2015 European Control Conference (ECC)
July 15-17, 2015. Linz, Austria

Further results on fault detection and isolation of malicious nodes in Multi-hop Control Networks

A. D'Innocenzo, F. Smarra and M. D. Di Benedetto

Abstract— **A Multi-hop Control Network (MCN) consists of a LTI system where the communication between sensors, actuators and computational units is supported by a (wireless) multi-hop communication network and data flow is performed using scheduling, routing and network coding of sensing and actuation data. In this paper we extend our previous results in [6] on the Fault Detection and Isolation (FDI) problem over a MCN where the plant is a MIMO system and the communication nodes are subject to permanent failures and malicious attacks. In particular, we provide necessary and sufficient conditions such that the FDI problem can be solved with much milder assumptions on the malicious signal.**

I. Introduction

Networked control systems (NCS) are distributed control systems where the communication between sensors, actuators, and computational units is supported by a possibly wireless communication network. The use of wireless NCS in industrial automation results in flexible architectures and generally reduces installation, debugging, diagnostic and maintenance costs with respect to wired networks (see e.g. [13], [12] and references therein). However modeling, analysis and design of wireless NCS are challenging open research problems since they require to take into account the joint dynamics of physical systems, communication protocols and network infrastructures. Recently, a huge effort has been made in scientific research on Networked Control Systems (NCSs), see e.g. [2], [26], [9], [16], [25], [15], [11], [18] and references therein for a general overview. Also, as can be inferred from the survey [10], fault tolerant control is one of the most challenging issues in NCSs, see e.g. [27], [20], [5], [3]. In general, the literature on NCSs addresses non–idealities (such as quantization errors, packets dropouts, variable sampling and delay and communication constraints) as aggregated network performance variables or disturbances, neglecting the dynamics introduced by the communication protocols. In particular in [1] a simulative environment of computer nodes and communication networks interacting with the continuous-time dynamics of the real world is presented. To the best of our knowledge the first formal model of a NCS that models the joint dynamics of a dynamical control system and of the MAC (scheduling) and Network (routing) layers of a time-triggered communication protocol over a shared multi-hop wireless network has been presented in [22]. This framework also models wireless industrial control protocols such as WirelessHART and ISA-100: note that many on-market engineering products are based on these protocols, see e.g. the Siemens SITRANS AW200, SITRANS AW210 and IE/WSN-PA Link.

In [7] we defined a Multi-hop Control Network (MCN) $\mathcal{M}$, that consists of a continuous-time SISO LTI plant $\mathcal{P}$ interconnected to a controller C via two multi-hop communication networks $\mathcal{R}$ (the controllability network) and $\mathcal{O}$ (the observability network), as illustrated in Figure 1. Because of wireless networking a MCN is subject to failures and/or malicious attacks in the communication nodes: we will call *malicious cluster* any set of communication nodes subject to a failure or malicious attack and denote by F the set of all malicious clusters. In [7] we addressed and solved, for SISO LTI systems, the problem of designing a set of controllers and the communication protocol parameters, so that it is possible, only using sensing and actuation data (i.e. without adding further data communication among nodes for fault detection), to detect and isolate on-the-fly the malicious cluster of the controllability and observability networks: this allows reconfiguring the scheduling only for the neighbor of the faulty or malicious nodes (which requires much less communication cost and time w.r.t. reconfiguring all nodes) and applying an appropriate controller to stabilize the reconfigured system. In particular, we proposed a solution to the above problem by solving the following two subproblems.

Problem 1: Assume that the current malicious cluster $f \in F$ is *known*, and let $\mathcal{M}_f$ be the corresponding MCN dynamics (which consists of the cascade of the controllability network, the plant and the observability network) due to a scheduling re-configuration that isolates the faulty nodes. We address the co-design problem of scheduling and routing so that, for any malicious cluster $f \in F$, the MCN $\mathcal{M}_f$ is controllable and observable.

If Problem 1 can be solved, then the existence and computation of a stabilizing controller C_f for any malicious cluster can be guaranteed using standard techniques.

Problem 2: Assume that the current malicious cluster is *unknown*. We address the problem of exploiting the MCN input and output signals to detect and isolate the current malicious cluster.

The above problem separation deliberately neglects the switching dynamics of the closed loop system since it is based on the assumption that failures/attacks occur with a time scale much greater than the sampling time, namely we only consider *permanent* failures and not *transient* failures (e.g. packet losses) [28].

In [6] we solved Problem 2 for MCN where the plant is a MIMO LTI system, by using some assumptions on

Center of Excellence DEWS, Dept. of Information Engineering Computer Science and Mathematics, Univ. of L'Aquila, Italy. {alessandro.dinnocenzo, francesco.smarra, mariadomenica.dibenedetto}@univaq.it.

978-3-9524269-3-7 ©2015 EUCA

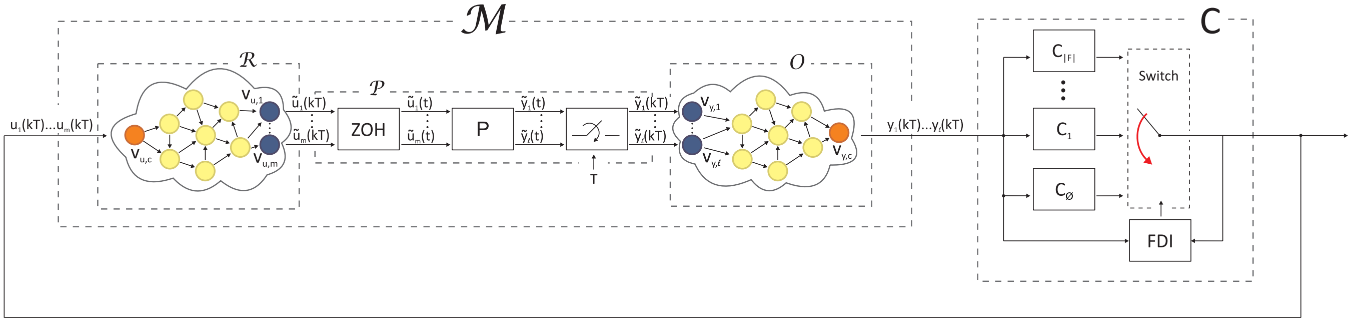

Fig. 1. Control scheme of a MCN subject to node failures and malicious attacks.

the failure/malicious attack signal. In this paper we solve Problem 2 strongly relaxing such assumption: in particular, while the results in [6] allowed to detect *only a zero-volume set of malicious signals*, the results in this paper allow to detect *all but a zero-volume set of malicious signals*.

Notation: We will denote by $\mathbb{N}, \mathbb{R}, \mathbb{R}^+, \mathbb{R}_0^+$ respectively the sets of natural, real, positive real and non-negative real numbers. Given $n \in \mathbb{N}$, we denote by $\mathbf{n}$ the set $\mathbf{n} \doteq \{1, 2, \ldots, n\}$. Given a finite set F and a subset $H \subseteq F$, we define $|H|$ and $|F|$ their cardinalities, $F \setminus H$ the difference set and 2^F the power set. Given a matrix A, we denote by $|A|$ its determinant. We denote by $diag(A_1, \ldots, A_n)$ the $n \times n$ diagonal matrix whose diagonal consists of the elements $A_1, \ldots, A_n$. We denote by $\mathbf{0}_{n \times m}$ the matrix of zeros with n rows and m columns and by $\mathbf{I}_n$ the identity matrix of dimension n. We denote by $F(z)$ the Z-transform of a signal $f(k)$ and by $diag(F_1(z), \ldots, F_n(z))$ the $n \times n$ diagonal transfer matrix whose diagonal consists of the scalar transfer functions $F_1(z), \ldots, F_n(z)$. Given a directed graph $(\mathcal{V}, \mathcal{E})$ we define *path* an alternating sequence of vertices and edges. A path is said to be *simple* if no vertices are repeated. A set of paths is said to be *vertex disjoint* if each two of them consist of disjoint sets of vertices. A vertex disjoint set of r simple paths from a set $\mathcal{V}_1 \subseteq \mathcal{V}$ to a set $\mathcal{V}_2 \subseteq \mathcal{V}$ is said to be an *r-linking* from $\mathcal{V}_1$ to $\mathcal{V}_2$.

II. Modeling of MCNs

Definition 1 below allows modeling time-triggered communication protocols that specify mixed TDMA, FDMA and/or CDMA (resp. Time, Frequency and Code Division Multiple Access) access to a shared communication resource, for a set of communication nodes interconnected by an arbitrary radio connectivity graph. In these standards the access to the shared communication channel is specified as follows: time is divided into slots of fixed duration Δ and groups of Π time slots are called frames of duration $T = \Pi\Delta$. For each frame, a communication scheduling allows each node to transmit data only in a specified time slot. The scheduling is periodic with period Π, i.e. it is repeated in all frames.

Definition 1: A MIMO Multi-hop Control Network is a tuple $\mathcal{M} = (P, G, W, \eta, \Delta)$ where:

P is a continuous-time MIMO LTI system, with n, m and ℓ respectively the dimensions of the internal state, input and output spaces.

$G = (G_{\mathcal{R}}, G_{\mathcal{O}})$. $G_{\mathcal{R}} = (V_{\mathcal{R}}, E_{\mathcal{R}})$ is an acyclic connected graph, where the vertices correspond to the communication nodes of the network and an edge from v to v' means that node v' can receive messages transmitted by node v through the wireless communication link (v, v'). We denote by v_c the special node of $V_{\mathcal{R}}$ that corresponds to the controller and by $v_{u,i} \in V_{\mathcal{R}}$, $i \in \mathbf{m}$, the special nodes that correspond to the actuators of the inputs u_i, $i \in \mathbf{m}$. $G_{\mathcal{O}} = (V_{\mathcal{O}}, E_{\mathcal{O}})$ is defined similarly to $G_{\mathcal{R}}$. We denote by v_c the special node of $V_{\mathcal{O}}$ that corresponds to the controller and by $v_{y,i} \in V_{\mathcal{O}}$, $i \in \boldsymbol{\ell}$, the special nodes that correspond to the sensors of the outputs y_i, $i \in \boldsymbol{\ell}$.

$W = (W_{\mathcal{R}}, W_{\mathcal{O}})$. $W_{\mathcal{R}} = \{W_{\mathcal{R}_i}\}_{i \in \mathbf{m}}$, where $W_{\mathcal{R}_i} : E_{\mathcal{R}} \to \mathbb{R}^+$ is a weight function for the input i that associates to each link a positive constant. $W_{\mathcal{O}} = \{W_{\mathcal{O}_i}\}_{i \in \boldsymbol{\ell}}$ is defined similarly to $W_{\mathcal{R}}$. The role of W will be clear in the following definition of η.

$\eta = (\eta_{\mathcal{R}}, \eta_{\mathcal{O}})$. $\eta_{\mathcal{R}} = \{\eta_{\mathcal{R}_i}\}_{i \in \mathbf{m}}$, where $\eta_{\mathcal{R}_i} : \mathbb{N} \to 2^{E_{\mathcal{R}}}$ is the controllability scheduling function for the input i that associates to each time slot of each frame a set of edges of the controllability radio connectivity graph $G_{\mathcal{R}}$. Since in this paper we only consider a periodic scheduling, that is repeated in all frames, we define the controllability scheduling functions by $\eta_{\mathcal{R}_i} : \{1, \ldots, \Pi\} \to 2^{E_{\mathcal{R}}}$. The integer constant Π is the period of the controllability scheduling. The semantics of $\eta_{\mathcal{R}_i}$ is that $(v, v') \in \eta_{\mathcal{R}_i}(h)$ if at time slot h of each frame the data associated to input i and contained in node v is transmitted to the node v', multiplied by the weight $W_{\mathcal{R}_i}(v, v')$. For any $\eta_{\mathcal{R}_i}$, we assume that each link can be scheduled only one time for each frame. $\eta_{\mathcal{O}} = \{\eta_{\mathcal{O}_i}\}_{i \in \boldsymbol{\ell}}$ is defined similarly to $\eta_{\mathcal{R}}$. We remark that the scheduling period of $\eta_{\mathcal{O}}$ is the same as that of $\eta_{\mathcal{R}}$.

Δ is the time slot duration. As a consequence, $T = \Pi\Delta$ is the frame duration.

Definition 2: Given a controllability graph $G_{\mathcal{R}}$ and scheduling $\eta_{\mathcal{R}_i}$, we define $G_{\mathcal{R}}(\eta_{\mathcal{R}_i}(h))$ the sub-graph of $G_{\mathcal{R}}$ induced by keeping the edges scheduled in the time slot h. We define $G_{\mathcal{R}}(\eta_{\mathcal{R}_i}) = \bigcup_{h=1}^{\Pi} G_{\mathcal{R}}(\eta_{\mathcal{R}_i}(h))$ the sub-graph of $G_{\mathcal{R}}$ induced by keeping the union of edges scheduled during the whole frame. We say that the pair $(G_{\mathcal{R}}, \eta_{\mathcal{R}_i})$ is *jointly connected* if there exists a path from the controller node v_c to the actuator node $v_{u,i}$ in $G_{\mathcal{R}}(\eta_{\mathcal{R}_i})$. Such definition can be given similarly for $G_{\mathcal{O}}$ and $\eta_{\mathcal{O}_i}$.

Designing a scheduling function in the above model induces a communication scheduling and a routing of the communication protocol. As in [19] our model exploits data redundancy, i.e. sending control data through multiple paths in the same frame and then merging these components according to the weight function, to render the MCN fault tolerant and to allow FDI. As illustrated in [6], the dynamics of a MCN $\mathcal{M}$ can be modeled by the interconnection of blocks as in Figure 1. The block $\mathcal{P}$ is characterized by the discrete-time transfer matrix $\mathcal{P}(z)$ obtained by discretizing the system P with sampling time $T = \Pi\Delta$. The block $\mathcal{R}$ models the dynamics introduced by the flow of the actuation data of all control inputs $u_i, i \in \mathbf{m}$ through the communication network represented by $G_{\mathcal{R}}$ according to the applied controllability scheduling functions $\eta_{\mathcal{R}_i}, i \in \mathbf{m}$. The reader is referred to [7] for a mathematical description of such dynamics. In particular, as a straightforward extension to the MIMO case of Proposition 1 of [7], and since the data flow semantics decouple the components of the input variables u and $\tilde{u}$, the transfer function that models the input/output behavior of $u_i(kT)$ with respect to $\tilde{u}_i(kT)$ can be expressed as follows, for any $i \in \mathbf{m}$:

$$\frac{U_i(z)}{\widetilde{U}_i(z)} \doteq \mathcal{R}_i(z) = \sum_{d=1}^{D_{\mathcal{R}_i}} \frac{\gamma_{\mathcal{R}_i}(d)}{z^d},\ \gamma_{\mathcal{R}_i}(d) = \sum_{\rho \in \chi_{\mathcal{R}_i}(d)} W_{\mathcal{R}_i}(\rho), \tag{1}$$

where $\chi_{\mathcal{R}_i}(d)$ is the set of paths of $G_{\mathcal{R}}(\eta_{\mathcal{R}_i})$ characterized by delay d, $D_{\mathcal{R}_i} \in \mathbb{N}$ is the maximum delay introduced by the paths of $G_{\mathcal{R}}(\eta_{\mathcal{R}_i})$ and $\forall d \in \{1, \ldots, D_{\mathcal{R}_i} - 1\}$, $\gamma_{\mathcal{R}_i}(d) \in \mathbb{R}_0^+$, $\gamma_{\mathcal{R}_i}(D_{\mathcal{R}_i}) \in \mathbb{R}^+$. By abuse of notation, we denote by $W_{\mathcal{R}_i}(\rho) = W_{\mathcal{R}_i}(v_c, v_1) \cdot W_{\mathcal{R}_i}(v_1, v_2) \cdot \ldots \cdot W_{\mathcal{R}_i}(v_{n-1}, v_n) \cdot W_{\mathcal{R}_i}(v_n, v_{u,i})$ the product of weights of all links that generate path $\rho = v_c, v_1, \ldots, v_n, v_{u,i}$ of $G_{\mathcal{R}}(\eta_{\mathcal{R}_i})$.

Given $\mathcal{R}_i(z)$, $i \in \mathbf{m}$, and since $\forall i \in \mathbf{m}, U_i(z) = \mathcal{R}_i(z)\widetilde{U}_i(z)$, the block $\mathcal{R}$ can be modeled by the transfer matrix $\mathcal{R}(z) = diag(\mathcal{R}_1(z), \ldots, \mathcal{R}_m(z))$.

Proposition 1: [23] The characteristic polynomial associated to $\mathcal{R}(z)$ is $\delta_{\mathcal{R}}(z) = z^{D_{\mathcal{R}}}$, where $D_{\mathcal{R}} = \prod_{i=1}^{m} D_{\mathcal{R}_i}$.

The transfer matrix $\mathcal{O}(z) = diag(\mathcal{O}_1(z), \ldots, \mathcal{O}_\ell(z))$ of the block $\mathcal{O}$ and the associated characteristic polynomial $\delta_{\mathcal{O}}(z) = z^{D_{\mathcal{O}}}$, with $D_{\mathcal{O}} = \prod_{i=1}^{\ell} D_{\mathcal{O}_i}$, can be derived similarly. The dynamics of a MIMO MCN $\mathcal{M}$ in the nominal case (i.e. when no node failures/attacks occur) can be modeled by the cascade of the transfer matrices $\mathcal{R}(z)$, $\mathcal{P}(z)$ and $\mathcal{O}(z)$, as illustrated in Figure 1, which is given by $\mathcal{M}(z) = \mathcal{O}(z)\mathcal{P}(z)\mathcal{R}(z)$.

We provide a modeling framework for the MCN dynamics induced by malicious clusters that have not yet been detected, which we will use to address Problem 2 in Section III; in particular, we will state (generic) conditions for detection and isolation of malicious clusters on a MCN exploiting the formalism of *structured systems* [14], which we now introduce. Given a system characterized by a state space representation $S = (A, B, C, D)$ we can define the associated structured system by defining the matrices $S_\lambda = (A_\lambda, B_\lambda, C_\lambda, D_\lambda)$ so that each entry is either zero (if the corresponding entry in the original matrix is zero) or a free parameter (if the corresponding entry of the original matrix is non-zero). For instance, consider a system S given by $A = [1, 2; 0, 0]$, $B = [0, 1]^\top$, $C = [1, 0]$: the corresponding structured system S_λ is given by $A = [\lambda_1, \lambda_2; 0, 0]$, $B = [0, \lambda_3]^\top$, $C = [\lambda_4, 0]$, where $\lambda_1, \ldots, \lambda_4$ are free parameters. A structured system can also be represented by a directed graph $(V_{S_\lambda}, E_{S_\lambda})$ whose vertices correspond to the input, state and output variables, and with an edge between two vertices if there is a non-zero free parameter λ_i relating the corresponding variables in the equations. The graph representation of the example above is given by $V_{S_\lambda} = \{u, x_1, x_2, y\}$, and $E_{S_\lambda} = \{(x_1, x_1), (x_2, x_1), (u, x_2), (x_1, y)\}$.

The model of a MCN $\mathcal{M}$ is the cascade of the blocks $\mathcal{R}$, $\mathcal{P}$ and $\mathcal{O}$, hence its structured graph representation $(V_{\mathcal{M}_\lambda}, E_{\mathcal{M}_\lambda})$ is given by the union of the structured graph representations of $\mathcal{R}_\lambda$, $\mathcal{P}_\lambda$ and $\mathcal{O}_\lambda$. Since failures/attacks are assumed to occur in the network nodes we denote by $(V_{\mathcal{P}_\lambda}, E_{\mathcal{P}_\lambda})$ the structured graph representation of the plant $\mathcal{P}$, where the sets of input and output nodes are respectively $\tilde{U} \doteq \{\tilde{u}_1, \ldots, \tilde{u}_m\}$ and $\tilde{Y} \doteq \{\tilde{y}_1, \ldots, \tilde{y}_\ell\}$.

An undetected failure/attack to a communication node $v \in V_{\mathcal{R}}$ can be modeled by a set of arbitrary signals $f_{v,i}(k)$, for any i such that there exist $v' \in V_{\mathcal{R}}$, $h \in \mathbf{\Pi}$ with $(v, v') \in \eta_{\mathcal{R}_i}(h)$, each summed to the i-th input component routed via node v. This general framework, as illustrated in [7] and [21], models several node failures (a node stops sending data or sends random data) as well as a wide set of malicious attacks (an arbitrary signal is injected that overrides/sums to the original data due to e.g. stealth, false-data injection and replay attacks). Following the same reasoning as in the definition of $\mathcal{R}_i(z)$, we can define the transfer function from $f_{v,i}(k)$ to $\tilde{u}_i(k)$ as follows:

$$T_{f_{v,i},\tilde{u}_i}(z) \doteq \frac{\widetilde{U}_i(z)}{F_{v,i}(z)} = \sum_{d=1}^{D_{v,i}} \frac{\gamma_{v,i}(d)}{z^d}, \tag{2}$$

where $D_{v,i} \in \mathbb{N}$ is the maximum delay introduced by the (routing) paths from v to the actuator node $v_{u,i}$ and $\forall d \in \boldsymbol{D_{v,i}}$, $\gamma_{v,i}(d) \in \mathbb{R}$, with $\gamma_{v,i}(D_{v,i}) \neq 0$. By the properties of $G_{\mathcal{R}}(\eta_{\mathcal{R}_i})$ it follows that $\forall v \in V_{\mathcal{R}}, D_{v,i} \leq D_{\mathcal{R}_i}$. The following proposition formalizes the structured graph representation of the block $\mathcal{R}$ when a set of malicious signals is applied to communication nodes.

Proposition 2: [6] Given $G_{\mathcal{R}}$, $\eta_{\mathcal{R}}$ and a set of faulty nodes $\bar{V} \subseteq V_{\mathcal{R}}$, the structured graph representation $(V_{\mathcal{R}_\lambda}, E_{\mathcal{R}_\lambda})$ of the block $\mathcal{R}$ is as follows:

$$V_{\mathcal{R}_\lambda} \doteq \{u_1, \ldots, u_m\} \cup \{\tilde{u}_1, \ldots, \tilde{u}_m\} \cup \bigcup_{i \in \boldsymbol{m},\, d \in \boldsymbol{D_{\mathcal{R}_i}}} \{x_{i,d}\} \cup \bigcup_{v \in \bar{V},\, i \in \boldsymbol{m}} \{f_{v,i}\},$$

$$\forall i \in \boldsymbol{m}, \forall d \in \boldsymbol{D_{\mathcal{R}_i}}, (u_i, x_{i,d}) \in E_{\mathcal{R}_\lambda} \Leftrightarrow \gamma_{\mathcal{R}_i}(d) \neq 0,$$
$$\forall i \in \boldsymbol{m}, \forall d \in \boldsymbol{D_{\mathcal{R}_i}}, \forall v \in \bar{V}, (f_{v,i}, x_{i,d}) \in E_{\mathcal{R}_\lambda} \Leftrightarrow \gamma_{v,i}(d) \neq 0,$$
$$\forall i \in \boldsymbol{m}, \forall d_1, d_2 \in \boldsymbol{D_{\mathcal{R}_i}}, (x_{i,d_1}, x_{i,d_2}) \in E_{\mathcal{R}_\lambda} \Leftrightarrow d_1 = d_2 + 1,$$
$$\forall i \in \boldsymbol{m}, (x_{i,1}, \tilde{u}_i) \in E_{\mathcal{R}_\lambda}.$$

Note that $f_{v,i}$ is a variable associated to a (malicious) signal $f_{v,i}(k)$ injected into the i-th component of the actuation data routed via node v. Also note that $x_{i,d}$ is a variable associated to the i-th input component that will be delivered with a delay d to the actuator node $v_{u,i}$,

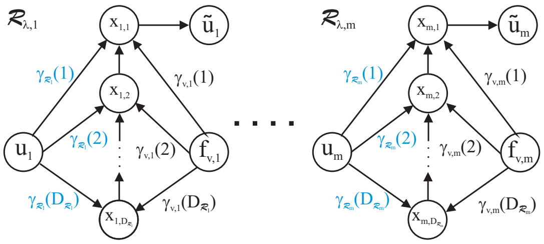
$\mathcal{R}_{\lambda,1}$
$x_{1,1}$
$\tilde{u}_1$
$\gamma_{\mathcal{R}_1}(1)$
$x_{1,2}$
$\gamma_{v,1}(1)$
u_1
$\gamma_{\mathcal{R}_1}(2)$
$\gamma_{v,1}(2)$
$f_{v,1}$
$\gamma_{\mathcal{R}_1}(D_{\mathcal{R}_1})$
$x_{1,D_{\mathcal{R}_1}}$
$\gamma_{v,1}(D_{\mathcal{R}_1})$
$\mathcal{R}_{\lambda,m}$
$x_{m,1}$
$\tilde{u}_m$
$\gamma_{\mathcal{R}_m}(1)$
$x_{m,2}$
$\gamma_{v,m}(1)$
u_m
$\gamma_{\mathcal{R}_m}(2)$
$\gamma_{v,m}(2)$
$f_{v,m}$
$\gamma_{\mathcal{R}_m}(D_{\mathcal{R}_m})$
$x_{m,D_{\mathcal{R}_m}}$
$\gamma_{v,m}(D_{\mathcal{R}_m})$

Fig. 2. Graph representation of $\mathcal{R}_\lambda$ when a failure/attack in the node v occurs.

and that $U \doteq \{u_1, \ldots, u_m\}$ and $\tilde{U} \doteq \{\tilde{u}_1, \ldots, \tilde{u}_m\}$ are respectively the sets of input and output variables. Figure 2 provides an example of the graph representation of $\mathcal{R}_\lambda$ when a failure/attack in the node v occurs. The edges of $E_{\mathcal{R}_\lambda}$ labeled with $\gamma_{\mathcal{R}_i}(d)$ (resp. $\gamma_{v,i}(d)$) are present in the graph if and only if the corresponding variables $\gamma_{\mathcal{R}_i}(d)$ in Equation (1) (resp. $\gamma_{v,i}(d)$ in Equation (2)) are not equal to 0. Note that $(V_{\mathcal{R}_\lambda}, E_{\mathcal{R}_\lambda})$ is composed by m weakly connected components, each associated to the data flow of the i-th input component. The same holds for defining the structured graph representation $(V_{\mathcal{O}_\lambda}, E_{\mathcal{O}_\lambda})$ of block $\mathcal{O}$, where the sets of input and output nodes are respectively $\tilde{Y} \doteq \{\tilde{y}_1, \ldots, \tilde{y}_\ell\}$ and $Y \doteq \{y_1, \ldots, y_\ell\}$. Note that in the structured graph representation $(V_{\mathcal{M}_\lambda}, E_{\mathcal{M}_\lambda})$ of a MCN $\mathcal{M}$ all nodes in $\tilde{U}$ and $\tilde{Y}$ are bridges since the removal of one of them increases the number of weakly connected components.

III. Fault Detection and Isolation on MCNs

In this section we provide a methodology to detect and isolate a malicious cluster in a MCN using the input signal of the controllability network and the output signal of the observability network. We first provide some basic definitions on FDI for structured systems, and then address the FDI problem for MCN.

Let a structured system S_λ be given in the form:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + E_1 d(k) + L_1 f(k), \\ y(k) &= Cx(k) + Du(k) + E_2 d(k) + L_2 f(k), \end{aligned}$$

where $d(k)$ is a vector of disturbance signals and $f(k) = [f_1(k), \ldots, f_r(k)]^\top$ is a vector of r malicious signals. In [4] necessary and sufficient conditions have been derived on the graph representation of S_λ that (generically) guarantee the existence of a bank of Luenberger observers, which takes as inputs $u(k), y(k)$, generates as output the *residual* signals vector $\hat{f}(k) = [\hat{f}_1(k), \ldots, \hat{f}_r(k)]^\top$ and is characterized by the diagonal transfer matrix

$$[\hat{F}_1(z), \ldots, \hat{F}_r(z)]^\top = diag\,(T_{11}(z), \ldots, T_{rr}(z))\,[F_1(z), \ldots, F_r(z)]^\top,$$

where $\forall i \in \boldsymbol{r}$, $T_{ii}(z) \neq 0$. Moreover, conditions in [4] are also given to guarantee that the transfer function from the disturbances to the residuals can be made zero, i.e. so that the disturbances do not affect the residuals at all. Characterizing the existence of such a bank of observers is called the *observer-based diagonal FDI problem*.

It is intuitive, and trivial to show from the conditions in [4], that in our setting the transfer function from disturbances applied to sensors and actuators to the outputs of the bank of residual generators cannot be made exactly zero: however, once the bank of residual generators has been designed using the methods in [17], such transfer function can be directly computed and classical methods for sensitivity analysis and disturbance rejection can be directly applied to our framework. As a consequence, in this paper we will not consider the disturbance (i.e. $E_1 = E_2 = \mathbf{0}$) and leave such generalization to further work. Also, it is well known [14] that the control input effects can be taken into account in the observer structure, therefore we will consider without loss of generality $B = D = 0$. The theorem below characterizes the observer-based diagonal FDI problem when there are no disturbances, and is a particular case of Theorem 3 in [4].

Theorem 3: [4] The observer-based diagonal FDI problem is generically solvable for a system S if and only if (i) S is structurally observable and (ii) $k = r$, where k is the maximum number of fault-output vertex disjoint paths in the graph representation of S_λ.

Note that, if the conditions of the above theorem are satisfied, there might be a zero-volume set of carefully chosen malicious signals that hide in the zero dynamics, being therefore undetectable: for example in [24], [21] the role of invariant zeros in detecting malicious attacks has been characterized for consensus networks.

In [23] we proved that, given a MIMO MCN $\mathcal{M}$ and if the plant $\mathcal{P}$ is controllable and observable, it is always possible to design a weight function W such that $\mathcal{M}$ is controllable and observable, thus we assume in this paper that $\mathcal{M}_\lambda$ is structurally observable - see [4] for the formal definition.

Given a MCN $\mathcal{M}$ subject to failures/attacks in the communication nodes, if the observer-based diagonal FDI problem is generically solvable for node failures/attacks in $\mathcal{M}_\lambda$, then the residual signals can be used to detect and isolate possibly simultaneous occurrence of node failures/attacks. In this section, in order to design the network configuration of a MCN $\mathcal{M}$ to enable FDI of malicious clusters, we state a formal relation between the network topology G and scheduling/routing η of $\mathcal{M}$, and the solvability conditions of the observer-based diagonal FDI problem for malicious clusters. We first show that if no assumptions are given on the malicious signals then the observer-based diagonal FDI problem is generically solvable for malicious clusters in $\mathcal{M}_\lambda$ only for trivial network topologies. Then we will raise a (weak) assumption on the malicious signals and show by means of necessary and sufficient conditions that, with this assumption, much more general network topologies guarantee FDI of malicious clusters.

In general, a failure/attack of a node $v \in V_\mathcal{R} \cup V_\mathcal{O}$ affects all the input components routed via v with possibly different signals $\{f_{v,i}(k)\}_{i \in \phi(v)}$, where

$$\phi(v) \doteq \begin{cases} \{i \in \boldsymbol{m} : (\exists v' \in V_\mathcal{R}, \exists h \in \boldsymbol{\Pi} : (v, v') \in \eta_{\mathcal{R}_i}(h))\} & \text{if } v \in V_\mathcal{R} \\ \{i \in \boldsymbol{\ell} : (\exists v' \in V_\mathcal{O}, \exists h \in \boldsymbol{\Pi} : (v, v') \in \eta_{\mathcal{O}_i}(h))\} & \text{if } v \in V_\mathcal{O} \end{cases}$$

represents the set components routed via node v. Since we aim at isolating malicious clusters, we are just interested in detecting whether at least one of the signals $\{f_{v,i}(k)\}_{i \in \phi(v)}$

is active. As a consequence the conditions for solvability of the observer-based diagonal FDI problem can be expressed as follows.

Lemma 4: Given a MCN $\mathcal{M}$ and a set $\bar{V} = \{v_1, \ldots, v_r\} \subseteq V_{\mathcal{R}} \cup V_{\mathcal{O}}$ of faulty nodes, define $\bar{F} = \bigcup_{v \in \bar{V}} \bigcup_{i \in \phi(v)} \{f_{v,i}\}$ the set of all malicious signals. The observer-based diagonal FDI problem is generically solvable for malicious clusters in $\mathcal{M}_\lambda$ if and only if, for any $\bar{v} \in \bar{V}$ and any $\bar{i} \in \phi(\bar{v})$, there exists an $(|\bar{F}| - |\phi(\bar{v})| + 1)$-linking from $\bar{F} \setminus \bigcup_{i \in \phi(\bar{v}) \setminus \{\bar{i}\}} \{f_{\bar{v},i}\}$ to Y in $(V_{\mathcal{M}_\lambda}, E_{\mathcal{M}_\lambda})$.

The above lemma shows that isolating failures/attacks of any 2 nodes belonging both to the controllability network is possible only for trivial network topologies, namely when all the graphs $G_{\mathcal{R}}(\eta_{\mathcal{R}_i}), i \in \boldsymbol{m}$ consist of a single communication node, and can be proven similarly for 2 nodes belonging to the observability graph.

Proposition 5: Given a MCN $\mathcal{M}$ and 2 faulty nodes $v_1, v_2 \in V_{\mathcal{R}}$, the observer-based diagonal FDI problem is generically solvable for malicious clusters in $\mathcal{M}_\lambda$ only if $\phi(v_1) \cap \phi(v_2) = \varnothing$.

This negative result depends of the fact that a malicious attack is assumed able to simultaneously inject into a node v some malicious signals $f_{v,i}(k)$ that are always zero and some that are non-zero for almost any k. To address this issue, we assumed in [6] that, when a failure/attack of a node v occurs, it affects all the input components routed via v with the same time signal, namely we assumed that $\forall i \in \phi(v), f_{v,i}(k) = f_v(k)$. Note that, with such assumption, a large class of malicious signals is ruled out.

In this paper we just assume that, when a failure/attack of a node v occurs, it affects all the input components routed via v with possibly different time signals that are all non-zero for almost all time instants.

Assumption 1: We assume that when a failure/attack of a node v occurs then, for all $i \in \phi(v)$, $f_{v,i}(k) \neq 0$ for *almost all* $k \in \mathbb{R}_0^+$.

Remark 1: We believe that the above assumption is quite weak because: (1) it allows modeling failures, because a malfunction of a node generically affects all components routed via that node injecting possibly different signals; (2) a malicious attack that does not satisfy Assumption 1 can be performed only if the attacker is aware of the scheduling function $\eta_{\mathcal{R}_i}, \forall i \in \boldsymbol{m}$ or $\eta_{\mathcal{O}_i}, \forall j \in \boldsymbol{\ell}$ associated to each component of the actuation/sensing signal, and of the internal structure of the payload of transmitted packets, in order to separately affect the bytes corresponding to specific components. Indeed, from a mathematical point of view, if $[f_{v,i_1}(k), \ldots, f_{v,i_{|\phi(v)|}}(k)]^\top$, where $\phi(v) = \{i_1, \ldots, i_{|\phi(v)|}\}$, is the malicious vector signal injected into node v, it is easy to see that Assumption 1 rules out a zero-volume set of malicious vector signals. On the other hand, we believe that the set of attack strategies of injecting non-zero signals only for some components of the routed data is not a zero-volume set with respect to the set of all attack strategies. As a consequence, it is controversial to state whether the results provided in the following of this paper hold generically or not. To meet the worst case we have made the choice to state that the following results, which provide necessary and sufficient conditions that guarantee FDI, hold only if Assumption 1 holds.

Let us define

$$\forall v \in V_{\mathcal{R}}, \forall i \in \boldsymbol{\ell},\ \Gamma_i(v) \doteq \{j \in \boldsymbol{m} : (\exists v' \in V_{\mathcal{R}}, \exists h \in \mathbf{\Pi} : (v, v') \in \eta_{\mathcal{R}_j}(h))$$
$$\text{and } \exists \text{ a path from } \tilde{u}_j \text{ to } \tilde{y}_i \text{ in } (V_{\mathcal{M}_\lambda}, E_{\mathcal{M}_\lambda})\},$$
$$\forall v \in V_{\mathcal{O}}, \forall i \in \boldsymbol{\ell},\ \Gamma_i(v) \doteq \begin{cases} \{i\} & \text{if } (\exists v' \in V_{\mathcal{O}}, \exists h \in \mathbf{\Pi} : (v, v') \in \eta_{\mathcal{O}_i}(h)) \\ \varnothing & \text{otherwise} \end{cases}$$

and let us model the effect of all node failures/attacks by applying, for any $i \in \boldsymbol{\ell}$, the function

$$f_i(k) \doteq \sum_{v \in V_{\mathcal{R}} \cup V_{\mathcal{O}}, j \in \Gamma_i(v)} t_{f_{v,j}, y_i}(k) *\ f_{v,j}(k), \qquad (3)$$

to the vertex y_i of the structured graph representation of $\mathcal{M}_\lambda$.

Remark 2: Note that $f_i(k)$ is given by a sum of convolution products (operator $*$) of each malicious signal $f_{v,j}$ with the inverse transform $t_{f_{v,j}, y_i}(k)$ of the transfer function $T_{f_{v,j}, y_i}(z)$ from $f_{v,j}(k)$ to $y_i(k)$. It is easy to see from Section II that if $v \in V_{\mathcal{O}}$ then $T_{f_{v,j}, y_i}(z)$ just corresponds to a delay transfer function, while if $v \in V_{\mathcal{R}}$ then $T_{f_{v,j}, y_i}(z)$ is the product of a pure delay transfer function and of the element i, j of the plant transfer function matrix $\mathcal{P}(z)$.

The main idea behind our approach for FDI on the basis of Assumption 1 can be summarized as follows: since it is not possible to detect and isolate certain malicious clusters (proven in Lemma 6), then we sum them up as in Equation (3), then we generate the residual signals $\hat{f}_i(k)$, and we finally exploit Assumption 1 to detect and isolate malicious clusters by applying a logic operator to $\hat{f}_i(k)$ being a zero or non-zero signal (proven in Lemma 3 and Corollary 8).

The following lemma shows that the observer-based diagonal FDI problem is not solvable for any set of malicious signals that models all node failures/attacks and is defined differently from Equation (3).

Lemma 6: The observer-based diagonal FDI problem is not solvable in $\mathcal{M}_\lambda$ for any set of malicious signals that models the effect of all malicious signals associated to all node failures/attacks, and that contains separately any two malicious signals $f_{v,j}(k), f_{v',j'}(k)$ such that $v, v' \in V_{\mathcal{R}} \cup V_{\mathcal{O}}, j \in \Gamma_i(v), j' \in \Gamma_i(v')$, for any $i \in \boldsymbol{\ell}$.

The following lemma proves that, if Assumption 1 holds, then $f_i(k)$ is non-zero for almost any k if and only if at least one node v such that $\Gamma_i(v) \neq \varnothing$ is faulty, and $f_i(k) = 0, \forall k \geq 0$ if and only if all nodes v such that $\Gamma_i(v) \neq \varnothing$ are not faulty.

Lemma 7: Let Assumption 1 hold, then the following properties hold generically:

1) $f_i(k)$ is non-zero for almost any k if and only if $\exists v, j : v \in V_{\mathcal{R}} \cup V_{\mathcal{O}}, j \in \Gamma_i(v)$ such that $f_{v,j}(k)$ is non-zero for almost any k.
2) $f_i(k) = 0, \forall k \geq 0$ if and only if for all $v, j : v \in V_{\mathcal{R}} \cup V_{\mathcal{O}}, j \in \Gamma_i(v)$, $f_{v,j}(k) = 0, \forall k \geq 0$.

The following corollary directly follows form Lemma 7.

Corollary 8: Given a set $\bar{V} \subseteq 2^{V_{\mathcal{R}} \cup V_{\mathcal{O}}}$ of admissible malicious clusters and any combination $\omega \in 2^{\boldsymbol{\ell}}$, the set

$\Omega(\omega) \subseteq 2^{\bar{V}}$ of malicious clusters that are possibly active when the signals $f_i(k), i \in \omega$ are non-zero and the signals $f_i(k), i \in \boldsymbol{\ell} \setminus \omega$ are zero, is given as follows:

$$\Omega(\omega) = \{\nu \in \bar{V} : (\forall i \in \omega, \exists \bar{v} \in \nu : \Gamma_i(\bar{v}) \neq \varnothing) \text{ and } \\ (\forall \bar{v} \in \nu, \forall i \in \boldsymbol{\ell} \setminus \omega, \Gamma_i(\bar{v}) = \varnothing)\}$$

The following theorem provides necessary and sufficient conditions that guarantee detection and isolation of malicious clusters on the basis of Assumption 1.

Theorem 9: Let a MCN $\mathcal{M}$ and a set $\bar{V} \subseteq 2^{V_{\mathcal{R}} \cup V_{\mathcal{O}}}$ of admissible malicious clusters be given. It is possible to exploit the solution of the observer-based diagonal FDI problem to detect and isolate each malicious cluster in $\bar{V}$ if and only if, for each $\nu \in \bar{V}$ there exists a set $\omega \in 2^{\boldsymbol{\ell}}$ such that $\nu = \Omega(\omega)$.

Remark 3: It is worth to remark that the more the number ℓ of outputs of the system, the easier to design a network topology, scheduling and routing that enable FDI of an increasing number of simultaneous failures/attacks. In particular, the set of possibly distinguishable malicious clusters increases with the cardinality of the set 2^{ℓ}, which grows very fast w.r.t. ℓ.

IV. Conclusions

In this paper we extend the results in [6] by relaxing the assumption on failure/malicious signals. As we have illustrated in Remark 1 we believe that our new assumption is quite weak since it rules out a zero-volume set of malicious vector signals. As future work we plan to improve the network model by introducing *transient* failures (e.g. packet losses) and to provide optimization techniques for optimal co-design of controller, network scheduling, routing and network coding that maximizes a performance metric and guarantees robust stability with respect to packet losses (see some initial results [8]).

References

[1] M. Andersson, D. Henriksson, A. Cervin, and K.-E. Årzén. Simulation of Wireless Networked Control Systems. In *Proceedings of the 44th IEEE CDC-ECC 2005. Seville, Spain*, pages 476 – 481, Dec. 2005.

[2] K. Aström and B. Wittenmark. *Computer-Controlled Systems: Theory and Design*. Prentice Hall, 1997.

[3] J. Chen and R. Kumar. Online failure diagnosis of stochastic discrete event systems. In *Proc. IEEE Conf. on Computer Aided Control System Design, Hyderabad, India*, pages 194–199, Aug 28-30 2013.

[4] C. Commault, J.-M. Dion, O. Sename, and R. Motyeian. Observer-based fault detection and isolation for structured systems. *IEEE Trans. on Automatic Control*, 47(12):2074–2079, Dec. 2002.

[5] C. Commault and J.-M. Dion. Sensor Location for Diagnosis in Linear Systems: A Structural Analysis. *IEEE Trans. on Automatic Control*, 52(2):155–169, Feb. 2007.

[6] A. D'Innocenzo, M.D. Di Benedetto, and F. Smarra. Fault detection and isolation of malicious nodes in MIMO Multi-hop Control Networks. In *Proc. of* 52^{st} *IEEE Conf. on Decision and Control, Firenze, Italy*, pages 5276–5281, December 10-13 2013.

[7] A. D'Innocenzo, M.D. Di Benedetto, and E. Serra. Fault tolerant control of multi-hop control networks. *IEEE Transactions on Automatic Control*, 58(6):1377–1389, June 2013.

[8] F. Smarra, A. D'Innocenzo, and M.D. Di Benedetto. Approximation methods for optimal network coding in a multi-hop control network with packet losses. In *Proceedings of the* 14^{th} *European Control Conference (ECC'15), Linz, Austria, July 15-17*, 2015.

[9] G.C. Walsh, Hong Ye, and L.G. Bushnell. Stability Analysis of Networked Control Systems. *IEEE Transactions on Control Systems Technology*, 10(3):438–446, 2002.

[10] R.A. Gupta and M.-Y. Chow. Networked Control System: Overview and Research Trends. *IEEE Transactions on Industrial Electronics*, 57(7):2527 –2535, July 2010.

[11] V. Gupta, A.F. Dana, J.P. Hespanha, R.M. Murray, and B. Hassibi. Data transmission over networks for estimation and control. *IEEE Transactions on Automatic Control*, 54(8):1807–1819, 2009.

[12] Song Han, Z. Xiuming, K.M. Aloysius, M. Nixon, T. Blevins, and D. Chen. Control over wirelesshart network. In *36th Annual Conf. on IEEE Industrial Electronics Society*, pages 2114–2119, 2010.

[13] I.F. Akyildiz and I.H. Kasimoglu. Wireless Sensor and Actor Networks: Research Challenges. *Ad Hoc Networks*, 2(4):351–367, 2004.

[14] J.-M. Dion, C. Commault, and J. van der Woude. Generic Properties and Control of Linear Structured Systems: a Survey. *Automatica*, 39(7):1125 –1144, July 2003.

[15] J.P. Hespanha, P. Naghshtabrizi, and Y. Xu. A Survey of Recent Results in Networked Control Systems. *Proceedings of the IEEE*, 95(1):138–162, January 2007.

[16] K.-E. Årzén, A. Bicchi, S. Hailes, K. H. Johansson, and J. Lygeros. On the design and control of wireless networked embedded systems. In *Proceedings of the 2006 IEEE Conference on Computer Aided Control Systems Design, Munich, Germany*, pages 440–445, October 2006.

[17] M.-A. Massoumnia, G.C. Verghese, and A.S. Willsky. Failure Detection and Identification. *IEEE Transactions on Automatic Control*, 34(3):316 –321, March 1989.

[18] M.C.F. Donkers, W.P.M.H. Heemels, Nathan van de Wouw, and Laurentiu Hetel. Stability Analysis of Networked Control Systems Using a Switched Linear Systems Approach. *IEEE Transactions on Automatic Control*, 56(9):2101 –2115, September 2011.

[19] M.D. Di Benedetto, A. D'Innocenzo, and E. Serra. Fault Tolerant Stabilizability of Multi-Hop Control Networks. In *Proceedings of the 18th IFAC World Congress, Milan, Italy*, pages 79–84, 2011.

[20] N. Meskin and K. Khorasani. Actuator Fault Detection and Isolation for a Network of Unmanned Vehicles. *IEEE Transactions on Automatic Control*, 54(4):835 –840, April 2009.

[21] F. Pasqualetti, A. Bicchi, and F. Bullo. Consensus computation in unreliable networks: A system theoretic approach. *IEEE Transactions on Automatic Control*, 57(1):90–104, 2012.

[22] R. Alur, A. D'Innocenzo, K.H. Johansson, G.J. Pappas, and G. Weiss. Compositional Modeling and Analysis of Multi-Hop Control Networks. *IEEE Transactions on Automatic Control, Special Issue on Wireless Sensor and Actuator Networks*, 56(10):2345–2357, 2011.

[23] F. Smarra, A. D'Innocenzo, and M.D. Di Benedetto. Fault Tolerant Stabilizability of MIMO Multi-Hop Control Networks. In *Proc.* 3^{rd} *IFAC Workshop on Estimation and Control of Networked Systems, Santa Barbara, CA*, pages 79–84, Sep 14-15 2012.

[24] S. Sundaram and C.N. Hadjicostis. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Trans. on Automatic Control*, 56(7):1495–1508, July 2011.

[25] M. Tabbara, D. Nešić, and A.R. Teel. Stability of Wireless and Wireline Networked Control Systems. *IEEE Transactions on Automatic Control*, 52(7):1615–1630, September 2007.

[26] W. Zhang, M.S. Branicky, and S.M. Phillips. Stability of Networked Control Systems. *IEEE Control Systems Magazine*, 21(1):84–99, Feb 2001.

[27] Y. Wang, S.X. Ding, H. Ye, and G. Wang. A New Fault Detection Scheme for Networked Control Systems Subject to Uncertain Time-Varying Delay. *IEEE Transactions on Signal Processing*, 56(10):5258 –5268, October 2008.

[28] G. Weiss, A. D'Innocenzo, R. Alur, K.H. Johansson, and G.J. Pappas. Robust Stability of Multi-Hop Control Networks. In *Proceedings of the 48th IEEE Conference on Decision and Control. Shangai, China*, pages 2210–2215, December 16-18, 2010.