}^T \mathbf{P}_{k|k}^{-1} \tilde{\mathbf{x}}_{k|k} \tag{2.27}$$

Under hypothesis H_0 that the filter is consistent and linear Gaussian assumption, ϵ_k is Chi-square distributed with n_x degrees of freedom, where n_x is the dimension of the system state $\mathbf{x}$, and

$$E\left[\epsilon_k\right] = n_x \tag{2.28}$$

Based on the Monte Carlo simulations with N independent samples $\epsilon_k^i, i = 1, ..., N$, the sample average of ϵ_k can be obtained,

$$\bar{\epsilon}_k = \frac{1}{N}\sum_{i=1}^{N} \epsilon_k^i \tag{2.29}$$

It can be shown that $N\bar{\epsilon}_k$ follows a Chi-square distribution with Nn_x degrees of freedom. The hypothesis of H_0 is accepted if

$$\bar{\epsilon}_k \in [r_1, r_2] \tag{2.30}$$

where the acceptance interval is determined such that

$$P\{\bar{\epsilon}_k \in [r_1, r_2] | H_0\} = 1 - \alpha \tag{2.31}$$

and α is the power of the test.

2.2 System Model

For a discrete-time linear dynamic system described in Section 2.1.1, let us assume that M sensors are used by the system. The measurement at time k collected by sensor i is

$$\mathbf{z}_{k,i} = \mathbf{H}_{k,i}\mathbf{x}_{k,i} + \mathbf{w}_{k,i} \tag{2.32}$$

with $\mathbf{H}_{k,i}$ being the measurement matrix, and $\mathbf{w}_{k,i}$ a zero-mean white Gaussian measurement noise with covariance matrix $E[\mathbf{w}_{k,i}\mathbf{w}_{k,i}^T] = \mathbf{R}_{k,i}$, for $i = 1, \cdots, M$. We further assume that the measurement noises are independent across sensors. The matrices $\mathbf{H}_{k,i}$ and $\mathbf{R}_{k,i}$ are assumed to be known with proper dimensions. In this dissertation, we assume that a bias $\mathbf{b}_{k,i}$ is injected by the adversary into the measurement of the ith sensor at time k intentionally. Therefore, the measurement equation (2.32) becomes

$$\mathbf{z}'_{k,i} = \mathbf{H}_{k,i}\mathbf{x}_k + \mathbf{w}_{k,i} + \mathbf{b}_{k,i} = \mathbf{z}_{k,i} + \mathbf{b}_{k,i} \tag{2.33}$$

where $\mathbf{z}'_{k,i}$ is the corrupted measurement, $\mathbf{b}_{k,i}$ is either an unknown constant or a random variable independent of $\{\mathbf{v}_{k,i}\}$ and $\{\mathbf{w}_{k,i}\}$. For compactness, let us denote the system sensor observation as $\mathbf{z}_k = [\mathbf{z}_{k1}^T, \cdots, \mathbf{z}_{kM}^T]^T$, which contains the observations from all the M sensors. Similarly, let us denote the system bias vector as $\mathbf{b}_k = [\mathbf{b}_{k1}^T, \cdots, \mathbf{b}_{kM}^T]^T$ which includes the biases at all the M sensors. Correspondingly, the measurement matrix becomes

$$\mathbf{H}_k = [\mathbf{H}_{k1}^T, \cdots, \mathbf{H}_{kM}^T]^T \tag{2.34}$$

With these notations, it is easy to convert (2.32) and (2.33) into the following equations respectively.

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k \tag{2.35}$$

and

$$\mathbf{z}'_k = \mathbf{z}_k + \mathbf{b}_k \tag{2.36}$$

Further, we have the measurement error covariance matrix corresponding to $\mathbf{w}_k$ is

$$\mathbf{R}_k = \begin{bmatrix} \mathbf{R}_{k,1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{k,M} \end{bmatrix} \tag{2.37}$$

which is obtained by using the assumption that measurement noises are independent across sensors.

2.3 Impact of False Information Injection

Let us first assume that the adversary attacks the system by injecting false information into the sensors while Kalman filter is unaware of such attacks. We start with the case where biases ($\mathbf{b}_k$) are continuously injected into the system starting from a certain time K. Note that single injection is just a special case of continuous injection when $\mathbf{b}_k$ are set to be nonzero at time K and zero otherwise. In the continuous injection case, Kalman filter' extra mean square error (EMSE), which is caused by the continuous bias injection alone, is derived in [32] and provided as follows.

Proposition 1 *When the bias sequence $\{\mathbf{b}_k\}$ is zero mean, random, and independent over time, the EMSE at time $K+N$ due to the biases injected at and after time K, denoted as $\mathbf{A}_{K+N}$, is*

$$\mathbf{A}_{K+N} = \sum_{m=0}^{N} \mathbf{D}_m \mathbf{\Sigma}_{K+N-m} \mathbf{D}_m^T \tag{2.38}$$

where $\mathbf{D}_m = \left(\prod_{i=0}^{m-1} \mathbf{B}_{K+N-i}\right) \mathbf{W}_{K+N-m}$, and $\mathbf{B}_K = (\mathbf{I} - \mathbf{W}_K \mathbf{H}_K)\, \mathbf{F}_{K-1}$. $\prod_{i=0}^{-1} \mathbf{B}_{K+N-i} = \mathbf{I}$ is an identity matrix, $\mathbf{W}_K$ is Kalman filter gain [31], and $\mathbf{\Sigma}_{K+N-m}$ is the covariance matrix of $\mathbf{b}_{K+N-m}$.

2.4 Attack Strategies from Trace Perspective

Firstly, we investigate the optimal attack strategy that an adversary can adopt to maximize the system estimator's estimation error. This problem can be formulated as a constrained optimization problem. Without loss of generality, let us consider that the attacker is interested in maximizing the system state estimation error at time K right after a single false bias is injected at time K. In this case, we are interested in designing the injected random bias' covariance matrix such that

$$\begin{aligned}&\max_{\mathbf{\Sigma}_K} \operatorname{Tr}\left[\mathbf{P}_{K|K} + \mathbf{A}_K(\mathbf{\Sigma}_K)\right] \\ &s.t. \quad \operatorname{Tr}(\mathbf{\Sigma}_K) = a^2\end{aligned} \tag{2.39}$$

where a is a constant, $\operatorname{Tr}(\cdot)$ is the matrix trace operator, and $\mathbf{P}_{K|K}$ is Kalman filter's state estimation error covariance matrix at time K in the absence of any false information. Note that it is meaningful to have a constraint on the trace of $\mathbf{\Sigma}_K$, since it can be deemed as the power of injected sensor bias $\mathbf{b}_K$, and a smaller power for $\mathbf{b}_K$ reduces the probability that the adversary is detected by the system estimator using an innovation based detector. Note that the optimization problem is equivalent to one that maximizes $\operatorname{Tr}(\mathbf{A}_K(\mathbf{\Sigma}_K))$, since $\mathbf{P}_{K|K}$ is not a function of $\mathbf{\Sigma}_K$, and trace is a linear operator. If one is more interested in the determinant of the estimation MSE matrix, a similar optimization problem can be easily formulated as follows.

$$\begin{aligned}&\max_{\mathbf{\Sigma}_K} \left|\mathbf{P}_{K|K} + \mathbf{A}_K(\mathbf{\Sigma}_K)\right| \\ &s.t. \quad \operatorname{Tr}(\mathbf{\Sigma}_K) = a^2\end{aligned} \tag{2.40}$$

To simplify the mathematical analysis, it is helpful to derive the equivalent sensor measurement, which is a linear combination of the observations from all the sensors, and is a sufficient statistic containing all the information about the systems state.

The equivalent sensor measurement vector and its corresponding covariance matrix should have much smaller dimensionality than the original measurement vector and its covariance, making the mathematical manipulation and derivation later in the dissertation much simpler. In a information filter recursion [31], which is equivalent to Kalman filter recursion, we have

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k \tag{2.41}$$

where $\hat{\mathbf{y}}_{k|k} = \mathbf{P}_{k|k}^{-1} \mathbf{x}_{k|k}$ and $\hat{\mathbf{y}}_{k|k-1} = \mathbf{P}_{k|k-1}^{-1} \mathbf{x}_{k|k-1}$. It is clear that $\hat{\mathbf{y}}_{k|k-1}$ represents the prior knowledge about the system state based on past sensor data, and the second term in (2.41) represents the new information from the new sensor data $\mathbf{z}_k$, which can be expanded by using (2.34) and (2.37) as follows.

$$\begin{aligned}
&\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{z}_k \\
&= [\mathbf{H}_{k1}^T, \cdots, \mathbf{H}_{kM}^T] \begin{bmatrix} \mathbf{R}_{k1}^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{R}_{kM}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{k1} \\ \vdots \\ \mathbf{z}_{kM} \end{bmatrix} \\
&= \sum_{i=1}^{M} \mathbf{H}_{ki}^T \mathbf{R}_{ki}^{-1} \mathbf{z}_{ki}
\end{aligned} \tag{2.42}$$

In the following derivations, we skip the time index k for simplicity. Our purpose is to find an equivalent measurement $\mathbf{z}_e$ such that

$$\mathbf{z}_e = \mathbf{H}_e \mathbf{x} + \mathbf{w}_e \tag{2.43}$$

where $\mathbf{w}_e \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_e)$, and

$$\mathbf{H}_e^T \mathbf{R}_e^{-1} \mathbf{z}_e = \sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i \tag{2.44}$$

Let us consider two cases. First, suppose all the $\mathbf{H}_i$s are the same $(\mathbf{H}_i = \mathbf{H})$, then it is natural to set $\mathbf{H}_e = \mathbf{H}$. Note that a sufficient condition for (2.44) to be true is

$$\mathbf{z}_e = \mathbf{R}_e \sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{z}_i \tag{2.45}$$

Finding the covariance on the both sides of (2.45), we get

$$\begin{aligned} \mathbf{R}_e &= \mathbf{R}_e \mathrm{cov} \left(\sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{z}_i \right) \mathbf{R}_e^T \\ &= \mathbf{R}_e \left[\sum_{i=1}^{M} \mathbf{R}_i^{-1} \mathbf{R}_i (\mathbf{R}_i^{-1})^T \right] \mathbf{R}_e^T \end{aligned} \tag{2.46}$$

This implies that

$$\mathbf{R}_e = \left(\sum_{i=1}^{M} \mathbf{R}_i^{-1} \right)^{-1} \tag{2.47}$$

In the second case, let us assume that the system state $\mathbf{x}$ is observable based on the observations from all the sensors, meaning that the Fisher information matrix $\sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i$ is invertible. In this case, by setting $\mathbf{H}_e = \mathbf{I}$, using (2.44), and following a similar procedure as in the first case, we have

$$\mathbf{z}_e = \mathbf{R}_e \sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i \tag{2.48}$$

and

$$\mathbf{R}_e = \left(\sum_{i=1}^{M} \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i \right)^{-1} \tag{2.49}$$

We have derived the optimal strategies concerning position-sensor case to maximize the trace of the state estimation MSE matrix as provided in the following two propositions [25].

Proposition 2 *For a system with M sensors, if the adversary injects independent random noises, the best strategy is to allocate all the power to the sensor with the* *smallest measurement noise variance.*

Proposition 3 *For a system with M sensors, the optimal strategy for the adversary is to inject dependent random noises with a pairwise correlation coefficient of* 1*. The noise power is allocated such that* $\sigma_{b_i} = \frac{c_i a}{\sqrt{\sum_{j=1}^{M} c_j^2}}, \quad i \in \{1, \cdots, M\}$*, where* σ_{b_i} *is the standard deviation (s.d.) of the noise injected to the* i*th sensor,* $c_i = \frac{1/\sigma_{w_i}^2}{\sum_{j=1}^{M}\left(1/\sigma_{w_j}^2\right)}$ *and* σ_{w_i} *is the* i*th position-only sensor's measurement noise s.d.*

As for the case where sensors measure both position and velocity of the target, the best attack strategy for single sensor and multiple sensors are studied and corresponding optimal attack strategies are shown as follow,

Proposition 4 *For a system with one sensor observing position and velocity of the target, the optimal strategy for the adversary is to inject random noise that has dependent position and velocity components. If* $w_{11}w_{12} + w_{21}w_{22} > 0$*, the correlation coefficient* ρ_{b_p,b_v} *should be set as* 1*, and the random bias power is allocated such that*

$$\begin{aligned}
&\sigma_{b_p} = a\sin(\theta^*) \qquad (2.50)\\
&\sigma_{b_v} = \frac{a}{T}\cos(\theta^*)\\
&\theta^* = \frac{\pi}{4} - \frac{\phi}{2}\\
&\phi = \arctan\left[\frac{\beta_2 - \beta_1 T^2}{2T(\alpha_1 + \alpha_2)}\right]\\
&w_{11}^2 + w_{21}^2 = \beta_1\\
&w_{12}^2 + w_{22}^2 = \beta_2\\
&w_{11}w_{12} = \alpha_1\\
&w_{21}w_{22} = \alpha_2
\end{aligned}$$

When $w_{11}w_{12} + w_{21}w_{22} < 0$*, we should set* $\rho_{b_p,b_v} = -1$ *and set* $\alpha_1 = -w_{11}w_{12}$ *and* $\alpha_2 = -w_{21}w_{22}$*. The rest of the equations in formula (2.50) remains the same.*

As for attack strategy for multiple position and velocity sensors, equivalent sensor is utilized to find the best attack strategy. Based on Proposition 1, we get the EMSE matrix,

$$\mathbf{A}_{K+N} = \sum_{m=0}^{N} \mathbf{D}_m \mathbf{\Sigma}_{K+N-m} \mathbf{D}_m^T$$

Suppose at time K, the adversary wants to attack the system continuously from time K to $K+N$, the weight for different time is $\alpha_m, m \in \{1, \cdots, N\}$, as shown below,

$$\begin{aligned} \mathbf{A}_K^{'} &= \alpha_0(\mathbf{D}_0 \mathbf{\Sigma}_K \mathbf{D}_0^T) \\ \mathbf{A}_{K+1}^{'} &= \alpha_1(\mathbf{D}_0 \mathbf{\Sigma}_{K+1} \mathbf{D}_0^T + \mathbf{D}_1 \mathbf{\Sigma}_K \mathbf{D}_1^T) \\ &... \\ \mathbf{A}_{K+N}^{'} &= \alpha_N(\mathbf{D}_0 \mathbf{\Sigma}_{K+N} \mathbf{D}_0^T + ... + \mathbf{D}_N \mathbf{\Sigma}_K \mathbf{D}_N^T) \end{aligned} \tag{2.51}$$

where $\sum_{m=0}^{N} \alpha_m = 1$. So the objective function in the multi-shot attack case is the trace of the weighted sum of the EMSE matrices at different time points that is $\sum_{m=0}^{N} \alpha_m \mathbf{A}_{K+m} = \sum_{m=0}^{N} \mathbf{A}_{K+m}^{'}$. It is equivalent to maximize the trace of the weighted sum of the MSE matrices of the state estimates, because once the system reaches its steady state, $\mathbf{P}_{K+m|K+m}$ becomes constant, and the weighted sum of $\mathbf{P}_{K+m|K+m}$ will remain the same. First we study the case where the system has position sensors which are being attacked, so all the items above are scalars. Using lower case d, σ_p^2 to denote $\mathbf{D}, \mathbf{\Sigma}$, we can formulate the optimization problem below,

$$\begin{aligned}
\max_{\sigma_{p_K},\cdots,\sigma_{p_{K+N}}} \quad & \sum_{m=0}^{N} \alpha_m \mathbf{A}_{K+m} = \sum_{m=0}^{N} \mathbf{A}'_{K+m} \\
& = \sigma_{p_K}^2(\alpha_0 d_0^2 + \alpha_1 d_1^2 + ... + \alpha_N d_N^2) \\
& + \sigma_{p_{K+1}}^2(\alpha_1 d_0^2 + \alpha_2 d_1^2 + ... + \alpha_N d_{N-1}^2) \\
& + \sigma_{p_{K+2}}^2(\alpha_2 d_0^2 + \alpha_3 d_1^2 + ... + \alpha_N d_{N-2}^2) \\
& + ... \\
& + \sigma_{p_{K+N}}^2(\alpha_N d_0^2) \\
s.t. \quad & \sum_{m=K}^{K+N} \sigma_{p_m}^2 \leq a^2 \\
& \sum_{m=0}^{N} \alpha_m = 1
\end{aligned} \tag{2.52}$$

The adversary can allocate the power based on the coefficients of the variance variables at different time. For example, if the weights $\alpha_m' s$ are all the same, the best strategy is to allocate all the power to the sensors at the first beginning (at time K) because the coefficient for $\sigma_{p_K}^2$ is the largest. Second, if the sensors measure both position and velocity, and the attacker aims to attack the system with position and velocity false information, the optimization problem can be characterized as below,

$$\begin{aligned}
\max_{\mathbf{\Sigma}_K,\cdots,\mathbf{\Sigma}_{K+N}} \quad & \mathrm{Tr}\left[\sum_{m=0}^{N}\alpha_m\mathbf{A}_{K+m}\right] = \mathrm{Tr}\left[\sum_{m=0}^{N}\mathbf{A}'_{K+m}\right] \\
& = \mathrm{Tr}\left[\mathbf{\Sigma}_K(\alpha_0\mathbf{D}_0^T\mathbf{D}_0 + ... + \alpha_N\mathbf{D}_N^T\mathbf{D}_N)\right] \\
& +\mathrm{Tr}\left[\mathbf{\Sigma}_{K+1}(\alpha_1\mathbf{D}_0^T\mathbf{D}_0 + ... + \alpha_N\mathbf{D}_{N-1}^T\mathbf{D}_{N-1})\right] \\
& +\mathrm{Tr}\left[\mathbf{\Sigma}_{K+2}(\alpha_2\mathbf{D}_0^T\mathbf{D}_0 + ... + \alpha_N\mathbf{D}_{N-2}^T\mathbf{D}_{N-2})\right] \\
& +... \\
& +\mathrm{Tr}\left[\mathbf{\Sigma}_{K+N}(\alpha_N\mathbf{D}_0^T\mathbf{D}_0)\right] \\
s.t. \quad & \sum_{m=K}^{K+N}\sigma_{p_m}^2 + T^2\sigma_{v_m}^2 \leq a^2 \\
& \sum_{m=0}^{N}\alpha_m = 1
\end{aligned} \tag{2.53}$$

where $\mathbf{\Sigma}_m$ and $\mathbf{D}_j^T\mathbf{D}_j$ are positive semidefinite matrices, so $\mathrm{Tr}\left[\mathbf{\Sigma}_m(\mathbf{D}_j^T\mathbf{D}_j)\right] \geq 0$ all the time. The trace function $\mathrm{Tr}(\cdot)$ is a monotonically increasing function of the positive semidefinite matrix. So the best strategy for the adversary to attack the system is to put all the power at the time with the largest positive semidefinite matrix.

2.5 Attack Strategies from Determinant Perspective

For the position-only sensors, we are interested in the effect of bias information on Kalman filter's MSE matrix from the determinant perspective as follows,

$$\begin{aligned}
|\mathbf{P}_{K|K} + \mathbf{A}_K| &= |\mathbf{P}_{K|K} + \Sigma_{eK}\mathbf{D}_0\mathbf{D}_0^T| \\
&= |\mathbf{P}_{K|K}||\mathbf{I} + \Sigma_{eK}\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T|
\end{aligned} \tag{2.54}$$

where $\mathbf{D}_0$ is defined in Proposition 1. As $\mathbf{P}_{K|K}$ is constant and positive definite, $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$ is positive semidefinite meaning that all the eigenvalues of $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$

are non-negative. First, let us denote $\mathbf{C}$ as a square matrix whose columns are the eigenvectors of $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$. Then through eigendecomposition, (2.54) can be written concisely as,

$$\begin{aligned} &|\mathbf{P}_{K|K}||\mathbf{CIC}^{-1} + \Sigma_{eK}\mathbf{C\Lambda C}^{-1}| \\ &= |\mathbf{P}_{K|K}||\mathbf{I} + \Sigma_{eK}\mathbf{\Lambda}| \end{aligned} \tag{2.55}$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of the $\mathbf{D}_0\mathbf{P}_{K|K}^{-1}\mathbf{D}_0^T$. So we just need to maximize Σ_{eK} in order to maximize the determinant of $\mathbf{P}_{K|K} + \mathbf{A}_K$. This is equivalent to maximizing the trace of $\mathbf{P}_{K|K} + \mathbf{A}_K$ as discussed in Section 2.4.

For the position-and-velocity sensors, we assume that the adversary knows the system model and the prior information $\mathbf{P}_{0|0}$ at time zero, so that he/she can calculate the offline Kalman filter gain matrix $\mathbf{W}_k$ recursively. The best attack strategy is the solution to the following optimization problem.

$$\begin{aligned} &\max_{\mathbf{\Sigma}_K} \left|\mathbf{P}_{K|K} + \mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T\right| \\ &s.t. \quad \sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = a^2 \\ &\qquad -1 \le \rho_{b_p,b_v} \le 1 \\ &\qquad \sigma_{b_p}, \sigma_{b_v} > 0 \end{aligned} \tag{2.56}$$

where $\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T = \mathbf{A}_K$, and

$$\mathbf{\Sigma}_K = \begin{bmatrix} \sigma_{b_p}^2 & \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} \\ \rho_{b_p,b_v}\sigma_{b_p}\sigma_{b_v} & \sigma_{b_v}^2 \end{bmatrix} \tag{2.57}$$

Using the properties of the determinant, we get the formula as follows.

$$\begin{aligned} & |\mathbf{P}_{K|K} + \mathbf{W}_K \boldsymbol{\Sigma}_K \mathbf{W}_K^T| \\ & = |\mathbf{P}_{K|K}||\mathbf{I}_n + \boldsymbol{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K| \end{aligned} \tag{2.58}$$

Since $\mathbf{P}_{K|K}$ is independent of $\boldsymbol{\Sigma}_K$, the optimization problem can be further written as:

$$\begin{aligned} \max_{\boldsymbol{\Sigma}_K} \quad & \left| \mathbf{I}_n + \boldsymbol{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \\ s.t. \quad & \sigma_{b_p}^2 + T^2 \sigma_{b_v}^2 = a^2 \\ & -1 \leq \rho_{b_p,b_v} \leq 1 \\ & \sigma_{b_p}, \sigma_{b_v} > 0 \end{aligned} \tag{2.59}$$

By defining

$$\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K = \begin{bmatrix} m_1 & m_2 \\ m_2 & m_3 \end{bmatrix} \tag{2.60}$$

and after simplifying (2.59), the objective function becomes

$$\begin{aligned} & \left| \mathbf{I}_n + \boldsymbol{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \\ & = 1 + (1 - \rho_{b_p,b_v}^2) \sigma_{b_p}^2 \sigma_{b_v}^2 (m_1 m_3 - m_2^2) \\ & + \sigma_{b_p}^2 m_1 + \sigma_{b_v}^2 m_3 + 2\rho_{b_p,b_v} \sigma_{b_p} \sigma_{b_v} m_2 \end{aligned} \tag{2.61}$$

The optimal solution to the problem will be the best strategy to attack the system.

We denote $\mathbf{\Sigma}_K = \mathbf{R}^T\mathbf{R}$ and since $\mathbf{\Sigma}_K$ is invertible, we have

$$\begin{aligned} & \left|\mathbf{I}_n + \mathbf{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K\right| \\ = & \left|\mathbf{I}_n + \mathbf{R}^T\mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K\right| \\ = & \left|\mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T\right| \end{aligned} \tag{2.62}$$

In order to obtain the optimal solution, two useful lemmas [33] are introduced as follows,

Lemma 1 *Suppose* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $n \times n$ *positive semidefinite matrices with eigendecomposition* $\mathbf{A} = \mathbf{\Psi_A \Sigma_A \Psi_A^T}$ *and* $\mathbf{B} = \mathbf{\Psi_B \Sigma_B \Psi_B^T}$*, the eigenvalues of* $\mathbf{A}$ *and* $\mathbf{B}$ *satisfy that* $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ *and* $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$*, then*

$$\Pi_{i=1}^n(\alpha_i + \beta_i) \leq \det(\mathbf{A} + \mathbf{B}) \leq \Pi_{i=1}^n(\alpha_i + \beta_{n+1-i}) \tag{2.63}$$

where the upper bound is achieved if and only if $\mathbf{\Psi_A} = \mathbf{\Psi_B \Theta}$*, the lower bound is achieved if and only if* $\mathbf{\Psi_A} = \mathbf{\Psi_B}$*, and* $\mathbf{\Theta}$ *is the matrix defined below,*

$$\begin{bmatrix} 0 & 0 & \cdots & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \tag{2.64}$$

Readers are referred to [33] for the proof of Lemma 1. The optimal solution to find the upper bound is the best strategy to attack the system with the most effect on Kalman filter system and the lower bound is the least attack effect the adversary can get.

Lemma 2 *Given a* $n \times n$ *matrix* $\mathbf{V}_1$ *and a* $n \times n$ *positive semidefinite matrix* Ξ_1 *with* $\mathbf{V}_1 \Xi_1 \mathbf{V}_1^T$ *being a diagonal matrix with diagonal elements in increasing order, it is* *always possible to find another $n \times n$ matrix $\bar{\mathbf{V}}_1$ such that $\bar{\mathbf{V}}_1 \mathbf{\Xi}_1 \bar{\mathbf{V}}_1^T = \beta \mathbf{V}_1 \mathbf{\Xi}_1 \mathbf{V}_1^T$ with $Tr(\mathbf{V}_1 \mathbf{V}_1^T) = Tr(\bar{\mathbf{V}}_1 \bar{\mathbf{V}}_1^T)$ where $\beta \geq 1$. $\bar{\mathbf{V}}_1$ can be written as $\Sigma_{\mathbf{\Xi}} \mathbf{\Psi}_1^T$, where $\mathbf{\Psi}_1$ is the unitary matrix whose columns are the eigenvectors corresponding to the eigenvalues of $\mathbf{\Xi}_1$ in increasing order, and $\mathbf{\Sigma}_{\mathbf{\Xi}}$ is a diagonal matrix.*

By combining the two lemmas together, we can get the final optimal solution to the optimization problem above. It is obvious that $\mathbf{I}_n$ and $\mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T$ are both positive semidefinite matrices, and their eigendecomposition can be written as follows,

$$\begin{aligned} &\mathbf{I}_n = \mathbf{\Psi}_1 \mathbf{\Sigma}_1 \mathbf{\Psi}_1^T \\ &\mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T = \mathbf{\Psi}_2 \mathbf{\Sigma}_2 \mathbf{\Psi}_2^T \end{aligned} \tag{2.65}$$

with identity matrix $\mathbf{\Sigma}_1 = diag([\sigma_{1,1}, \cdots, \sigma_{1,n}])$ and $\mathbf{\Sigma}_2 = diag([\sigma_{2,1}, \cdots, \sigma_{2,n}])$, where $\sigma_{2,i}, i \in \{1, \cdots, n\}$ is the diagonal element of the matrix $\mathbf{\Sigma}_2$. Based on Lemma 1, we can get,

$$\left| \mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T \right| \leq \Pi_{i=1}^n (\sigma_{2,i} + 1) \tag{2.66}$$

where $\mathbf{\Psi}_1 = \mathbf{\Psi}_2 \mathbf{\Theta}$.

$$\begin{aligned} &|\mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T| \\ &= |\mathbf{\Psi}_1^T||\mathbf{I}_n + \mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T||\mathbf{\Psi}_1| \\ &= |\mathbf{I}_n + \mathbf{\Psi}_1^T \mathbf{R}\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}^T \mathbf{\Psi}_1| \end{aligned} \tag{2.67}$$

Set $\mathbf{R}_1 = \mathbf{\Psi}_1^T \mathbf{R}$ and $\mathbf{\Sigma}_3 = \mathbf{\Theta} \mathbf{\Sigma}_2 \mathbf{\Theta}^T$ with the eigenvalues of increasing order and $Tr(\mathbf{R}\mathbf{R}^T) = Tr(\mathbf{R}_1 \mathbf{R}_1^T)$. So the optimization problem can be written as below,

$$\begin{aligned} \max \quad & |\mathbf{I}_n + \mathbf{R}_1 \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}_1^T| \\ s.t. \quad & Tr(\mathbf{R}_1 \mathbf{R}_1^T) \leq a^2 \\ & \mathbf{R}_1 \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \mathbf{R}_1^T = \mathbf{\Sigma_3} \end{aligned} \tag{2.68}$$

Setting $\mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K = \tilde{\mathbf{\Xi}}$, we have $\mathbf{R}_1 \tilde{\mathbf{\Xi}} \mathbf{R}_1^T = \mathbf{\Sigma}_3$. Based on Lemma 2, we can surely find a matrix $\bar{\mathbf{R}}$ such that $\bar{\mathbf{R}}_1 \tilde{\mathbf{\Xi}} \bar{\mathbf{R}}_1^T = \beta \mathbf{R}_1 \tilde{\mathbf{\Xi}} \mathbf{R}_1^T$, with $\beta \geq 1$. Note that $\det(\cdot)$ is a monotonic increasing function of the positive semidefinite matrix. So

$$|\mathbf{I}_n + \mathbf{R}_1 \tilde{\mathbf{\Xi}} \mathbf{R}_1^T| \leq |\mathbf{I}_n + \bar{\mathbf{R}}_1 \tilde{\mathbf{\Xi}} \bar{\mathbf{R}}_1^T| \tag{2.69}$$

So the optimal solution $\bar{\mathbf{R}}$ should be in the form of $\bar{\mathbf{V}}$. The eigendecompostion of $\tilde{\mathbf{\Xi}}$ is as follows,

$$\tilde{\mathbf{\Xi}} = \mathbf{V}_\Xi \mathbf{\Sigma}_\Xi \mathbf{V}_\Xi^T \tag{2.70}$$

where $\mathbf{\Sigma}_\Xi = diag([\sigma_{\xi,1}, \sigma_{\xi,2}, \cdots, \sigma_{\xi,n}])$ in increasing order. $\mathbf{V}_\Xi$ is a unitary matrix whose column vectors corresponds to the eigenvalues of $\tilde{\mathbf{\Xi}}$. The problem can be written as

$$\begin{aligned} \max_{\sigma_{b,i}^2} \quad & \sum_{i=1}^{n} log(\sigma_{b,i}^2 \sigma_{\xi,i} + 1) \\ s.t. \quad & \sum_{i=1}^{n} (\sigma_{b,i}^2) \leq a^2 \end{aligned} \tag{2.71}$$

The objective function above is a concave and increasing function. The optimal solution is achieved through Lagrangian multipliers yielding the water-filling strategy,

$$\sigma_{b,i}^2 = \left(\frac{1}{\lambda} - \frac{1}{\sigma_{\xi,i}} \right)^+ \tag{2.72}$$

where the value of λ can be obtained by solving

$$\sum_{i=1}^{n} \left(\frac{1}{\lambda} - \frac{1}{\sigma_{\xi,i}} \right)^+ = a^2 \tag{2.73}$$

The solution is

$$\mathbf{R}^{opt} = \mathbf{\Psi}_1 [\mathbf{\Sigma_b^{1/2}}]^{\mathbf{T}} \mathbf{V_\Xi^T} \tag{2.74}$$

Finally, the optimal solution of (2.59) is,

$$\mathbf{\Sigma}_K = \mathbf{V}_{\mathbf{\Xi}} \mathbf{\Sigma}_b \mathbf{V}_{\mathbf{\Xi}}^T \tag{2.75}$$

2.6 Numerical Results

Some numerical results are presented in this section to illustrate the theoretical results.

2.6.1 System with Position Sensors

The parameters used in the target tracking example are provided below. The system sampling interval is $T = 1$. The adversary injects bias information to two sensors with $\sigma_{w_1}^2 = 3$ and $\sigma_{w_2}^2 = 4$, respectively. The variance of the system process noise is $\sigma_v^2 = 0.25$. The biases b_is are zero-mean Gaussian random variables with variances $\sigma_{b_i}^2$s. For the power constraint we discussed earlier, we set the sum of $\sigma_{b_i}^2$ to be 3000.

The effect of the bias injection on Kalman filter is measured by a Chi-squared test. More specifically, we use the sum of the normalized MSE over N_m Monte-Carlo runs

$$q_k = \sum_{j=1}^{N_m} \left[\hat{\mathbf{x}}_{k|k}^{\prime j} - \mathbf{x}_k^j \right]^T \mathbf{P}_{k|k}^{-1} \left[\hat{\mathbf{x}}_{k|k}^{\prime j} - \mathbf{x}_k^j \right] \tag{2.76}$$

where at time k, $\mathbf{P}_{k|k}$ is the nominal state covariance matrix calculated by Kalman filter, $\hat{\mathbf{x}}_{k|k}^{\prime j}$ is the state estimate, and $\mathbf{x}_k^j$ is the true state, during the jth Monte-Carlo run. First, if the random biases injected to different sensors are independent, we should allocate all the bias power to the sensor with the smallest measurement noise variance. This is clearly true as demonstrated in Fig. 1, where allocating all the power to sensor 1 causes the maximum mean squared estimation error. In Fig. 2, three

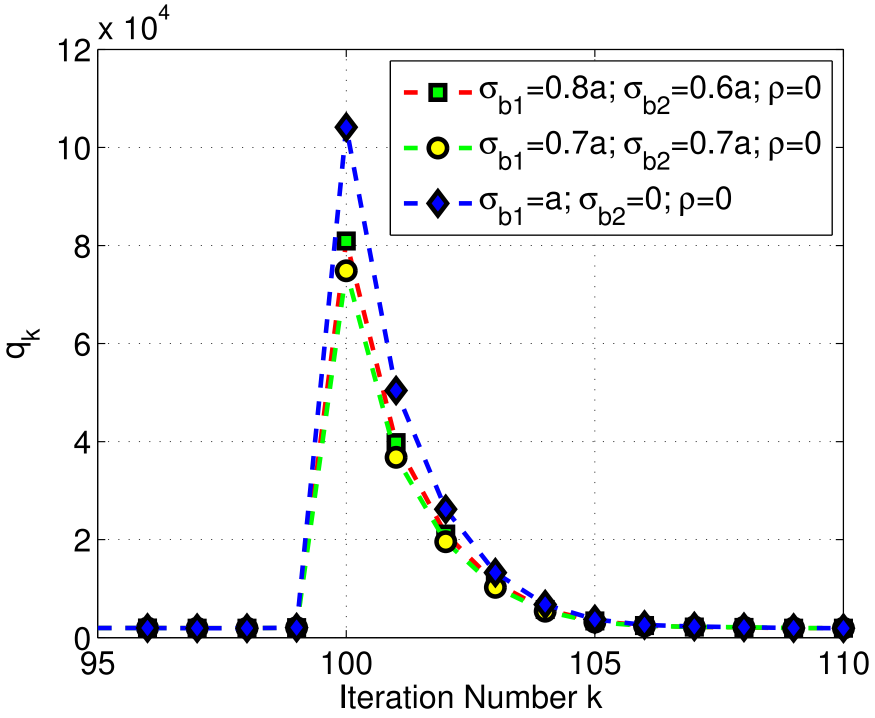

Fig. 1. The normalized MSE for independent biases. $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2$ for each case.

dependent-noise attack strategies are compared, including the optimal one according to Proposition 3, allocating the power equally among the sensors, and allocating all the power to the sensor with smallest measurement error variance. It is clear that the optimal solution has the largest impact on the estimation performance, and it outperforms the best independent-noise attack strategy significantly.

2.6.2 Systems with Position and Velocity Sensors

We now consider the case where the adversary attacks Kalman filtering system with a vector sensor observation containing both position and velocity measurements. We first consider a single-sensor system, and the sensor has a position measurement variance of 3 and a velocity measurement variance of 4. We set the sum of $\sigma_{b_{p_1}}^2$ and $T^2\sigma_{b_{v_1}}^2$ to be 3000. In this particular case, $w_{11}w_{12} + w_{21}w_{22} > 0$, so the optimal

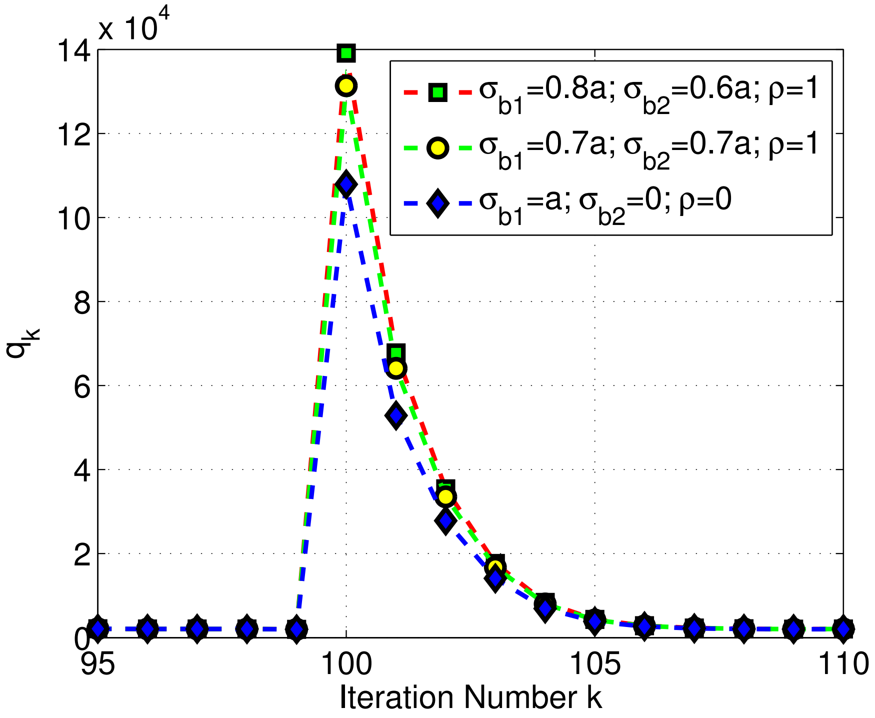

Fig. 2. The normalized MSE for dependent biases. $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2$ for each case.

choice is $\rho_{b_p,b_v} = 1$. Based on Theorem 4, the best strategy is to set $\sigma_{b_p} = 52.3$ and $\sigma_{b_v} = 16.2$. It is clear from Fig. 3 that the strategy provided in Theorem 4 maximizes the MSE of Kalman filter system by injecting vector bias information.

Next we consider a system with two sensors. The first sensor is the same as the one described above, and the second one is with position measurement variance 4 and velocity measurement variance 5. In this particular case, again we have $w_{11}w_{12} + w_{21}w_{22} > 0$, so all the ρs in s_1, s_2, and s_3 should be set as 1. We first use a systematic grid search to find an approximate globally optimal solution and then we use the FMINCON function in Matlab, a local search algorithm, to refine this approximate globally optimal solution. The optimal solution we have obtained is $\sigma_{b_{p_1}}^2 = 1826, \sigma_{b_{p_2}}^2 = 1023, \sigma_{b_{v_1}}^2 = 81, \sigma_{b_{v_2}}^2 = 68$. For comparison purposes, we also implement an attack strategy that allocates power equally among the observation

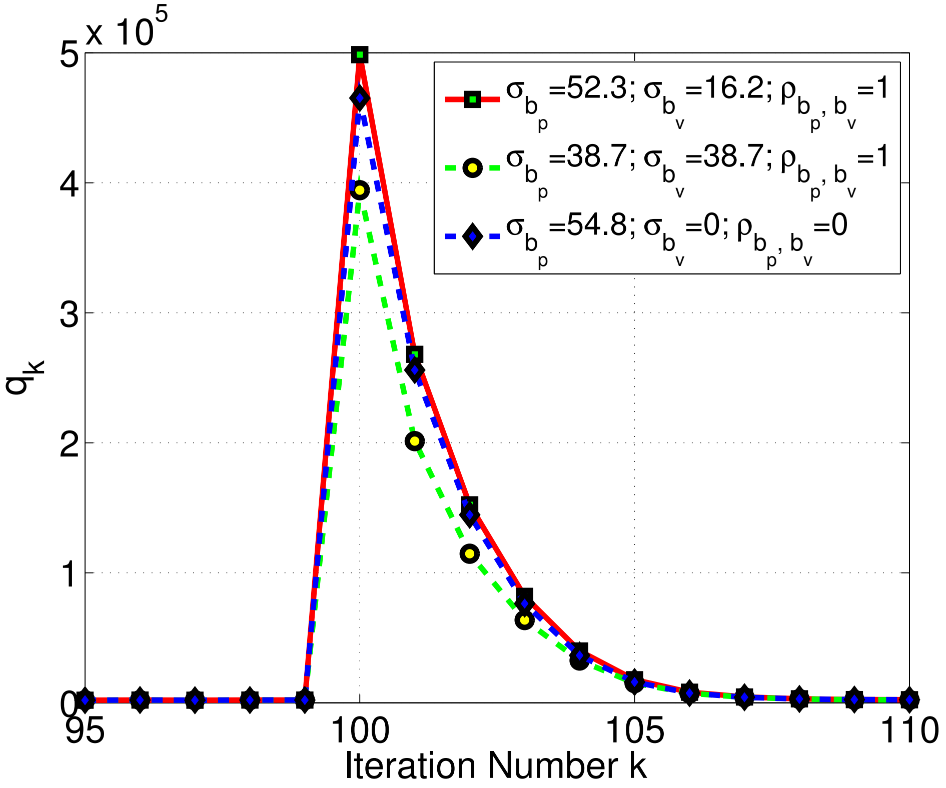

Fig. 3. The normalized MSE for a system with a single sensors. $\sigma_{p_1}^2 + T^2\sigma_{v_1}^2 = a^2$ for each case.

components and among the two sensors, which is $\sigma_{b_{p_1}}^2 = \sigma_{b_{p_2}}^2 = \sigma_{b_{v_1}}^2 = \sigma_{b_{v_2}}^2 = 750$. The simulation result is shown in Fig. 4. As we can see, the optimal attack strategy has a much greater impact than the one that allocates power equally. Based on the optimal solution, we can find that allocating more power to the measurement with lower variance will have a greater effect on Kalman filter system.

2.6.3 Determinant Perspective

Numerical results are presented in this section to illustrate the effectiveness of the proposed attack strategies. Assuming that the injected bias noise $\mathbf{b}_k$ is zero-mean and Gaussian distributed, we can show that the posterior probability density

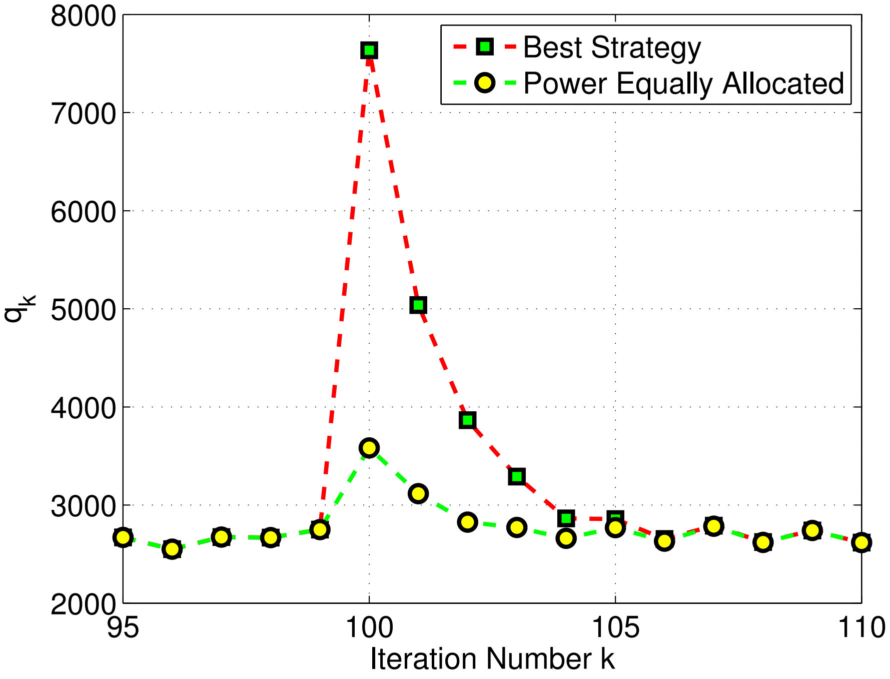

Fig. 4. The normalized MSE for a system with two sensors. $\sigma_{p_1}^2 + \sigma_{p_2}^2 + T^2\sigma_{v_1}^2 + T^2\sigma_{v_2}^2 = a^2$ for each case.

function (PDF) of the target state conditioned on the past observations and the current corrupted observation is

$$p(\mathbf{x}_K|\mathbf{z}_{1:K-1}, \mathbf{z}'_K) = \mathcal{N}(\hat{\mathbf{x}}_{K|K}, \mathbf{P}_{K|K} + \mathbf{A}_K) \tag{2.77}$$

where $\hat{\mathbf{x}}_{K|K}$ is the updated state estimate calculated by Kalman filter, which is unaware of the presence of the injected false information. Then the target state $\mathbf{x}_K$ will be in the following confidence region (or error ellipse)

$$\left\{\mathbf{x} : (\mathbf{x} - \hat{\mathbf{x}}_{K|K})^T(\mathbf{P}_{K|K} + \mathbf{A}_K)^{-1}(\mathbf{x} - \hat{\mathbf{x}}_{K|K}) \leq \gamma\right\} \tag{2.78}$$

with probability determined by the threshold γ [34]. The volume of the confidence region defined by (2.78) corresponding to the threshold γ is

$$V(K) = c_{n_x}|\gamma(\mathbf{P}_{K|K} + \mathbf{A}_K)|^{1/2} \tag{2.79}$$

where n_x is the dimension of the target state $\mathbf{x}$,

$$c_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)} \tag{2.80}$$

and $\Gamma(\cdot)$ is the gamma function. First, let us consider a single-sensor case, where the sensor has a position measurement with noise variance of 3, which is independent of the velocity measurement with noise variance of 4. We set the bias noise power constraint as $\sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = 3000$. We solve the optimization problem formulated in Section 2.5 numerically, and the optimal solution to (2.56) is $\sigma_{b_p}^2 = 1500, \sigma_{b_v}^2 = 1500, \rho_{b_{p,v}} = 0.063$. In Fig. 5, error ellipsis for different attack strategies are plotted. For all the different attack strategies, we set $\rho_{b_{p,v}} = 0.063$. As we can see, under

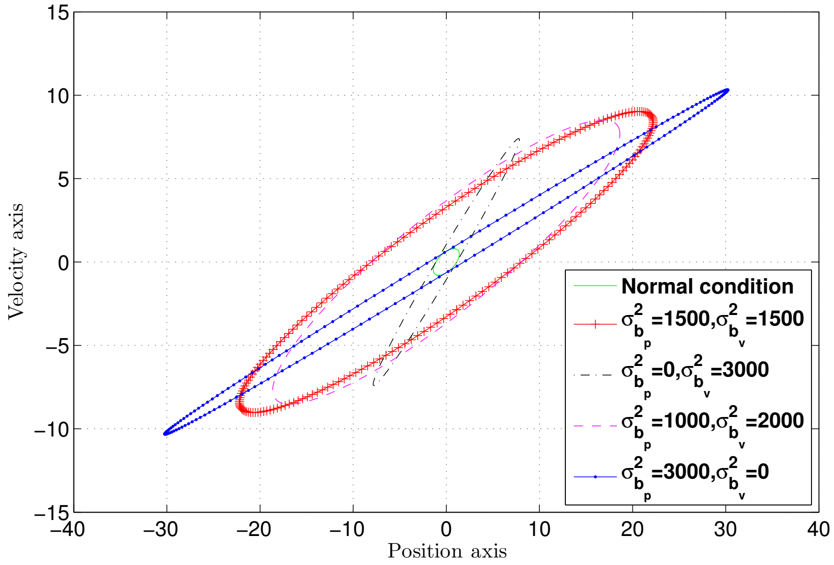

Fig. 5. Error ellipsis for different power allocation strategies

normal condition without false information injection, the error ellipse has the smallest area, while the optimal attack strategy leads to an error ellipse with the largest area. In Figs. 6 and 7, the volume (area) of the error ellipse is provided as a function of $\rho_{b_{p,v}}$ and the ratio $\kappa = \frac{\sigma_{b_p}}{\sigma_{b_v} T}$. We can see that when the $\kappa = \frac{\sigma_{b_p}}{\sigma_{b_v} T} = 1$, the area of the ellipse is maximized. Also from Figs. 6 and 7, it is clear that the area of ellipse increases as the absolute value of ρ decreases. In Fig. 8, the trend of the error ellipsis as the ρ changes from -1 to $+1$ is illustrated.

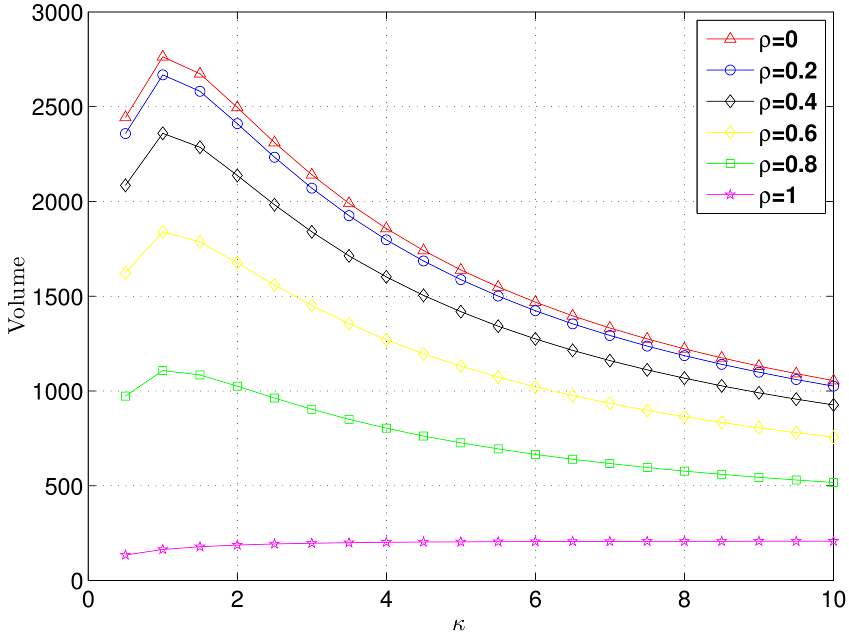

Fig. 6. Error ellipse volume for the positive correlation case

In this particular case, since $\sigma_{b_p}^2 + T^2\sigma_{b_v}^2 = 3000$, $\mathbf{\Sigma}_K$ is large and in (2.56) the second term ($\mathbf{W}_K\mathbf{\Sigma}_K\mathbf{W}_K^T$) dominates. Therefore, in (2.61) the identity matrix in the objective function is relatively small comparing to the second item, and approximately

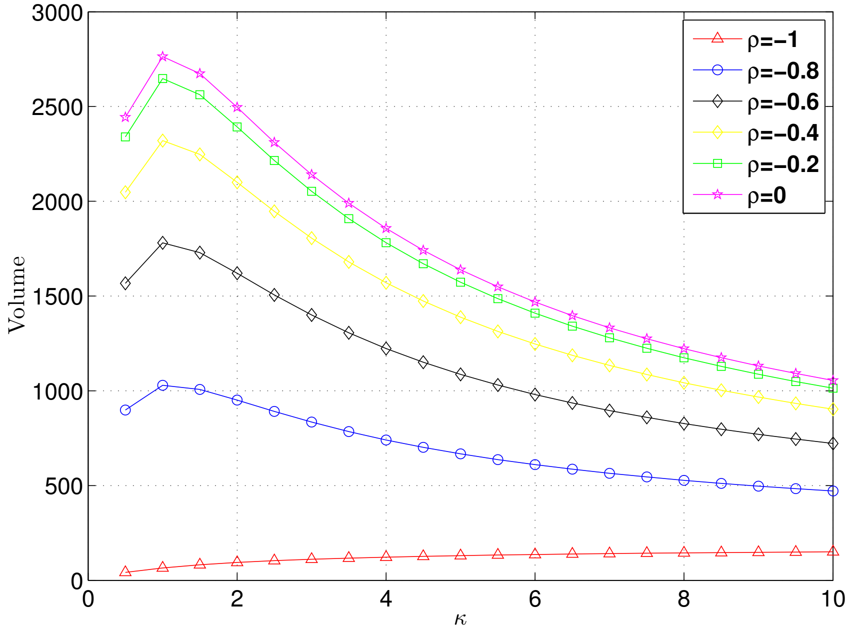

Fig. 7. Error ellipse volume for the negative correlation case

we have

$$\begin{aligned} & \left| \mathbf{I}_n + \boldsymbol{\Sigma}_K \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \\ \approx & \left| \boldsymbol{\Sigma}_K \right| \left| \mathbf{W}_K^T \mathbf{P}_{K|K}^{-1} \mathbf{W}_K \right| \end{aligned} \tag{2.81}$$

The second term in the second line of the above equation is a constant. Hence, in order to get the maximum determinant, we should set $\sigma_{b_p}^2 = \sigma_{b_v}^2 T^2$ and $\rho_{b_p,b_v} = 0$. This is almost the same solution as we have obtained numerically. Next we consider a system with two sensors. The first sensor is the same as the one described above, and the second one is with position measurement variance 4 and velocity measurement variance 5. To solve the optimization problem formulated in (2.56), we first use a systematic grid search to find an approximate globally optimal solution and then we use the FMINCON function in Matlab, a local search algorithm, to refine

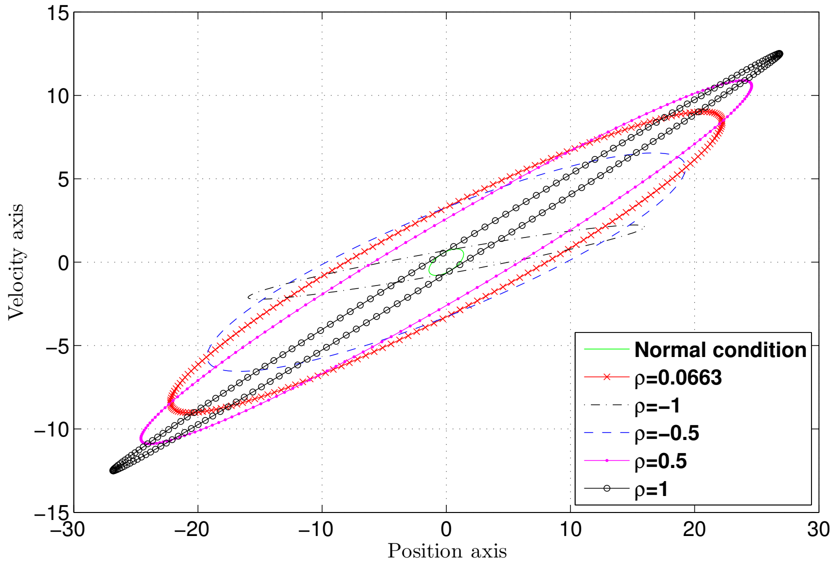

Fig. 8. Error ellipsis for different ρs

this approximate globally optimal solution. The optimal solution we have obtained is $\sigma^2_{b_{p_1}} = 1100$, $\sigma^2_{b_{p_2}} = 600$, $\sigma^2_{b_{v_1}} = 750$, $\sigma^2_{b_{v_2}} = 550$, $\rho_{b_{p_1,p_2}} = 0.99$, $\rho_{b_{p_1,v_1}} = -0.83$, $\rho_{b_{p_1,v_2}} = 0.75$, $\rho_{b_{v_1,p_2}} = 0.89$, $\rho_{b_{p_2,v_2}} = -0.23$, $\rho_{b_{v_1,v_2}} = 0.95$. For comparison purpose, we introduce three sub-optimal attack strategies: Strategy I with all the ρs being 0s, and $\sigma^2_{b_{p_1}} = 1100, \sigma^2_{b_{p_2}} = 600, \sigma^2_{b_{v_1}} = 750, \sigma^2_{b_{v_2}} = 550$; Strategy II with all the ρs being 1s, and $\sigma^2_{b_{p_1}} = 1100, \sigma^2_{b_{p_2}} = 600, \sigma^2_{b_{v_1}} = 750, \sigma^2_{b_{v_2}} = 550$; and Strategy II with the ρs being the same as those for the optimal strategy, and $\sigma^2_{b_{p_1}} = \sigma^2_{b_{p_2}} = \sigma^2_{b_{v_1}} = \sigma^2_{b_{v_2}} = 750$. The numerical results are shown in Fig. 9. As we can see, the optimal attack strategy has a greater impact than those sub-optimal attack strategies, resulting in the largest error ellipse.

2.7 Conclusion

In this chapter, we derived the EMSE due to the injected random biases for a Kalman filter in a discrete-linear dynamic system. This allows us to find how to

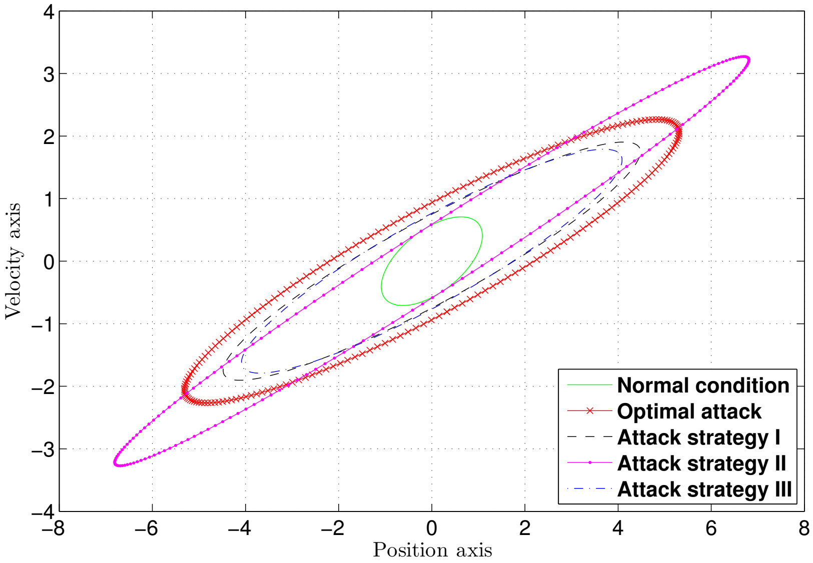

Fig. 9. Error ellipsis for different power allocation strategies

allocate the bias power among multiple sensors in order to maximize the effect of the false information on Kalman filter from two perspectives: trace and determinant. A concrete example of multi-sensor target tracking system has been provided. In this example, we investigated both the case where the sensors provide position measurements and the case where they collect both position and velocity measurements. Further, many closed-form results have been provided for the optimal attack strategies.

CHAPTER 3

FALSE INFORMATION DETECTION WITH MINIMUM MEAN SQUARED ERRORS

The problem of false information detection has not been discussed. In this chapter, the optimal Bayesian detector minimizing the average system estimation error will be investigated. For a Bayesian estimator whose sensors could be attacked by false information injected by an adversary, we investigate the strategies for the Bayesian estimator to detect the false information and defend itself from such attacks. We assume that the adversary attacks the system with certain probability, and that he/she adopts the worst possible strategy which maximizes the MSE if the attack is undetected. The defender's goal is to minimize the average system estimation MSE instead of minimizing the probability of error, as a conventional Bayesian detector typically does. The cost functions are based on the traces of the MSE matrices of the estimation error. Numerical results show that the new detection-estimation structure outperforms the traditional detectors such as the conventional Bayesian detector and the chi-squared detector significantly in terms of the average MSE. One proposed detection-estimation strategy, discarding sensor data when the presence of attack is declared, is very robust even when the attacker uses an attack strategy significantly different from the one assumed by the defender.

3.1 System Model

For a general linear and Gaussian system, the measurement $\mathbf{z}$ is supposed to be

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{3.1}$$

where $\mathbf{H}$ is the measurement matrix, $\mathbf{x}$ is the $n_x \times 1$ system state vector and $\mathbf{w}$ is the measurement noise which is supposed to be white and Gaussian. In this dissertation, we assume that a bias $\mathbf{b}$ is injected by the adversary into the sensor measurement intentionally. Therefore, the measurement equation (3.1) becomes

$$\mathbf{z}' = \mathbf{H}\mathbf{x} + \mathbf{w} + \mathbf{b} = \mathbf{z} + \mathbf{b} \tag{3.2}$$

where $\mathbf{z}'$ is the corrupted measurement, $\mathbf{b}$ is a random variable independent of $\mathbf{w}$ and $\mathbf{x}$. Therefore, the two hypotheses can be modeled as follows.

$$\begin{aligned} &H_0 : \mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{w} \\ &H_1 : \mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{w} + \mathbf{b} \end{aligned} \tag{3.3}$$

where H_0 denotes that there is no attack with prior probability $P(H_0) = p_0$, H_1 denotes the alternative hypothesis with probability $P(H_1) = p_1$. Let us suppose that the following prior information is known: $\mathbf{x} \sim \mathcal{N}(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{P}_{xx})$, $\bar{\mathbf{x}} = E(\mathbf{x})$, $\mathbf{w} \sim \mathcal{N}(\mathbf{w}; \mathbf{0}, \mathbf{P}_{ww})$, and $\mathbf{b} \sim \mathcal{N}(\mathbf{b}; \mathbf{0}, \mathbf{P}_{bb})$. The cost function is defined as follows.

$$\begin{aligned} c = {} & P(H_1)\left[P(D_1|H_1)c_1 + P(D_0|H_1)c_2\right] \\ & + P(H_0)\left[P(D_1|H_0)c_3 + P(D_0|H_0)c_4\right] \end{aligned} \tag{3.4}$$

where c is the total cost and $c_i, i \in \{1, 2, 3, 4\}$ are the cost functions which are the traces of the MSE matrices of the estimator in different scenarios: correct detection of the attack, missed detection of the attack, false alarm, and correct rejection of the attack hypothesis. $D_j|H_i, \quad i, j \in \{0, 1\}$ denotes that the detector decides D_j when the true underlying hypothesis is H_i. It is easy to show that under H_1, the probabilities of detection and miss are

$$P(D_1|H_1) = \int_{R_1} p(\mathbf{z}|H_1) d\mathbf{z} \tag{3.5}$$

$$P(D_0|H_1) = 1 - P(D_1|H_1) \tag{3.6}$$

respectively. R_1 is the decision region for D_1, and

$$p(\mathbf{z}|H_1) = |2\pi \mathbf{P}_{zz,H_1}|^{-1/2} e^{-\frac{1}{2}(\mathbf{z}-\bar{\mathbf{z}})^T \mathbf{P}_{zz,H_1}^{-1}(\mathbf{z}-\bar{\mathbf{z}})} \tag{3.7}$$

where $\bar{\mathbf{z}} = E(\mathbf{z}) = \mathbf{H}\bar{\mathbf{x}}$.

$$\begin{aligned} \mathbf{P}_{zz,H_1} &= E[(\mathbf{z}-\bar{\mathbf{z}})(\mathbf{z}-\bar{\mathbf{z}})^T] \\ &= \mathbf{H}\mathbf{P}_{xx}\mathbf{H}^T + \mathbf{P}_{ww} + \mathbf{P}_{bb} \end{aligned} \tag{3.8}$$

Similarly, under H_0, the probabilities of false alarm and its complement are

$$P(D_1|H_0) = \int_{R_1} p(\mathbf{z}|H_0) d\mathbf{z} \tag{3.9}$$

$$P(D_0|H_0) = 1 - P(D_1|H_0) \tag{3.10}$$

respectively.

$$p(\mathbf{z}|H_0) = |2\pi \mathbf{P}_{zz,H_0}|^{-1/2} e^{-\frac{1}{2}(\mathbf{z}-\bar{\mathbf{z}})^T \mathbf{P}_{zz,H_0}^{-1}(\mathbf{z}-\bar{\mathbf{z}})} \tag{3.11}$$

and

$$\mathbf{P}_{zz,H_0} = \mathbf{H}\mathbf{P}_{xx}\mathbf{H}^T + \mathbf{P}_{ww}$$

Therefore, (3.4) can be rewritten as:

$$\begin{aligned} c = p_1 c_2 + p_0 c_4 + \int_{R_1} [p_1(c_1 - c_2) p(\mathbf{z}|H_1) \\ + p_0(c_3 - c_4) p(\mathbf{z}|H_0)] d\mathbf{z} \end{aligned} \tag{3.12}$$

Clearly, in order to minimize the cost function, we should include $\mathbf{z}$ in R_1 if the integrand is negative for that value of $\mathbf{z}$.

Theorem 1 *For the problem formulated above, the optimal Bayesian detector that minimizes the average MSE, c, is*

$$p_1(c_1 - c_2)p(\mathbf{z}|H_1) + p_0(c_3 - c_4)p(\mathbf{z}|H_0) \underset{D_1}{\overset{D_0}{\gtrless}} 0. \tag{3.13}$$

where $c_i, i \in \{1, 2, 3, 4\}$ are the traces of the estimator MSE matrices in different scenarios respectively.

In the following, we consider two defending strategies and derive the optimal detector when the system adopts each strategy to defend itself.

3.1.1 Discarding Sensor Data after Detection

In this defense strategy, once the defender declares an attack either in the case of $D_1|H_1$ or $D_1|H_0$, sensor data will be discarded. Hence, the estimator is left with only the prior information about the state $\mathbf{x}$, and the trace of the MSE matrix in these two cases is

$$c_1 = c_3 = Tr(\mathbf{P}_{xx}) \tag{3.14}$$

Under $D_0|H_1$ when the system fails to detect the false information, the MSE has been derived in [25] and provided below:

$$\begin{aligned} c_2 &= Tr[\mathbf{P}_{xx} - \mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}\mathbf{P}_{zx,H_0} \\ &+\mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}\mathbf{P}_{bb}\mathbf{P}_{zz,H_0}^{-1}\mathbf{P}_{zx,H_0}] \\ &= Tr\left[\mathbf{P}_{xx} - \mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}(\mathbf{I} - \mathbf{P}_{bb}\mathbf{P}_{zz,H_0}^{-1})\mathbf{P}_{zx,H_0}\right] \end{aligned} \tag{3.15}$$

where $\mathbf{P}_{xz,H_0} = \mathbf{P}_{zx,H_0}^T = \mathbf{P}_{xx}\mathbf{H}^T$. When the defender declares no attack under H_0, which is the best case for the defender, we have

$$c_4 = Tr(\mathbf{P}_{xx} - \mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}\mathbf{P}_{zx,H_0}) \tag{3.16}$$

It is easy to show that $c_3 > c_4$ always holds. In the case where $c_2 > c_1$, we get $c_2 > c_1 = c_3 > c_4$, and the optimal detector is based on normalized distance squared, which is provided in Corollary 1. In the case where $c_2 < c_1$, the term on the left hand side of the inequality in (3.13) will always be positive leading the system to declare no attack. This is a very interesting result, which basically means that since the cost of missing the detection of the false information (c_2) is smaller than that of correctly detecting the false information (c_1), the detector will always declare D_0, even under hypothesis H_1. The derived optimal Bayesian detector for the strategy of discarding sensor data once D_1 is declared is provided in the following corollary. We name this detection-estimation strategy optimal Bayesian Detection and Discarding corrupted sensor data (OBDD).

Corollary 1 *For the defending strategy of discarding sensor data after declaring the presence of false information, under the condition $c_1 < c_2$, or equivalently $Tr\left[\mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}(\mathbf{I} - \mathbf{P}_{bb}\mathbf{P}_{zz,H_0}^{-1})\mathbf{P}_{zx,H_0}\right] < 0$, the optimal Bayesian detector that minimizes the average MSE is,*

$$(\mathbf{z} - \bar{\mathbf{z}})^T(\mathbf{P}_{zz,\mathbf{H_0}}^{-1} - \mathbf{P}_{zz,\mathbf{H_1}}^{-1})(\mathbf{z} - \bar{\mathbf{z}}) \underset{D_0}{\overset{D_1}{\gtrless}} \alpha \tag{3.17}$$

where α is

$$\alpha = 2\ln\frac{p_0(c_3 - c_4)|\mathbf{P}_{zz,H_1}|^{1/2}}{p_1(c_2 - c_1)|\mathbf{P}_{zz,H_0}|^{1/2}} \tag{3.18}$$

When $c_1 > c_2$, the optimal Bayesian detector is to always declare no attack (D_0).

Clearly, when $c_2 < c_1$, the derived optimal detector is no longer a LRT based detector.

3.1.2 Incorporating Sensor Data after Detection

In this strategy, once the defender declares the presence of false information, instead of discarding the sensor data, it will take advantage of the information from the sensor for estimation by changing the sensor model from (3.1) to (3.2). In the case of $D_1|H_1$, we have,

$$c_1 = Tr(\mathbf{P}_{xx} - \mathbf{P}_{xz,H_1}\mathbf{P}_{zz,H_1}^{-1}\mathbf{P}_{zx,H_1}) \tag{3.19}$$

where $\mathbf{P}_{xz,H_1} = \mathbf{P}_{zx,H_1}^T = \mathbf{P}_{xx}\mathbf{H}^T = \mathbf{P}_{xz,H_0}$. But this strategy will also incur more error when the system wrongly declares D_1 when H_0 is true ($D_1|H_0$), in which case we have

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{P}_{xz,\mathbf{H_1}}P_{zz,\mathbf{H_1}}^{-1}(\mathbf{z} - \bar{\mathbf{z}}) \tag{3.20}$$

and the MSE is,

$$\begin{aligned} c_3 &= Tr\left(E[(\mathbf{x}-\hat{\mathbf{x}})(\mathbf{x}-\hat{\mathbf{x}})^T]\right) \qquad (3.21)\\ &= Tr(\mathbf{P}_{xx} + \mathbf{P}_{xz,\mathbf{H_1}}\mathbf{P}_{zz,\mathbf{H_1}}^{-1}\mathbf{P}_{zz,\mathbf{H_0}}\mathbf{P}_{zz,\mathbf{H_1}}^{-1}\mathbf{P}_{zx,\mathbf{H_1}}\\ &-2\mathbf{P}_{xz,\mathbf{H_0}}\mathbf{P}_{zz,\mathbf{H_1}}^{-1}\mathbf{P}_{zx,\mathbf{H_0}}) \end{aligned}$$

The cost functions c_2 and c_4 will remain the same as in Subsection 3.1.1. Hence, we have

$$\begin{aligned} &c_2 - c_1 = \qquad (3.22)\\ &Tr\left[\mathbf{P}_{xz,H_0}(\mathbf{P}_{zz,H_1}^{-1} - \mathbf{P}_{zz,H_0}^{-1} + \mathbf{P}_{zz,H_0}^{-1}\mathbf{P}_{bb}\mathbf{P}_{zz,H_0}^{-1})\mathbf{P}_{zx,H_0}\right] \end{aligned}$$

Because $\mathbf{P}_{bb}$ is a positive semidefinite matrix, there exists a matrix $\mathbf{K}$ such that

$$\mathbf{P}_{zz,\mathbf{H_1}} = \mathbf{P}_{zz,\mathbf{H_0}} + \mathbf{P}_{bb} = \mathbf{P}_{zz,\mathbf{H_0}} + \mathbf{K}\mathbf{K}^T \tag{3.23}$$

Based on (3.23), denoting $\mathbf{A} = \mathbf{P}^{-1}_{zz,\mathbf{H_1}}$, we have

$$\begin{aligned}&\mathbf{A} = \\ &\mathbf{P}^{-1}_{zz,H_0} - \mathbf{P}^{-1}_{zz,H_0}\mathbf{K}(\mathbf{I} + \mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0}\mathbf{K})^{-1}\mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0}\end{aligned} \tag{3.24}$$

Denoting $\mathbf{B} = \mathbf{P}^{-1}_{zz,H_0} - \mathbf{P}^{-1}_{zz,H_0}\mathbf{K}\mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0}$, then we have

$$\mathbf{A} - \mathbf{B} = \mathbf{P}^{-1}_{zz,H_0}\mathbf{K}[\mathbf{I} - (\mathbf{I} + \mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0}\mathbf{K})^{-1}]\mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0} \tag{3.25}$$

Since $\mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0}\mathbf{K}$ is a positive semidefinite matrix, according to the spectral theorem, there exist an orthogonal matrix $\mathbf{U}$ and a real diagonal matrix $\mathbf{\Lambda}$ such that

$$\begin{aligned}&\mathbf{A} - \mathbf{B} = \mathbf{P}^{-1}_{zz,H_0}\mathbf{K}[\mathbf{I} - (\mathbf{I} + \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T)^{-1}]\mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0} \\ &= \mathbf{P}^{-1}_{zz,H_0}\mathbf{K}\mathbf{U}[\mathbf{I} - (\mathbf{I} + \mathbf{\Lambda})^{-1}]\mathbf{U}^T\mathbf{K}^T\mathbf{P}^{-1}_{zz,H_0}\end{aligned} \tag{3.26}$$

where $\mathbf{I} - (\mathbf{I} + \mathbf{\Lambda})^{-1}$ has positive diagonal entries. We can see from the formula above, $\mathbf{A} - \mathbf{B}$ is still a positive semidefinite matrix, so $Tr(\mathbf{A} - \mathbf{B})$ is positive, and $c_2 - c_1 > 0$. Now let us consider the sign of $c_3 - c_4$. It can be shown that

$$\begin{aligned}&c_3 - c_4 \\ &= Tr(\mathbf{P}_{xz,H_0}\mathbf{P}^{-1}_{zz,H_0}\mathbf{P}_{zx,H_0} \\ &+\mathbf{P}_{xz,\mathbf{H_1}}\mathbf{P}^{-1}_{zz,\mathbf{H_1}}\mathbf{P}_{zz,\mathbf{H_0}}\mathbf{P}^{-1}_{zz,\mathbf{H_1}}\mathbf{P}_{zx,\mathbf{H_1}} \\ &-2\mathbf{P}_{xz,\mathbf{H_0}}\mathbf{P}^{-1}_{zz,\mathbf{H_1}}\mathbf{P}_{zx,\mathbf{H_0}}) \\ &= Tr[\mathbf{P}_{xz,H_0}(\mathbf{P}^{-1}_{zz,H_0} + \mathbf{P}^{-1}_{zz,\mathbf{H_1}}\mathbf{P}_{zz,\mathbf{H_0}}\mathbf{P}^{-1}_{zz,\mathbf{H_1}} \\ &-2\mathbf{P}^{-1}_{zz,\mathbf{H_1}})\mathbf{P}_{zx,H_0}] \\ &= Tr(\mathbf{P}_{xz,\mathbf{H_0}}\mathbf{P}^{-1}_{zz,\mathbf{H_1}}\mathbf{P}_{bb}\mathbf{P}^{-1}_{zz,\mathbf{H_0}}\mathbf{P}_{bb}\mathbf{P}^{-1}_{zz,\mathbf{H_1}}\mathbf{P}_{zx,\mathbf{H_0}})\end{aligned} \tag{3.27}$$

Clearly, the matrix inside the trace operator in (3.27) is positive semidefinite, and we have $c_3 > c_4$. The Optimal Bayesian Detector for the strategy of detection and Incorporating sensor data (OBDI) is provided in the following corollary.

Corollary 2 *For the defending strategy of incorporating sensor data after declaring the presence of false information, knowing that $c_3 > c_4$ and $c_2 > c_1$, the optimal Bayesian detector that minimizes the average MSE is,*

$$(\mathbf{z}-\bar{\mathbf{z}})^T(\mathbf{P}_{zz,\mathbf{H_0}}^{-1}-\mathbf{P}_{zz,\mathbf{H_1}}^{-1})(\mathbf{z}-\bar{\mathbf{z}}) \underset{D_0}{\overset{D_1}{\gtrless}} \alpha \tag{3.28}$$

where α is

$$\alpha = 2\ln\frac{p_0(c_3-c_4)|\mathbf{P}_{zz,H_1}|^{1/2}}{p_1(c_2-c_1)|\mathbf{P}_{zz,H_0}|^{1/2}} \tag{3.29}$$

3.2 Minimum Mean Square Error (MMSE) Estimator

Given all the system information, we can also derive the MMSE estimator of the system state. Using Bayes' rule, it could be shown that the MMSE estimator, or the conditional mean is

$$\begin{aligned} E(\mathbf{x}|\mathbf{z}) &= \int \mathbf{x}p(\mathbf{x}|\mathbf{z})d\mathbf{x} \qquad (3.30)\\ &= \frac{p_0 p(\mathbf{z}|H_0)}{p(\mathbf{z})}\left[\bar{\mathbf{x}}+\mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}(\mathbf{z}-\bar{\mathbf{z}})\right]\\ &\quad+\frac{p_1 p(\mathbf{z}|H_1)}{p(\mathbf{z})}\left[\bar{\mathbf{x}}+\mathbf{P}_{xz,H_1}\mathbf{P}_{zz,H_1}^{-1}(\mathbf{z}-\bar{\mathbf{z}})\right] \end{aligned}$$

where

$$p(\mathbf{z}) = p_0 p(\mathbf{z}|H_0) + p_1 p(\mathbf{z}|H_1) \tag{3.31}$$

$p(\mathbf{z}|H_0) = \mathcal{N}(\mathbf{z};\mathbf{H}\bar{\mathbf{x}},\mathbf{P}_{zz,H_0})$, and $p(\mathbf{z}|H_1) = \mathcal{N}(\mathbf{z};\mathbf{H}\bar{\mathbf{x}},\mathbf{P}_{zz,H_1})$.

We show later in the chapter that even the MMSE estimator gives the best estimation performance in terms of MSE, it is not robust and does not perform very well when the true system parameters deviate from the nominal parameters. On the other hand, the OBDD approach provides robust performance even when there is a mismatch between the nominal and the actual parameters.

3.3 Numerical Results

In this section, the optimal Bayesian detectors are applied both in a one-dimensional tracking system and a static parameter estimation system to detect false information. They are compared with other widely used detection/estimation strategies, such as the CBD that minimizes the probability of error (P_e), the chi-square detector (CSD), and the MMSE estimator.

3.3.1 Target Tracking Example

The state equation used in Kalman filter target tracking system is [31]:

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{v}_k \tag{3.32}$$

where $\mathbf{x}_k = [\xi \quad \dot{\xi}]$ is the system state vector at time k, and ξ and $\dot{\xi}$ represent the target's position and velocity along the ξ-axis at time k respectively. The state transition matrix is

$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{3.33}$$

where $T = 1\ s$ is the time interval between measurements. The process noise is $\mathbf{v}_k = \mathbf{\Gamma} v_k$, where v_k is a zero mean white acceleration noise, with variance σ_v^2. In this example, we set $\sigma_v^2 = 0.25$, and the vector gain multiplying the scalar process noise is given by $\mathbf{\Gamma}^T = [T^2/2 \quad T]$. The covariance matrix of the process noise is therefore

$\mathbf{Q} = \sigma_v^2 \mathbf{\Gamma}\mathbf{\Gamma}^T$. Under H_1, the measurement equation is provided as follows.

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k + \mathbf{b}_k \tag{3.34}$$

where $\mathbf{H}_k$ is a 2×2 identity matrix. Suppose that Kalman filter runs 200 iterations and the false information $\mathbf{b}_k$ is injected to Kalman filter system at time $k = 100$ with probability $p_1 = 0.85$, and the total bias power for the position and velocity takes different values in the following range: $\sigma_{b_p}^2 + T^2 \sigma_{b_v}^2 = a^2 \in [7.5, 120]$. The false information $\mathbf{b}$ is zero-mean Gaussian random noise. The optimal covariance matrix $\mathbf{P}_{bb}$ for $\mathbf{b}_k$ has been derived in [25], which maximizes c_2, i.e.

$$Tr\left[\mathbf{P}_{xx} - \mathbf{P}_{xz,H_0}\mathbf{P}_{zz,H_0}^{-1}(I - \mathbf{P}_{bb}\mathbf{P}_{zz,H_0}^{-1})\mathbf{P}_{zx,H_0}\right] \tag{3.35}$$

as given in (3.15). We assume that the adversary uses this $\mathbf{P}_{bb}$ to attack Kalman filter system. The effect of the bias injection on Kalman filter is measured by the average MSE over $N_m = 10000$ Monte-Carlo runs,

$$\frac{\sum_{j=1}^{N_m} [\mathbf{x} - \hat{\mathbf{x}}]^T [\mathbf{x} - \hat{\mathbf{x}}]}{N_m} \tag{3.36}$$

The optimal Bayesian detectors either discarding or incorporating sensor data after they make a decision of D_1 are compared with the CSD and the CBD. The CSD we use is

$$(\mathbf{z} - \hat{\mathbf{z}})^T \mathbf{P}_{zz,H_0}^{-1} (\mathbf{z} - \hat{\mathbf{z}}) \underset{D_0}{\overset{D_1}{\gtrless}} 5.99 \tag{3.37}$$

with a false alarm rate of $P_{fa} = 0.05$, and the CBD, which minimizes the probability of error, is

$$\frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{D_0}{\overset{D_1}{\gtrless}} \frac{p_0}{p_1} \tag{3.38}$$

Similar to the OBDD, in both the CBD and CSD, once a decision D_1 is declared, the sensor data will be discarded,

It is clear from Figs. 10, 11, and 12, the OBDI leads to the smallest MSE. This is because it takes advantage of all the sensor data even when D_1 is declared.

When the false information power is low, the OBDD has a smaller MSE than the CBD, even though the former has a larger P_e than the latter. The reason for the OBDD's larger P_e is because when the false information power is small, it will always declare no attack (D_0) to minimize the MSE instead of P_e. This is also clear from Fig. 11, in which for small false information power, both P_{fa} and P_d (probability of detection) for OBDD are zeros. The CSD gives the worst performance in terms of the average MSE, when the false information power is large, this is because it does not use the prior information of p_0 and p_1, or the information about $\mathbf{P}_{bb}$, and it has a poorer P_d than other detectors, when the false information power is large. With a large prior probability $p_1 = 0.85$ for hypothesis H_1, to minimize P_e, the CBD always declares the presence of an attack (D_1) in this particular example.

3.3.2 Parameter Estimation Example

The second example involves a static parameter estimation system. The prior information about a parameter $\mathbf{x}$ is $\bar{\mathbf{x}} = [10, \;\; 5]^T$, $\mathbf{P}_{xx} = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}$, the measurement matrix $\mathbf{H} = \mathbf{I}$ is a 2×2 identity matrix, and $\mathbf{P}_{ww} = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$. The false information $\mathbf{b}$ is injected to the system with $p_1 = 0.85$, and $\mathbf{P}_{bb} = \begin{bmatrix} \sigma_{b_1}^2 & 0 \\ 0 & \sigma_{b_2}^2 \end{bmatrix}$. The false information power is $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2 \in [0, 400]$. It can be shown that the optimal attack strategy that maximizes c_2 in (3.15) is $\mathbf{P}_{bb} = \begin{bmatrix} a^2 & 0 \\ 0 & 0 \end{bmatrix}$, which is used by the adversary to attack the system. From Fig. 13, it is clear that the MMSE estimator leads to the smallest MSE, the OBDI has a performance which is very close to the MMSE estimator, the OBDD provides the third smallest MSE. Again, the OBDD scarifies P_e performance to achieve a smaller MSE than the CBD when the false information power is small. The CSD provides a better MSE performance than the CBD when the false information power is small, but a larger MSE when the false information power becomes larger.

3.3.3 Robustness Analysis

In this subsection, we assume that the setting is almost the same as that in Subsection 3.3.2. Let us suppose that the defender uses the nominal $\mathbf{P}_{bb}$ to design the various detectors or the MMSE estimator, assuming that the adversary puts all the power to the measurement with the smaller variance. However, the adversary's actual power allocation strategy is just the opposite by injecting all the power to the other measurement. The false information power is $\sigma_{b_1}^2 + \sigma_{b_2}^2 = a^2 \in [0, 400]$. Simulation results show that the OBDD has the best performance in this case, as illustrated in Fig. 14. This is because the OBDD will discard the sensor data once it declares D_1, which makes it less susceptible to the mismatch in the system model. As for the CBD and CSD, since they will discard the sensor data once they declare the presence of an attack, their performance will not affected much by the model mismatch either. Their results are not provided in Fig. 14 for the ease of presentation. On the other hand, since both the MMSE estimator and the OBDI try to incorporate the sensor data even when D_1 is declared, their performances are significantly degraded as shown in Fig. 14.

3.4 Conclusion

For a Bayesian estimation system whose sensors are attacked by false information injected by an adversary, we have derived the optimal Bayesian detection strategies which help the system achieve the smallest average estimation MSE. The proposed Bayesian detectors minimize the average MSE instead of the probability of error, and they may not be the LRT based detectors any more. Different defending scenarios cases: either discarding or taking advantage of sensor data declared to be compromised by the false information were investigated. Numerical results show that the derived Bayesian detectors lead to significantly smaller average MSE than the traditional detectors, such as the conventional Bayesian detector and chi-squared detector. In addition, the optimal Bayesian detector coupled with the defending strategy of discarding sensor data once the presence of an attack is declared, proves to be very robust to the mismatch between the model assumed by the defender and that actually adopted by the attacker.

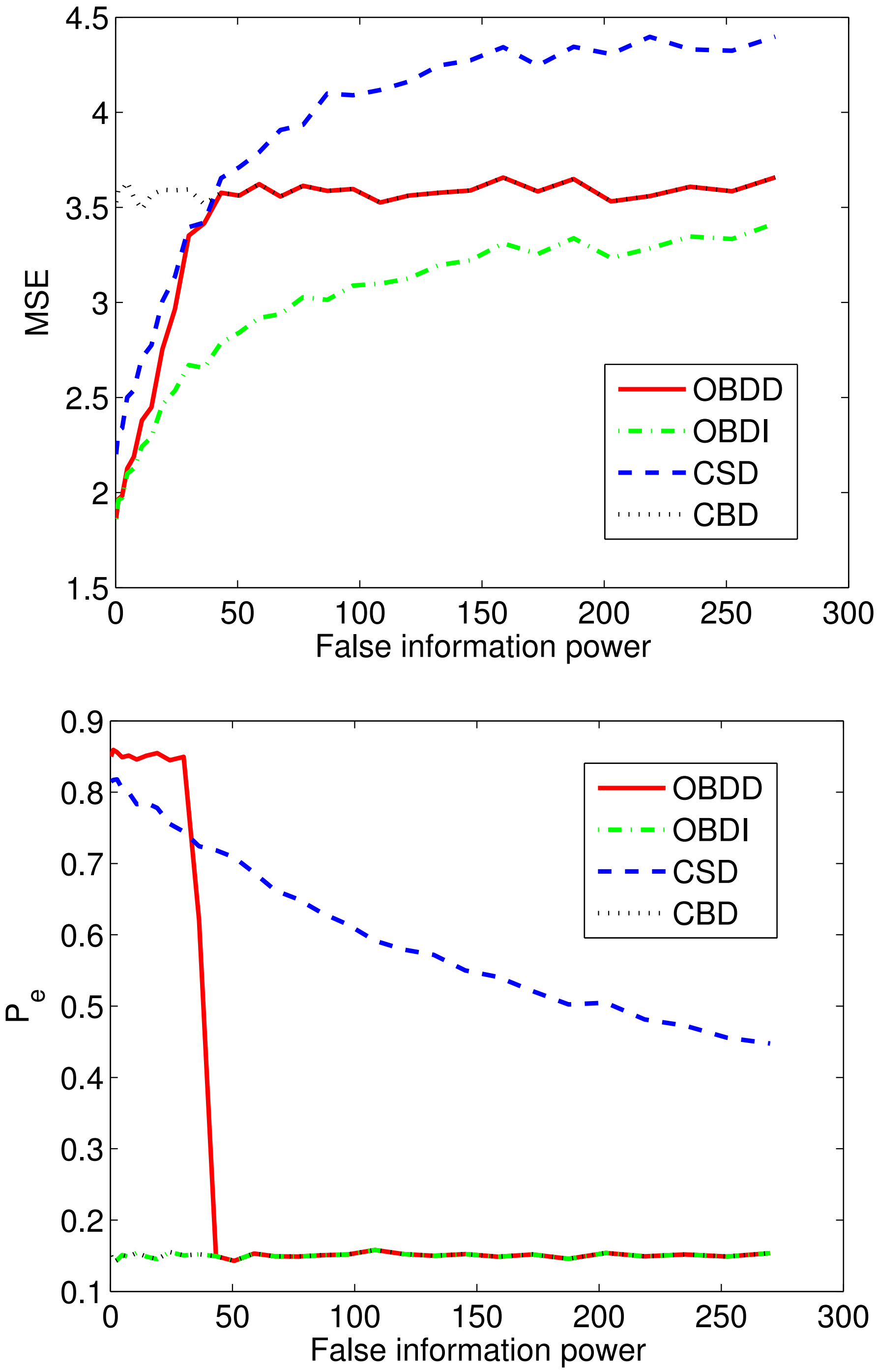

Fig. 10. Performances of different detectors

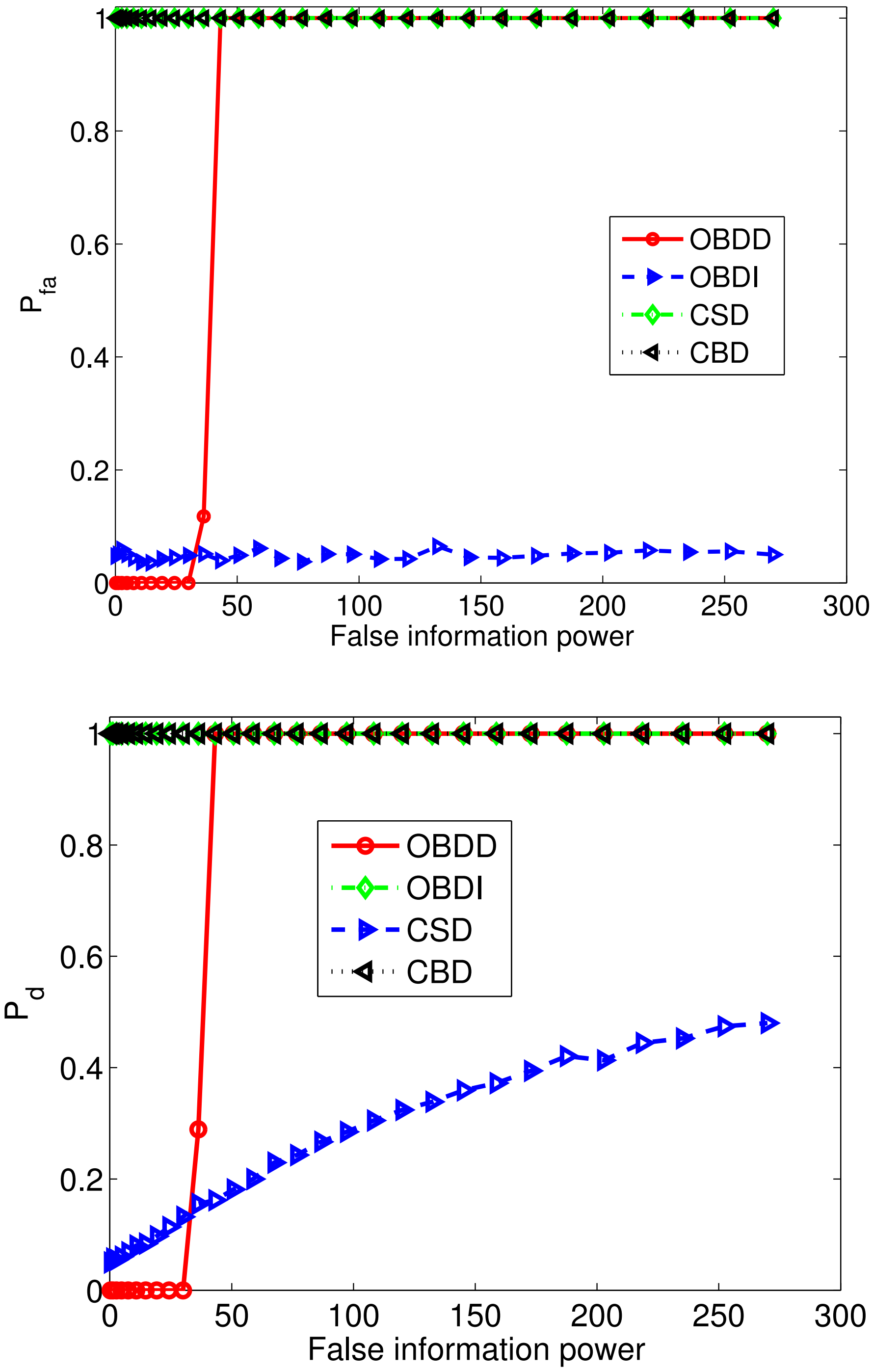

Fig. 11. Probabilities of false alarm and detection

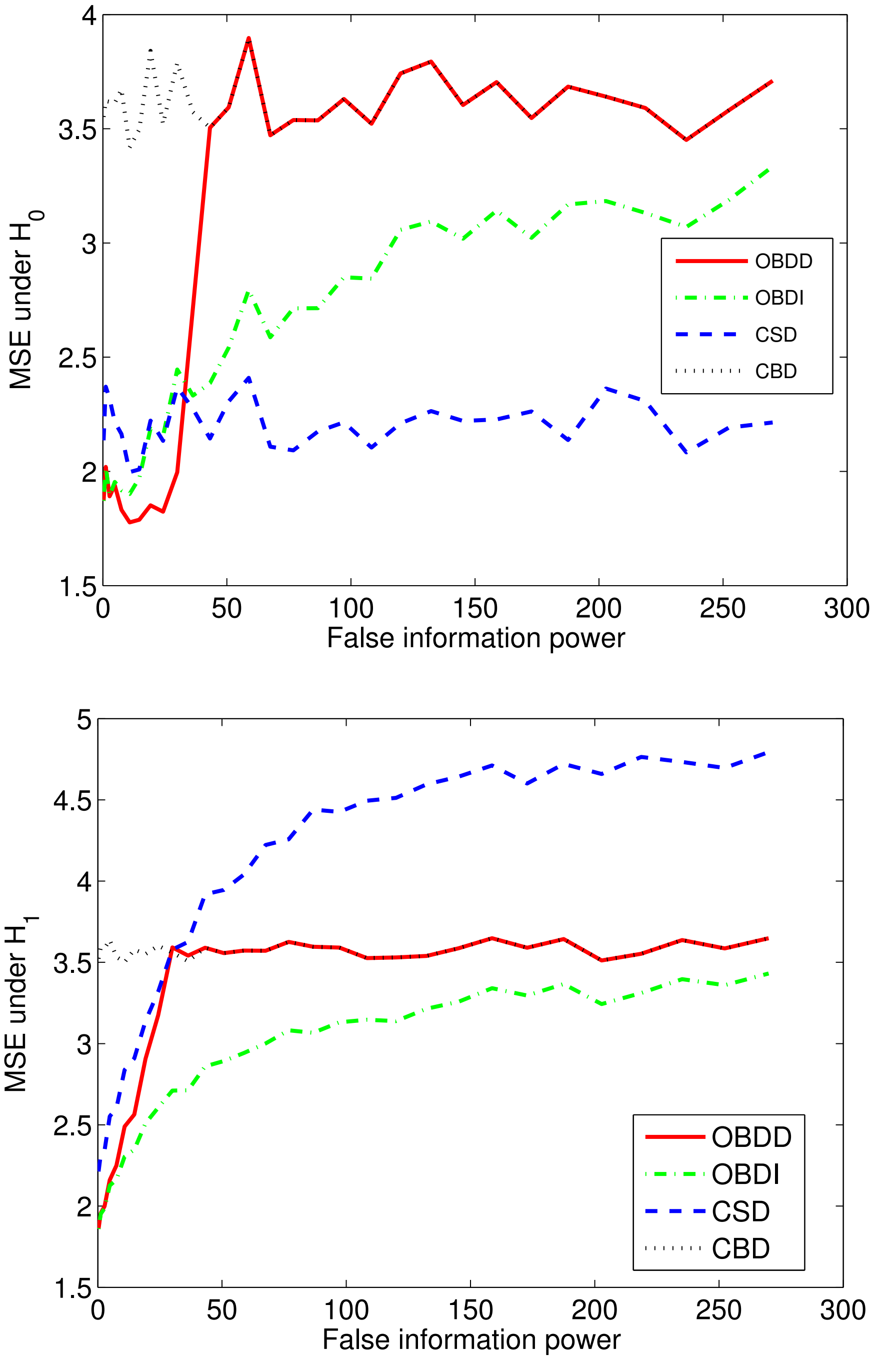

Fig. 12. MSEs under no attack ($\mathbf{H_0}$) and attack ($\mathbf{H_1}$)

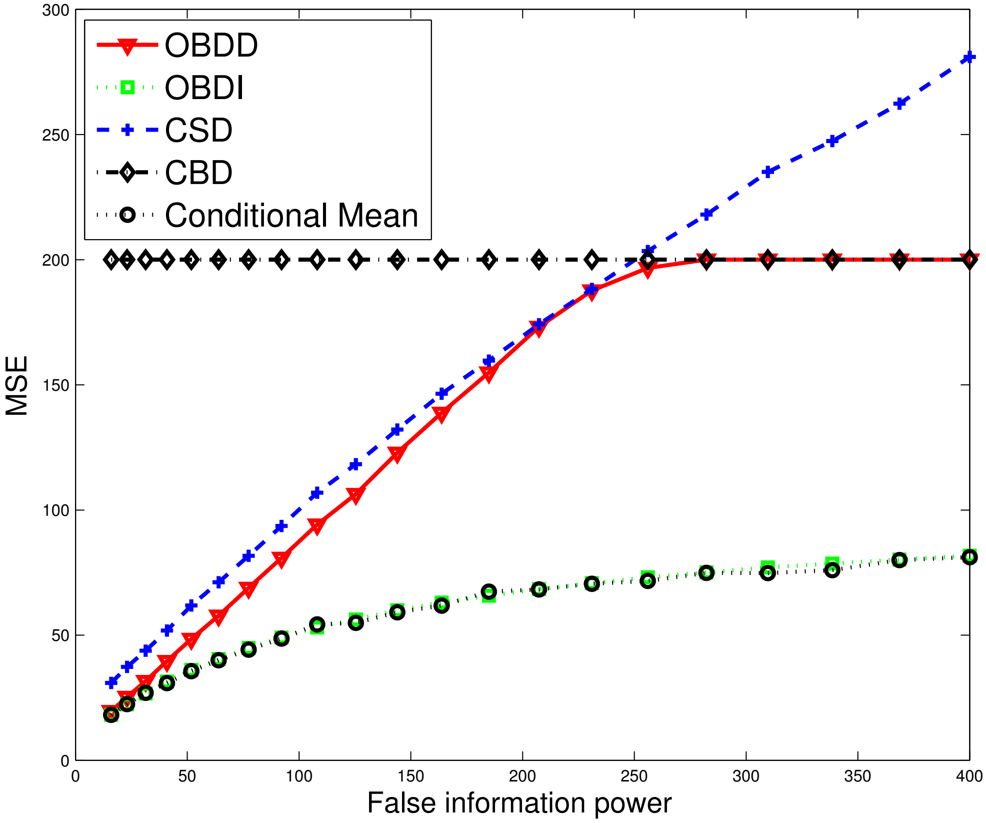

Fig. 13. Performances of different detector-estimation strategies

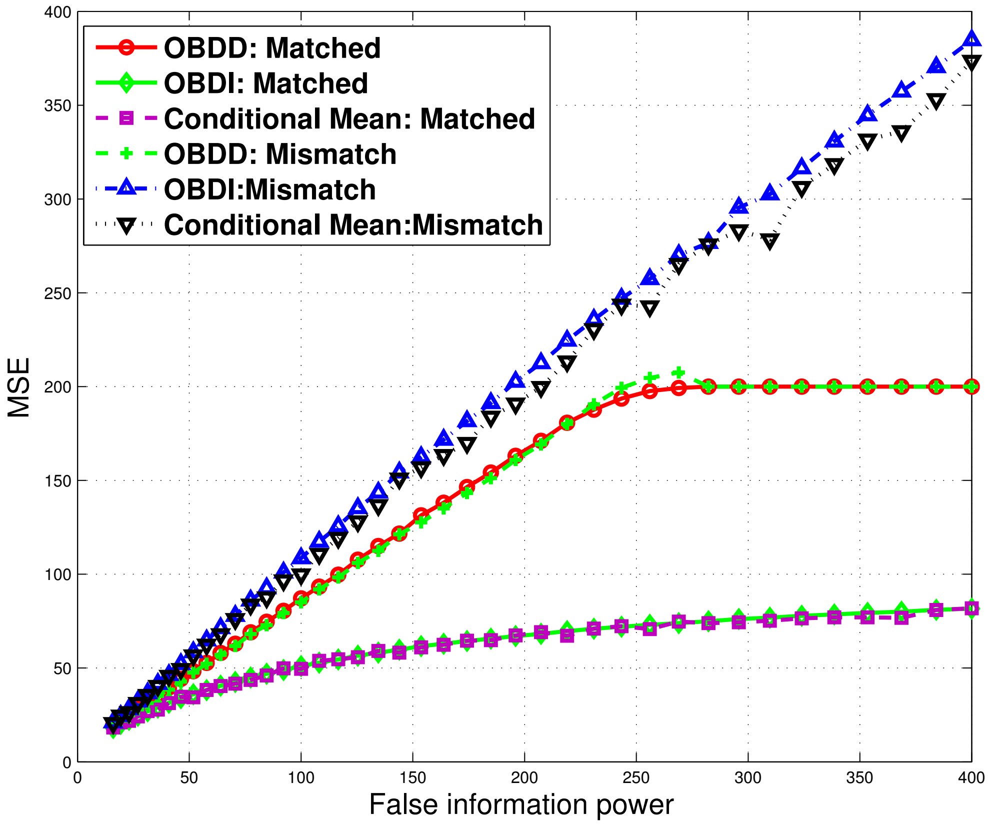

Fig. 14. Robustness analysis

CHAPTER 4

SPARSE ATTACK STRATEGIES

So far we have discussed the optimal strategies the adversary can adopt to attack the system under the assumption that the control center is unaware of the existence of the false information. In this chapter, it is assumed that the system can perfectly detect and remove sensors once they are corrupted by false information injected by an adversary. The adversary aims to maximize the covariance matrix of the system state estimate by the end of the attack period under the constraint that the adversary can only attack the system a few times over time and over sensors, which leads to an integer programming problem.

4.1 Problem Formulation

Based on notations introduced in Chapter 2 and using the information filter form [31] for Kalman filter, the state prediction covariance at time $k+1$, denoted as $\mathbf{P}_{k+1|k}$, is shown below,

$$\mathbf{P}_{k+1|k} = \mathbf{F}_k \mathbf{P}_{k|k} \mathbf{F}_k^T + \mathbf{Q}_k \tag{4.1}$$

and the updated state covariance matrix at time $k+1$, denoted as $\mathbf{P}_{k+1|k+1}$, can be obtained as

$$\mathbf{P}_{k+1|k+1}^{-1} = \mathbf{P}_{k+1|k}^{-1} + \sum_{i=1}^{M} \mathbf{H}_{k+1,i}^T \mathbf{R}_{k+1,i}^{-1} \mathbf{H}_{k+1,i} \tag{4.2}$$

It is assumed that the system has perfect detection of the existence of the false information, that is to say, its false information detector's probability of false alarm is 0 and probability of detection is 1. The adversary needs to decide when and which sensors to attack under the sparsity constraint, which leads to an integer programming problem. It is assumed that the system has M sensors and the adversary attacks the system from time $K+1$ to time $K+N$. The active sensor set, which includes the sensors being attacked by the adversary at time $k \in \{K+1, \cdots, K+N\}$, is denoted as A_k, where $0 \leq |A_k| \leq M$, and $|\cdot|$ is the cardinality of a set. $A = \cup A_k$ is the sensor set that includes the sensors attacked by the adversary over time. The active set A_k is designed in order to maximize the system estimation error under the sparsity constraint $|A| = c$. Base on the perfect detection assumption, if one sensor is attacked at certain time, Kalman filter will not use the measurement from that sensor at that time to perform system state estimation. Define the sensor set $D = \{s_1, ..., s_M\}$, where s_i denotes the ith sensor. For each time $k \in \{K+1, \cdots, K+N\}$, the inverse of the updated state covariance matrix is provided as follows

$$\mathbf{P}_{k|k}^{-1} = \mathbf{P}_{k|k-1}^{-1} + \sum_{i \in D \backslash A_k} \mathbf{H}_{k,i}^T \mathbf{R}_{k,i}^{-1} \mathbf{H}_{k,i} \tag{4.3}$$

The adversary aims to maximize the state estimation error covariance matrix $\mathbf{P}_{K+N|K+N}$ by the end of the attack period, and the problem can be formulated as follows,

$$\begin{aligned} &\max_{A} \mathit{\Phi} \left(\mathbf{P}_{K+N|K+N} \right) \\ &s.t. \quad |A| = c \end{aligned} \tag{4.4}$$

where function $\mathit{\Phi}(\cdot)$ could be either trace or determinant of a matrix and $\mathbf{P}_{K+N|K+N}$ is calculated iteratively using (4.1) and (4.3). That is to say, a subset is chosen out of the whole option set so that the object function will be maximized, leading to the largest estimation error. The optimal solution can be obtained by using the exhaustive search to check all the different sensor combinations. For each candidate sensor combination, N iterations have to be performed to evaluate Kalman filter's covariance matrix over time, and for each iteration, there are roughly M matrix additions as shown in (4.3), leading to a complexity of $n = MN$. The complexity for the exhaustive search algorithm is therefore

$$\varphi_1(n) = n\frac{n!}{(n-c)!c!} = n\prod_{i=1}^{c}\frac{n+1-i}{i} \tag{4.5}$$

4.1.1 Greedy Search Based Approaches

Concerning the high complexity of the exhaustive algorithm, it will be infeasible to find the optimal solution as the size of problem increases. Some suboptimal algorithms, including sequential forward selection (SFS), sequential backward selection (SBS), and SFS improved by the simplex approach (SFS-SS) [35, 36] are proposed to find the attack strategies. The SFS starts with an empty set for A, and one sensor is added at each iteration, whose elimination from the system will lead to the maximum MSE. This process terminates when $|A|$ reaches c. The pseudo code of the SFS algorithm is provided in Algorithm 1. The complexity of the SFS is provided below

$$\begin{aligned}\varphi_2(n) &= n\left[n + (n-1) + ... + (n-c+1)\right] \\ &= \frac{2cn^2 - c(c-1)n}{2}\end{aligned} \tag{4.6}$$

which has a complexity of $O(n^2)$.

SBS solves the problem in the opposite direction. The SBS starts with a set A containing all the sensors over all the time steps, and one sensor is reduced at each iteration, whose addition to the system will lead to the minimum reduction in the state estimation's MSE. This process terminates when $|A|$ reaches c. The complexity of SBS is $O(n^3)$:

$$\begin{aligned} \varphi_3(n) &= n\left[n + n - 1 + ... + c + 1\right] \\ &= \frac{n^3 + n^2 - (c^2 + c)n}{2} \end{aligned} \tag{4.7}$$

Comparing (4.7) to (4.6), it is clear that the SFS is preferable in terms of computational complexity.

Algorithm 1 Sequential Forward Selection

```
1: $A^0$=$\emptyset$; $j = 0$
2: $ind^+ = arg\max_{ind \notin A^j} \Phi\left(A^j \cup \{ind\}\right)$
3: update $A^{j+1} = A^j \cup \{ind^+\}$
4: $j = j + 1$
5: if $j < c$, go to 2
6: end
```

As for SFS-SS, it tries to improve the suboptimal solution found by the SFS. SFS-SS works by checking whether replacing a sensor in the active set with a sensor in the inactive set will increase the system estimation error. The index of active set $A_{initial}$ achieved from SFS is sorted in the order the sensors are selected by the SFS. The SFS-SS starts from the $(c-1)_{th}$ sensor in the active set and checks whether replacing this sensor with any sensor in the inactive set will increase the system estimation error. If no improvement is found, the next sensor in the active set will be checked. Otherwise, the sensor in $A_{initial}$ is replaced with the sensor found from the inactive set and the c_{th} sensor is to be checked in the next iteration. Once the first sensor in $A_{initial}$ is checked and no more improvement is found, the algorithm terminates. The pseudo code of the SFS-SS is provided in Algorithm 2.

4.1.2 Dynamic Programming

Another computationally tractable suboptimal solution to the formulated integer programming problem is dynamic programming (DP). Consider the cases where once

Algorithm 2 Simplex Improved SFS

```
1: $A$=$A_{initial}$
2: $i = c - 1$
3: $s^+ = arg\max_{ind \notin A} \mathit{\Phi}\left((A \backslash \{i\}) \cup \{ind\}\right)$
4: **if** $\mathit{\Phi}\left((A \backslash \{i\}) \cup \{s^+\}\right) > \mathit{\Phi}(A)$ **then**
5:     Update $A = (A \backslash \{i\}) \cup \{s^+\}$, $i = c$
6: **else**
7:     $i = i - 1$
8: if $i > 0$, go to 3
9: end
```

$\mathbf{P}_{k|k-1}$ and $|A_k|$ are given, the optimal attack strategy at snapshot k is certain, i.e., A_k can be determined without enumerating all the M-choosing- $|A_k|$ combinations. Such cases are common for systems with scalar-valued measurements, where the optimal attack strategy is to attack the A_k sensors with the smallest measurement variances. For such systems, we can develop a DP algorithm, which can also be performed in polynomial time, and is optimal if the system state is scalar-valued.

First, note that the predicted state covariance matrix $\mathbf{P}_{k|k-1}$ in (4.2) or (4.3) satisfies

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}, \tag{4.8}$$

which is only a function of $\mathbf{P}_{k-1|k-1}$. Combining (4.8) and (4.3), we can see that $\mathbf{P}_{k|k}$ is a function of $\mathbf{P}_{k-1|k-1}$ and $|A_k|$, namely

$$\mathbf{P}_{k|k} = \mathcal{KF}(\mathbf{P}_{k-1|k-1}, |A_k|). \tag{4.9}$$

The validity of DP lies on a straightforward but important nature of dynamic state estimation systems — the uncertainty $\mathit{\Phi}(\mathbf{P}_{k|k})$ of the current estimation, is generally increasing with that of previous estimation, $\mathit{\Phi}(\mathbf{P}_{k-1|k-1})$. It indicates that when A_k is fixed, in order to maximize $\mathit{\Phi}(\mathbf{P}_{k|k})$ subject to $|A_{K+1}| + |A_{K+2}| + \cdots + |A_k| = s$, one first needs to solve a subproblem that maximizes $\mathit{\Phi}(\mathbf{P}_{k-1|k-1})$ subject to $|A_{K+1}| + |A_{K+2}| + \cdots + |A_{k-1}| = s - |A_k|$. Then among all the feasible choices of

$|A_k|$, we choose the one that corresponds to the largest $\varPhi(\mathbf{P}_{k|k})$, i.e.,

$$\begin{aligned}&\max_{\sum_{j=K+1}^{k}|A_j|=s} \varPhi(\mathbf{P}_{k|k}) = \\ &\max_{|A_k|}\mathcal{KF}\left(\max_{\sum_{j=K+1}^{k-1}|A_j|=s-|A_k|} \varPhi(\mathbf{P}_{k-1|k-1}), |A_k|\right).\end{aligned} \tag{4.10}$$

Denote

$$\begin{aligned}&\Phi_{DP}(s,k) = \max_{\sum_{j=K+1}^{k}|A_j|=s} \varPhi(\mathbf{P}_{k|k}), \\ &\forall s \in \{0,1,\cdots,c\}, k \in \{K+1,\cdots,K+N\},\end{aligned} \tag{4.11}$$

equation (4.10) becomes

$$\begin{aligned}&\Phi_{DP}(s,k) = \max_{0\le r\le s}\mathcal{KF}\left(\Phi_{DP}(s-r,k-1),r\right), \\ &\forall s \in \{0,1,\cdots,c\}, k \in \{K+1,\cdots,K+N\}.\end{aligned} \tag{4.12}$$

By definition, the optimal solution of (4.4) is $\Phi_{DP}(c, K+N)$. From (4.12) we can see that in order to obtain $\Phi_{DP}(c, K+N)$, we need to compute and store all the $(c+1) \times N$ values of $\Phi_{DP}(s,k)$. Furthermore, in order to trace back the optimal attack strategy, we also need to store another $(c+1) \times N$ numbers, which are given by

$$\begin{aligned}&C(s,k) = \arg\max_{0\le r\le s}\mathcal{KF}\left(\Phi_{DP}(s-r,k-1),r\right), \\ &\forall s \in \{0,1,\cdots,c\}, k \in \{K+1,\cdots,K+N\}.\end{aligned} \tag{4.13}$$

In this way, once $\Phi_{DP}(c, K+N)$ is obtained, the best attack strategy can be found by

$$|A_k| = C(c-|A_{k+1}|,k), \forall k \in \{K+1,\cdots,K+N-1\}, \tag{4.14}$$

where

$$|A_{K+N}| = C(c, K+N). \tag{4.15}$$

Therefore, the memory cost of DP is proportional to $2N(c+1) = O(Nc)$.

Run-time Complexity: It is clear that calculating the first column in Φ_{DP}, i.e., $\Phi_{DP}(s, K+1)$, $\forall s \in \{0, 1, \cdots, c\}$, has complexity on the order of $M(c+1)$. This is because given the steady state $\mathbf{P}_{K|K}$, it only takes $c+1$ repetitions of (4.8) and (4.3) to calculate the first column of Φ_{DP}. When $k > K+1$, there are $s+1$ possible values of $|A_k|$ for determining $\Phi_D P(s, k)$, which needs totally $1+2+\cdots+(c+1) = (c+1)(c+2)/2$ repetitions of (4.8) and (4.3). Therefore, the run-time complexity of DP is

$$\phi_4(n) = M(c+1) + (N-1)M\frac{(c+1)(c+2)}{2} = O(nc^2). \tag{4.16}$$

As we can see, if c keeps constant, then DP is a linear-time algorithm and has smaller complexity than greedy algorithms.

4.2 Numerical Results

Numerical results for a target tracking example are presented in this section to illustrate the effectiveness of the proposed suboptimal solutions. Two cases involving position sensors and position-velocity sensors are presented to show the attack strategies of the adversary under different sensor configurations.

4.2.1 System with Position Sensors

For the system with position sensors, the parameters used in the target tracking example are provided below. The system has $M = 3$ position sensors with sampling interval $T = 1$. The system input $\mathbf{u}_k = \mathbf{0}$. The system state transition matrix is

$$\mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{4.17}$$

The measurement matrix for each sensor is

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{4.18}$$

The standard deviation (s.d.) of the system process noise is $\sigma_v = 0.02$. The s.d.s of the measurement noise for the three sensors are $\sigma_{w_1} = 0.2$, $\sigma_{w_2} = 0.4$, and $\sigma_{w_3} = 0.5$, respectively. The sparsity constraint for the adversary is $c = 5$, meaning that the adversary has to choose 5 spots to attack the system over $M = 3$ sensors and over $N = 6$ time steps in order to maximize the trace of the state covariance matrix by the end of the attack period.

To begin with, SFS is used to find the suboptimal solution. Table 1 shows the found attack strategy. The numbers shown in the table denote the order of sensors for the adversary to attack. The reason why the adversary attacks the first sensor is Sensor 1 has the smallest measurement variance. Examing (4.3), the second item is a diagonal matrix, with only position variance on the diagonal. In order to maximize the trace of $\mathbf{P}_{K+6|K+6}$, in each iteration, it is better to minimize the matrix $\mathbf{P}_{k|k}^{-1}$. The result shows that the adversary attacks Sensor 1 from time $K+3$ to time $K+6$. The interesting thing for this method is that it also provides the adversary with an attack strategy if he/she wants to attack the system less than c times because of the greedy nature of the SFS. Another observation is that the attacker tends to attack sensors in the times near the end, which is due to the "forgetting" property of Kalman filter, implying that the sensor data in the past will become less and less important as time goes on.

Table 1. Attack Strategy for Position Sensors

$Time/Sensor$	K+1	K+2	K+3	K+4	K+5	K+6
Sensor 1			5	3	2	1
Sensor 2						4
Sensor 3						

For the same parameter setup, different optimization algorithms including DP and exhaustive search (EXS) are tested, and the simulation results are shown in Table 2. $Tr(\cdot)$ denotes the trace operator for a matrix. The number of sensors is 3, the problem size (MN) is enlarged by increasing the attack time period from 6 to 20. From Table 2, it is clear that SFS and SFS-SS have a lower complexity than the SBS and the EXS. DP is a linear-time algorithm, which has the least computational complexity. As the size of the problem increases, it will be not feasible to get the optimal solution using EXS. In this example, all the approaches can find the global optimum at least when the EXS is still feasible.

For the case the adversary attacks the system from $K+1$ to $K+10$, the system parameters are set the same as above. The results for the optimal attack strategy $(10,1),(9,1),(8,1),(10,2),(7,1)$, the strategy to attack backwards $(9,1),(9,2)$, $(10,1),(10,2),(10,3)$, and the strategy to attack the best sensor $(10,1),(9,1),(8,1)$, $(7,1),(6,1)$ are shown in Fig. 17, where (k,i) denotes that the adversary attacks sensor i at time k. It is clear that the maximal system estimation error is achieved by using the optimal attack strategy.

4.2.2 System with Position and Velocity Sensors

For the system with position and velocity sensors, transition matrix $\mathbf{F}$ and input $\mathbf{u}_k$ are set the same as in Section 4.2.1. The measurement matrix $\mathbf{H}$ for each sensor

Table 2. Performance of Different Algorithms

Alg.	Size	Time (s)	$Tr(P_{K+N\|K+N})$
SFS	18	0.025	0.033
	30	0.064	0.033
	60	0.249	0.033
SBS	18	0.045	0.033
	30	0.239	0.033
	60	1.982	0.033
SFS-SS	18	0.047	0.033
	30	0.128	0.033
	60	0.495	0.033
DP	18	0.015	0.033
	30	0.025	0.033
	60	0.049	0.033
EXS	18	2.269	0.033
	30	79.383	0.033
	60	–	–

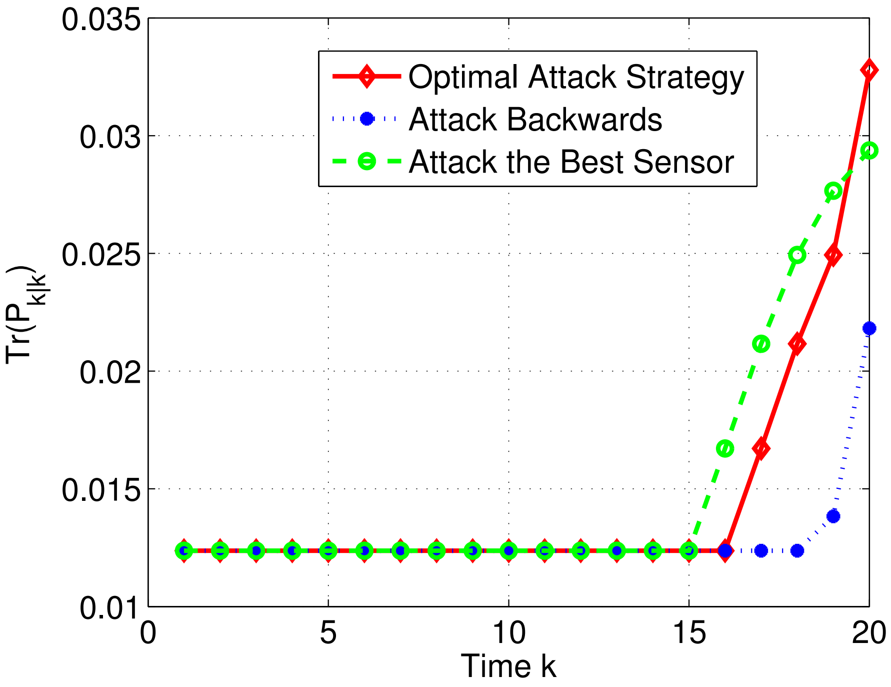

Fig. 17. Trace of MSE for the system with three sensors

is a 2×2 identity matrix. In this subsection, the determinant of the state covariance matrix is used as the objective function. Here we investigate three cases with different system parameters. In Case I, we set $\sigma_v = 0.02$, and the correlation coefficients between position and velocity measurements for the 3 sensors are $\rho_1 = 0.5, \rho_2 = 0, \rho_3 = -0.5$, $\sigma_{w_{1_p}} = \sigma_{w_{1_v}} = 0.5$, $\sigma_{w_{2_p}} = \sigma_{w_{2_v}} = 0.5$, and $\sigma_{w_{3_p}} = \sigma_{w_{3_v}} = 0.5$. Using SFS, the optimal attack strategy is shown in Table 3. The first item in (4.3) is a positive semidefinte matrix with negative off-diagonal elements. The information from Sensor 3 $\mathbf{R}_3^{-1}$ will enlarge the diagonal elements and lower the off-diagonal elements, leading to smaller determinant of $\mathbf{P}_{k|k}$, so the adversary will attack Sensor 3 first. For Sensors 1 and 2, the inverse of covariance matrices are $\mathbf{R}_1^{-1} = \begin{bmatrix} 5.3 & -2.7 \\ -2.7 & 5.3 \end{bmatrix}$ and

Table 3. Attack Strategy for Case I

Time/Sensor	K+1	K+2	K+3	K+4	K+5	K+6
Sensor 1					5	2
Sensor 2						4
Sensor 3					3	1

Table 4. Attack Strategy for Case II

Time/Sensor	K+1	K+2	K+3	K+4	K+5	K+6
Sensor 1					4	2
Sensor 2						
Sensor 3				5	3	1

$\mathbf{R}_2^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$. Comparing with Sensor 2, Sensor 1 will make $\mathbf{P}_{k|k}^{-1}$ larger. Thus the adversary will attack Sensor 1 next instead of Sensor 2. In Case II, we set $\sigma_v = 0.001$, all the other parameters are set the same as in Case I, and the attack strategy is shown in Table 4. From Table 4, it is clear that as the variance of the state process noise decreases, the adversary will attack the sensors with correlated measurements. In Case III, we set $\sigma_{w_{2_p}} = \sigma_{w_{2_v}} = 0.2$, all the other parameters are set the same as in Case II, and the optimal attack strategy is shown in Table 5. In this case, instead of attacking the sensors with correlated measurements, the adversary will attack the sensor with the smallest covariance.

To compare the greedy approach and DP, we consider the case that only involves position sensors under different configurations. The standard deviation σ_v of the system process noise varies from 0.01 to 1. The sparsity constraint for the adversary is $c = 10$, meaning that the adversary has to choose 10 spots to attack the system

Table 5. Attack Strategy for Case III

Time/Sensor	K+1	K+2	K+3	K+4	K+5	K+6
Sensor 1						
Sensor 2		5	4	3	2	1
Sensor 3						

over $M = 10$ sensors and over $N = 6$ time steps in order to maximize the trace of the state covariance matrix by the end of the attack period. Furthermore, the attacker uses the trace of covariance matrix as the uncertainty measure. For each value of σ_v, we run 100 Monte-Carlo trials, and in each trial we set the standard deviations σ_w of the measurement noise for the ten sensors by drawing ten values from uniform $(0, 1)$ distribution. Then, we count the number of simulations where DP gives the same, larger, and smaller $Tr(\mathbf{P}_{K+N|K+N})$ compared with the greedy approach SFS.

In Fig. 18, we compare the attack performance of greedy approach and DP under different σ_v's. It can be shown that for a small σ_v, the two algorithms mostly give the same results, and greedy approach is performing better; however, as σ_v increases up to 0.1, DP begins to outperform greedy approach significantly. The underlying reason comes from the different frameworks of SFS and DP. In each round of SFS, attacker chooses one sensor-time pair to attack by running Kalman filter through all the snap shots $K + 1$ up to $K + N$. On the other hand, DP determines $\Phi_{DP}(s, k)$ only based on the previous $c + 1$ states at snap shot $k - 1$. Therefore, when σ_v is small, or the system state evolves smoothly, SFS will have a better sense of "global view" than DP; however, for a large σ_v where the prediction gives little information, DP gives more credit for the current measurement which is more informative, and hence outperforms SFS significantly.

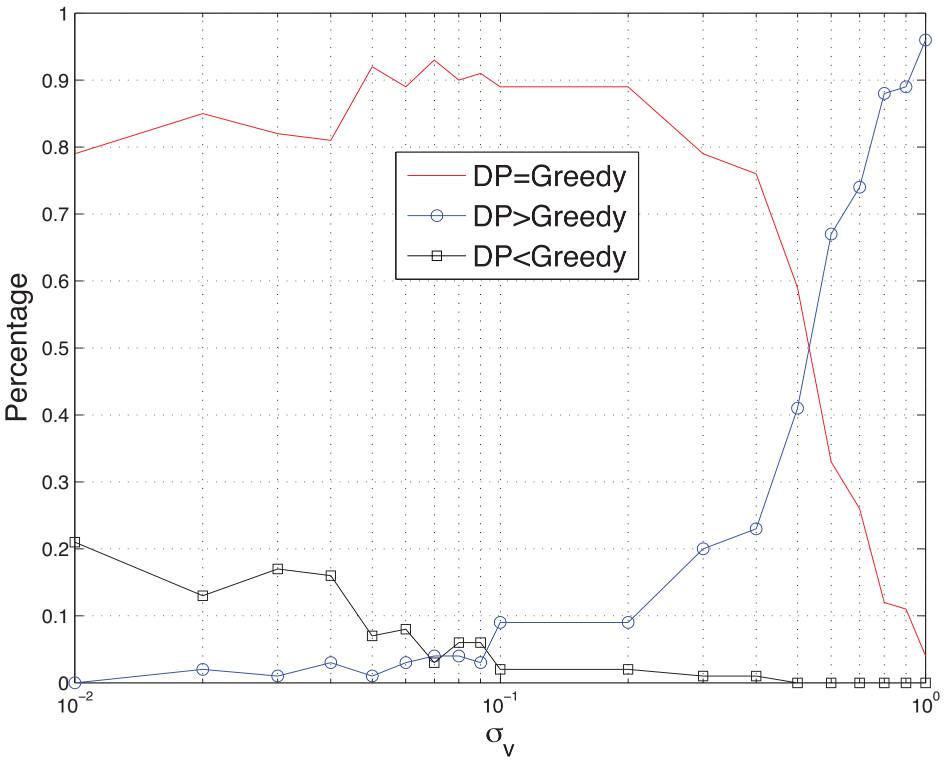

Fig. 18. Comparison between greedy approach (SFS) and DP.

4.3 Conclusion

In this chapter, sparse attack strategies in multi-sensor dynamic systems have been studied from the adversary's point of view. By assuming that the system defender can perfectly detect and remove the sensors attacked by the adversary, this becomes an integer programming problem. As the size of the problem increases, it will be infeasible to find the optimal solution. Different suboptimal algorithms: SFS, SBS, SFS-SS, and DP have been studied and corresponding attack strategies were developed. Their computational complexities have been analyzed and their performances have been evaluated and compared based on simulations. All the proposed suboptimal solutions can provide very good performance (in some examples they lead to the optimal solution) with significantly lower complexities. It has been shown that the greedy approach outperforms DP when the system process noise is small, since it has a more long-term view of the problem. On the other hand, DP performs better when the process noise is large and the state is more unpredictable.

CHAPTER 5

A GAME BETWEEN STATE ESTIMATION AND MALICIOUS ATTACKS

We have studied attack strategies from the adversary perspective. In this chapter, the problem of false information attack on and Kalman filter's defense of state estimation in dynamic multi-sensor systems is investigated from a game theoretic perspective. The relationship between Kalman filter and the adversary can be regarded as a two-person zero-sum game. Under which condition both sides of the game will reach the Nash equilibrium is investigated in this chapter. The multi-sensor Kalman filter system and the adversary are supposed to be rational players. Kalman filter and the adversary have to choose their respective subsets of sensors to perform system state estimation and false information injection. It is shown how both sides pick their strategies in order to gain more and lose less.

5.1 A Target Tracking Example

In this section, we give a concrete target tracking example, which is also discussed in [37]. Assume that the target moves in a one-dimensional space according to a discrete white noise acceleration model [31], which can still be described by the plant and measurement equations provided in (2.1) and (2.2). In such a system, the state is defined as $\mathbf{x_k} = [\xi_k \;\; \dot{\xi}_k]^T$, where ξ_k and $\dot{\xi}_k$ denote the target's position and velocity at time k respectively. The input $\mathbf{u}_k$ is a zero sequence. The state transition matrix is

$$\mathbf{F}_k = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad \forall k \tag{5.1}$$

where T is the sensor sampling interval. The process noise is $\mathbf{v}_k = \mathbf{\Gamma} v_k$, where v_k is a zero mean white acceleration noise, with variance σ_v^2, and the vector gain multiplying the scalar process noise is given by $\mathbf{\Gamma} = [T^2/2 \quad T]^T$. The covariance matrix of the process noise is therefore $\mathbf{Q} = \sigma_v^2 \mathbf{\Gamma}\mathbf{\Gamma}^T$. The observation matrix is given as

$$\mathbf{H}_{k,i} = [1 \;\; 0], \;\; \forall k, \; i \tag{5.2}$$

Once the system model is known, it is straightforward for both Kalman filter and the adversary to calculate Kalman filter's state covariance matrix $\mathbf{P}_{K|K}$ as in [31]. Using Proposition 1, we can obtain the trace of the total state estimation MSE matrix:

$$\mathrm{Tr}(\mathrm{MSE}) = \mathrm{Tr}(\mathbf{P}_{K|K} + \mathbf{W}_K \mathbf{\Sigma}_K \mathbf{W}_K^T) \tag{5.3}$$

5.2 Noncooperative Two-Person Zero-Sum Game

In a noncooperative two-person zero-sum game [38], we assume that there are two players, referred to as Players 1 and 2, and an $m \times n$ payoff matrix $\mathbf{L} = \{l_{ij}\}$. Each entry of the matrix is an outcome of the game corresponding to a particular pair of decisions made by both players. Player 1 gets m rows of the matrix as his/her strategy set, while for Player 2, the strategy set is the corresponding n columns of the same matrix.

In our problem, suppose there are totally M sensors, Kalman filter and the adversary can choose any non-empty subsets of sensors to perform state estimation and attack respectively, which means $m = n = 2^M - 1$. $\mathbf{L}$ is a square matrix of the size $(2^M - 1) \times (2^M - 1)$. The payoff in the game between Kalman filter system and the adversary will be the trace of the state estimation MSE matrix. For each set of sensors he/she chooses to attack, the adversary is under a total injected noise power constraint. The Nash equilibrium between Kalman filter and the adversary is achieved by solving the minimax optimization problem.

Let {row i, column j} be a pair of strategies adopted by the players, and the corresponding outcome (payoff) be l_{ij}, which means that Player 1 should pay Player 2 the amount of l_{ij}. If $l_{i^*j} \leq l_{i^*j^*} \leq l_{ij^*}$, for all $i = 1, \ldots, m$ and all $j = 1, \ldots, n$, the pair $\{i^*, j^*\}$ is said to constitute a saddle-point equilibrium, and the game is said to have a saddle point in pure strategy. On the other hand, if the pair of inequalities does not exist, one can derive the mixed strategy to obtain the equilibrium. A mixed strategy is a probability distribution on the space of the player's pure strategies. A mixed strategy allows for a player to select a pure strategy randomly with a certain probability. In this case, the utility function u is defined as

$$u(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i l_{ij} y_j = \mathbf{x}^T \mathbf{L} \mathbf{y} \tag{5.4}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the probability distribution vectors for the mixed strategies. Also, $\mathbf{x} \in X$, $\mathbf{y} \in Y$, where the set $X = \{\mathbf{x} \in R^m : \mathbf{x} \geq \mathbf{0},\ \sum_{i=1}^{m} x_i = 1\}$, and Y is defined in the same way. Kalman filter playing as defender is trying to minimize the utility function $u(\mathbf{x}, \mathbf{y})$ by choosing the best defending strategy, while the attacker wants to maximize the utility function by choosing the best attack strategy. For the payoff matrix $\mathbf{L}$ of size $m \times n$, a vector of $\mathbf{x}^*$ is the best mixed strategy for Kalman filter if

$$\overline{U}_m(\mathbf{L}) = \max_{\mathbf{y} \in Y} (\mathbf{x}^*)^T \mathbf{L} \mathbf{y} \leq \max_{\mathbf{y} \in Y} \mathbf{x^T} \mathbf{L} \mathbf{y}, \mathbf{x} \in X \tag{5.5}$$

The $\overline{U}_m(\mathbf{L})$ is known as the average security level (loss ceiling) of the defender, the average security level (gain-floor) of the attacker $\underline{U}_m$ can also be defined as below,

$$\underline{U}_m(\mathbf{L}) = \min_{\mathbf{x}\in X} \mathbf{x}^T\mathbf{L}\mathbf{y}^* \geq \min_{\mathbf{x}\in X} \mathbf{x}^T\mathbf{L}\mathbf{y}, \mathbf{y} \in Y \tag{5.6}$$

It always holds that $\overline{U}_m(\mathbf{L}) = \underline{U}_m(\mathbf{L})$ for mixed strategies in noncooperative two-person zero-sum game. The saddle point in the mixed strategies is defined when the two bounds are equal to each other, which can be found by solving the following linear programming problem [38]:

$$\begin{aligned} \min_{\mathbf{x}\in X} \quad & b_u \\ \text{s.t.} \quad & \mathbf{L}^T\mathbf{x} \leq b_u\mathbf{1} \\ & \mathbf{x}^T\mathbf{1} = 1 \\ & \mathbf{x} \geq 0 \end{aligned} \tag{5.7}$$

where b_u denotes a constant upper bound. For the attacker, the formula is the other way around,

$$\begin{aligned} \max_{\mathbf{y}\in Y} \quad & b_l \\ \text{s.t.} \quad & \mathbf{L}\mathbf{y} \geq b_l\mathbf{1} \\ & \mathbf{y}^T\mathbf{1} = 1 \\ & \mathbf{y} \geq 0 \end{aligned} \tag{5.8}$$

where b_l denotes a constant lower bound. From the formulation above, it is easy to see that (5.8) is the dual form of the optimization problem (5.7). The optimal function for the two problems are the same. Interested readers are referred to [38] for more details.

5.3 Game With Incomplete Information

So far, it has been assumed that the defender has the knowledge on what type of false information has been injected by the adversary to the sensors. In practice, this knowledge may not be readily available to the defender. So let us suppose that Kalman filter does not know the type of false information the adversary uses to attack the system. If the adversary has S types of false information like independent or dependent false information as discussed earlier in the chapter, there should be S $(m \times n)$ payoff matrices $\mathbf{L}_k = \{l_{k_{ij}}\}$, and $\sum_{k=1}^{S} p_k = 1$, $k \in S$. Based on p_k, the prior probability of the kth false information type, the Bayesian equilibrium can be achieved by solving the following linear optimization problem.

$$
\begin{aligned}
\min_{\mathbf{x} \in X} \quad & b_u \\
\text{s.t.} \quad & \sum_{k=1}^{S} p_k \mathbf{L}_k{}^T \mathbf{x} \leq b_u \mathbf{1} \\
& \mathbf{x}^T \mathbf{1} = 1 \\
& \mathbf{x} \geq 0
\end{aligned}
\tag{5.9}
$$

where b_u denotes a constant upper bound.

5.4 Numerical Results

In the example, for simplicity and ease of presentation, we assume that there are three sensors denoted as $\{z_1, z_2, z_3\}$ in the system having independent measurement noises with noise standard deviations $\sigma_{w_1} = 3$, $\sigma_{w_2} = 4$, $\sigma_{w_3} = 5$. The system process noise s.d. is $\sigma_v = 0.5$, sensors' sampling interval is $T = 1s$, and the system initial

Table 6. Payoff Matrix (Independent Case)

KF/At	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
z_1	25.4	4.7	4.7	25.4	25.4	4.7	25.4
z_2	7.2	23.5	7.2	7.2	7.2	23.5	7.2
z_3	10	10	23.6	10	10	10	10
z_1z_2	13.5	6.6	3.4	13.5	13.5	6.6	13.5
z_1z_3	16.4	3.8	5.4	16.4	16.4	3.8	16.4
z_2z_3	5.0	12.4	8.0	5.0	5.0	12.4	5.0
$z_1z_2z_3$	10.2	5.2	3.9	10.2	10.2	5.2	10.2

state $\mathbf{x}_0$ is assumed to follow a $\mathcal{N}(\hat{\mathbf{x}}_{0|0},\ \mathbf{P}_{0|0})$ distribution, where $\hat{\mathbf{x}}_{0|0} = [1\ \ 1]^T$ and

$$\mathbf{P}_{0|0} = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.5 \end{bmatrix}.$$

The adversary can choose any combination of sensors from the set $P_1 = \{z_1, z_2, z_3, z_1z_2, z_1z_3,\ z_2z_3, z_1z_2z_3\}$ to attack with the power constraint of $\sum_1^3 {\sigma_{b_i}}^2 = 100$, where σ_{b_i} is the s.d. of the random noise injected to Sensor i. Likewise, the defender can choose any combination of sensors to perform state estimation, and its strategy set is the same: $P_2 = P_1$. The game is played as below: if the defender uses data from Sensors i and j for state estimation, while the adversary attacks Sensors i and k, then system state estimation is affected by the false information from the ith sensor only.

In this game, the trace of the state estimation MSE matrix is regarded as the payoff of the game. In the games of the independent and dependent attacks, the system is attacked according to the strategies provided in Propositions 2 and 3 respectively. Let us assume that the adversary attacks the sensors at time $k = 100$, and the payoff matrix is given in Tables 6 and 7. From Tables 6 and 7, we can see that there is no pure strategy Nash Equilibrium. Instead, we use mixed strategies to find the Nash Equilibrium. In order to obtain the optimal probability vector, we solve the optimization problem formulated in (5.7). The solution to (5.7) is the optimal probability vector for the defender, and the dual solution is the optimal mixed strategy

Table 7. Payoff Matrix (Dependent Case)

KF/At	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
z_1	25.4	4.7	4.7	13.2	15.9	4.7	10.3
z_2	7.2	23.5	7.2	9.3	7.2	13.3	8.6
z_3	10	10	23.6	10	11.0	12.1	10.5
z_1z_2	13.5	6.6	3.4	16.7	12.4	5.6	15.6
z_1z_3	16.4	3.8	5.4	15.0	18.1	4.2	15.0
z_2z_3	5.0	12.4	8.0	6.8	5.3	15.5	8.2
$z_1z_2z_3$	10.2	5.2	3.9	12.5	11.1	6.2	13.4

for the attacker. The optimal solutions for independent- and dependent-attack cases are shown in Tables 8 and 9 respectively.

For the independent case, we can see from Table 6 that $(6, 6)$ and $(7, 7)$ elements of the payoff matrix ($\mathbf{L}$) are the smallest among the seven diagonal elements. This means that in the worst cases for the KF when its chosen sensor combination happens to be the same as that being attacked by the adversary, the strategies z_2z_3 and $z_1z_2z_3$ will lead to the smallest state estimation MSEs. In addition, for the KF, the values of last two rows are relatively small. As a result, for the KF, the probabilities of the last two strategies (z_2z_3 and $z_1z_2z_3$) are much larger than those of other strategies, which are shown in Table 8.

In the dependent case, for the KF, the probabilities for the last two pure strategies (z_2z_3 and $z_1z_2z_3$) are relatively large as shown in Table 9. This can be explained similarly as in the independent case. In $\mathbf{L}$, the entries of the rows corresponding to z_3, z_1z_2, and z_1z_3 are relatively large, so the KF assigns nearly zero probabilities to these three strategies. In the first two rows of $\mathbf{L}$, even though the diagonal elements are large, the rest of the elements are relatively small, so strategies z_1 and z_2 are assigned significant probabilities for the KF as shown in Table 9.

When the information is incomplete, the defender is not sure whether independent or dependent false information will be injected by the adversary. Table 10 shows

Table 8. Optimal Strategy Probabilities (Independent Case)

Player	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
KF	0.00	0.00	0.00	0.00	0.00	0.40	0.60
Attacker	0.14	0.22	0.00	0.14	0.14	0.22	0.24

Table 9. Optimal Strategy Probabilities (Dependent Case)

Player	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
KF	0.16	0.14	0.00	0.00	0.00	0.37	0.33
Attacker	0.14	0.02	0.00	0.00	0.00	0.34	0.50

Kalman filter's best defending strategy under the condition that Kalman filter can be attacked by the two types of false information equally probably. This result could be explained by the results for the case with complete information about the type of attacks. Clearly from Tables 8 and 9, the defender assigns significant probabilities to the last two strategies (z_2z_3 and $z_1z_2z_3$). As a result, the optimal strategy for the case with incomplete knowledge of the type of the false information also assigns most probabilities to the last two strategies.

We also provide a simulation result to demonstrate the optimality of the derived strategy for the independent case. In this example, four different scenarios are explored: 1) there is no attack and the KF uses all the sensors' data; 2) the KF uses the optimal mixed strategy; 3) the KF uses a mixed strategy to pick each pure strategy with an equal probability 1/7; 4) the KF always chooses the first sensor to do the system estimation. In Scenarios 2)-4), the attacker injects false information according to his/her optimal mixed strategy to the sensors at time $k = 100$. The resulting position estimation MSEs are plotted in Fig. 19. It is clear that the optimal mixed strategy provides the best defense against the attacker, with the minimum increase in the

Table 10. Optimal Strategy Probabilities (Incomplete Information Case)

Player	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
KF	0.00	0.00	0.00	0.00	0.00	0.45	0.55

MSE after the attack. Fig. 20 shows the case when the defender's knowledge about the attack false information is incomplete. Let us suppose that Kalman filter can be attacked by independent and dependent false information with equal probability. The results corresponding to three different defending strategies, the best defending strategy for independent attacks, the best defending strategy for dependent attacks, and the best defending strategy for incomplete information case, are shown in Fig. 20. It is clear that the optimal defending strategy derived from (5.9) leads to smaller MSE after the attack than the other two defending strategies corresponding to the independent and dependent attacks respectively.

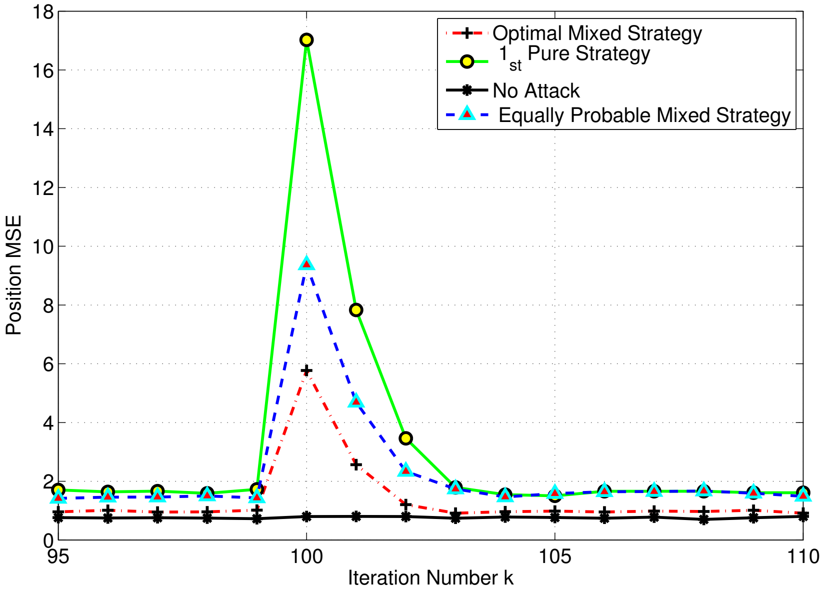

Fig. 19. Optimal Mixed Strategy vs. Other Options

5.5 Conclusion

The relationship between Kalman filter and the adversary has been investigated in a two-person zero-sum game. Kalman filter (defender) tries to achieve more accu-

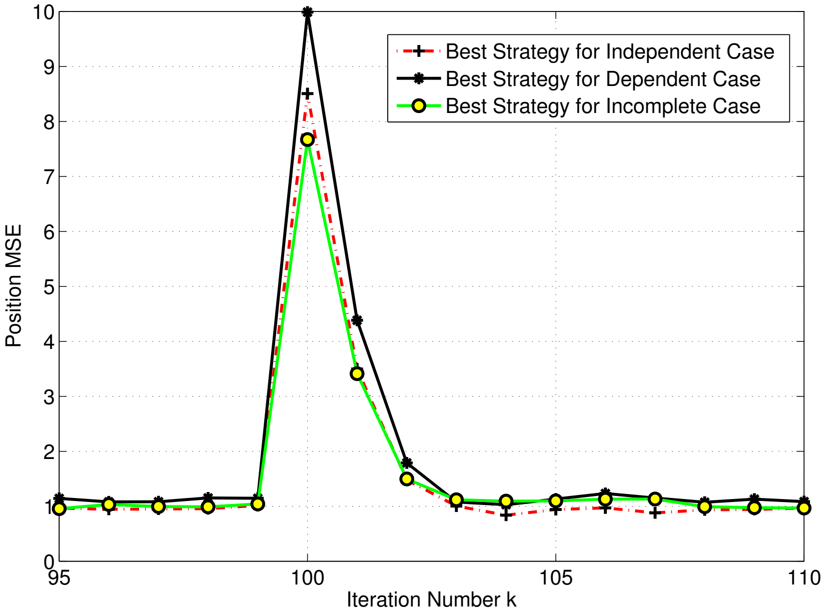

Fig. 20. Optimal Mixed Strategy vs. Other Options with Incomplete Information about Attacks

rate system state estimation and avoid being attacked by the adversary. The adversary tries to mislead Kalman filter as much as possible. Both sides of the game will reach a Nash Equilibrium through the mixed strategies. Using minimax techniques, we found the mixed strategy saddle point in the game.

CHAPTER 6

CONCLUSION

Overall, resilient dynamic state estimation in the presence of false information injection attacks has been studied from different aspects. Under the assumption that Kalman filter is not aware of the existence of false information, many optimal attack strategies that maximize the system estimation error were derived from either trace or determinant perspective. For the case where the sensors only measure the position, the adversary could either inject independent or dependent false information into the sensor readings, it was proved that under the same power constraint, the dependent false information will incur more system estimation error than the independent one does in terms of the trace of the state estimation MSE matrix. For the case where the sensors measure both the position and velocity, the optimal attack strategies were derived and demonstrated via simulations. The optimal attack strategy, which maximizes the determinant of Kalman filter's MSE matrix, has been studied and the closed-form optimal solution was provided. The impact of the correlation coefficients between different components of the injected bias noise on the volume of the error ellipse for Kalman filter's state estimation was also studied through simulations. As for the multi-sensor case, an equivalent sensor was utilized in order to simplify the problem.

In order to detect the false information injected by the adversary, an optimal Bayesian detector that minimizes the average state estimation MSE was derived. This detector can be coupled with different defending strategies. In this dissertation, defending strategies of discarding data after detection and incorporating data after detection were studied. The Minimum mean square error estimator was also investigated given all the system information. But the minimum mean square error estimator is not robust when there is a mismatch between the attack strategy assumed by the defender and the actual one adopted by the attacker. The optimal Bayesian detector coupled with defending strategy of discarding data after detection proved to be robust through simulations.

Sparse attack strategies were also investigated under the assumption that the defender can perfectly detect the false information and remove the sensors once they are corrupted by the false information injected by the adversary. The adversary aims to maximize the MSE matrix of the system state estimate by the end of the attack period under the constraint that he/she can only attack the system a few times over the sensors and over the time. Greedy search and dynamic programming based approaches were utilized to obtain suboptimal attack strategies since the optimal exhaustive search becomes intractable even when the problem size increases moderately. As for the greedy search, SFS, SBS, and SFS-SS were studied and evaluated via simulations. The performances of greedy search and dynamic programming were also compared. The greedy search outperforms dynamic programming when the system process noise is small. Meanwhile, dynamic programming outperforms the greedy search when the system process noise is larger. This is because the greedy search has a better global view than dynamic programming regarding the system estimation.

The relation between the defender and the adversary was further studied using the game theory. It was supposed that the defender and the adversary are both rational player. It was shown how both sides choose strategies in order to gain more and lose less. Two cases where the defender either has complete information or incomplete information about attack strategies, which the adversary adopts, were both studied in a two-person-zero-sum game. A Nash equilibrium was finally achieved by solving a linear programming problem. For the defender, different defending strategies were compared in the simulations. The defending strategies obtained via the Nash equilibrium help the system maintain a much better estimation performance.

ABBREVIATIONS

CBD	Conventional Bayesian Detection
CSD	Chi-Square Detection
DOE	Department of Energy
DP	Dynamic Programming
EMC	Energy Management Centers
EMSE	Extra Mean Square Error
EXS	Exhaustive Search
FAA	Federal Aviation Administration
GPS	Global Positioning System
KF	Kalman Filter
LRT	Likelihood Ratio Test
MIMO	Multiple Input Multiple Output
MSE	Mean Square Error
OBDD	Optimal Bayesian Detection and Discarding Sensor Data
OBDI	Optimal Bayesian Detection and Incorporating Sensor Data
PDF	Probability Density Function
QCQP	Quadratically Constrained Quadratic Program
RVA	Richmond Virginia
SBS	Sequential Backward Selection
SFS	Sequential Forward Selection
SFS-SS	Simplex Improved Sequential Forward Selection
UAV	Unmanned Aerial Vehicle
VCU	Virginia Commonwealth University

REFERENCES

[1] F. F. Wu, K. Moslehi, and A. Bose. “Power System Control Centers: Past, Present, and Future”. In: *Proceedings of the IEEE* 93.11 (2005), pp. 1890–1908.

[2] M.N. Mladenovic, M. Abbas, and T. McPherson. “Development of socially sustainable traffic-control principles for self-driving vehicles: The ethics of anthropocentric design”. In: *the 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering.* 2014, pp. 1–8.

[3] E. Mills. “Hackers broke into FAA air traffic control system”. In: *The Wall Street Journal* (2009), A6.

[4] N. Leavitt. “Researchers Fight to Keep Implanted Medical Devices Safe from Hackers”. In: *Computer* 43.8 (2010), pp. 11–14.

[5] S. Sridhar, A. Hahn, and M. Govindarasu. “Cyber-Physical System Security for the Electric Power Grid”. In: *Proceedings of the IEEE* 100.1 (2012), pp. 210–224.

[6] National Energy Technology Laboratory. “A Systems View of the Modern Grid”. In: *U.S. Department of Energy.* 2007.

[7] Y. Huang et al. “Bad data injection in smart grid: attack and defense mechanisms”. In: *IEEE Communications Magazine* 51.1 (2013), pp. 27–33.

[8] Y. Liu, M.K. Reiter, and P. Ning. “False data injection attacks agianst state estimation in electric power grids”. In: *Proc. the 16th ACM Conference on Computer and Communications Security.* Chicago, IL, 2009.

[9] X. Song, P. Willett, and S. Zhou. "Jammer detection and estimation with MIMO radar". In: *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. 2012, pp. 1312–1316.

[10] H. Chen and B. Himed. "Analyzing and improving MIMO radar detection performance in the presence of cybersecurity attacks". In: *2016 IEEE Radar Conference (RadarConf)*. 2016, pp. 1–4.

[11] D. M. Wolpert Z. Ghahramani. "Computational principles of movement neuroscience". In: *Nature Neuroscience* 3 (2000), pp. 1212–1217.

[12] L. Jia, R.J. Thomas, and L. Tong. "Malicious data attack on real-time electricity market". In: *Proc. International Conference on Acoustics, Speech, and Signal Processing*. Prague, Czech Republic, 2011, pp. 5952–5955.

[13] O. Kosut et al. "Malicious Data Attack on Smart Grid State Estimation: Attack Strategies and Countermeasures". In: *Proc. First IEEE International Conference on Smart Grid Communications (SmartGridComm)*. Gaithersburg, MD, 2010, pp. 220–225.

[14] L. Jia, R. J. Thomas, and L. Tong. "On the nonlinearity effects on malicious data attack on power system". In: *Power and Energy Society General Meeting*. San Diego, CA, 2012, pp. 1–8.

[15] M. A. Rahman and H. Mohsenian-Rad. "False data injection attacks with incomplete information against smart power grids". In: *Proc. Global Communications Conference*. San Diego, CA, 2012, pp. 3153–3158.

[16] J. Kim, L. Tong, and R. J. Thomas. “Data Framing Attack on State Estimation”. In: *IEEE Trans. on Aerospace and Electronic Systems* 49.3 (2013), pp. 1637–1653.

[17] X. Song et al. “The MIMO Radar and Jammer Games”. In: *IEEE Trans. on Signal Processing* 60.2 (2012), pp. 687–699.

[18] R. Niu and L. Huie. “System State Estimation in the Presence of False Information Injection”. In: *Statistical Signal Processing Workshop (SSP)*. Ann Arbor, MI, 2012, pp. 385–388.

[19] W. Xiong, A. Mukherjee, and H. M. Kwon. “MIMO Cognitive Radio User Selection With and Without Primary Channel State Information”. In: *IEEE Transactions on Vehicular Technology* 65.2 (2016), pp. 985–991.

[20] W. Xiong et al. “Hybrid onboard and ground based digital channelizer beamforming for SATCOM interference mitigation and protection”. In: *Proc. SPIE Sensors and Systems for Space Applications IX*. Vol. 9838. 2016.

[21] N. Zhang S. K. Das K. Kant. *Handbook on Securing Cyber-physical Critical Infrastructure: Foundations and Challenges*. Elsevier, 2012.

[22] C. Yang, L. Kaplan, and E. Blasch. “Performance Measures of Covariance and Information Matrices in Resource Management for Target State Estimation”. In: *IEEE Trans. on Aerospace and Electronic Systems* 48.3 (2012), pp. 2594–2612.

[23] C. Yang et al. “Optimal Placement of Heterogeneous Sensors for Targets with Gaussian Priors”. In: *IEEE Trans. on Aerospace and Electronic Systems* 49.3 (2013), pp. 1637–1653.

[24] X. Lin and Y. Bar-Shalom. "Multisensor target tracking performance with bias compensation". In: *IEEE Trans. Aerosp. Electron. Syst.* 42.3 (2006), pp. 1139–1149.

[25] J. Lu and R. Niu. "False Information Injection Attack on Dynamic State Estimation in Multi-Sensor Systems". In: *Proc. of the 17th International Conference on Information Fusion*. Salamanca, Spain, 2014.

[26] J. Lu and R. Niu. "Malicious Attacks on State Estimation in Multi-Sensor Dynamic Systems". In: *Proc. IEEE Workshop on Information Forensics and Security*. Atlanta, Georgia, USA, 2014.

[27] J. Lu and R. Niu. "False Information Detection with Minimum Mean Squared Errors for Bayesian Estimation". In: *Proc. The 49th Annual Conference on Information Systems and Sciences*. Baltimore, Maryland, 2015.

[28] J. Lu and R. Niu. "Sparse Attacking Strategies in Multi-Sensor Dynamic Systems Maximizing State Estimation Errors". In: *Proc. 41st IEEE International Conference on Acoustics, Speech and Signal Processing*. Shanghai, China, 2016.

[29] J. Lu, R. Niu, and P. Han. "Optimal space-time attacks on system state estimation under a sparsity constraint". In: *Proc. SPIE Sensors and Systems for Space Applications IX*. 2016.

[30] J. Lu and R. Niu. "A State Estimation and Malicious Attack Game in Multi-Sensor Dynamic Systems". In: *Proc. the 18th International Conference on Information Fusion*. Washington, USA, 2015.

[31] Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. New York: Wiley, 2001.

[32] R. Niu. *Dynamic System State Estimation in the Presence of Continuous False Information Injection*. Tech. rep. Extension Grant from Visiting Faculty Research Program, Air Force Research Laboratory Information Directorate, 2012.

[33] B. Tang, J. Tang, and Y. Peng. "MIMO Radar Waveform Design in Colored Noise Based on Information Theory". In: *IEEE Trans. on Signal Processing* 58.9 (2010), pp. 4684 –4697.

[34] Y. Bar-Shalom, P.K. Willett, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT: YBS Publishing, 2011.

[35] L.M. Kaplan. "Global node selection for localization in a distributed sensor network". In: *IEEE Transactions on Aerospace and Electronic Systems* 42.1 (2006), pp. 113–135.

[36] P. Pudil, J. Novoviov, and J. Kittler. "Floating search methods in feature selection". In: *Pattern Recognition Letters* 15.11 (1994), pp. 1119–1125.

[37] C. Miller et al. "Estimation of mobile vehicle range amp and position using the tobit Kalman filter". In: *IEEE 53rd Annual Conference on Decision and Control (CDC)*. 2014, pp. 5001–5007.

[38] T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999.

VITA

Jingyang Lu was born in Jiamusi, Heilongjiang Province, China, 1988. He attended Jiamusi NO.1 Middle School.

Jingyang Lu got his Bachelor's Degree of Science in Harbin Institute of Technology in August 2012. He enrolled in the doctoral program in the Department of Electrical and Computer Engineering at Virginia Commonwealth University in 2012. His research interests are in the areas of statistical signal processing and its applications, including tracking, estimation, detection, and information fusion. During his Ph.D. program, he got his Master's Degree in Electrical and Computer Engineering in 2015. He published 6 papers on dynamic system state estimation under the false information injection attacks.

Jingyang Lu likes playing basketball. During his college, he played for the Department of Engineering. When he was in VCU, he usually joined the intramural.