Consensus-based Distributed State Estimation of Biofilm in Reverse Osmosis Membranes by WSNs

D. Alonso-Román, C. Asensio-Marco and B. Beferull-Lozano
Dept. of Information & Communication Technology, WISENET Lab, University of Agder
Jon Lilletunsvei 3, Grimstad, 4879, Norway
{daniel.alonso,cesar.asensio,baltasar.beferull}@uia.no

ABSTRACT

The appearance of biofilm has become a serious problem in many reverse osmosis based systems such as the ones found in water treatment and desalination plants. In these systems, the use of traditional techniques such as pretreatment or dozing biocides are not effective when the biofilm reaches an irreversible attachment phase. In this work, we present a framework for the use of a WSN as an estimator of the biofilm evolution in a reverse osmosis membrane so that effective solutions can be applied before the irreversible phase is attained. This design is addressed in a complete distributed and decentralized fashion, and subject to realistic constraints where cooperation between nodes is performed under unreliable links.

CCS CONCEPTS

•**Computing methodologies** →*Distributed algorithms;*

KEYWORDS

Biofilm, distributed, estimation, tracking, WSNs

ACM Reference format:
D. Alonso-Román, C. Asensio-Marco and B. Beferull-Lozano. 2016. Consensus-based Distributed State Estimation of Biofilm in Reverse Osmosis Membranes by WSNs. In *Proceedings of The 3rd International Workshop on Cyber-Physical Systems for Smart Water Networks, Pittsburgh, PA USA, April 2017 (CySWATER 2017),* 5 pages.
DOI: http://dx.doi.org/10.1145/3055366.3055369

1 INTRODUCTION

A biofilm is an accumulation of free-floating microorganisms, that become attached to a specific surface. If the microorganisms are not immediately separated and removed, they might start taking some nutrients from the bulk liquid and the subsequent growing process may lead to the appearance of biofilm. In water treatment and distribution plants based on reverse osmosis (RO), the formation of biofilm might result in a decrease of the quality of the water, and a reduction of their performance [1]. The biofilm formation affects the flux of water that is processed through the membranes, which requires an increase in the pressure applied by the pumps to maintain the flux, resulting in a higher energy consumption. It has been shown in [2] that the application of pretreatment or dozing of biocides, such as chlorine, is not an effective solution once the biofilm reaches an irreversible attachment phase and then the RO membrane replacement is needed.

Continuous monitoring of the water in the plant may allow to accurately estimate the biofilm evolution, enabling the application of control measures before it reaches an irreversible attachment phase. This potentially reduces the cost in the use of chemicals, membrane replacements and energy consumption, in addition to a more correct operation of the plant. In this context, wireless sensor networks (WSNs) have emerged as a powerful technology for the monitoring and control of industrial environments. These networks can be deployed without disrupting the normal operation of the system in areas that are inaccessible to wired and bulky devices, making possible to establish correlation between different measurements taken in different parts of the structure.

A key property of these networks is that the sensor devices, which may have limited computational and storage resources, are able to solve complex tasks by means of cooperation between them. A widely-used distributed technique is the consensus algorithm [3], which its goal is to obtain, in a distributed way, a common value by iterative processing the measurements collected by sensor nodes. These algorithms avoid the need of performing all the computations at one or more sink nodes, thus, reducing congestion around them and incrementing the robustness of the network against nodes failures and attacks. However, since cooperation between nodes is performed under random phenomena such as interferences, packet losses and fading, the different communication topologies that arise during the process are completely random and, in general, non symmetric. Under these conditions, the consensus value becomes a random variable, which is, in general, different from the desired value and consequently suboptimal [4].

In this work, given the equations that model the evolution of the biofilm, we show how the Kalman filter, a centralized tool for the estimation of the biofilm growing process, can be decomposed in a consensus-based basis. We explain how the link failures affect the filter accuracy, and how the error can be mitigated by redesigning the estimation process. We show how our new scheme is able to mitigate the error, and how its overall performance approximates the one of a centralized approach. Finally, we present numerical results that shows clearly the stated results.

2 PROBLEM FORMULATION

A wireless sensor network is deployed inside the pipeline and before the RO membrane, as it is shown in Figure 1. Along the pipeline, and under some assumptions, the one-dimensional dependence between the value of the flux of nutrient into the biofilm $j_S(L_F)$,

CySWATER 2017, Pittsburgh, PA USA

DOI: http://dx.doi.org/10.1145/3055366.3055369

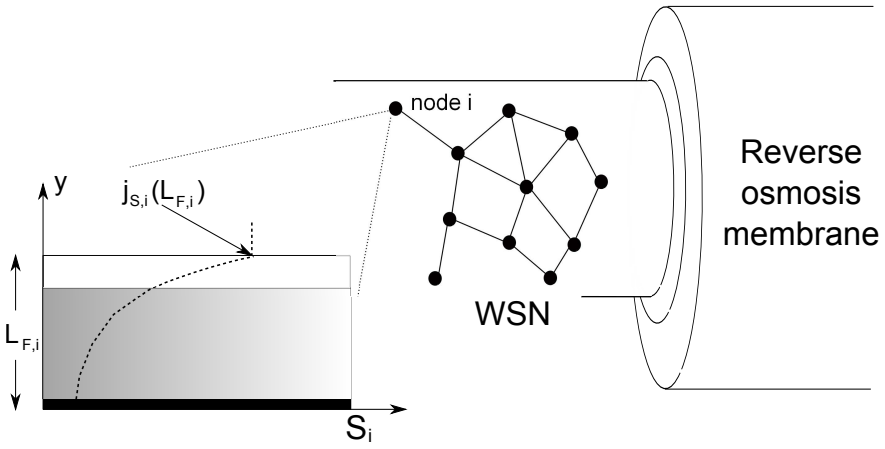


Figure 1: Deployment of a wireless sensor network in RO membrane in order to track the biofilm evolution. Each sensor i in the network obtains a local measurement of the biofilm thickness $L_{F,i}$ and a noisy observation of the flux of substrate in the biofilm surface $j_{S,i}(L_{F,i})$. We assume a multi-one-dimension model, where this flux changes at each sensor but remains constant in its local domain. Nodes can communicate with neighbors within their wireless range.

its thickness L_F and the biomass concentration X_F is given by [2]:

$$j_S(L_F) = -D\frac{\partial S}{\partial y}\Big|_{y=L_F} = q_{\max} L_F X_F \tag{1}$$

where S is the substrate concentration and $q_{\max}$ is the maximum specific substrate conversion rate.

If we assume a flat and homogeneous biofilm, it can be shown [7] that the discrete evolution in time of the biomass concentration is:

$$X_F(k+1) = \left[1 + \alpha\left(q_{\max} - \frac{\partial u_F}{\partial y}\right)\right] X_F(k) \tag{2}$$

We consider that the substrate concentration S changes along the pipeline and is different at each sensor location (S_i). However, we assume that in the range of each node S_i only changes along the y-coordinate and remains constant along the other two coordinates (see Figure 1). In addition, since all the nodes are affected by the same nutrient, we assume that all the nodes have the same biomass concentration. Each sensor i has a flat substratum where the biofilm may appear, and two sensors to measure the flux of substrate in the biofilm surface $j_{S,i}(L_{F,i})$ and the thickness of the biofilm L_i at its location. While the measurement of the flux involves an observation noise, for the sake of simplicity of our model, we assume that the biofilm thickness is observed without error (an additional observation error could be introduced). Based on that, we consider that each node i obtains a noisy observation φ_i of the substrate flux into the biofilm at its location given by:

$$\varphi_i = j_{S,i}(L_{F,i}) + w_i$$

where the observation noise w_i is assumed to be zero mean Gaussian with variance σ_w^2 and spatially uncorrelated. By applying equation (1), this expression becomes:

$$\varphi_i = q_{\max} L_{F,i} X_F + w_i$$

By taking $h_i = q_{\max} L_{F,i}$, we have the following observation equation in matrix form:

$$\boldsymbol{\varphi} = \mathbf{h} X_F + \mathbf{w} \tag{3}$$

which expresses the substrate flux into the biofilm in all nodes of the network as a noisy and distorted version of the biomass concentration X_F. Similarly, we also assume that the time evolution of the biomass concentration in (2) is affected by a zero mean Gaussian noise $v(k)$ of variance σ_v^2 and time uncorrelated. By doing $a = \left[1 + \alpha\left(q_{\max} - \frac{\partial u_F}{\partial y}\right)\right]$, previous expression becomes:

$$X_F(k+1) = aX_F(k) + v(k) \tag{4}$$

Expressions (3) and (4) are the linear and Gaussian state-space representation of our biofilm system and are used by the nodes of the WSN to track the biomass concentration in the biofilm. Each node i is assumed to know h_i, a, σ_w^2 and σ_v^2.

Given the observations (3) and the biomass concentration evolution (4), it is well known that the recursive scheme defined by the Kalman filter [5] provides the optimal estimate of the system state. Whenever a new set of observations $\boldsymbol{\varphi}(k)$ is obtained, the filter refines its *a priori* estimation $\hat{X}_F^-(k)$ to yield the *a posteriori* estimation $\hat{X}_F^+(k)$, as follows:

$$\hat{X}_F^+(k) = \hat{X}_F^-(k) + \mathbf{g}^T(k)\left[\boldsymbol{\varphi}(k) - \mathbf{h}\hat{X}_F^-(k)\right] \tag{5}$$

where $\mathbf{g}^T$ is the gain of the filter. Then, the filter projects over time this estimation to get the next *a priori* estimation:

$$\hat{X}_F^-(k+1) = a\hat{X}_F^+(k) \tag{6}$$

If we denote the estimation error variance of $\hat{X}_F^-(k)$ and $\hat{X}_F^+(k)$ by $p_x^-(k)$ and $p_x^+(k)$ respectively, then:

$$p_x^+(k) = p_x^-(k)\left[1 - \mathbf{g}^T(k)\mathbf{h}\right]^2 + \sigma_w^2 \mathbf{g}^T(k)\mathbf{g}(k)$$

and the next *a priori* variance is given by:

$$p_x^-(k+1) = a^2 p_x^+(k) + \sigma_w^2 \tag{7}$$

The optimal filter gain that minimizes the variance of the error, at each time k, is given by:

$$\mathbf{g}(k) = p_x^-(k)\mathbf{h}\left[p_x^-(k)\mathbf{h}^T\mathbf{h} + \sigma_v^2\right]^{-1} \tag{8}$$

Since $\mathbf{h}$, σ_v^2 and σ_w^2 are time-invariant, both the error variance and the filter gain can be pre-computed by running the filter off-line. Thus, the filter only needs to compute expressions (5) and (6) during the estimation process.

3 CONSENSUS-BASED DISTRIBUTED STATE ESTIMATION

The Kalman filter [5] can be easily implemented in a centralized way, where all the necessary information is available at a central entity that computes (5) at each time step. In a distributed scenario, each node i has only access to its own observation $\varphi_i(t)$, however, still must be able to compute expression in (5) by means of exchanging information with one-hop neighbors. The corresponding sequence of communication steps can be modeled as a time-varying graph $\mathcal{G}(n) = (\mathcal{V}, \mathcal{E}(n))$, consisting of a constant set $\mathcal{V}$ of N nodes and a set $\mathcal{E}(n)$ of directed links that changes at each step n. A directed link from any node i to any other node j is denoted by e_{ij}. The random and non-symmetric $N \times N$ matrix $\mathbf{A}(n)$ is the adjacency matrix, whose entry $[\mathbf{A}(n)]_{ij}$ is equal to 1 if $e_{ij} \in \mathcal{E}(n)$ and 0 otherwise. The random set of neighbors of a node i at time n is defined as $\Omega_i(n) = \{j \in \mathcal{V} : e_{ij} \in \mathcal{E}(n)\}$. The degree matrix $\mathbf{D}(n)$ is a diagonal matrix whose entries are $[\mathbf{D}]_{ii} = |\Omega_i(n)|$, and the instantaneous Laplacian matrix is defined as $\mathbf{L}(n) = \mathbf{D}(n) - \mathbf{A}(n)$. We denote by $\mathbf{1}$ and $\mathbf{I}$ the all-one N-dimension vector and the NxN identity matrix, respectively.

At each time k, every node i computes $x_i(0) = Ng_i(k)\varphi_i(k)$, where $\varphi_i(k)$ is its current observation and $g_i(k)$ the i-th entry of $\mathbf{g}(k)$. Then, by denoting $\mathbf{x}(0) = [x_1(0) \dots x_N(0)]$, nodes exchange their values based on the iterative process:

$$\mathbf{x}(n) = \mathbf{W}(n) \dots \mathbf{W}(0)\mathbf{x}(0) = \mathbf{M}(n)\mathbf{x}(0) \tag{9}$$

where $\mathbf{W}(n)$ is the weight matrix described in [2]. If both conditions stated in [3] hold, then $\lim_{n\to\infty} \mathbf{M}(n) = \frac{1}{N}\mathbf{1}\mathbf{1}^T$, namely the nodes asymptotically achieve consensus[1] on the common value $\sum_{i=1}^{N} g_i(k)\varphi_i(k) = \mathbf{g}^T(k)\boldsymbol{\varphi}(k)$. In this case, expression in (5) can be computed by all nodes in a distributed way, and each node can run a local version of the filter identical to the optimal one.

Nevertheless, in a real setting where the connectivity is random due to interferences and packet losses, each instantaneous underlying topology defined by $\mathbf{A}(n)$ is also random, and so is the corresponding weight matrix $\mathbf{W}(n)$. Therefore, $\lim_{n\to\infty} \mathbf{M}(n) = \frac{1}{N}\mathbf{1}\mathbf{m}^T$ and then $\mathbf{m} = (m_1 \dots m_N)$ is a random vector, which is, in general, different from $\frac{1}{N}\mathbf{1}$. This implies that expression (5) becomes:

$$\hat{X}_F^+(k) = \hat{X}_F^-(k) + \mathbf{g}^T(k)\boldsymbol{\Delta}_m\boldsymbol{\varphi}(k) - \mathbf{g}^T(k)\mathbf{h}\hat{X}_F^-(k) \tag{10}$$

where $\boldsymbol{\Delta}_m = \text{diag}(m_1 \dots m_N)N$. If we evaluate the *a posteriori* estimation error of (10), we have the following [6]:

$$\begin{aligned}\epsilon^+(k) &= X_F(k) - \hat{X}_F^+(k) \\ &= \left[1 - \mathbf{g}^T(k)\mathbf{h}\right] a\epsilon^+(k-1) - \mathbf{g}^T(k)\left[\boldsymbol{\Delta}_m - \mathbf{I}\right]\mathbf{h}aX_F(k-1) \\ &+ \left[1 - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{h}\right] w(k-1) - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{v}(k)\end{aligned}$$

[1]We assume that between k and $k+1$ the network can iterate such that a consensus process as the one described in (9) asymptotically converges.

By taking expectations at both sides, we obtain the estimation error mean:

$$\begin{aligned}\mathbb{E}\left[\epsilon^+(k)\right] = \left[1 - \mathbf{g}^T(k)\mathbf{h}\right] a\mathbb{E}\left[\epsilon^+(k-1)\right] - \\ \mathbf{g}^T(k)\left(\mathbb{E}\left[\boldsymbol{\Delta}_m\right] - \mathbf{I}\right)\mathbf{h}a\mathbb{E}\left[X_F(k-1)\right]\end{aligned}$$

where we have considered that $\boldsymbol{\Delta}_m$ is independent of X_F and $w(k)$ and $\mathbf{v}(k)$ are zero-mean processes. Therefore, the expectation of the error at each step depends not only on the mean at the previous step, but also on a scaled version of the expectation of the state. Given that, in general, none of the two conditions in [3] are met, then $\mathbb{E}\left[\boldsymbol{\Delta}_m\right] \neq \mathbf{I}$. It entails that, unless $\mathbb{E}[X_F(0)] = 0$, this second term never vanishes, hence the error mean is always different from zero. Consequently, this filter is a biased estimator. In addition, the general gain computed in (8) is not optimal for this filter, since it has been obtained by considering expression (5). Then, the error variance is not minimized either.

In the sequel, we explain how the distributed design can be reformulated in order to approximate the performance of the centralized approach. First, we force the filter to be unbiased. Then, and by considering the statistical properties of the consensus, we compute the optimal gain that minimizes the error variance. In order to do that, we first propose the following alternative equation for the state update:

$$\hat{X}_F^+(k) = \hat{X}_F^-(k) + \mathbf{g}^T(k)\boldsymbol{\Delta}_m\boldsymbol{\varphi}(k) - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{h}\hat{X}_F^-(k) \tag{11}$$

that is, an additional consensus process is applied for the nodes to compute in a distributed way the term $\mathbf{g}^T(k)\mathbf{h}\hat{\boldsymbol{\varphi}}^-(k)$. Again, the evaluation of the *a posteriori* estimation error leads us to the following recursive expression:

$$\begin{aligned}\epsilon^+(k) = &\left[1 - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{h}\right] a\epsilon^+(k-1) \\ &+ \left[1 - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{h}\right] w(k-1) - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{v}(k)\end{aligned} \tag{12}$$

which depends only on the error at the previous step, and on the process and observation noises. By taking expectations at both sides, we obtain the estimation error mean:

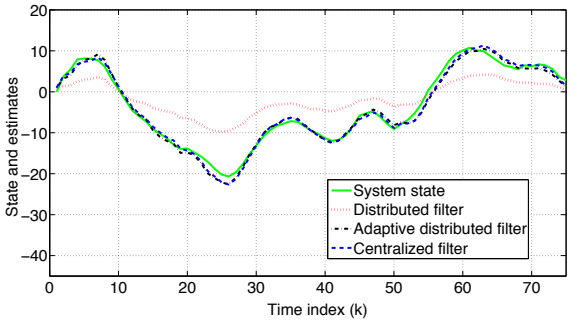
$$\mathbb{E}\left[\epsilon^+(k)\right] = \left[1 - \mathbf{g}^T(k)\mathbb{E}\left[\boldsymbol{\Delta}_m\right]\mathbf{h}\right] a\mathbb{E}\left[\epsilon^+(k-1)\right]$$

Therefore, if we make $\hat{X}_F^+(0) = \mathbb{E}[X_F(0)]$, then $\mathbb{E}[\epsilon^+(k)] = 0$, for all k. Consequently, as long as the initial estimate of X_F is equal to the expected value of $X_F(0)$, the expected value of $\hat{X}_F^+(k)$ is equal to $X_F(k)$, and the proposed estimator is unbiased regardless the value of $\mathbf{g}(k)$.

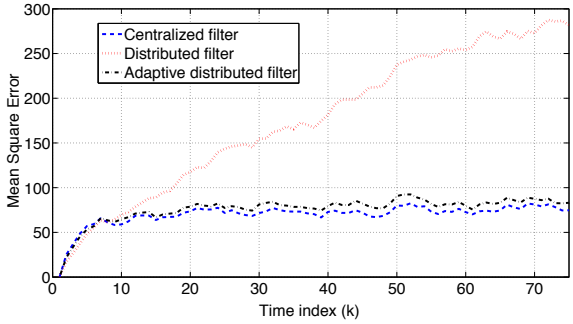
In order to compute the optimal value of the filter gain $\mathbf{g}(k)$, we aim at minimizing the trace of the *a posteriori* error variance $p^+(k) = \mathbb{E}\left[\left(\epsilon^+(k)\right)^2\right]$. By using (12), $p^+(k)$ can be expressed as:

$$\begin{aligned}p^+(k) = &\mathbb{E}\left[\left(1 - \mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{h}\right)^2 p^-(k)\right] \\ &+ \mathbb{E}\left[\left(\mathbf{g}^T(k)\boldsymbol{\Delta}_m\mathbf{v}(k)\right)^2\right]\end{aligned} \tag{13}$$

where we have applied the whiteness of both $w(k)$ and $\mathbf{v}(k)$, their independence of $\epsilon^+(k)$, and the projection of the variance in (7). If we derivate the previous expression with respect to the gain $\mathbf{g}(k)$,



(a) System state and estimations



(b) MSE

Figure 2: State estimation of a scalar parameter via three different versions of the Kalman filter. In (a), the parameter evolution and three estimators are shown for a single realization. It can be seen how the tracking capability of the proposed adaptive filter approaches the optimal one. The square error averaged over $R_2 = 200$ realizations is depicted in (b), for the three estimators.

we have that:

$$\frac{\partial p^+(k)}{\partial \mathbf{g}(k)} = 2\mathbf{g}^T(k)\mathbb{E}\left[\boldsymbol{\Delta}_m \mathbf{h}\mathbf{h}^T \boldsymbol{\Delta}_m p^-(k) + \sigma_v^2 \boldsymbol{\Delta}_m^2\right] - 2p^-(k)\mathbf{h}^T \mathbb{E}\left[\boldsymbol{\Delta}_m\right]$$

If we define the matrix $\mathbf{C}_m = N^2\mathbb{E}\left[\mathbf{m}\mathbf{m}^T\right]$, equal to zero, and solve for $\mathbf{g}(k)$, we obtain the following optimal filter gain:

$$\mathbf{g}(k) = p^-(k)\mathbb{E}[\boldsymbol{\Delta}_m]\left(\mathbf{C}_m \circ \mathbf{h}\mathbf{h}^T p^-(k) + \sigma_v^2 \mathbb{E}[\boldsymbol{\Delta}_m^2]\right)^{-1} \quad (14)$$

where $\circ$ stands for the Hadamard product, namely $(\mathbf{A} \circ \mathbf{B})_{i,j} = (\mathbf{A})_{i,j}(\mathbf{B})_{i,j}$. This expression gives the optimal filter gain as a function of the first and second order statistics of the consensus process.

Finally, we insert in (13) the optimal value of the gain in (14), to obtain the evolution of the error variance:

$$p^+(k) = \left(1 - \mathbf{g}^T(k)\mathbb{E}[\boldsymbol{\Delta}_m]\mathbf{h}\right) p^-(k) \quad (15)$$

The expressions in (14) and (15) become the ones for the centralized case if both conditions in [3] hold.

Therefore, our consensus based Kalman filter is completely defined by the expressions for the state update in (11), the optimal filter gain in (14), and the error variance evolution in (15). Provided that the nodes know the statistical moments of the consensus process, they are able to distributively compute expressions (14) and (15) in a consensus basis. Furthermore, the gain of the filter is adapted to minimize the error variance at each step.

4 NUMERICAL RESULTS

Our Matlab setup includes $N = 40$ nodes randomly deployed in a square area of $L = 50$ meters side with an average number of neighbors equal to 15. We assume that the power of the signal is $\mathcal{E} = 1$. Without loss of generality, the consensus process is based on the Laplacian matrix, such that $\mathbf{W}(n) = \mathbf{I} - \epsilon\mathbf{L}(n)$, with a value of the step size $\epsilon = 1/N$ to ensure convergence [2]. We also assume that the initial value of the signal is $X_F(0) = 0$, and the process noise has a variance $\sigma_v^2 = 10^{-1}$. The observation at each node i is attenuated by a factor $h_i < 1$ and the variance of the observation noise is $\sigma_w^2 = 10$.

Figure 2 shows the result of applying three different Kalman filters: the optimal filter based on a centralized setting, the general consensus-based distributed filter expressed in (10), and the distributed adaptive filter proposed in this work. More precisely, the tracking capabilities can be appreciated in Figure 2(a), where the evolution of the system, as well as the different estimates, are depicted for a single realization of the experiment. Although the centralized solution shows the best performance, the consensus-based distributed filter proposed in the present work exhibits a very similar behavior. It implies that, due to its unbiasedness and its adaptive nature, this filter compensates the error introduced by the random consensus and approximates the optimal estimator. Oppositely, the biased and non-adaptive distributed filter is barely able to track the evolution of the parameter. This is more easily appreciated in Figure 2(b), which shows the mean square error of the three estimators for $R_2 = 200$ realizations of the experiment, and where it can be seen that the biased filter diverges.

ACKNOWLEDGMENTS

This work was supported by the PETROMAKS "Smart-Rig" grant 244205/E30 and the SFI "Offshore Mechatronics" grant 237896/O30, both from the Research Council of Norway.

REFERENCES

[1] Hydrobionets: Autonomous control of large-scale water treatment plants based on self-organized wireless biomem sensor and actuator net- works. http://www.hydrobionets.eu.
[2] A. Radu, J. Vrouwenvelder, M. Van Loosdrecht, and C. Picioreanu, Modeling the effect of biofilm formation on reverse osmosis performance: flux, feed channel pressure drop and solute passage, Journal of Membrane Science, vol. 365, no. 1, pp. 1-15, 2010.
[3] L. Xiao and S. Boyd, Fast linear iterations for distributed averaging, Systems and Control Letters, vol. 53, no. 1, pp. 65-78, 2004.
[4] A. Tahbaz-Salehi and A. Jadbabaie, Consensus over ergodic stationary graph processes, IEEE Trans. on Automatic Control, vol. 55, no. 1, pp. 225-230, 2010.
[5] G. Welch and G. Bishop, An introduction to the Kalman filter, 1995.
[6] B. Beferull-Lozano et al., Adaptive consensus-based distributed kalman filter for wsns with random link failures, International Conference on Distributed Computing in Sensor Systems, pp. 187-192, 2016.
[7] H. Eberl, Mathematical modeling of biofilms. Number 18. IWA publishing, 2006.