Author/s:
Farokhi, F

Title:
Feedback control using a strategic sensor

Date:
2021-01-01

Citation:
Farokhi, F. (2021). Feedback control using a strategic sensor. International Journal of Control, 94 (1), pp.1-6. https://doi.org/10.1080/00207179.2019.1575983.

Persistent Link:
https://hdl.handle.net/11343/285054

ARTICLE TEMPLATE

Feedback Control Using a Strategic Sensor

Farhad Farokhi[a]

[a]F. Farokhi is with the CSIRO's Data61 and the Department of Electrical and Electronic Engineering at the University of Melbourne, Australia

ARTICLE HISTORY
Compiled January 16, 2019

ABSTRACT
A dynamic estimation and control problem with a strategic sensor is considered. The strategic sensor may provide corrupted messages about the state measurements of a discrete-time linear time-invariant dynamical system to the system operator (or the controller). The system operator then uses this information to construct an estimate of the state of the system (and perhaps private variables of the sensor). The estimate is used to control the system to achieve the operator's desired objective. The problem is formulated as a game, which might be conflicting to that of the strategic sensor. An equilibrium of the game is computed and its properties are investigated.

KEYWORDS
Game theory; Estimation; Optimal Control.

1. Introduction

Traditional infrastructure operators most often rely on sensors installed and maintained by themselves for estimating the states of the system. The constant connectedness provided by networked smart devices has enabled the use of emerging technologies, such as crowd sensing, in systems. This helps to improve the quality of service provided by existing infrastructure with low investment costs. For instance, in transportation systems, the infrastructure operator can monitor the movements of recruited participants to estimate the traffic flow (as in, e.g., Waze). The challenges associated with these non-traditional sensors arise from the limited authority of the operator over the behaviour of the users. The undesirable effects can only be worsened by the fact that the crowd might have different objectives to that of the infrastructure operator (e.g., generating personal income rather than guaranteeing smooth operation of the system). Further, noting that the infrastructure is a valued target for adversaries and malicious agents, the crowd-sensing systems can be infiltrated by hackers with possibly devastating consequences for the society. Therefore, there is a need for developing a comprehensive framework for investigating the problem of dynamic estimation with strategic sensors and the use of such state estimates in the closed-loop performance of the system.

In this paper, we consider the dynamic estimation and control problem in Figure 1, where a strategic sensor $\mathcal{S}$ provides possibly dishonest messages about the state measurements of a discrete-time linear time-invariant dynamical system $\mathcal{P}$ to the system

CONTACT F. Farokhi. Email: farhad.farokhi@data61.csiro.au, farhad.farokhi@unimelb.edu.au

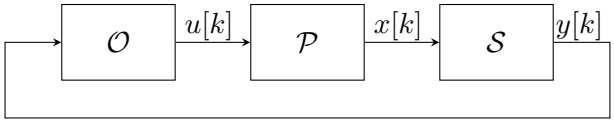

Figure 1. Interconnection pattern between the dynamical system $\mathcal{P}$, the sensor $\mathcal{S}$, and the system operator $\mathcal{O}$.

operator $\mathcal{O}$ (which may be also referred to as the controller). The system operator then uses this information to construct an estimate of the state of the system (and other hidden or private variables of the sensor). The state estimate is subsequently used to control the system to achieve the desired objectives of the system operator. Noting that the sensor and the system operator might have conflicting objectives, the problem is formulated as a game (see , e.g., (Başar and Olsder, 1998)). An equilibria of the game is computed and its properties are investigated.

The problem of strategic communication between a receiver and a better informed sender, known as the cheap-talk game, has been studied in the economics literature (Crawford and Sobel, 1982; Battaglini, 2002). This problem has more recently attracted attention in the engineering community, where it has application in privacy-constrained communication and cyber-security (Farokhi et al., 2015, 2017) and estimation (Dobakhshari et al., 2016; Westenbroek et al., 2017). These studies mostly consider static estimation problems. In (Farokhi et al., 2017), a dynamic estimation setup is considered; however, the sensor and the receiver act myopically (i.e., they do not consider the effects of their actions in the future). This could be reasonable for monitoring purposes but conservative in control formulations (as the sensor would not consider the effects of its actions in the future). It is worth mentioning that the problem of dynamic cheap-talk games has been recently studied in the economics literature (Golosov et al., 2014). In contrast with (Golosov et al., 2014), in this paper, we restrict the problem formulation to Gaussian random variables and quadratic cost functions. Thus, our model is much closer to the assumptions typically made in the control and estimation community. Further, this allows us to explicitly compute the equilibrium and study its properties. Finally, note that all the above-mentioned studies only focus on estimation issues and the effects of control are mostly unexplored.

The problem studied in this paper is closer to the the study of security of networked control systems (Mo and Sinopoli, 2009; Sandberg et al., 2010; Hendrickx et al., 2014; Teixeira et al., 2013; Pasqualetti et al., 2013; Cardenas et al., 2008). Those however mostly consider the case where the attacker wants to disrupt the operation of the system rather than steering it towards its selfish goals.

The rest of the paper is organized as follows. A motivating example is provided in Section 2. The problem formulation and the results are, respectively, presented in Sections 3 and 4. A numerical example is provided in Section 5. Finally, some concluding remarks are presented in Section 6.

2. Motivating Example

Consider the simplest robot in a two dimensional space modelled as

$$x[k+1] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x[k] + u[k] + w[k],\ x[0] = 0,$$

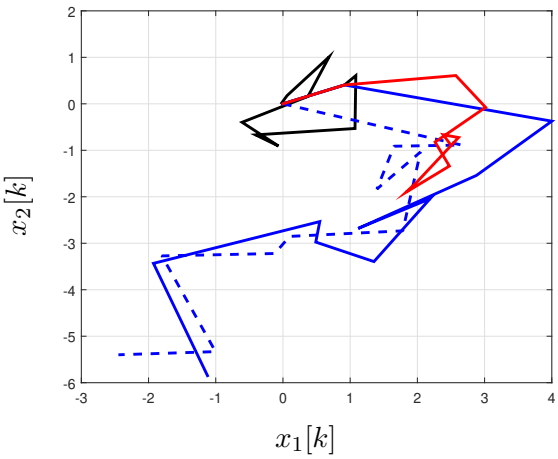

Figure 2. The closed-loop trajectory of the robot with a honest sensor (black), strategic sensor without compensation (blue), and strategic sensor compensated using the method of this paper (red). The path that the strategic sensor wants the robot to follows is shown by the dashed blue curve.

where $x[k] \in \mathbb{R}^n$ is the position of the robot in the chosen two-dimensional coordinate frame, $u[k] \in \mathbb{R}^m$ is the trust generated in each direction, and $w[k] \in \mathbb{R}^n$ is the unknown effect of the environment, e.g., wind, modelled by an i.i.d.[1] zero-mean Gaussian random variable with variance $W = 0.5I$. The operator wishes that the robot stays around the origin over the horizon $\{0, \dots, T\}$ with $T = 10$. Noting that the robot is only marginally stable if it is not controlled, it would aimlessly drift in the space (i.e., the model of the robot resembles that of random walk). Assume that the robot request (or asks) its position measurements from a sensor and uses a linear quadratic Gaussian (LQG) controller with the design parameters from its quadratic cost function given by $Q = 10I$ and $R = 0.1I$. The black curve in Figure 2 illustrates the trajectory of the robot with a honest sensor as a reference. Evidently, in this case, the robot stays fairly close to the origin. Now, assume that the sensor acts strategically. In fact, the sensor wants the robot to deviate from the origin (i.e., its safe place) so that the sensor (perhaps cooperating with an adversary) can capture it[2]. The solid blue curve in Figure 2 shows the case where the strategic sensor can send any arbitrary message to the system operator and the operator does not compensate for the strategic nature of the sensor. This way, the sensor can make sure that the robot closely follows its desired path shown by the dashed blue curve in Figure 2, which is generated randomly, to get far from the origin. However, if the receiver reacts "optimally" (in a specific sense appropriately defined in the next section) to the transmitted messages of the sensor, it will follow the path illustrated by the red curve in Figure 2. This path clearly outperforms the blue one; however, it is not as good as the case with a honest sensor. This degradation in closed-loop performance is the price of strategicness of the sensor. In the remainder of this paper, this problem is formally introduced and its solution is presented.

[1] i.i.d. stands for independently and identically distributed.

[2] A story reminiscent of `https://www.wired.com/2011/12/iran-drone-hack-gps/`

3. Problem Formulation

Consider the discrete-time linear time-invariant dynamical system, denoted by $\mathcal{P}$ in Figure 1, of the form

$$x[k+1] = Ax[k] + Bu[k] + w[k],\ x[0] = 0, \tag{1}$$

where $x[k] \in \mathbb{R}^n$ is the state of system, $u[k] \in \mathbb{R}^m$ is the control input, and $w[k] \in \mathbb{R}^n$ is the process noise. Here, the pair A and B are model matrices of the system (i.e., matrices containing the parameter of the system modelling its state transition). The process noise models the unknown effects of the environment or the uncertainties of the system and is assumed to be a sequence of i.i.d. zero-mean Gaussian random variable with variance $W \in \mathcal{S}^n_{++}$. The system operator, denoted by $\mathcal{O}$ in Figure 1, wishes to minimize the cost function

$$J_{\mathrm{c}} = \mathbb{E}\Big\{ \sum_{k=0}^{T} x[k]^\top Q x[k] + \sum_{k=0}^{T-1} u[k]^\top R u[k] \Big\}, \tag{2}$$

where $Q \in \mathcal{S}^n_+$ and $R \in \mathcal{S}^n_{++}$ with $\mathcal{S}^n_+$ and $\mathcal{S}^n_{++}$, respectively, denoting the sets of positive semi-definite and positive definite matrices. To achieve this task, the system operator $\mathcal{O}$ deploys a possibly strategic sensor, denoted by $\mathcal{S}$ in Figure 1, to report the measurements of the states of the dynamical system $\mathcal{P}$. The sensor measures the state $x[k] \in \mathbb{R}^n$ directly and sends a message $y[k] \in \mathbb{R}^p$ to the operator. The assumption that the sensor has access to the perfect full state measurements of the state is rather conservative; however, the results based on this assumption provide the worst case analysis of the performance of the closed-loop system in the presence of a strategic sensor. This can be used as a stepping stone towards understanding the problem and relaxing this assumption in the future. The cost function of the sensor is assumed to be of the form

$$J_{\mathrm{s}} = \mathbb{E}\Big\{ \sum_{k=0}^{T} \begin{bmatrix} x[k] \\ \theta[k] \end{bmatrix}^\top Q_{\mathrm{s}} \begin{bmatrix} x[k] \\ \theta[k] \end{bmatrix} + \sum_{k=0}^{T-1} u[k]^\top R_{\mathrm{s}} u[k] \Big\}, \tag{3}$$

where $Q_{\mathrm{s}} \in \mathcal{S}^{n+n_{\mathrm{s}}}_+$ and $R_{\mathrm{s}} \in \mathcal{S}^m_+$ are weighting matrices and $\theta[k] \in \mathbb{R}^{n_{\mathrm{s}}}$ is the private state of the sensor governed by

$$\theta[k+1] = A_{\mathrm{s}}\theta[k] + Fx[k] + B_{\mathrm{s}}u[k] + v[k],\ \theta[0] = 0, \tag{4}$$

where $v[k] \in \mathbb{R}^{n_{\mathrm{s}}}$ denotes the process noise, which is assumed to be a sequence of i.i.d. zero-mean Gaussian random variable with variance $V \in \mathcal{S}^{n_{\mathrm{s}}}_{++}$. The sensor is interested in minimizing the cost function in (3). For instance, if the sensor $\mathcal{S}$ has the desire that the states of the system $\mathcal{P}$ follows its private state $\theta[k]$ (evidently in the case where $n_s = n$), it must select

$$Q_{\mathrm{s}} = \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}, R_{\mathrm{s}} = 0, F = 0, B_{\mathrm{s}} = 0.$$

Alternatively, if the sensor is benevolent, it must select

$$Q_{\mathrm{s}} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}, R_{\mathrm{s}} = R, A_{\mathrm{s}} = 0, F = 0, B_{\mathrm{s}} = 0. \tag{5}$$

Clearly, many other cases can fit the above description and are not discussed.

The timing of the game is as follows. The sensor $\mathcal{S}$ uses the conditional probability density function $\gamma_k(y[k]|x[0], \dots, x[k], \theta[0], \dots, \theta[k], y[1], \dots, y[k-1])$ to determine its message $y[k]$ at time step $k \in \{1, \dots, T-1\}$. Note that the sensor does not transmit any message at $k = 0$ as $x[0]$ is deterministically known to be zero and thus the system operator would not listen to it anyhow. Let $\gamma := (\gamma_1, \dots, \gamma_{T-1})$ denote the tuple of the conditional probability density functions over the entire horizon. The set of all such tuples of conditional probability density functions is denoted by Γ. At time step $k \in \{0, \dots, T-1\}$, the system operator $\mathcal{O}$ uses the conditional probability density function $\xi_k(u[k]|x[0], \theta[0], y[1], \dots, y[k])$ to determine its control action $u[k]$. Note that the ultimate goal of the system operator is to control the system but, as a by-product, it also extracts an estimate of the state of the system. Similarly, let $\xi := (\xi_0, \dots, \xi_{T-1})$ and denote the set of all such tuples of conditional probability density functions by Ξ. Clearly, both cost functions in (2) and (3) are functions of the policies of the sensor γ and the system operator ξ. Considering possibly randomized control policies is an unorthodox choice as, most often in the literature, controllers are assumed to be fixed and deterministic. However, it should be noted that deterministic control policies are a subset of randomized ones (deterministic policies select the preferred control signal with probability one and selecting the rest of the options with probability zero). In fact, as later proved in this paper, the equilibrium consist of only deterministic policies. This observation extends the earlier results from estimation with strategic sensors in which it was also proved that the optimal policies are deterministic (Farokhi et al., 2017). Therefore, the notations $J_{\mathrm{c}}(\gamma, \xi)$ and $J_{\mathrm{s}}(\gamma, \xi)$ is used to denote these cost functions. Let Ξ^{Γ} denote the sets of all mappings from Γ to Ξ. Hence, for any $\psi \in \Xi^{\Gamma}$ and any given γ, $\psi(\gamma)$ is a bundle of conditional probability density functions in Ξ. Now, we are ready to define the equilibria of the game.

Definition 1. *A tuple $(\psi^*, \gamma^*) \in \Xi^{\Gamma} \times \Gamma$ constitutes an equilibrium if*

$$\psi^* \in \underset{\psi \in \Xi^{\Gamma}}{\mathit{argmin}}\, J_{\mathrm{c}}(\gamma^*, \psi(\gamma^*)), \tag{6a}$$

$$\gamma^* \in \underset{\gamma \in \Gamma}{\mathit{argmin}}\, J_{\mathrm{s}}(\gamma, \psi^*(\gamma)). \tag{6b}$$

The equilibrium describes the best way that the sensor $\mathcal{S}$ can persuade the system operator $\mathcal{O}$ to achieve its goals and the best strategy of the system operator to counteract the strategic nature of the sensor. It should be noted that the equilibrium notion in Definition 1 corresponds to a pure strategy Nash equilibrium rather than a Stackelberg equilibrium. In this paper, instead of the usual decision space of control laws Ξ, the system operator selects a function from Ξ^{Γ} (i.e., its action space is much larger and thus can be more robust to the strategic sensor). This function can be seen as a response to the sensor's policy because the implemented control law $\psi(\gamma)$ is in fact a function of γ. For the Stackelberg equilibrium, the response of the system operator must be optimal for all choices of γ, i.e., $\psi^* \in \mathrm{argmin}_{\psi \in \Xi^{\Gamma}} J_{\mathrm{c}}(\gamma, \psi(\gamma)), \forall \gamma \in \Gamma$, rather than optimality for just γ^* as in (6a). This observation places the notion of equilibrium in Definition 1 between the Nash equilibrium (where the action space of the system

operator is only the set of permissible control laws) and the Stackelberg equilibrium. With these definitions in hand, we are ready to construct an equilibrium of the game and study its properties.

4. Main Results

Similar to other signalling games, e.g., (Farokhi et al., 2015; Sobel, 2012), the described game can also admit a babbling equilibrium, i.e., an equilibrium in which the sensor does not provide any useful information (i.e., $y[k]$ is statistically independent of $x[1], \ldots, x[k]$) and the system operator completely ignores the sensor ($u[k]$ is statistically independent of $y[1], \ldots, y[k]$). We are not interested in any of those equilibria. In the next theorem, we construct a non-trivial equilibrium for the game.

Theorem 1. *Let $K[k] \in \mathbb{R}^{m\times n}$, $L[k] \in \mathbb{R}^{(n+n_{\mathrm{s}})\times p}$, $C[k] \in \mathbb{R}^{p\times(n+n_{\mathrm{s}})}$, $X[k] \in \mathcal{S}_+^n$, $Y[k] \in \mathcal{S}_+^{n+n_{\mathrm{s}}}$, and $P[k] \in \mathcal{S}_+^{2n+2n_{\mathrm{s}}}$ be a joint solution (if any exists) of the following set of non-linear equations:*

$$K[k] = -(BX[k+1]B^\top + R)^{-1}B^\top X[k+1]A, \tag{7a}$$

$$X[k] = Q + A^\top X[k+1]A - A^\top X[k+1]B(BX[k+1]B^\top + R)^{-1}B^\top X[k+1]A, \tag{7b}$$

$$X[T] = Q, \tag{7c}$$

$$L[k] = -\overline{Y}[k]C[k]^\top (C[k]\overline{Y}[k]C[k]^\top)^{-1} \tag{7d}$$

$$\overline{Y}[k] = \begin{bmatrix} A+BK[k-1] & 0 \\ F & A_{\mathrm{s}} \end{bmatrix}^\top Y[k] \begin{bmatrix} A+BK[k-1] & 0 \\ F & A_{\mathrm{s}} \end{bmatrix} + W[k], \tag{7e}$$

$$Y[k+1] = (I + L[k]C[k])\overline{Y}[k], \quad Y[1] = 0, \tag{7f}$$

$$C[k] = \overline{B}[k]^\top P[k+1]\overline{A}[k] \begin{bmatrix} I_{n+n_{\mathrm{s}}} \\ 0_{(n+n_{\mathrm{s}})\times(n+n_{\mathrm{s}})} \end{bmatrix}, \tag{7g}$$

$$P[k] = \overline{Q}[k] + \overline{A}[k]^\top P[k+1]\overline{A}[k] - \overline{A}[k]^\top P[k+1]\overline{B}[k](\overline{B}[k]^\top P[k+1]\overline{B}[k])^\dagger \times \overline{B}[k]^\top P[k+1]\overline{A}[k], \ P(T) = \overline{Q}[T], \tag{7h}$$

with $\overline{B}[k], \overline{A}[k], \overline{Q}[k]$ defined as in

$$\overline{A}[k] := \left[\begin{array}{cc|c} A & 0 & \begin{bmatrix} B \\ B_{\mathrm{s}} \end{bmatrix} K[k] \begin{bmatrix} I + L_x[k]C_x[k] & L_x[k]C_\theta[k] \end{bmatrix} \\ F & A_{\mathrm{s}} & \\ \hline 0 & 0 & I + L_x[k]C_x[k] \quad L_x[k]C_\theta[k] \\ 0 & 0 & L_\theta[k]C_x[k] \quad I + L_\theta[k]C_\theta[k] \end{array}\right] \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & A & 0 \\ 0 & 0 & F & A_{\mathrm{s}} \end{bmatrix}, \tag{8a}$$

$$\overline{B}[k] := -\begin{bmatrix} BK[k]L_x[k] \\ B_{\mathrm{s}}K[k]L_x[k] \\ L[k] \end{bmatrix}, \tag{8b}$$

$$\overline{Q}[k] := \left[\begin{array}{c|c} Q_{\mathrm{s}} & 0 \\ \hline 0 & \begin{bmatrix} (I + L_x[k]C_x[k])^\top \\ C_\theta[k]^\top L_x[k]^\top \end{bmatrix} K[k]^\top R_{\mathrm{s}} K[k] \begin{bmatrix} (I + L_x[k]C_x[k])^\top \\ C_\theta[k]^\top L_x[k]^\top \end{bmatrix}^\top \end{array}\right]. \tag{8c}$$

Then, there exists an equilibrium (γ^*, ψ^*) in which the system operator employs the

policy ψ^* such that, with probability one, the control signal is given by

$$u[k]{=}K[k]\hat{x}[k], \tag{9a}$$

$$\begin{bmatrix}\hat{x}[k]\\ \hat{\theta}[k]\end{bmatrix}{=}(I+L[k]C[k])\left(\begin{bmatrix}A & 0\\ F & A_{\mathrm{s}}\end{bmatrix}\begin{bmatrix}\hat{x}[k-1]\\ \hat{\theta}[k-1]\end{bmatrix} + \begin{bmatrix}B\\ B_{\mathrm{s}}\end{bmatrix}u[k-1]\right){+}\begin{bmatrix}L_x[k]\\ L_\theta[k]\end{bmatrix}y[k],\ \begin{bmatrix}\hat{x}[-1]\\ \hat{\theta}[-1]\end{bmatrix}{=}\begin{bmatrix}0\\ 0\end{bmatrix} \tag{9b}$$

and the sensor employs the policy γ^* such that, with probability one, the transmitted message is given by

$$y[k] = C[k]\begin{bmatrix}x[k]\\ \theta[k]\end{bmatrix}. \tag{10}$$

Further, all pairs of policies (γ^*, ψ^*) constructed by scaling the transmitted message $y[k]$ using an invertible matrix $\Phi[k] \in \mathbb{R}^{p\times p}$ with probability one are still equilibria of the game.

Proof. If the system operator follows the policy in the statement of the theorem, we get

$$\begin{bmatrix}x[k+1]\\ \theta[k+1]\\ \hat{x}[k]\\ \hat{\theta}[k]\end{bmatrix}{=}\overline{A}[k]\begin{bmatrix}x[k]\\ \theta[k]\\ \hat{x}[k-1]\\ \hat{\theta}[k-1]\end{bmatrix}{+}\overline{B}[k]y[k],\ \begin{bmatrix}x[0]\\ \theta[0]\\ \hat{x}[-1]\\ \hat{\theta}[-1]\end{bmatrix}{=}\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix},$$

where $\overline{A}[k]$ and $\overline{B}[k]$ are given in (8). Therefore, the problem of finding the optimal message $y[k]$ for the sensor becomes an optimal control problem of the system above with the cost function

$$\sum_{k=0}^{T}\begin{bmatrix}x[k]\\ \theta[k]\\ \hat{x}[k-1]\\ \hat{\theta}[k-1]\end{bmatrix}^{\top}\overline{Q}[k]\begin{bmatrix}x[k]\\ \theta[k]\\ \hat{x}[k-1]\\ \hat{\theta}[k-1]\end{bmatrix},$$

where $\overline{Q}[k]$ is given in (8). Note that a full state feedback policy can be calculated as the sensor can construct $\hat{x}[k]$ and $\hat{\theta}[k]$ as it has access to more information as the system operator (i.e., it knows at least what the system operator knows). The solution to this problem is a linear state feedback policy of the form

$$y[k] = \underbrace{\begin{bmatrix}C_x[k] & C_\theta[k] & C_{\hat{x}}[k] & C_{\hat{\theta}}[k]\end{bmatrix}}_{:=\overline{C}[k]}\begin{bmatrix}x[k]\\ \theta[k]\\ \hat{x}[k-1]\\ \hat{\theta}[k-1]\end{bmatrix},$$

where

$$\overline{C}[k] = -(\overline{B}[k]^{\top}P[k+1]\overline{B}[k])^{\dagger}\overline{B}[k]^{\top}P[k+1]\overline{A}[k],$$

with $X^\dagger$ denoting the Moore-Penrose pseudoinverse of any square matrix X with the convention that $0^\dagger = 0$ and $P[k]$ defined using the reverse discrete Riccati equation

$$\begin{aligned} P[k] =& \overline{Q}[k] + \overline{A}[k]^\top P[k+1]\overline{A}[k] - \overline{A}[k]^\top P[k+1]\overline{B}[k](\overline{B}[k]^\top P[k+1]\overline{B}[k])^\dagger \\ & \times \overline{B}[k]^\top P[k+1]\overline{A}[k],\ P(T) = \overline{Q}[T]. \end{aligned}$$

The fact that the matrix $\overline{B}[k]^\top P[k+1]\overline{B}[k]$ is potentially not invertible does not create an issue in our setup as we are only considering the finite horizon problem and stability is not a problem (in fact a similar approach as in (Ferrante and Ntogramatzidis, 2015) to our problem). Noting the system operator can always subtract $C_{\hat{x}}[k]\hat{x}[k-1] + C_{\hat{\theta}}[k]\hat{\theta}[k-1]$ from the transmitted message, the sensor without the loss of generality can transmit

$$y[k] = \begin{bmatrix} C_x[k] & C_\theta[k] \end{bmatrix} \begin{bmatrix} x[k] \\ \theta[k] \end{bmatrix}.$$

Clearly, in response to this message, the system operator employs the control law

$$u[k] = K[k]\hat{x}[k],$$

where $\hat{x}[k] = \mathbb{E}\{x[k]|y[0], \ldots, y[k]\}$ and

$$K[k] = -(BX[k+1]B^\top + R)^{-1}B^\top X[k+1]A,$$

with $X[k]$ defined using the reverse discrete Riccati equation

$$X[k]=Q + A^\top X[k+1]A - A^\top X[k+1]B(BX[k+1]B^\top + R)^{-1}B^\top X[k+1]A,$$

with $X[T] = Q$. To estimate the state, the system operator uses the *Kalman filter* of the form

$$\begin{aligned} \begin{bmatrix} \hat{x}[k] \\ \hat{\theta}[k] \end{bmatrix} =& \left(I + \begin{bmatrix} L_x[k] \\ L_\theta[k] \end{bmatrix} \begin{bmatrix} C_x[k] & C_\theta[k] \end{bmatrix} \right) \\ & \times \left(\begin{bmatrix} A & 0 \\ F & A_{\mathrm{s}} \end{bmatrix} \begin{bmatrix} \hat{x}[k-1] \\ \hat{\theta}[k-1] \end{bmatrix} + \begin{bmatrix} B \\ B_{\mathrm{s}} \end{bmatrix} u[k-1] \right) + \begin{bmatrix} L_x[k] \\ L_\theta[k] \end{bmatrix} y[k], \quad \begin{bmatrix} \hat{x}[-1] \\ \hat{\theta}[-1] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

where

$$\begin{aligned} \begin{bmatrix} L_x[k] \\ L_\theta[k] \end{bmatrix} =& -\overline{Y}[k] \begin{bmatrix} C_x[k]^\top \\ C_\theta[k]^\top \end{bmatrix} \left(\begin{bmatrix} C_x[k]^\top \\ C_\theta[k]^\top \end{bmatrix}^\top \overline{Y}[k] \begin{bmatrix} C_x[k]^\top \\ C_\theta[k]^\top \end{bmatrix} \right)^{-1}, \\ \overline{Y}[k] =& \begin{bmatrix} A & 0 \\ F & A_{\mathrm{s}} \end{bmatrix}^\top Y[k] \begin{bmatrix} A & 0 \\ F & A_{\mathrm{s}} \end{bmatrix} + W[k], \\ Y[k+1] =& \left(I + \begin{bmatrix} L_x[k] \\ L_\theta[k] \end{bmatrix} \begin{bmatrix} C_x[k] & C_\theta[k] \end{bmatrix} \right) \overline{Y}[k]. \end{aligned}$$

This shows that both the sensor and the system operator are responding optimally to each other's policies. Therefore, none of them have an incentive to deviate from their respective policies. This concludes the proof. □

Remark 1 (Numerical Complexities)**.** *The set of nonlinear equations in* (7) *can be solved using the Newton's method given that the algorithm can be initialized in the vicinity of the solution. We can also switch between solving the set of linear equations in* (7a)-(7f) *(when fixing* $C[k]$*) and the set of linear equations in* (7g)-(7h) *(when fixing* $K[k]$ *and* $L[k]$*). This method is easier to implement but does not have any convergence guarantees in general. For the case where* $J_{\rm s} = -J_{\rm c}$ *(i.e., the game is zero sum), the convergence of this method may not be guaranteed because the convexity-concavity assumption required for the convergence of the best response learning methods (see, e.g., (Hofbauer and Sorin, 2006)) no longer holds. In general, convergence requires existence of a convex potential function, which might hold for the setup of this paper. We have used this method in the numerical examples of the paper.*

Remark 2 (Benevolent Sensor)**.** *It can be easily shown that if the sensor is benevolent, e.g., see* (5)*, the equilibrium in Theorem 1 takes a specific form that, with probability one,* $y[k] = x[k]$ *and* $u[k] = K[k]y[k]$*.*

Now that we can calculate the equilibrium, we are ready to investigate some of its numerical properties in the following section.

5. Numerical Example

In this section, we revisit the motivating example in Section 2 and further investigate the properties of the equilibrium for various objectives of the sensor. To be complete, we first discuss some aspects of the example that were not described at that stage (due to its preliminary nature). Firstly, in this example, we consider the following model

$$\theta[k+1] = \begin{bmatrix} 1.1 & 0 \\ 0 & 1.1 \end{bmatrix} \theta[k] + v[k],\ \theta[0] = 0, \tag{11}$$

where $\theta[k]$ denotes the path that the sensor wants the robot to follow and $v[k]$ is a sequence of i.i.d. zero-mean Gaussian random variable with variance $V = I$. The fact that the sensor wants the robot to follow its private path of $\theta[k]$ is captured by the cost function in (3) with

$$Q_{\rm s} = \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}, \quad R_{\rm s} = 0.$$

Due to the symmetry of the problem, there exists an equilibrium at which, $C[k]$ in (7g) takes the special form

$$C[k] = \begin{bmatrix} c_1[k] & 0 & c_2[k] & 0 \\ 0 & c_1[k] & 0 & c_2[k] \end{bmatrix}.$$

The gains $c_1[k]$ and $c_2[k]$ are shown in Figure 3. Evidently, the strategic sensor never flat out lies as, in that case, the system operator would ignore its messages completely and it achieves nothing towards minimizing its cost. Therefore, it is best for the sensor to combine its hidden agenda with the truth to convince the system operator.

The performance degradation caused by the strategic nature of the sensor can be captured by the ratio of the cost of the system operator if the sensor was completely

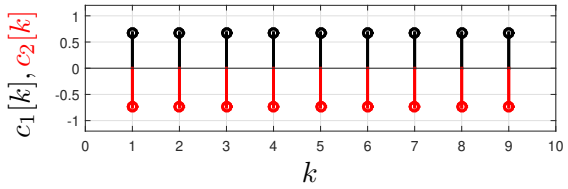

Figure 3. The gains $c_1[k]$ (black colour) and $c_2[k]$ (red colour) versus k.

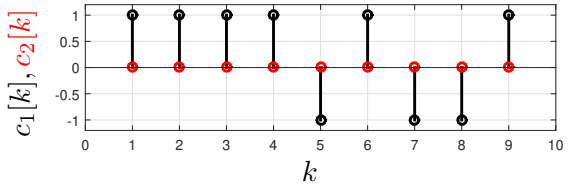

Figure 4. The gains $c_1[k]$ (black colour) and $c_2[k]$ (red colour) versus k.

honest versus the cost of the system operator when it follows the policy of Theorem 1. This ratio in this example (for the described setup) is 6.25. However, this is far better than the case where the system operator does not follow the strategy in Theorem 1; see Figure 2 for comparison.

Now, we can consider the case where the sensor is benevolent, e.g., see (5). As expected, the equilibrium in Theorem 1 (calculated numerically) takes a specific form that, with probability one, $y[k] = x[k]$ and $u[k] = K[k]y[k]$. This is shown in Figure 4.

6. Conclusions and Future Work

This paper considered the problem of feedback control using a strategic sensor. The sensor can manipulate the state measurements transmitted to the system operator to steer the system towards its goals. The problem is formulated as a game. An equilibrium of the game is constructed and its properties are investigated. Interestingly, the strategic sensor does not completely lie so as the system operator does not ignore its messages. Future work can focus on the case where several strategic sensors with possibly conflicting objectives communicate with the system operator.

7. Acknowledgments

The work was supported by a McKenzie Fellowship, a grant (MyIP: ID6874) from Defence Science and Technology Group (DSTG), and VESKI Victoria Fellowship by the Victorian State Government.

References

Başar, T. and Olsder, G. J. (1998). *Dynamic noncooperative game theory*. SIAM.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica*, 70(4):1379–1401.

Cardenas, A. A., Amin, S., and Sastry, S. (2008). Secure control: Towards survivable cyber-physical systems. In *Proceedings of the 28th International Conference on Distributed Computing Systems Workshops*, pages 495–500.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451.

Dobakhshari, D. G., Li, N., and Gupta, V. (2016). An incentive-based approach to distributed estimation with strategic sensors. In *Proceedings of the 55th IEEE Conference on Decision and Control*, pages 6141–6146.

Farokhi, F., Sandberg, H., Shames, I., and Cantoni, M. (2015). Quadratic Gaussian privacy games. In *Proceedings of the 54th IEEE Annual Conference on Decision and Control*, pages 4505–4510.

Farokhi, F., Teixeira, A. M. H., and Langbort, C. (2017). Estimation with strategic sensors. *IEEE Transactions on Automatic Control*, 62(2):724–739.

Ferrante, A. and Ntogramatzidis, L. (2015). A note on finite-horizon LQ problems with indefinite cost. *Automatica*, 52:290–293.

Golosov, M., Skreta, V., Tsyvinski, A., and Wilson, A. (2014). Dynamic strategic information transmission. *Journal of Economic Theory*, 151:304–341.

Hendrickx, J. M., Johansson, K. H., Jungers, R. M., Sandberg, H., and Sou, K. C. (2014). Efficient computations of a security index for false data attacks in power networks. *IEEE Transactions on Automatic Control*, 59(12):3194–3208.

Hofbauer, J. and Sorin, S. (2006). Best response dynamics for continuous zero-sum games. *Discrete and Continuous Dynamical Systems Series B*, 6(1):215.

Mo, Y. and Sinopoli, B. (2009). Secure control against replay attacks. In *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing*, pages 911–918.

Pasqualetti, F., Dörfler, F., and Bullo, F. (2013). Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729.

Sandberg, H., Teixeira, A., and Johansson, K. H. (2010). On security indices for state estimators in power networks. In *Proceedings of the 1st Workshop on Secure Control Systems (CPSWEEK 2010)*.

Sobel, J. (2012). Signaling games. In Meyers, R. A., editor, *Computational Complexity: Theory, Techniques, and Applications*, pages 2830–2844. Springer New York, New York, NY.

Teixeira, A., Sou, K. C., Sandberg, H., and Johansson, K. H. (2013). Quantifying cyber-security for networked control systems. In Tarraf, D. C., editor, *Control of Cyber-Physical Systems*, volume 449 of *Lecture Notes in Control and Information Sciences*, pages 123–142. Springer International Publishing.

Westenbroek, T., Dong, R., Ratliff, L. J., and Sastry, S. S. (2017). Statistical estimation in competitive settings with strategic data sources. In *Proceedings of the 56th IEEE Confefence on Decision and Control*.