Cyberattack Detection for Nonlinear Leader-Following Multi-Agent Systems Using Set-Membership Fuzzy Filtering

Mahshid Rahimifard, Amir M. Moradi Sizkouhi, and Rastko R. Selmic, Senior Member, IEEE

Abstract—This paper is concerned with cyberattack detection in discrete-time, leader-following, nonlinear, multi-agent systems subject to unknown but bounded (UBB) system noises. The Takagi–Sugeno (T-S) fuzzy model is employed to approximate the nonlinear systems over the true value of the state. A distributed cyberattack detection method, based on a new fuzzy set-membership filtering method, which consists of two steps, namely a prediction step and a measurement update step, is developed for each agent to identify two types of cyberattacks at the time of their occurrence. The attacks are replay attacks and false data injection attacks affecting the leader-following consensus. We calculate an estimation ellipsoid set by updating the prediction ellipsoid set with the current sensor measurement data. Two criteria are provided to detect cyberattacks based on the intersection between the ellipsoid sets. If there is no intersection between the prediction set and the estimation set of an agent at the current time instant, a cyberattack on its sensors is declared. Control signal or communication signal data of an agent are under a cyberattack if its prediction set has no intersection with the estimation set updated at the previous time instant. Recursive algorithms for solving the consensus protocol and calculating the two ellipsoid sets for detecting attacks are proposed. Simulation results are provided to demonstrate the effectiveness of the proposed method.

I. INTRODUCTION

Multi-Agent Systems (MAS) have wide range of applications such as internet of things (IoT), electrical grids, water distribution systems, transportation systems, autonomous vehicles and Unmanned Aerial Vehicles (UAVs) [1]. Reaching consensus in a distributed manner is a fundamental problem in MAS. Some distributed and decentralized methods for attack detection have been proposed [2]–[5]. The agents transmit their data to neighboring agents through communication channels in distributed consensus protocols, and these channels are vulnerable to cyberattacks.

In [2] a distributed method to detect attacks in the communication network for the distributed control of interconnected systems has been proposed. In this study, only the local knowledge of the system is needed. However, through this method, stealthy attacks cannot be detected. In [6]–[10], secure designs against Denial of Service (DoS) attacks in a centralized setting have been studied. In [11], an investigation on a distributed event-triggered secure cooperative control of linear multi-agent systems under DoS attacks has been conducted. This paper studies how to achieve average consensus in the presence of DoS attacks and does not detect the attacks.

The problem of distributed simultaneous fault detection and leader-following consensus control for multi-agent systems has been investigated in [12]. In this study, a single module is used that conducts both tasks of fault detection and control objectives, simultaneously. Also, the proposed fault isolation method detects the faulty agent as well as the kind of fault. However, this method is restricted to the actuator and sensor faults to the linear systems with undirected topology. The authors of [13] have studied the leader-following consensus problem for heterogeneous multi-agent systems subject to both sensor and actuator attacks. In this study, the system is linear and there have been considered some restricting conditions.

In [14] a state-dependent event-triggered control strategy for time-varying MASs over a finite horizon has been designed for the first time. The only attack considered in this study is the false data injection attack to the linear multi-agent system with undirected topology.

Most of attack detection approaches, which are based on the state estimation, necessitate systems noises to be in a stochastic framework, and this leads to a probabilistic state estimation. For many real-world applications, accuracy in the state estimation is crucial. However, estimation based on probabilistic approach, such as Kalman filtering method, necessitates the use of mean and variance to describe the state distributions modelled as random variables (usually white and Gaussian perturbations). Consequently, considering unknown but bounded (UBB) noises is a much more appropriate approach to modeling state distributions.

Additionally, a common attack detection method, called the performance index test (χ²-detector), uses a residual signal to determine if the estimated behavior differs from that predicted by a model. Due to the nature of the Kalman filtering technique, the estimated and predicted states are single vectors and as a result, they cannot guarantee that a state is included in some region. Also, as the resulting UBB noises are sub-optimal for Kalman-type filtering, the reliability of attack detection is decreased. As a result of the need for set-valued estimation, the ellipsoidal state estimation technique was developed [15]. This method, known as the set-membership or set-valued state estimation filtering approach, has been extensively studied in filtering problems [16]–[22] and provides a set of state estimates in state space that contains the system’s true state [23], [24]. By using convex optimization approaches, an optimal ellipsoid with minimal size can be determined for set-membership estima-
tion, improving state estimation and detection performance.

The authors of [25] have studied a cyberattack detection method for the linear networked control systems through which for the first time, simultaneously, using the set-membership filtering for the purpose of the attack detection and distinguishing attacks on control signals from attacks on measurement outputs have been considered. However, they only considered the attack detection problem and there is no approach to the control of the system and the system is a single agent.

Except a few publications [26], [27], most research on set-membership filtering considers linear systems [25], [28]–[30]. Linearization should best fit the nonlinear functions over a state estimate set rather than a state estimate point when we use the set-membership framework. The authors of [26], linearized the nonlinear dynamics around the current estimate, then bounded the remaining terms by using interval mathematics and finally incorporated the remaining bounds as additions to the process or measurement noise bounds. Due to linearization around the estimated value of the state rather than the true value, the above approximations, bring a base point error [31].

There are few works on the detection of replay attacks and to the best of our knowledge, all the existing works have been only done on linear systems. Therefore, detection of these attacks for the nonlinear systems are of prime importance since real-world systems are mostly nonlinear. The fuzzy model of Takagi-Sugeno (T-S) is an effective and universal approximator for a certain class of nonlinear dynamic systems.

Therefore, in this paper, we use it to approximate nonlinear systems [32], [33]. We linearize the nonlinear systems over the true value of state and eliminate the base point error. Our objective is to design a simultaneous distributed attack detection strategy and leader-following consensus control based on a new two-step fuzzy set-membership filtering approach in a distributed framework. By utilizing the fuzzy modeling approach and the S-procedure technique [34], we determine bounding ellipsoidal sets for each agent by a recursive algorithm in state-space which guarantee the always enclosing of the system’s true state [23], [24], regardless of UBB noises, assuming no attacks are being made on the agent.

Each agent has a prediction and a measurement update step in its state estimation algorithm. The following two criteria are then used to detect cyberattacks:

1) When a cyberattack violates the control signal of any agent, the prediction ellipsoid set of that agent and its estimation ellipsoid set, updated with the previous measurement output, do not intersect.
2) When a cyberattack violates the sensor signal of any agent, the prediction ellipsoid set of that agent and its estimation ellipsoid set, updated at the current time instant, do not intersect.

Comparing with the previous works, the contributions of our work are as follows:

- To the best of our knowledge, we studied the attack detection problem of nonlinear multi-agent systems subject to replay attacks for the first time in the literature.
- We developed the fuzzy set-membership filtering approach for detection of the attacks.
- We considered false data injection attacks on the control signal and communication networks, as well as replay attacks on the sensor measurement data.
- We can distinguish attacks on control signals from attacks on measurement outputs.
- We also are able to mitigate the effects of the attacks and recover the system performance.
- Moreover, we ensure that we achieve the control goal which is achieving all the agent states to the leader-following consensus.

II. PROBLEM FORMULATION

Interaction and communication is modelled as a connected directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, $\mathcal{V} = \{1, 2, ..., N\}$, $\mathcal{E} = \{(i, j), (i, j) \in \mathcal{V}\}$ and $\mathcal{A} = (a_{ij}) \in \mathbb{R}^{N \times N}$ which are the vertex set, the directed edge set and the weighted adjacency matrix of $\mathcal{G}$, respectively. The weights are defined as $a_{ij} > 0$, if $(j, i) \in \mathcal{E}$ and $a_{ij} = 0$, otherwise. A node from which an edge goes to node i is a neighbor of node i. The set of the neighbors of node i are indicated by $N_i$, where $N_i = \{j \mid (j, i) \in \mathcal{E}\}$. Moreover, the Laplacian matrix $L = (l_{ij}) \in \mathbb{R}^{N \times N}$ is defined as $L = D - \mathcal{A}$ and $D = \text{diag}(\{d_i\})$, with $d_i = \sum_{j=1}^{N} a_{ij}$.

Consider a discrete-time nonlinear multi-agent system with N agents, and the dynamics of agent, $i, i \in \{1, ..., N\}$ is given as

$$
\begin{align*}
&x_i(k+1) = f_i(x_i(k)) + G_iu_i(k) + I_i(x_i)\omega_i(k) \\
&y_i(k) = h_i(x_i(k)) + F_i(x_i)v_i(k),
\end{align*}
$$

(1)

where $x_i(k) \in \mathbb{R}^n$, $u_i(k) \in \mathbb{R}^m$ and $y_i(k) \in \mathbb{R}^p$ represent state variables, control inputs and measurable output, respectively. The functions $f_i(x_i(k))$, $I_i(x_i(k))$, $h_i(x_i(k))$, and $F_i(x_i(k))$ are the functions of $x_i(k)$ with $f_i(0) = 0$, $I_i(0) = 0$, $h_i(0) = 0$, and $F_i(0) = 0$ and $G_i$’s are known matrices. A process uncertainty is denoted by $\omega_i(k) \in \mathbb{R}^m$, and $v_i(k) \in \mathbb{R}^p$ as a measurement noise which are assumed to be confined to specified ellipsoidal sets.

Definition 1. An ellipsoidal set has the form $\mathcal{X} \triangleq \{\zeta : \zeta = c + \Xi \xi, ||\xi|| \leq 1\}$, where $c \in \mathbb{R}^n$ is the center and $\Xi \in \mathbb{R}^{n \times m}$ with rank($\Xi$) = $m \leq n$ is its shape matrix. Assume that $\Xi$ is a lower triangular matrix whose diagonal elements all are positive. According to a Cholesky factorization, it can be seen that $P = \Xi \Xi^T > 0$ and $\zeta^T \zeta = (\zeta - c)^T P^{-1} (\zeta - c) \leq 1$. Consequently, the ellipsoidal set can also be represented as $\mathcal{X} \triangleq \{\zeta : (\zeta - c)^T P^{-1} (\zeta - c) \leq 1\}$. The size of the ellipsoid is dependent on the squares shape matrix $P$ and can be calculated as $\text{Tr}(P)$, which is the sum of the squared semi-axes lengths [30].

Assumption 1. The process noise $w_i(k)$ is UBB, which is
assumed to belong to the following specified ellipsoidal sets:
\[ \Psi_i(\mu) = \{ w_i(\mu) : w_i(\mu)R_i^{-1}w_i(\mu) \leq 1 \} \]
where \( R_i = Q_i^T > 0 \) and \( R_i = Q_i^T > 0 \) are known matrices with compatible dimensions.

The system model for the \( i \)th agent is presented by fuzzy IF-THEN rules.

Plant Rule \( l_i : \) IF \( \theta_{i,1}(k) = \mu_{l_i,1} \) and \( \theta_{i,2}(k) = \mu_{l_i,2} \ldots \) and \( \theta_{i,q}(k) = \mu_{l_i,q} \), THEN
\[
\begin{align*}
\dot{x}_i(k+1) &= A_i x_i(k) + B_i u_i(k) + M_i \omega_i(k) \\
y_i(k) &= C_i x_i(k) + D_i \nu_i(k),
\end{align*}
\]
where \( l_i = 1, \ldots, r \) (\( r \) stands for the total number of plant IF-THEN rules), \( \mu_{l_i,1}, \ldots, \mu_{l_i,q} \) are fuzzy sets, \( \theta_{i}(k) = [ \theta_{i,1}(k) \theta_{i,2}(k) \ldots \theta_{i,q}(k) ]^T \) denotes the premise variable, \( A_i, B_i, M_i, C_i, D_i \) are the system matrices with appropriate dimensions. The above-mentioned system can be inferred as follows:
\[
\begin{align*}
x_i(k+1) &= \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))A_i x_i(k) \\
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))B_i u_i(k) \\
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))M_i \omega_i(k) \\
y_i(k) &= \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))C_i x_i(k) \\
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))D_i \nu_i(k)
\end{align*}
\]
where \( g_{l_i}(\theta_{i}(k)) = \psi_{l_i}(\theta_{i}(k)) / \sum_{l=1}^{r} \psi_{l_i}(\theta_{i}(k)) \) is the normalized weight for each rule with \( \psi_{l_i}(\theta_{i}(k)) = \Pi_{i=1}^{n} \mu_{l_i}(\theta_{i}(k)) \geq 0 \) and \( \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k)) = 1 \), where \( \mu_{l_i}(\theta_{i}(k)) \) is the grade of membership of \( \theta_{i}(k) \) in \( \mu_{l_i,q} \).

By considering fuzzy model as an interpolation of \( r \) linear systems through the membership function \( g_{l_i}(\theta_{i}(k)) \), we can approximate the nonlinear system. Therefore, the nonlinear multi-agent system can be described as
\[
\begin{align*}
x_i(k+1) &= f_i(x_i(k)) + G_i u_i(k) + l_i(x_i(k))
\end{align*}
\]
\[
\begin{align*}
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))A_i x_i(k) + \Delta f_i(x_i(k)) \\
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))B_i u_i(k) \\
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))M_i \omega_i(k) \\
&\quad + \Delta l_i(x_i(k)) \omega_i(k) \\
y_i(k) &= \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))C_i x_i(k) + \Delta l_i(x_i(k)) \\
&\quad + \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))D_i \nu_i(k) + \Delta f_i(x_i(k)) \nu_i(k),
\end{align*}
\]
where
\[
\begin{align*}
&\Delta f_i(x_i(k)) = f_i(x_i(k)) - \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))A_l x_l(k) \\
&\Delta l_i(x_i(k)) = l_i(x_i(k)) - \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))M_l \omega_l(k) \\
&\Delta h_i(x_i(k)) = h_i(x_i(k)) - \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))C_l x_l(k) \\
&\Delta f_i(x_i(k)) = f_i(x_i(k)) - \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))D_l \nu_l(k)
\end{align*}
\]
denote the approximation (or interpolation) errors between the nonlinear system and the fuzzy model.

**Assumption 2.** According to [27], we assume
\[
\begin{align*}
&\Delta f_i(x_i(k)) = H_{i,1} \Delta x_i(k) + \Delta \theta_{i}(k) E_{i,1} x_i(k) \\
&\Delta l_i(x_i(k)) = H_{i,2} \Delta x_i(k) + \Delta l_i(x_i(k)) E_{i,2} x_i(k) \\
&\Delta h_i(x_i(k)) = H_{i,3} \Delta x_i(k) + \Delta h_i(x_i(k)) E_{i,3} x_i(k) \\
&\Delta f_i(x_i(k)) = H_{i,4} \Delta x_i(k) + \Delta f_i(x_i(k)) E_{i,4} x_i(k)
\end{align*}
\]
where \( H_{i} \) and \( E_{i} \) are known matrices, and \( \Delta \) is unknown but bounded with \( \| \Delta \| \leq 1 \).

We are interested in constructing the fuzzy-based leader following consensus protocol, which utilizes the estimated state instead of the full system state. First, consider the leader agent's dynamic by the following IF-THEN rules.

Plant Rule \( l_i : \) IF \( \theta_{i,1}(k) = \mu_{l_i,1} \) and \( \theta_{i,2}(k) = \mu_{l_i,2} \ldots \) and \( \theta_{i,q}(k) = \mu_{l_i,q} \), THEN
\[
\begin{align*}
x_i'(k+1) &= A'_{l_i} x_i'(k),
\end{align*}
\]
where \( x_i'(k) \in \mathbb{R}^{n_i} \) is the state of the leader, and \( A'_{l_i} \) are the system matrices with appropriate dimensions. It is assumed that the leader's dynamics are not subject to UBB process noise. The above-mentioned system can be inferred as follows:
\[
\begin{align*}
x_i'(k+1) &= \sum_{l=1}^{r} g_{l_i}(\theta_{i}(k))A'_{l_i} x_i'(k).
\end{align*}
\]

**Assumption 3.** The initial states \( x_i(0) \) and \( x_i'(0) \) are assumed to belong to a given ellipsoid
\[
\begin{align*}
&P_0(\theta_{i}(0)) = \{ x_i(0) : (x_i(0) - \hat{x}_i(0 | 0))^T P_i(0 | 0)^{-1} (x_i(0) - \hat{x}_i(0 | 0)) \leq 1 \} \\
&U_i(0) = \{ x_i(0) - x_i'(0) : (x_i(0) - x_i'(0))^T U_i(0)^{-1} (x_i(0) - x_i'(0)) \leq 1 \}
\end{align*}
\]
where \( \hat{x}_i(0 | 0) \) is the given estimate of \( x_i(0) \), and \( P_i(0 | 0) = P_i(0 | 0)^T > 0 \) and \( U_i(0 | 0) = U_i(0 | 0)^T > 0 \) are known matrices.

In this paper, we consider two kinds of attacks on the system.

**A. False Data Injection Attacks**

The original data packets are replaced by false ones when they are transferred from controllers to actuators or from another agent via communication channels.
Process noise
Sensor noise

\[ u_i(k) \]
\[ v_i(k) \]
\[ \omega_i(k) \]
\[ \nu_i(k) \]

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{The structure of a leader-following MAS with a fuzzy-based set-membership filtering detection method.}
\end{figure}

\begin{align}
\begin{cases}
  u_i^c(k) = u_i(k) + u_i^f(k) \\
  \bar{x}_i^a(k) = \bar{x}_i(k) + \phi_i^a(k)\bar{e}_i^a(k),
\end{cases}
\end{align}

where \( u_i(k) \in \mathbb{R}^m, u_i^c(k) \) and \( u_i^f(k) \) are the uncompromised control input, unknown false data injected to the actuator of agent \( i \), and the compromised input available to agent \( i \). Moreover, \( \bar{x}_i^a(k) \) is the corrupted neighbouring data, and in the presence of an attack on neighbouring channel \( \phi_j^a \) is 1", otherwise it is "0".

\subsection*{B. Replay Attacks}

A successful replay attack does not need a priori knowledge of the system components. It is assumed that the attacker can record sensor’s measurement data from \( k_0 \) until \( k_r \), with the window size \( \tau = k_r - k_0 \) in the first phase. Then, in the second phase, the attacker replays the recorded data to the system from \( k = k_r + d \) until the end of the attack at \( k = k_f \), where \( d \) is the delay between the recording time and replaying time. We model this attack according to [25] as

\[ a_i^v(k) = y_i(k - \tau) - y_i(k). \]

Thus, the sensor’s data affected by the attack is

\[ \tilde{y}_i(k) = y_i(k) + a_i^v(k). \]

We propose a distributed attack detector to detect the aforementioned types of attacks. The modules are tasked to detect attacks as well as ensure that the desired control specifications are satisfied. Also, the method can recover the system performance and mitigate the effects of the attacks. The structure of the system with the detector is shown in Fig. 1.

\section{III. Consensus Protocol and Fuzzy-Based Two-Step Set-Membership Estimation Method}

\subsection*{A. Prediction Step}

First, the prediction filter is considered in the form of Plant Rule \( I_i: \) IF \( \hat{\theta}_i,1(k) \) is \( \mu_{i,1} \) and \( \hat{\theta}_i,2(k) \) is \( \mu_{i,2} \) and \( \hat{\theta}_i,q(k) \) is \( \mu_{i,q} \); THEN

\[ \hat{x}_i(k + 1 | k) = \hat{A}_i \hat{x}_i(k | k), \]

where \( \hat{x}_i(k | k) \) is the estimation of the state \( x_i(k) \), \( \hat{A}_i \) is the fuzzy filter parameter to be determined and \( \hat{\theta}_i(k) = \{ \hat{\theta}_i,1(k), \hat{\theta}_i,2(k), \ldots, \hat{\theta}_i,q(k) \} \) are premise variables, which maybe functions of the state estimates. The overall fuzzy filter can be written from (14) as [35], [36]

\[ \hat{x}_i(k + 1 | k) = \sum_{l=1}^{c_i} g_{i,l}(\hat{\theta}_i(k)) \hat{A}_i \hat{x}_i(k | k). \]

(15)

For the given state estimation ellipsoid set \( \mathcal{E}_i(k | k) \) with the center \( \bar{x}_i(k | k) \) and the shape matrix \( \mathcal{Z}_i(k | k) \), the real state \( x_i(k) \) can be described by

\[ x_i(k) = \bar{x}_i(k | k) + \mathcal{Z}_i(k | k) z_i. \]

(16)

Then, our goal is to obtain the prediction ellipsoid set

\[ \mathcal{E}_i(k + 1 | k) = \{ x_i(k + 1) : (x_i(k + 1) - \bar{x}_i(k + 1 | k) + \mathcal{Z}_i(k + 1 | k) z_i \}. \]

(17)

Note that the state \( x_i(k + 1) \) belongs to such an ellipsoid set for any value of the system noises in their specified sets.

\subsection*{B. Measurement Update Step}

The update based on the current measurement is considered for the system (5), which is in the form of Plant Rule \( I_i: \) IF \( \hat{\theta}_i,1(k) \) is \( \mu_{i,1} \) and \( \hat{\theta}_i,2(k) \) is \( \mu_{i,2} \) and \( \hat{\theta}_i,q(k) \) is \( \mu_{i,q} \); THEN

\[ \hat{x}_i(k + 1 | k + 1) = \bar{x}_i(k + 1 | k) + L_i (y_i(k + 1) - \tilde{y}_i(k + 1 | k)), \]

where \( L_i \) is the filter parameter to be determined. The overall fuzzy update can be written from (18) as [17], the state \( x_i(k + 1) \) can be written as

\[ x_i(k + 1) = \bar{x}_i(k + 1 | k) + \mathcal{Z}_i(k + 1 | k) z_i. \]

Our objective is to update this prediction set with the one yielding from the current measurement \( y_i(k + 1) \). In other words, we look for an updated ellipsoid set \( \mathcal{E}_i(k + 1 | k + 1) \) with the center \( \hat{x}_i(k + 1 | k + 1) \) and the shape matrix \( \mathcal{Z}_i(k + 1 | k + 1) \) for the state \( x_i(k + 1) \), given by the current measurement information at the time instant \( k + 1 \). Thus, the updated ellipsoid set should satisfy the condition

\[ (x_i(k + 1) - \hat{x}_i(k + 1 | k + 1))^T P_i^{-1}(k + 1 | k + 1) \times (x_i(k + 1) - \hat{x}_i(k + 1 | k + 1)) \leq 1, \]

(20)

where \( P_i(k + 1 | k + 1) \) is the fuzzy process noise.
whenever the output constraint
\[ y_i(k + 1) = \sum_{l=1}^{r} g_{l1}(\theta_i(k)) C_l \left( \hat{x}_i(k + 1 \mid k) + \Xi_l(k + 1 \mid k) z_i \right) + H_{i,3} \Delta_{i,3} E_{i,3} \left( \hat{x}_i(k + 1 \mid k) + \Xi_l(k + 1 \mid k) z_i \right) + \sum_{l=1}^{r} g_{l1}(\theta_i(k)) D_l + H_{i,4} \Delta_{i,4} E_{i,4} \right) u_i(k + 1) \] (22)
holds for some \( \| z_i \| \leq 1 \).

C. Leader Following Consensus Protocol

The distributed observer-based leader following consensus protocol [37] is
\[ u_i(k) = K_{li} \left( \sum_{j \in \mathcal{N}_i} a_{ij} \left( \hat{x}_i(k \mid k) - \hat{x}_j(k \mid k) \right) + \lambda_i \left( \hat{x}_i(k \mid k) - x^l(k) \right) \right) \] (23)
where \( K_{li} \) are constant matrices to be designed, \( a_{ij} \) is a nonnegative element of the weighted adjacency matrix \( \mathcal{A} = [a_{ij}] \in \mathbb{R}^{N \times N} \). The adjacency matrix of the topology is selected as a binary matrix, where \( a_{ij} = 1 \) if follower \( i \) can receive information from follower \( j \), otherwise \( a_{ij} = 0 \). The leader-following multi-agent system (1), (6) achieves set-membership leader-following consensus under protocol (23) and two-step filter (15), (19), if the existence of desired gain sequences \( K_{li}, A_{li}, \) and \( L_{li} \) can guarantee that the one step ahead states \( x_i(k + 1), \forall i \in \mathcal{V} \) for all the followers reside in a leader state ellipsoid \( \mathcal{V}_l(k + 1) \) always enclosing all the followers’ true states, where
\[ \mathcal{V}_l(k + 1) \triangleq \left\{ x_i(k + 1) : \begin{array}{c} \left( x_i(k + 1) - x^l(k + 1) \right)^T U^{-1}(k + 1) \\ \left( x_i(k + 1) - x^l(k + 1) \right) \leq 1 \end{array} \right\} \] (24)
For the given leader ellipsoid set \( \mathcal{V}_l(k) \) with the center \( x^l(k) \) and the shape matrix \( \xi_l(k) \), the state \( x_i(k) \) can be described by
\[ x_i(k) = x^l(k) + \xi_l(k) z_i. \] (25)

IV. ATTACK DETECTION USING SET-MEMBERSHIP FUZZY FILTERING

The proposed cyberattack detection problem is addressed in this section by developing a set-membership filter. First, we develop the prediction ellipsoidal sets based on the leader following consensus protocol (23) and then update the prediction ellipsoidal set with the current measurement. Also, we develop the leader ellipsoidal set based on the leader following consensus protocol (23). Finally, convex optimization problems and one algorithm are provided to expose the cyberattack diagnosis scheme.

A. The Prediction Ellipsoid Set Design Based on Leader Following Consensus

From the system model (5) and (7), and the filter (15) and (16), the prediction error \( x_i(k + 1) - \hat{x}_i(k + 1 \mid k) \) can be written as
\[ x_i(k + 1) - \hat{x}_i(k + 1 \mid k) = \left( \sum_{l=1}^{r} g_{l1}(\theta_i(k)) A_l - \sum_{j=1}^{r} g_{j1}(\theta_i(k)) A_{lj} \right) \hat{x}_i(k \mid k) + \sum_{l=1}^{r} g_{l1}(\theta_i(k)) B_l \Delta_{i,l} + \sum_{j \neq i} a_{ij} \left( \hat{x}_i(k \mid k) - \hat{x}_j(k \mid k) \right) + \lambda_i \left( \hat{x}_i(k \mid k) - x^l(k) \right) \] (26)
where
\[ q_{i,1} = \Delta_{i,1} E_{i,1} \xi_l(k \mid k), \]
\[ q_{i,2} = \Delta_{i,2} E_{i,2} \xi_l(k \mid k) z_i, \]
\[ q_{i,3} = \Delta_{i,3} E_{i,3} \xi_l(k \mid k). \]

Denoting
\[ \eta_{i,1}(k) = \begin{bmatrix} 1 & z_i & a_{i,1} & q_{i,1} & q_{i,2} & q_{i,3} \end{bmatrix}^T, \] (28)
and considering the fact that \( \sum_{l=1}^{r} g_{l1}(\theta_i(k)) = 1 \), we can write (26) in a compact form as
\[ x_i(k + 1) - \hat{x}_i(k + 1 \mid k) = \sum_{l=1}^{r} g_{l1}(\theta_i(k)) \sum_{j=1}^{r} g_{j1}(\theta_i(k)) B_l \Delta_{i,l} + \sum_{j \neq i} a_{ij} \left( \hat{x}_i(k \mid k) - \hat{x}_j(k \mid k) \right) \] (29)
By denoting
\[ P_{i,1,i,j} = (A_{li} - \hat{A}_{lj}) \xi_l(k \mid k) - B_{li} \Delta_{i,l} x^l(k) \] (30)
we have
\[ \Gamma_{i,1,i,j} = \begin{bmatrix} P_{i,1,i,j} & A_{li} \Xi_l(k \mid k) & M_{li} & H_{i,1} & H_{i,2} \end{bmatrix}, \] (31)
where \( \Delta = Z + \Lambda = \begin{bmatrix} \Lambda_i \end{bmatrix}_{N \times N} \) and \( \Lambda = \text{diag} \{ \Lambda_1, \Lambda_2, \ldots, \Lambda_N \} \).
According to (29), we can write
\[ (x_i(k + 1) - \hat{x}_i(k + 1 \mid k))^T P_{i,1}^{-1}(k + 1 \mid k) \] (32)
Therefore, we can achieve
\[
(x_i(k + 1) - \hat{x}_i(k + 1 | k))^T P_i^{-1}(k + 1 | k) \times (x_i(k + 1) - \hat{x}_i(k + 1 | k)) \\
\leq \sum_{l=1}^{r} g_l(\theta_i(k)) \sum_{j=1}^{r} g_j(\hat{\theta}_i(k)) \\
\times \eta_i^{-1}(k) \Gamma_i^{-1}(k + 1 | k) \times \Gamma_i^{T}(k + 1 | k) \eta_i(1) \leq 0.
\]

The condition in (17) can be written as
\[
\eta_i^{-1}(k) \left[ \sum_{l=1}^{r} g_l(\theta_i(k)) \sum_{j=1}^{r} g_j(\hat{\theta}_i(k)) \\
\times \Gamma_i^{T}(k + 1 | k) \right] \eta_i(1) \leq 0.
\]

With $||\Delta_i|| \leq 1$, we can infer from (27) that
\[
\begin{align*}
q_i^T q_i + \xi_i^T \Gamma_i Q_i^{-1} \xi_i &\leq 0, \\
q_i^T q_i - \xi_i^T \Gamma_i Q_i^{-1} \xi_i &\leq 0.
\end{align*}
\]

From (2), (16) and (35), the unknown variables $z_i, \omega_i(k), q_i, q_i, q_i, q_i$ and $g_i$ satisfy the following constraints:
\[
\begin{align*}
||z_i|| &\leq 1, \\
w_i^T (k) Q_i^{-1} \xi_i &\leq 0, \\
q_i^T q_i + \xi_i^T \Gamma_i Q_i^{-1} \xi_i &\leq 0, \\
q_i^T q_i - \xi_i^T \Gamma_i Q_i^{-1} \xi_i &\leq 0.
\end{align*}
\]

Applying S-procedure [34] to (38) and (37), we can conclude that the inequality (34) holds if there exist non-negative scalars $\tau_1(1), \tau_2(1), \tau_3(1), \tau_4(1)$ and $\tau_5(1)$ such that
\[
\begin{align*}
\Gamma_i^{T}(k + 1 | k) \Gamma_i^{T}(k + 1 | k) - \text{diag}(1, 0, 0, 0, 0) &\leq 0, \\
\tau_1(1) &\leq 0, \\
\tau_2(1) &\leq 0, \\
\tau_3(1) &\leq 0, \\
\tau_4(1) &\leq 0, \\
\tau_5(1) &\leq 0.
\end{align*}
\]

Inequality (35) can be written in the following compact form
\[
\begin{align*}
\Gamma_i^{T}(k + 1 | k) \Gamma_i^{T}(k + 1 | k) - \text{diag}(1, 0, 0, 0, 0) &\leq 0, \\
\tau_1(1) &\leq 0, \\
\tau_2(1) &\leq 0, \\
\tau_3(1) &\leq 0, \\
\tau_4(1) &\leq 0, \\
\tau_5(1) &\leq 0.
\end{align*}
\]

Finally, denoting
\[
\Theta_i(1) = \text{diag}(1 - \tau_i(1), \tau_2(k))
\]

we can write (29) as
\[
\begin{align*}
\Gamma_i^{T}(k + 1 | k) \Gamma_i^{T}(k + 1 | k) - \text{diag}(1, 0, 0, 0, 0) &\leq 0.
\end{align*}
\]

By using Schur complements, (31) is equivalent to
\[
\begin{align*}
-\Gamma_i^{T}(k + 1 | k) \Gamma_i^{T}(k + 1 | k) - \Theta_i(1) &\leq 0.
\end{align*}
\]

Moreover, from the system model (2), (5), and (25), and by considering (27) the error $x_i(k + 1) - x_i^T(k + 1 | k)$ can be written as
\[
\begin{align*}
x_i(k + 1) - x_i^T(k + 1 | k) &= \sum_{l=1}^{r} g_l(\theta_i(1)) A_l \left( x_i^T(k) + \xi_i(\hat{z}) \right) + \sum_{l=1}^{r} g_l(\theta_i(1)) B_l K_l \\
&\times \left( \sum_{j=1}^{N} \alpha_{ij} \left( \hat{x}_i(k) - \hat{x}_j(k) \right) \right) + \Lambda \left( \hat{x}_i(k) - x_i^T(k) \right) \\
&+ H_i q_i + H_i q_i + H_i q_i.
\end{align*}
\]

Considering the fact that $\sum_{l=1}^{r} g_l(\theta_i(1)) = 1$, we can write (43) in a compact form as
\[
\begin{align*}
x_i(k + 1) - x_i^T(k + 1 | k) &= \sum_{l=1}^{r} g_l(\theta_i(1)) \sum_{j=1}^{r} \Gamma_i^{T}(k + 1 | k) \eta_i(1).
\end{align*}
\]

By denoting
\[
\begin{align*}
I_{i, j} &= \left( A_l + A_l B_l K_l \right) x_i^T(k) \\
&+ B_l K_l \sum_{j=1}^{N} \hat{z}_j(k),
\end{align*}
\]

we have
\[
\begin{align*}
\Gamma_i^{T}(k + 1 | k) = \left[ I_{i, j} \right] \left[ \begin{array}{c}
\begin{array}{c}
A_l \xi_i(1) \\
M_l \\
H_i \\
H_i \\
H_i
\end{array}
\end{array} \right].
\end{align*}
\]

According to (44), we can write
\[
\begin{align*}
(x_i(k + 1) - x_i^T(k + 1 | k))^T U_i^{-1}(k + 1) \\
\times (x_i(k + 1) - x_i^T(k + 1 | k))^T U_i^{-1}(k + 1) \\
= \sum_{l=1}^{r} g_l(\theta_i(1)) \sum_{j=1}^{r} g_j(\hat{\theta}_i(1)) \\
\times \eta_i^{-1}(k) \Gamma_i^{T}(k + 1 | k) \eta_i(1).
\end{align*}
\]
Therefore, we can achieve
\[(x_t(k+1) - x^*(k+1))^T U_t^{-1}(k+1) \times (x_t(k+1) - x^*(k+1)) \leq 1\]  \hspace{1cm} (48)

Therefore, the condition in (24) can be written as
\[\eta^T_{1,l}(k) \sum_{i=1}^r g_{li}(\theta_i(k)) \Gamma_{i,2,l}^{-1}(k+1) \Gamma_{i,2,l} - \text{diag}[1,0,0,0,0,0] \eta_{1,l}(k) \leq 0. \]  \hspace{1cm} (49)

Applying S-procedure to (47) and (49), we can conclude that the inequality (49) holds if there exist nonnegative scalars \(\tau_6(k), \tau_7(k), \tau_8(k), \tau_9(k)\) and \(\tau_{10}(k)\) such that
\[\Gamma_{i,2,l}^{-1}(k+1) \Gamma_{i,2,l} - \text{diag}[1,0,0,0,0,0] - \tau_6(k) \text{diag}[1,0,0,0,0] - \tau_7(k) \text{diag}[1,0,0,0,0,0] - \tau_8(k) \text{diag}(-x^T(k) \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, k) - \tau_9(k) \text{diag}(0, -x^T(k) \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, k) - \tau_{10}(k) \text{diag}[0,0,0,0,0,0] \leq 0. \]  \hspace{1cm} (50)

Inequality (50) can be written in the following compact form
\[\Gamma_{i,2,l}^{-1}(k+1) \Gamma_{i,2,l} - \text{diag}[1-\tau_6(k) - \tau_7(k) - \tau_8(k) \tilde{E}_1 E_{i,1} E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, k] - \tau_9(k) \text{diag}(0, -x^T(k) \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, k) - \tau_{10}(k) \text{diag}[0,0,0,0,0,0] \leq 0. \]  \hspace{1cm} (51)

Finally, denoting \(\Theta_{i,2}(k) = \text{diag}(1-\tau_6(k) - \tau_7(k) - \tau_8(k) \tilde{E}_1 E_{i,1} E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, k) - \tau_9(k) \text{diag}(0, -x^T(k) \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, \tilde{E}_1 E_{i,1}, k) - \tau_{10}(k) \text{diag}[0,0,0,0,0,0]\), we can write (51) as
\[\Gamma_{i,2,l}^{-1}(k+1) \Gamma_{i,2,l} - \Theta_{i,2}(k) \leq 0. \]  \hspace{1cm} (53)

By using Schur complements, (53) is equivalent to
\[\begin{bmatrix} -U_t(k+1) & \Gamma_{i,2,l} \\ \Gamma_{i,2,l}^T & -\Theta_{i,2}(k) \end{bmatrix} \leq 0. \]  \hspace{1cm} (54)

**Theorem 1.** Consider the leader following multi-agent system (1) that satisfies Assumption [1] and Assumption [2] and Assumption [3] Suppose that the state \(x_t(k)\) belongs to its state estimation ellipsoid \((x_t(k) - \hat{x}_t(k | k))^T P_t^{-1}(k) (x_t(k) - \hat{x}_t(k | k)) \leq 1\) and leader state ellipsoid \((x_t(k) - x^*(k))^T U_t^{-1}(k) (x_t(k) - x^*(k)) \leq 1\), then the one-step ahead state \(x_t(k+1)\) will reside in its state prediction ellipsoid \((x_t(k+1) - \hat{x}_t(k+1 | k))^T P_t^{-1}(k+1 | k) (x_t(k+1) - \hat{x}_t(k+1 | k)) \leq 1\) as well as leader state ellipsoid \((x_t(k+1) - x^*(k+1))^T U_t^{-1}(k+1) (x_t(k+1) - x^*(k+1)) \leq 1\), if there exist \(P_t(k+1 | k) > 0, U_t(k+1) > 0, A_t, \tilde{E}_i, \tilde{E}_j, \gamma_m(k) \geq 0, \) for \(m = 1, \ldots, 10,\) such that the linear matrix inequalities (LMI) (42) and (43) hold for all \(i_l, j_l = 1, \ldots, r\).
By denoting
\[ P_{i,2,j,h} = (I - L_j C_{ij}) \Xi_i (k + 1 | k), \]
we have
\[ \Gamma_{i,3,i,j,h} = \begin{bmatrix} 0 & P_{i,2,j,h} & -L_j D_{hi} & -L_j H_{i,3} & -L_j H_{i,4} & -L_j H_{i,5} \end{bmatrix}. \]

(65)

Taking (63) into account, we can write
\[
\begin{align*}
(x_i(k+1) - \hat{x}_i(k+1 | k+1))^T P_i^{-1} (k+1 | k+1) \\
& \times (x_i(k+1) - \hat{x}_i(k+1 | k+1)) \\
& = \sum_{l=1}^{r} g_{i,l} (\theta_i(k)) \sum_{j=1}^{r} g_{i,j} (\hat{\theta}_i(k)) \\
& \sum_{m=1}^{r} g_{i,m} (\theta_i(k)) \sum_{n=1}^{r} g_{i,n} (\hat{\theta}_i(k)) \\
& \eta_{i,2}^T (k+1) \Gamma_{i,3,i,j,h} \Gamma_{i,3,i,j,h} \eta_{i,2} (k+1) \\
\end{align*}
\]

Therefore, we can achieve
\[ \begin{align*}
(x_i(k+1) - \hat{x}_i(k+1 | k+1))^T P_i^{-1} (k+1 | k+1) \\
& \times (x_i(k+1) - \hat{x}_i(k+1 | k+1)) \\
& \leq \sum_{l=1}^{r} g_{i,l} (\theta_i(k)) \sum_{j=1}^{r} g_{i,j} (\hat{\theta}_i(k)) \\
& \eta_{i,2}^T (k+1) \Gamma_{i,3,i,j,h} \Gamma_{i,3,i,j,h} \eta_{i,2} (k+1) \\
\end{align*} \]

Therefore, the condition (21) in Section III-B can be described as
\[ \begin{align*}
\eta_{i,2}^T (k+1) \left[ \sum_{l=1}^{r} g_{i,l} (\theta_i(k)) \sum_{j=1}^{r} g_{i,j} (\hat{\theta}_i(k)) \right] \\
\times P_i^{-1} (k+1 | k+1) \Gamma_{i,3,i,j,h} \\
- \text{diag} \{1, 0, 0, 0, 0, 0\} \eta_{i,2} (k+1) \leq 0
\end{align*} \]

(68)

On the other hand, from (20), and (63) the unknown variables \( z_i, r_i(k+1), q_{i,4}, q_{i,5}, \) and \( q_{i,6} \) satisfy the following constraints:

\[ \| z_i \| \leq 1 \]
\[ v_i^T (k+1) R_i^{-1} (k+1) v_i (k+1) \leq 1 \]
\[ q_{i,4}^T q_{i,4} - \tilde{x}_i^T (k+1 | k) E_{i,3}^T E_{i,3} \tilde{x}_i (k+1 | k) \leq 0 \]
\[ q_{i,5}^T q_{i,5} - \tilde{z}_i^T \Xi_i (k+1 | k) \leq 0 \]
\[ q_{i,6}^T q_{i,6} - \tilde{v}_i^T (k+1) E_{i,4}^T E_{i,4} \tilde{v}_i (k+1) \leq 0 \]

which can be written in \( \eta_{i,2} (k+1) \) as:
\[ \begin{align*}
\eta_{i,2}^T (k+1) \text{diag} \{1, 0, 0, 0, 0, 0\} \eta_{i,2} (k+1) \leq 0 \\
\eta_{i,2}^T (k+1) \text{diag} \{-1, 0, 0, 0, 0, 0\} \eta_{i,2} (k+1) \leq 0 \\
\eta_{i,2}^T (k+1) \text{diag} \{1, 0, 0, 0, 0, 0\} \eta_{i,2} (k+1) \leq 0 \\
\eta_{i,2}^T (k+1) \text{diag} \{1, 0, 0, 0, 0, 0\} \eta_{i,2} (k+1) \leq 0 \\
\end{align*} \]

(69)

By applying S-procedure to (68) and (70), we can conclude that the inequality (68) holds if there exist nonnegative scalars \( \tau_{i,11}(k), \tau_{i,12}(k), \tau_{i,13}(k), \tau_{i,14}(k), \) and \( \tau_{i,15}(k) \) such that
\[ \begin{align*}
\Gamma_{i,3,i,j,h}^T P_i^{-1} (k+1 | k+1) \Gamma_{i,3,i,j,h} \\
- \text{diag} \{1, 0, 0, 0, 0, 0\} - \tau_{i,11}(k) \text{diag} \{1, 0, 0, 0, 0, 0\} \\
- \tau_{i,12}(k) \text{diag} \{-1, 0, R_i^{-1}(k+1), 0, 0, 0\} \\
- \tau_{i,13}(k) \text{diag} \{-\tilde{x}_i^T (k+1 | k) E_{i,3}^T E_{i,3} \tilde{x}_i (k+1 | k), \\
0, 0, I, 0, 0\} \\
- \tau_{i,14}(k) \text{diag} \{0, -\tilde{z}_i^T \Xi_i (k+1 | k), 0, 0, I, 0\} \\
- \tau_{i,15}(k) \text{diag} \{0, 0, -E_{i,4}^T E_{i,4}, 0, 0, I\} \\
\end{align*} \]

(71)

Inequality (71) can be written in the following compact form
\[ \begin{align*}
\Gamma_{i,3,i,j,h}^T P_i^{-1} (k+1 | k+1) \Gamma_{i,3,i,j,h} - \text{diag} \{1, 0, 0, 0, 0, 0\} \\
- \tau_{i,11}(k) - \tau_{i,12}(k) - \tau_{i,13}(k) \\
- \tau_{i,14}(k) - \tau_{i,15}(k) \\
\end{align*} \]

Finally denoting
\[ \Theta_{i,3}(k) = \text{diag} \{1, 0, 0, 0, 0, 0\} \]
\[ \tau_{i,11}(k), \tau_{i,12}(k), \tau_{i,13}(k), \tau_{i,14}(k), \tau_{i,15}(k) \]
we can write (72) as
\[ \begin{align*}
\Gamma_{i,3,i,j,h}^T P_i^{-1} (k+1 | k+1) \Gamma_{i,3,i,j,h} - \Theta_{i,3}(k) \leq 0
\end{align*} \]

(74)

Now, we deal with the output constraint (22) in Section III-B. First, it can be described by
\[ \Gamma_{k,i,j} (\hat{x}_i(k+1 | k)) \eta_{i,2} (k+1) = 0. \]

(75)

By denoting
\[ P_{i,y,1} = C_i \hat{x}_i (k+1 | k) - y_i (k+1) \]
\[ P_{i,y,2} = C_i \Xi_i (k+1 | k) \]
we have
\[ \begin{align*}
\Gamma_{k,i,j} (\hat{x}_i(k+1 | k)) \left[ \begin{array}{c}
P_{i,y,1} \\
P_{i,y,2}
\end{array} \right] = \left[ \begin{array}{c}
P_{i,y,1} \\
P_{i,y,2}
\end{array} \right] - \Theta_{k,4}(k) \leq 0
\end{align*} \]

(77)

By virtue of Finsler’s lemma [38], the inequality (68) under constraint (75) holds if there exists a \( Z_i(k+1) \) such that
\[ \begin{align*}
\Gamma_{i,3,i,j,h}^T P_i^{-1} (k+1 | k+1) \Gamma_{i,3,i,j,h} - \Theta_{k,4}(k) \\
+ Z_i (k+1) \Gamma_{k,i,j} (\hat{x}_i(k+1 | k)) \\
\end{align*} \]

(78)

For the purpose of simplicity, denote
\[ \Theta_{k,4}(k) = \Theta_{k,4}(k) - Z_i (k+1) \Gamma_{k,i,j} (\hat{x}_i(k+1 | k)) \]
\[ - \Gamma_{k,i,j} (\hat{x}_i(k+1 | k)) Z_i (k+1) \leq 0. \]

(79)

Then, by using Schur complements, (78) is equivalent to
\[ \left[ \begin{array}{cc}
P_i^{-1} (k+1 | k) & \Gamma_{i,3,i,j,h} \\
-\Theta_{k,4}(k) & 0
\end{array} \right] \leq 0. \]
Theorem 2. Consider the leader following multi-agent system (1, 8) that satisfies Assumption 1, Assumption 2 and Assumption 3. If the state $x_l(k+1)$ belongs to its state prediction ellipsoid $(x_l(k+1) - \hat{x}_l(k+1 | k))P_l^{-1}(k+1 | k) (x_l(k+1) - \hat{x}_l(k+1 | k)) \leq 1$, then such a state also resides in its updated state estimation ellipsoid $(x_l(k+1) - \hat{x}_l(k+1 | k+1) P_l^{-1}(k+1 | k+1) (x_l(k+1) - \hat{x}_l(k+1 | k+1)) \leq 1$ with the center determined by (18), where $P_l(k+1 | k+1) > 0$ satisfies matrix inequality (80) with other decision variables $L_{l_j}(k+1)$, $N_l(k+1)$, and $\tau_{m,l}(k) \geq 0$ for $m=1,11,..,15$ for all $l_j, j=1,..,r$, then we have

$$\sum_{l=1}^{r} g_l(\theta_l(k)) \sum_{j=1}^{r} g_j(\hat{\theta}_j(k)) \eta_{i,2}(k+1 | k+1) \leq 1. \tag{81}$$

From (67), we obtain

$$\begin{align*}
(\hat{x}_l(k+1) - \hat{x}_l(k+1 | k+1)) P_l^{-1}(k+1 | k+1) \times (\hat{x}_l(k+1) - \hat{x}_l(k+1 | k+1)) \leq 1,
\end{align*} \tag{82}$$

which completes the proof.

Now, the convex optimization approach is applied to determine an optimal ellipsoid with the minimal size. Therefore, $P_l(k+1 | k+1)$ is obtained by solving the following optimization problem:

$$\begin{align*}
\min_{P_l(k+1 | k+1), L_{l_j}(k+1), N_l(k+1), \tau_{m,l}(k)} \text{Tr}(P_l(k+1 | k+1)) \\
\text{s.t.} \quad P_l(k+1 | k+1) = \Xi(k+1 | k+1),
\end{align*} \tag{83}$$

Algorithm 1 Recursive State Estimation

1. Initialization:

   Given an initial ellipsoid $\mathcal{X}_l(0 | 0)$, $\mathcal{Y}_l(0)$, recursive times $T_l$, and set $k=0$. Let $x_l(k) = x_l(0)$, $\hat{x}_l(k | k) = \hat{x}_l(0 | 0)$, $\Xi_l(k) = \Xi_l(0 | 0)$, $x_l(k) = x_l(0)$, and $\hat{\xi}_l(k) = \hat{\xi}_l(0)$.

2. Prediction:

   1) Calculate $P_l(k+1 | k), L_{l_j}(k+1), \hat{x}_l(k), k_i$ by solving the optimization problem (83).

2) Obtain the minimum $\Xi_i(k+1 | k), \hat{\xi}_l(k+1 | k | k+1)$ according to $P_l(k+1 | k) = \Xi_l(k+1 | k+1) P_l^{T}(k+1 | k+1)$, and $L_{l_j}(k+1) = \hat{\xi}_l(k+1 | k+1)$.\tag{19}

3) Calculate the center of the prediction ellipsoid $\hat{x}_l(k+1 | k+1)$ by (19).

3. Attack Detection: Control Signal Data or Communication Signal Data Cyber Attack Diagnosis

   1) If $\mathcal{X}_l(k | k) \cap \mathcal{Y}_l(k+1 | k) \neq \emptyset$, there is no attack and go to step 5.

2) If $\mathcal{X}_l(k | k) \cap \mathcal{Y}_l(k+1 | k) = \emptyset$, data is subject to attack and go to step 4.

4. Recovery Step and Attack Mitigation:

   Set $\mathcal{Y}_l(k+1 | k) = \mathcal{X}_l(k | k)$, $\mathcal{Y}_l(k+1) = \mathcal{Y}_l(k+1 | k)$, $u_l(k) = u_l(k-1)$, and go to step 5.

5. Measurement Update:

   1) Calculate $P_l(k+1 | k) + L_{l_j}(k+1)$ by solving the optimization problem (83).

2) Obtain the new $\Xi_i(k+1 | k+1)$ according to $P_l(k+1 | k+1) = \Xi_l(k+1 | k+1) P_l^{T}(k+1 | k+1)$.

3) Calculate the center of the updated prediction ellipsoid $\hat{x}_l(k+1 | k+1)$ by (19).

6. Attack Detection: Sensor Measurement Data Cyber Attack Diagnosis

   1) If $\mathcal{X}_l(k+1 | k) \cap \mathcal{Y}_l(k+1 | k) \neq \emptyset$, there is no attack and go to step 8.

2) If $\mathcal{X}_l(k+1 | k+1) \cap \mathcal{Y}_l(k+1 | k) = \emptyset$, data is under attack and go to step 7.

7. Recovery Step and Attack Mitigation:

   Set $\mathcal{Y}_l(k+1 | k+1) = \mathcal{X}_l(k+1 | k), y_l(k+1) = y_l(k)$ and go to step 8.

8. Loop

   If $k = T_l$ then Exit, Else $k <- k$ and go to step 2.

V. Simulation Results

Consider the following multi-agent, discrete-time, nonlinear system:

$$x_{1,1}(k+1) = 0.2 x_{1,1}(k) - 0.3 \left( x_{1,2}(k) - x_{2,1}(k) \right) + u_1(k) + w_k$$

$$x_{1,2}(k+1) = 0.3 x_{1,1}(k) + 0.2 \left( x_{1,2}(k) - x_{2,1}(k) \right) + 0.3 u_{1,1}(k) + 0.9 u_{1,2}(k) + \eta_k$$

$$y_k = x_{1,1}(k) + 0.1 x_{2,1}(k) + x_{1,2}(k) + v_k$$

$$x_{2,1}(k+1) = 0.5 x_{2,1}(k) - 0.1 \left( x_{2,2}(k) - x_{2,1}(k) \right) + 0.9 u_{2,1}(k) + 0.2 u_{2,2}(k) + v_k$$

$$x_{2,2}(k+1) = 0.9 x_{2,1}(k) + 0.5 \left( x_{2,2}(k) - x_{2,1}(k) \right) + u_2(k) + \omega_k$$

$$y_k = x_{2,1}(k) + 0.1 x_{2,1}(k) + x_{2,2}(k) + v_k$$

where the state $x_l(k) = \begin{bmatrix} x_{1,1}(k) & x_{1,2}(k) \end{bmatrix}^T$. Now, we construct the following fuzzy models to approximate the above nonlinear multi-agent system for each agent:

Agent 1:

- Rule 1: IF $x_{1,1}(k)$ is about 1, THEN

$$x_1(k+1) = A_{1,1} x_1(k) + B_{1,1} u_1(k) + M_{1,1} \omega_1(k),$$

$$y_1(k) = C_{1,1} x_1(k) + D_{1,1} v_1(k) \tag{85}$$

- Rule 2: IF $x_{1,1}(k)$ is about 0, THEN

$$x_1(k+1) = A_{1,2} x_1(k) + B_{1,2} \omega_1(k),$$

$$y_1(k) = C_{1,2} x_1(k) + D_{1,2} v_1(k) \tag{86}$$

where

$$A_{1,1} = \begin{bmatrix} 0.5 & -0.3 \\ 0.1 & 0.2 \end{bmatrix}, \quad B_{1,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$C_{1,1} = \begin{bmatrix} 1.1 \\ 1.1 \end{bmatrix}, \quad D_{1,1} = 1$$

$$A_{1,2} = \begin{bmatrix} 0.2 & -0.3 \\ 0.3 & 0.2 \end{bmatrix}, \quad B_{1,2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$C_{1,2} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}, \quad D_{1,2} = 1.$$

Agent 2:

- Rule 1: IF $x_{2,1}(k)$ is about 1, THEN

$$x_2(k+1) = A_{2,1} x_2(k) + B_{2,1} u_2(k) + M_{2,1} \omega_1(k),$$

$$y_2(k) = C_{2,1} x_2(k) + D_{2,1} v_2(k) \tag{87}$$

- Rule 2: IF $x_{2,1}(k)$ is about 0, THEN

$$x_2(k+1) = A_{2,2} x_2(k) + B_{2,2} \omega_2(k),$$

$$y_2(k) = C_{2,2} x_2(k) + D_{2,2} v_2(k) \tag{88}$$

where

$$A_{2,1} = \begin{bmatrix} 0.5 & -0.3 \\ 0.1 & 0.2 \end{bmatrix}, \quad B_{2,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$C_{2,1} = \begin{bmatrix} 1.1 \\ 1.1 \end{bmatrix}, \quad D_{2,1} = 1$$

$$A_{2,2} = \begin{bmatrix} 0.2 & -0.3 \\ 0.3 & 0.2 \end{bmatrix}, \quad B_{2,2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$C_{2,2} = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}, \quad D_{2,2} = 1.$$
- Rule 2: IF $x_{21}(k)$ is about 0, THEN

$$
\begin{align*}
    x_2(k+1) &= A_{22}x_2(k) + B_{22}o_2(k) \\
    y_2(k) &= C_{22}x_2(k) + D_{22}y_2(k)
\end{align*}
$$

where

$$
\begin{align*}
    A_{21} &= \begin{bmatrix} 0.6 & -0.1 \\ 0.4 & 0.5 \end{bmatrix} ,
    B_{21} &= \begin{bmatrix} 1 \end{bmatrix} \\
    C_{21} &= \begin{bmatrix} 1.1 \\ 1.1 \end{bmatrix} ,
    D_{21} = 1 \\
    A_{22} &= \begin{bmatrix} 0.5 & -0.1 \\ 0.9 & 0.5 \end{bmatrix} ,
    B_{22} = \begin{bmatrix} 1 \end{bmatrix} \\
    C_{22} &= \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix} ,
    D_{22} = 1.
\end{align*}
$$

For the convenience of simulation, triangular membership functions are used for Rule 1 and Rule 2 in this example.

In the above fuzzy models, the approximation errors between the nonlinear system and the fuzzy models are assumed to satisfy (7), where

$$
\begin{align*}
    H_{11} &= \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} ,
    H_{21} &= \begin{bmatrix} 0.3 \\ 0.3 \end{bmatrix} \\
    E_{11} &= \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} ,
    E_{21} = \begin{bmatrix} 0 \\ 0.6 \end{bmatrix} \\
    H_{12} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
    H_{13} &= 0.1 \\
    H_{14} &= 0.1 \\
    H_{22} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
    E_{i1} &= \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} \\
    E_{i2} &= 0.1 \\
    E_{i3} &= 0.1 \\
    E_{i4} &= 0.1
\end{align*}
$$

The leader matrices described in (9) are defined as follow:

$$
\begin{align*}
    A_{l1} &= \begin{bmatrix} 0.5 & -0.6 \\ 0.2 & 0.7 \end{bmatrix} ,
    A_{l2} = \begin{bmatrix} 0.5 \\ -0.4 \end{bmatrix} \\
    A_{l3} &= \begin{bmatrix} 0.2 \\ 0.7 \end{bmatrix} \\
    A_{l4} &= \begin{bmatrix} -0.6 \\ 0.7 \end{bmatrix}
\end{align*}
$$

In the simulation, $o_l(k)$ and $v_l(k)$ are chosen as 0.5$\sin(2k)$ and 0.5$\sin(20k)$, respectively. The initial state is set as $x_i(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}^T$, which belongs to the ellipsoids $(x_i(0) - \bar{x}_i(0))^T P_i^{-1}(0 \mid 0) (x_i(0) - \bar{x}_i(0)) \leq 1$ and $(x_i(0) - \bar{x}_i(0))^T U_i^{-1}(0) (x_i(0) - \bar{x}_i(0)) \leq 1$, where $\bar{x}_i(0) = x_i(0)$, $P_i(0 \mid 0) = U_i(0) =$

$$
\begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix} ,
Q_i(k) = 1 - k/50, \quad \text{and} \quad R_i(k) = 1 - k/50.
$$

The communication between the agents and the leader is modeled as Fig. 2.

We obtained the simulation results under MATLAB 9.8 with YALMIP and SDPT3. We considered the following scenarios in 50 sampling steps.

A. Attack Free System

In this case, the prediction ellipsoid set and the updated estimation ellipsoid set must always have the intersection. Fig. 3(a) and Fig. 3(b) show the existence of the intersection between these sets for agent 1 and agent 2, respectively.

B. Replay Attacks on Sensor Data

We consider adding the replay attack on the sensor measurement data based on the definition in the Subsection II-B. Therefore, we assume that the attack records the data from k=5 to k=10 and replaces the sensor data at k=20 to k=25 with them. Therefore, Fig. 4(a) and Fig. 4(b) confirm that the prediction ellipsoid set and the updated measurement set for the next iteration do not have the intersection during this attack period.
C. False Data Injection Attacks on Control Signal

In this case, the original data packets are replaced by false ones when they are transferred from controllers to actuators via communication channels. Therefore, we consider that the attacker replaces the control signal via targeting the communication channel between the controller and actuator. In the simulation, the attack vector in (11) is modeled as \( \mathbf{a}_i(k) = [4 \ 4]^T \) for both agents from step \( k = 20 \) to \( k = 25 \). Therefore, as it is shown in Fig. 5 (a) and Fig. 5 (b), when the attack on the control signal occurs, the prediction ellipsoid set and the estimation ellipsoid set updated with the previous time instant has no intersection.

\[
\mathbf{x}_{i}(k + 1 | k) = \begin{bmatrix} 4 \\ 4 \end{bmatrix}
\]

\[
\mathbf{x}_{i}(k | k) = \begin{bmatrix} 5 \\ 5 \end{bmatrix}
\]

\[\mathbf{R}(k) = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.1 \end{bmatrix}\]

Finally, Fig. 7 illustrates that we can guarantee the leader following consensus in the attack free system and in the presence of the attacks for all the above scenarios.

D. False Data Injection Attacks on Communication Channel

In this case, the original data packets were replaced by false ones when they were transferred between two agents via communication channels. Therefore, we consider that the attacker replaces the signal transferring from agent 2 to agent 1 via communication channel. In the simulation, the attack vector in (11) is modeled as \( \mathbf{a}_2(k) = [5 \ 5]^T \) from step \( k=20 \) to \( k=25 \). Therefore, as it is shown in Fig. 6 for agent 1, the prediction ellipsoid set and the estimation ellipsoid set updated with the previous time instant has no intersection. Also, as there is no attack on the communication channel from agent 1 to agent 2, the prediction ellipsoid set and the estimation ellipsoid set updated with the previous time instant for agent 2 has an intersection.

Finally, Fig. 7 illustrates that we can guarantee the leader following consensus in the attack free system and in the presence of the attacks for all the above scenarios.

VI. CONCLUSION

This paper deals with the problem of cyberattack detection in discrete-time leader-following nonlinear multi-agent systems subject to unknown but bounded system noises. For the approximation of the nonlinear systems over the true value of the state, the T-S fuzzy model has been employed. A new fuzzy set-membership filtering method consisting of two steps has been developed for each agent to detect two types of cyberattacks at the time of their occurrence. It has considered the detection of replay attacks and false data injection attacks affecting the leader-following consensus. We proposed recursive algorithms for achieving the leader-following consensus protocol and finding the two ellipsoid sets for detecting attacks. The cyberattacks have been detected based on two criteria about the intersections between the ellipsoid sets. Finally, simulation results have been provided to demonstrate the effectiveness of the proposed method.

REFERENCES

[1] J. Shamma, Cooperative control of distributed multi-agent systems. John Wiley & Sons, 2008.
[2] F. Boem, A. J. Gallo, G. Ferrari-Trecate, and T. Parisini, “A distributed attack detection method for multi-agent systems governed by consensus-based control,” in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), pp. 5961–5966, IEEE, 2017.
[3] A. Barbini, H. Rezaee, F. Boem, and T. Parisini, “Detection of covert cyber-attacks in interconnected systems: A distributed model-based approach,” IEEE Transactions on Automatic Control, vol. 66, no. 9, pp. 3728–3741, 2020.
[4] A. Mousavi, K. Aryanhashemi, and R. R. Selmic, “Cyber-attack detection in discrete-time nonlinear multi-agent systems using neural networks,” in 2021 IEEE Conference on Control Technology and Applications (CCTA), pp. 911–916, IEEE, 2021.
[5] F. Pasqualetti, F. Dörfler, and F. Bullo, “Attack detection and identification in cyber-physical systems,” IEEE transactions on automatic control, vol. 58, no. 11, pp. 2715–2729, 2013.
[6] H. Zhang, P. Cheng, L. Shi, and J. Chen, “Optimal denial-of-service attack scheduling against linear quadratic gaussian control,” in 2014 American Control Conference, pp. 3996–4001, IEEE, 2014.
[7] H. Zhang, P. Cheng, L. Shi, and J. Chen, “Optimal denial-of-service attack scheduling with energy constraint,” IEEE Transactions on Automatic Control, vol. 60, no. 11, pp. 3023–3028, 2015.
Fig. 7. Leader-following consensus achievements of the agents in all scenarios.
[32] Y.-Y. Cao and P. M. Frank, “Robust $h_{\infty}$ disturbance attenuation for a class of uncertain discrete-time fuzzy systems,” *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 4, pp. 406–415, 2000.

[33] F. Delmotte, T. M. Guerra, and A. Kraszewski, “Discrete takagi–sugeno’s fuzzy models: reduction of the number of lmi in fuzzy control techniques,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 5, pp. 1423–1427, 2008.

[34] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[35] Z. Gao, X. Shi, and S. X. Ding, “Fuzzy state/disturbance observer design for t-s fuzzy systems with application to sensor fault estimation,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 3, pp. 875–880, 2008.

[36] C.-S. Tseng, “Robust fuzzy filter design for nonlinear systems with persistent bounded disturbances,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 4, pp. 940–945, 2006.

[37] X. Ge, Q.-L. Han, and F. Yang, “Event-based set-membership leader-following consensus of networked multi-agent systems subject to limited communication resources and unknown-but-bounded noise,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 5045–5054, 2016.

[38] R. E. Skelton, T. Iwasaki, and K. Grigoriadis, “A unified algebraic approach to linear control design,” 2013.