

Distributed Faulty Node Detection in DTNs

Wenjie Li, Laura Galluccio, Michel Kieffer, Francesca Bassi

▶ **To cite this version:**

Wenjie Li, Laura Galluccio, Michel Kieffer, Francesca Bassi. Distributed Faulty Node Detection in DTNs. International Conference on Computer Communication and Networks, ICCCN 2016, Aug 2016, Hawaii, United States. 10.1109/icccn.2016.7568511 . hal-01327802

HAL Id: hal-01327802
https://hal.science/hal-01327802

Submitted on 7 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed Faulty Node Detection in DTNs

Wenjie Li*, Laura Galluccio†, Michel Kieffer*‡§, and Francesca Bassi*¶
*Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506) CNRS-CentraleSupelec-Université Paris-Sud
3, rue Joliot Curie 91192 Gif-sur-Yvette, France
†Dipartimento di Ingegneria Elettrica, Elettronica e Informatica, University of Catania, 95125 Catania, Italy
‡LTCI Telecom ParisTech, 75013 Paris, France
§Institut Universitaire de France, 75005 Paris, France
¶ESME-Sudria, 94200 Ivry-sur-Seine, France

***Abstract*—Due to their inherent feature of exhibiting frequent disconnections, propagation of faulty data in Delay Tolerant Networks can be a critical aspect to counteract. Indeed the rare meeting events require that nodes are effective and efficient in propagating the correct information. Accordingly mechanisms to rapidly identify possible faulty or misbehaving nodes should be searched. Distributed fault detection has been addressed in the literature in the context of sensor and vehicular networks, but unfortunately these solutions suffer for long delays in identifying and isolating misbehaving nodes. In this paper instead we propose a fully distributed, easily implementable, and fast convergent approach to allow each DTN node to rapidly identify whether its sensors are producing outliers. The behavior of the proposed algorithm is described by some continuous-time state equation, whose equilibrium is characterized. Detection and false alarm rates are estimated by comparing both theoretical and simulation results. Numerical results assess the effectiveness of the proposed solution and can give guidelines in the design of the algorithm.**[1]

I. Introduction

Delay Tolerant Networks (DTN) are challenging networks which have dynamic topology with frequent disconnectivity [1]. Due to this sparse and intermittent connectivity, inference and learning over DTNs is much more complicated than in traditional networks (see, *e.g.*, [2], [3], [4], [5]).

This paper considers the problem of distributed defective node detection in DTNs. A node is defective when at least one of its sensors frequently reports erroneous measurements. The identification of such defective nodes is very important to save communication resources and to prevent erroneous measurements polluting the estimates provided by the DTN. We assume, as in [6], that nodes are not aware of the status (good or defective) of their sensors, while their computation and communication capabilities remain fine, even if their sensors are defective. Moreover, each node of the DTN is assumed to behave in a rational way and is willing to know the status of its sensors.

Distributed fault detection (DFD) is a well-investigated topic when considering Wireless Sensor Networks (WSNs) (see, *e.g.*,[7], [8], [9]). WSNs considered in most of the works are dense and have a static topology. DFD in DTNs is much less investigated. Classical DFD algorithms mainly consist of two phases. First, a local outlier detection test (LODT) is performed using data collected from neighboring nodes. The LODTs usually employ different techniques such as majority voting [10], median [11] and mean [12] of the measurements, the modified three-sigma edit test [13] with the aim to decide which data is erroneous. Then, outcomes of LODT are disseminated to improve decision accuracy. When few measurements are available, classical LODTs are not very efficient. This is a typical situation in DTNs when two nodes meet, take measurements and share these measurements.

In the past, few papers on DFD for delay tolerant networks appeared. In [14], a VDTN is considered where a large number of sensor nodes are fixed and some vehicles, called mobile carriers (MCs) collect data from these sensors. Once a defective node is deemed defective by a MC, it is added to its blacklist. The MCs provide information to sensors about their status. MCs also exchange their blacklists to accelerate the faulty node detection.

In [15], the related problem of distributed misbehavior/malware detection in DTN is addressed, but the main drawback of this approach is the long time required to identify and isolate misbehaving nodes. Misbehavior detection in DTN is also considered in [4], where the DTN is perturbed by routing misbehavior caused by selfish or malicious nodes. Nevertheless, the identification is not distributed.

This paper presents a fully distributed and easily implementable algorithm to allow each node of a DTN to determine whether its own sensors produce outliers. As in [6], a LODT is assumed to be able to detect the presence of outliers in a set of measurements, without necessarily being able to determine which are the outliers. Generic LODTs are considered, only characterized by their probabilities of detection and false alarm. When two nodes meet, they exchange their local measurements and perform the same LODT. This LODT result helps both nodes to update the estimate of the status of their sensors. When, for a given node, the proportion of meetings during which the LODT outcome is that there are outliers is larger than some threshold, this node decides its sensors may be defective. In this case, it becomes silent, *i.e.* it does not transmit its measurements to its neighbors, but keeps collecting measurements from nodes met and updating the estimate of the status of its sensors. It may then have the opportunity to update its estimate and to communicate again if it then realizes not to be defective.

[1]This work has been partly supported by the NEWCOM# NoE and by the iCODE Institute, research project of the IDEX Paris-Saclay.

This paper shows that the behavior of the proposed DFD algorithm can be analyzed using continuous-time state equations. For this purpose, the belief of each node about the status of its sensors is quantized. The evolution of these quantized beliefs is then shown to follow two Markov chains. Given the characteristics of the LODT, approximations of the detection rate (DR), *i.e.*, proportion of nodes which have effectively identified their sensor as defective, and of the false alarm rate (FAR), *i.e.*, proportion of nodes which believe that their good sensor is defective while it is not, are also obtained. These theoretical results provide guidelines to properly choose the parameters of the algorithm.

The rest of the paper is organized as follows. Section II presents the system model and basic assumptions. Section III details the DFD algorithm for DTNs. Section IV discusses the transition probabilities between the states while Section V develops the theoretical analysis of the macroscopic evolution of the proportion of nodes in different states and presents a proof of the convergence of the algorithm. Section VI provides an analysis of the DTN state equations while Section VII introduces some approximations for the equilibrium; Section VIII provides some numerical results and Section IX concludes this paper.

II. System Model

Consider a set $\mathcal{S}$ of N_S moving nodes equipped with sensors. A subset $\mathcal{D} \subset \mathcal{S}$ of nodes is equipped with defective sensors producing *outliers*, *i.e.*, measurements corrupted by a noise which has characteristics significantly different from those of the noise corrupting measurements provided by good sensors. The true *status* of Node i is $\theta_i(t) = 0$ if its sensors are all good and $\theta_i(t) = 1$ if at least one of them is defective. Let $\widehat{\theta}_i$ denote the estimated status[2] of Node i. The proportion of nodes with good and defective sensors are p_0 and p_1, respectively, with $p_0 + p_1 = 1$. Each node is not aware of the status of its sensors. In what follows, we assume that over the time horizon of the observation, the status of sensors does not change, *i.e.*, $\theta_i(t) = \theta_i$. Our aim is to design a distributed algorithm so that each Node i evaluates an accurate estimate $\widehat{\theta}_i$ of its status θ_i, in the lowest time.

A. Communication model

Nodes can exchange information only during the limited time interval in which they are in proximity. As in [16], [5], we assume that the time interval between two successive meetings follows an exponential distribution with an inter-contact rate λ. Moreover, we assume that, during a meeting, at most two nodes meet and we do not consider meetings of more than two nodes. When more than two nodes meet at the same time instant, processing is performed pair-by-pair. These assumptions facilitate the analysis of the proposed DFD algorithm.

[2]We are assuming that each node is equipped with multiple sensors. This set of sensors at a single node is considered as a whole and described through a unique variable θ_i.

B. Local outlier detection test

As in [6], we consider a family of LODTs able to detect the presence of outliers in a set of n data (measurements, measurements with associated regressors or experimental conditions) $\mathcal{M} = \{m_1, \ldots, m_n\}$. LODTs can take various forms, see, *e.g.*, [6]. A LODT t takes a binary value: $t(m_{1,\ldots,m_n}) = 1$ when an outlier is detected, and $t(m_{1,\ldots,m_n}) = 0$ otherwise. The LODT is assumed to be characterized by a detection probability $q_{\text{D}}(n_0, n_1)$ and a false alarm probability $q_{\text{FA}}(n_0)$, where n_0 is the number of data produced by good sensors and n_1 is the number of data coming from defective sensors. Each node performing a LODT on a set of data has not to know n_0 and n_1, but the performance of the LODT will depend on the actual values of n_0 and n_1, which are used in the analysis of the DFD algorithm.

C. Detection scenario

We assume that during each meeting of a pair of nodes $(i, j) \in \mathcal{S}$, the nodes collect data through their sensors. Each node then may transmit its data to the other node. If a node has received data from its neighbor, it may run a LODT involving its own data and those received from its neighbor. We assume that the spatial and temporal correlation between data is such that only data collected during the meeting of two nodes can be exploited by a LODT. Therefore, previously collected data are not used. As a consequence, differently from [6] where n_0 and n_1 may be large, in the DTN scenario, a LODT will have $n_0 + n_1 = 2$ only.

III. DFD algorithm

In the proposed DFD algorithm, each node manages two counters $c_{\text{m},i}(t)$ and $c_{\text{d},i}(t)$ initialized at 0 at $t = 0$. Using $c_{\text{m},i}(t)$, Node i counts the number of *meetings* during which it has received data from its neighbor, and was able to perform a LODT. Using $c_{\text{d},i}(t)$, it counts the number of LODT tests it has performed resulting in the *detection* of outliers. Thus, Node i consider itself as carrying defective sensors, *i.e.* it estimates $\widehat{\theta}_i(t) = 1$, when the proportion of LOTDs leading to outlier detection is larger than some decision threshold ν, *i.e.* when $c_{\text{d},i}(t)/c_{\text{m},i}(t) \geqslant \nu$.

Nodes which currently estimate $\widehat{\theta}_i(t) = 1$ still acquire measurements when they meet another node, but they do not send the data to the node they meet to avoid infecting the network with outliers. A node, upon receiving data from another node, performs a LODT and updates $c_{\text{m},i}(t)$ and $c_{\text{d},i}(t)$. A node which meets a node considering itself as defective simply transmits its data, but since it does not receive any data, it does not update $c_{\text{m},i}(t)$ and $c_{\text{d},i}(t)$ at the end of the meeting.

The vector $\mathbf{x}_i(t) = (\theta_i, c_{\text{m},i}(t), c_{\text{d},i}(t))$ represents the (microscopic) *state* of each Node i. As $t \to \infty$, one has $c_{\text{m},i}(t) \to \infty$ if all the LODT results obtained in the previous meetings are considered, which leads to an infinite number of possible values for the state variable $\mathbf{x}_i(t)$ and the global (macroscopic) behavior of the algorithm is difficult to analyze. To limit the number of possible states, one has chosen to

consider the evolution of $c_{\mathrm{m},i}(t)$ and $c_{\mathrm{d},i}(t)$ over a sliding variable-length time interval containing the time instants of the last M meetings during which Node i has performed a LODT. Algorithm 1 summarizes the proposed DFD technique for an arbitrary reference Node i.

Algorithm 1 Sliding-Window DFD algorithm for Node i

1) Initialize $t_i^0 = 0$, $\widehat{\theta}_i\left(t_i^0\right) = 0$, $c_{\mathrm{m},i}(t_i^0) = c_{\mathrm{d},i}(t_i^0) = 0$, $\kappa = 1$, and $\mu = 0$.
2) Do

$$\begin{cases} \widehat{\theta}_i\left(t\right) = \widehat{\theta}_i\left(t_i^{\kappa-1}\right), \\ c_{\mathrm{m},i}\left(t\right) = c_{\mathrm{m},i}\left(t_i^{\kappa-1}\right), \\ c_{\mathrm{d},i}\left(t\right) = c_{\mathrm{d},i}\left(t_i^{\kappa-1}\right), \end{cases} \tag{1}$$

$$t = t + \delta t \tag{2}$$

until the κ-th meeting occurs at time t_i^κ with Node $j^\kappa \in \mathcal{S}$.
3) Perform local measurement of data $m_i\left(t_i^\kappa\right)$.
4) If $\widehat{\theta}_i\left(t_i^\kappa\right) = 0$, then transmit $m_i\left(t_i^\kappa\right)$ to Node j^κ.
5) If data m_{j^κ} have been received from Node j^κ, then
   a) $\mu = \mu + 1$. Perform a LODT with outcome y_i^μ.
   b) Update $c_{\mathrm{m},i}$ and $c_{\mathrm{d},i}$ as

$$\begin{cases} c_{\mathrm{m},i}(t_i^\kappa) = \min\left\{\mu, M\right\} \\ c_{\mathrm{d},i}(t_i^\kappa) = \sum_{m=\max\{1,\mu-M+1\}}^{\mu} y_i^m \end{cases} \tag{3}$$

   c) Update $\widehat{\theta}_i$ as follows

$$\widehat{\theta}_i(t_i^\kappa) = \begin{cases} 1 & \text{if } c_{\mathrm{d},i}(t_i^\kappa)/c_{\mathrm{m},i}(t_i^\kappa) \geqslant \nu, \\ 0 & \text{else} \end{cases} \tag{4}$$

6) $\kappa = \kappa + 1$.
7) Go to 2.

In the next sections we will detail the sliding-window DFD procedure illustrated in Algorithm 1.

IV. Evolution of the state of a node

At time t, among the nodes with status $\theta \in \{0,1\}$, let $X_\theta^{\ell,k}\left(t\right)$ be the proportion of nodes in state $\mathbf{x}_i(t) = (\theta, \ell, k)$, *i.e.*, with $c_{\mathrm{m},i}\left(t\right) = \ell$ and $c_{\mathrm{d},i}\left(t\right) = k$. Since $\ell \in \{0, 1, \ldots, M\}$ and $k \in \{0, \ldots, \ell\}$, the number of values that may be taken by the state of a node is $\sum_{\ell=0}^{M}\left(\ell+1\right) = \left(M+1\right)\left(M+2\right)/2$. The evolution of the state of Node i, conditioned by its status θ_i, follows a Markov model.

In particular, there are two chains, one conditioned by $\theta_i = 0$ and the other conditioned by $\theta_i = 1$. Both are characterized by a transient phase for states with $\ell < M$, corresponding to a number of meetings resulting in an LODT $c_{\mathrm{m},i}(t)$ less than M. Then, a permanent regime starts when $c_{\mathrm{m},i}(t) = M$. The transitions from State (θ, ℓ, k) to State (θ, ℓ', k') are analyzed in what follows.

Assume some reference Node i meets a random Node J at time t and define the random event

$$\mathcal{E}_1(t) = \left\{\widehat{\theta}_J\left(t\right) = 0\right\}, \tag{5}$$

representing the event that the node met believes its status is good. According to (4), among the nodes with status θ, the proportion of nodes that believe themselves as good ($\widehat{\theta}\left(t\right) = 0$) is [3]

$$p^{\theta 0}\left(t\right) = X_\theta^{0,0}\left(t\right) + \sum_{\ell,k,k/\ell<\nu} X_\theta^{\ell,k}\left(t\right). \tag{6}$$

Assuming that the nodes are randomly spread, when meeting Node J, one has

$$p^{\theta 0}\left(t\right) = \mathbb{P}\left\{\widehat{\theta}_J\left(t\right) = 0 | \theta_J\left(t\right) = \theta\right\}, \tag{7}$$

and

$$\mathbb{P}\left\{\mathcal{E}_1(t)\right\} = p_0 p^{00}\left(t\right) + p_1 p^{10}\left(t\right). \tag{8}$$

Similarly, introducing $\overline{\mathcal{E}_1}(t) = \left\{\widehat{\theta}_J\left(t\right) = 1\right\}$, among the nodes with sensors in status θ, the proportion of nodes with $\widehat{\theta}_j\left(t\right) = 1$ is

$$p^{\theta 1}\left(t\right) = \sum_{\ell,k,k/\ell\geqslant\nu} X_\theta^{\ell,k}\left(t\right), \tag{9}$$

leading to

$$\mathbb{P}\left\{\overline{\mathcal{E}_1}(t)\right\} = p_0 p^{01}\left(t\right) + p_1 p^{11}\left(t\right). \tag{10}$$

Since Node i performs an LODT only when it meets a node J with $\widehat{\theta}_J\left(t\right) = 0$, one introduces the random event

$$\mathcal{E}_2^\theta(t) = \left\{Y_i\left(t\right) = 1 \mid \theta_i = \theta, \widehat{\theta}_J\left(t\right) = 0\right\}, \tag{11}$$

for the reference node with actual status θ. From Section II-B, the statistical property of the outcome of the LODT $Y_i\left(t\right)$ depends only on θ_i and θ_j. For example, when Node i has good sensors, one has

$$\begin{aligned} \mathbb{P}\left\{\mathcal{E}_2^0(t)\right\} &= \mathbb{P}\left\{Y_i\left(t\right) = 1 \mid \theta_i = 0, \widehat{\theta}_J\left(t\right) = 0\right\} \\ &\overset{(a)}{=} \sum_{\varphi=0}^{1} \mathbb{P}\left\{Y_i\left(t\right) = 1 \mid \theta_i = 0, \theta_J = \varphi\right\} \mathbb{P}\left\{\theta_J = \varphi \mid \widehat{\theta}_J\left(t\right) = 0\right\} \\ &\overset{(b)}{=} \frac{p_0 q_{\mathrm{FA}}\left(2\right) p^{00}\left(t\right) + p_1 q_{\mathrm{D}}\left(1,1\right) p^{10}\left(t\right)}{p_0 p^{00}\left(t\right) + p_1 p^{10}\left(t\right)}. \end{aligned} \tag{12}$$

In (12-a), one uses the fact that the Y_i is not influenced by $\widehat{\theta}_J$ and θ_J does not depend on θ_i. In (12-b),

$$\begin{cases} \mathbb{P}\left\{Y_i\left(t\right) = 1 \mid \theta_i = 0, \theta_J = 0\right\} = q_{\mathrm{FA}}\left(2\right). \\ \mathbb{P}\left\{Y_i\left(t\right) = 1 \mid \theta_i = 0, \theta_J = 1\right\} = q_{\mathrm{D}}\left(1,1\right). \end{cases}$$

If Node i has defective sensors, (11) can be expressed as

$$\mathbb{P}\left\{\mathcal{E}_2^1(t)\right\} = \frac{p_0 q_{\mathrm{D}}\left(1,1\right) p^{00}\left(t\right) + p_1 q_{\mathrm{D}}\left(0,2\right) p^{10}\left(t\right)}{p_0 p^{00}\left(t\right) + p_1 p^{10}\left(t\right)}. \tag{13}$$

Similarly, one may introduce the random events

$$\mathcal{E}_3^\theta(t) = \left\{Y_i\left(t\right) = 0 \mid \theta_i = \theta, \widehat{\theta}_J\left(t\right) = 0\right\}, \tag{14}$$

[3]For the sake of simplicity, the dependency of $p^{\theta 0}\left(t\right)$ in ν is omitted, as ν is constant during the DFD algorithm.

and show that

$$\mathbb{P}\left\{\mathcal{E}_3^{\theta}(t)\right\} = \begin{cases} \frac{p_0(1-q_{\text{FA}}(2))p^{00}(t)+p_1(1-q_{\text{D}}(1,1))p^{10}(t)}{p_0p^{00}(t)+p_1p^{10}(t)}, & \text{if } \theta=0, \\ \frac{p_0(1-q_{\text{D}}(1,1))p^{00}(t)+p_1(1-q_{\text{D}}(0,2))p^{10}(t)}{p_0p^{00}(t)+p_1p^{10}(t)}. & \text{if } \theta=1. \end{cases} \quad (15)$$

Define $\pi_{\theta}^{\delta_{\text{m}},\delta_{\text{d}}}$ as the transition probability from State (θ,ℓ,k) to State $(\theta,\ell+\delta_{\text{m}},k+\delta_{\text{d}})$, where $\theta\in\{0,1\}$. One has $\delta_{\text{m}}\in\{0,1\}$ since ℓ may either increase $(\delta_{\text{m}}=1)$ in the transient regime or remain constant $(\delta_{\text{m}}=0)$ in the permanent regime. One has $\delta_{\text{d}}\in\{-1,0,1\}$, depending on the value of the last LODT outcome and on the value of the $M+1$-th last LODT outcome, which is no more considered in the permanent regime. Thus, $(\delta_{\text{m}},\delta_{\text{d}})\in\{(0,0),(0,1),(0,-1)(1,0),(1,1),(1,-1)\}$. Note that $\pi_{\theta}^{\delta_{\text{m}},\delta_{\text{d}}}$ depends on the current state of the reference node, but also on the current proportion of active (good and defective) nodes. Therefore, the transition probabilities are denoted as $\pi_{\theta}^{\delta_{\text{m}},\delta_{\text{d}}}(t,\ell,k)$, where t is the time instant, $c_{\text{m},i}(t)=\ell$, and $c_{\text{d},i}(t)=k$. Depending on the value of ℓ, two different cases are considered in Sections IV-A and IV-B, respectively corresponding to the transient and permanent regimes.

A. Case I, $\ell < M$

In the transient regime, when $c_{\text{m},i}(t)=\ell<M$ and when a Node J with $\widehat{\theta}_J(t)=0$ is met, $c_{\text{m},i}(t)$ and $c_{\text{d},i}(t)$ are updated according to (3). The only possibility that leads to $\delta_{\text{m}}=0$ is the event $\overline{\mathcal{E}_1}$, *i.e.*, Node i meets Node J with $\widehat{\theta}_J(t)=1$. As a consequence, no LODT is performed by Node i. Therefore, for any $\theta\in\{0,1\}$,

$$\pi_{\theta}^{0,0}(t,\ell,k)=\mathbb{P}\left\{\overline{\mathcal{E}_1}(t)\right\}=p_0p^{01}(t)+p_1p^{11}(t), \quad (16)$$

where $p^{\theta 1}(t)$ is defined by (9).

A state transition occurs with $(\delta_{\text{m}},\delta_{\text{d}})=(1,1)$ when Node i with status $\theta_i=\theta$ meets Node J with $\widehat{\theta}_J(t)=0$ *and* when the LODT yields $y_i(t)=1$. One has then

$$\begin{aligned}\pi_{\theta}^{1,1}(t,\ell,k)&=\mathbb{P}\left\{Y_i(t)=1,\widehat{\theta}_J(t)=0|\theta_i=\theta\right\}\\&=\mathbb{P}\left\{\mathcal{E}_1\right\}\mathbb{P}\left\{\mathcal{E}_2^{\theta}\right\},\end{aligned} \quad (17)$$

since both events are independent. Depending on the value of θ_i, using (8), (12), and (13), one may rewrite (17) as

$$\pi_{\theta}^{1,1}(t,\ell,k)=\begin{cases}p_0q_{\text{FA}}(2)p^{00}(t)+p_1q_{\text{D}}(1,1)p^{10}(t), & \text{if } \theta=0,\\ p_0q_{\text{D}}(1,1)p^{00}(t)+p_1q_{\text{D}}(0,2)p^{10}(t), & \text{if } \theta=1.\end{cases} \quad (18)$$

Finally, $\pi_{\theta}^{1,0}(t,\ell,k)=\mathbb{P}\left\{Y_i(t)=0,\widehat{\theta}_J(t)=0|\theta_i=\theta\right\}$ is obtained similarly from (15)

$$\pi_{\theta}^{1,0}(t,\ell,k)=\begin{cases}p_0(1-q_{\text{FA}}(2))p^{00}(t)+p_1(1-q_{\text{D}}(1,1))p^{10}(t), & \text{if } \theta=0,\\ p_0(1-q_{\text{D}}(1,1))p^{00}(t)+p_1(1-q_{\text{D}}(0,2))p^{10}(t), & \text{if } \theta=1.\end{cases} \quad (19)$$

B. Case II, $c_{m,i}(t)=M$

In the permanent regime, $c_{\text{m},i}(t)=M$ and does not increase any more, thus $\delta_{\text{m}}=0$. In Algorithm 1, μ is the number of LODTs performed by Node i till time t. When $\mu\geqslant M$, only the last M LODT outcomes are considered. To determine the value taken by $\delta_{\text{d}}\in\{-1,0,1\}$ after the μ-th LODT, consider an arbitrary $y\in\{0,1\}$ and the random event

$$\mathcal{E}_4^y(t)=\left\{Y_i^{\mu-M}=y\middle|\sum_{m=\mu-M}^{\mu-1}Y_i^m=k\right\}, \quad (20)$$

which corresponds to a situation where one knows that k LODTs are positive among the last M tests and $Y_i^{\mu-M}=y$ will be ignored once the new LODT outcome is available. $\mathbb{P}\left\{\mathcal{E}_4^1(t)\right\}$ is relatively complex to evaluate, since $\mathbb{P}\{Y_i^n=y\}$ is time-varying according to (12-13). In what follows, we assume that LODT outcomes with $Y_i^m=y$, are independently distributed over the time horizon corresponding to $m=\mu-M,\ldots,\mu-1$. One obtains then $\mathbb{P}\left\{\mathcal{E}_4^1(t)\right\}=k/M$ and $\mathbb{P}\left\{\mathcal{E}_4^0(t)\right\}=1-k/M$. This approximation is exact in steady-state, when the $X_{\theta}^{\ell,k}$s do not vary any more.

Assume that the $(\mu-M)$-th LODT performed by Node i occurred at time $\tilde{t}$, then $y_i^{\mu-M}$ can also be denoted as $y_i(\tilde{t})$ and the transition related to $c_{\text{d},i}$ is such that

$$\delta_{\text{d}}=y_i(t)-y_i(\tilde{t})\in\{-1,0,1\}. \quad (21)$$

To have $(\delta_{\text{m}},\delta_{\text{d}})=(0,1)$, three independent events have to occur: 1) the encountered Node J believes it is good at time t, *i.e.*, $\mathcal{E}_1(t)$; 2) $y_i(t)=1$, *i.e.*, $\mathcal{E}_2(t)$; 3) $y_i(\tilde{t})=0$, *i.e.*, $\mathcal{E}_4^0(t)$. Thus the transition probability

$$\pi_{\theta}^{0,1}(t,M,k)=\mathbb{P}\left\{\mathcal{E}_1(t)\right\}\mathbb{P}\left\{\mathcal{E}_2^{\theta}(t)\right\}\mathbb{P}\left\{\mathcal{E}_4^0(t)\right\}. \quad (22)$$

Using (8), (12), and (13) in (22), one gets

$$\pi_{\theta}^{0,1}(t,M,k)=\begin{cases}\left(p_0q_{\text{FA}}(2)p^{00}(t)+p_1q_{\text{D}}(1,1)p^{10}(t)\right)\frac{M-k}{M}, & \text{if } \theta=0,\\ \left(p_0q_{\text{D}}(1,1)p^{00}(t)+p_1q_{\text{D}}(0,2)p^{10}(t)\right)\frac{M-k}{M}, & \text{if } \theta=1.\end{cases} \quad (23)$$

Consider now $(\delta_{\text{m}},\delta_{\text{d}})=(0,-1)$. To have such transition, the three following independent events should occur: 1) $\mathcal{E}_1(t)$; 2) $y_i(t)=0$, *i.e.*, $\mathcal{E}_3(t)$; 3) $y_i(\tilde{t})=1$, *i.e.*, $\mathcal{E}_4^1(t)$. Thus, the transition probability is

$$\pi_{\theta}^{0,-1}(t,M,k)=\begin{cases}\left(p_0(1-q_{\text{FA}}(2))p^{00}(t)+p_1(1-q_{\text{D}}(1,1))p^{10}(t)\right)\frac{k}{M}, & \text{if } \theta=0,\\ \left(p_0(1-q_{\text{D}}(1,1))p^{00}(t)+p_1(1-q_{\text{D}}(0,2))p^{10}(t)\right)\frac{k}{M}, & \text{if } \theta=1.\end{cases} \quad (24)$$

Considering the last transition $(\delta_{\text{m}},\delta_{\text{d}})=(0,0)$, two situations have to be considered. First, the encountered node believes itself is defective. This corresponds to the event $\overline{\mathcal{E}_1}(t)$ where no LODT is performed. Second, one may have $\mathcal{E}_1(t)$ and $Y_i(t)=Y_i(\tilde{t})$. An alternative way to obtain $\pi_{\theta}^{0,0}(t,M,k)$ is by

$$\pi_{\theta}^{0,0}(t,M,k)=1-\pi_{\theta}^{0,1}(t,M,k)-\pi_{\theta}^{0,-1}(t,M,k). \quad (25)$$

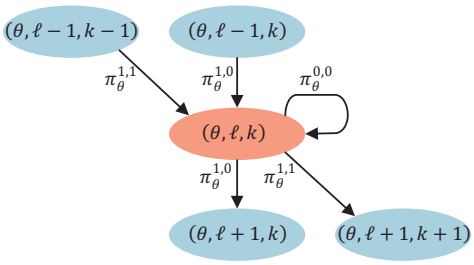


Fig. 1. Possible state transitions from and to State (θ,ℓ,k) when $0<\ell<M$ and $0<k<\ell$

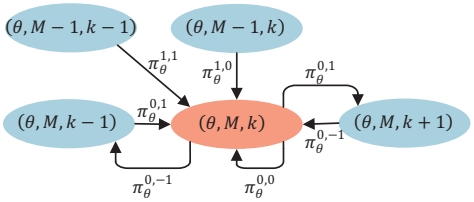


Fig. 2. Possible state transitions from and to State (θ,M,k) when $0<k<M$

Introducing (23-24) into (25), the expression of $\pi_\theta^{0,0}$ can be obtained.

V. Macroscopic evolution

All node state transition probabilities evaluated in Section IV are now used to determine the evolution of the state of the DTN, *i.e.*, of the various proportions $X_0^{\ell,k}(t)$ and $X_1^{\ell,k}(t)$ of nodes in the corresponding states, with $\ell=0,\dots,M$ and $k\leqslant\ell$.

Considering an inter-contact rate λ and a well-mixed population of nodes, during a short time interval $[t,t+\delta t]$, the number of nodes with state (θ,ℓ,k) that will meet another node can be estimated as $\lambda p_\theta N_S X_\theta^{\ell,k}(t)\delta t$. When $0<\ell<M$ and $0<k<\ell$, these nodes will switch to the states $(\theta,\ell+\delta_{\rm m},k+\delta_{\rm d})$, with $(\delta_{\rm m},\delta_{\rm d})\in\{(0,0),(1,0),(1,1)\}$ with a probability $\pi_\theta^{\delta_{\rm m},\delta_{\rm d}}(t,\ell,k)$. Moreover, nodes in the states $(\theta,\ell-1,k-1)$ and $(\theta,\ell-1,k)$ that have met another node in the time interval $[t,t+\delta t]$ may reach state (θ,ℓ,k), respectively with a probability $\pi_\theta^{1,1}(t,\ell-1,k-1)$ and $\pi_\theta^{1,0}(t,\ell-1,k)$ (see Figure 1). The evolution of $X_\theta^{\ell,k}(t)$ is then described by the following differential equation, where the time dependency is omitted to lighten notations

$$\begin{aligned}\frac{dX_\theta^{\ell,k}}{dt}&=-\lambda X_\theta^{\ell,k}\left(\pi_\theta^{1,0}(\ell,k)+\pi_\theta^{1,1}(\ell,k)\right)\\&+\lambda X_\theta^{\ell-1,k-1}\pi_\theta^{1,1}(\ell-1,k-1)+\lambda X_\theta^{\ell-1,k}\pi_\theta^{1,0}(\ell-1,k).\end{aligned}\quad(26)$$

When $\ell=M$ and $0<k<M$, nodes in state (θ,M,k) will switch to the states $(\theta,M,k+\delta_{\rm d})$, $\delta_{\rm d}\in\{-1,0,1\}$ with a probability $\pi_\theta^{0,\delta_{\rm d}}(t,M,k)$. Nodes in the states $(\theta,M-1,k-1)$ and $(\theta,M-1,k)$ that have met an other node in the time interval $[t,t+\delta t]$ may reach state (θ,M,k), respectively with a probability $\pi_\theta^{1,1}(t,M-1,k-1)$ and $\pi_\theta^{1,0}(t,M-1,k)$, see Figure 1. As a consequence, the evolution of $X_\theta^{M,k}(t)$ can be described by

$$\begin{aligned}\frac{dX_\theta^{M,k}}{dt}&=-\lambda X_\theta^{M,k}\left(\pi_\theta^{0,1}(M,k)+\pi_\theta^{0,-1}(M,k)\right)\\&+\lambda X_\theta^{M-1,k-1}\pi_\theta^{1,1}(M-1,k-1)+\lambda X_\theta^{M-1,k}\pi_\theta^{1,0}(M-1,k)\\&+\lambda X_\theta^{M,k-1}\pi_\theta^{0,1}(M,k-1)+\lambda X_\theta^{M,k+1}\pi_\theta^{0,-1}(M,k+1).\end{aligned}\quad(27)$$

Similar derivations can be made for the remaining DTN state components to obtain

$$\begin{cases}\frac{dX_\theta^{0,0}}{dt}&\overset{(a)}{=}-\lambda X_\theta^{0,0}\left(\pi_\theta^{1,0}(0,0)+\pi_\theta^{1,1}(0,0)\right),\\\frac{dX_\theta^{\ell,0}}{dt}&\overset{(b)}{=}\lambda\Big(-X_\theta^{\ell,0}\left(\pi_\theta^{1,0}(\ell,0)+\pi_\theta^{1,1}(\ell,0)\right)\\&\qquad+X_\theta^{\ell-1,0}\pi_\theta^{1,0}(\ell-1,0)\Big),\\\frac{dX_\theta^{\ell,\ell}}{dt}&\overset{(c)}{=}\lambda\Big(-X_\theta^{\ell,\ell}\left(\pi_\theta^{1,0}(\ell,\ell)+\pi_\theta^{1,1}(\ell,\ell)\right)\\&\qquad+X_\theta^{\ell-1,\ell-1}\pi_\theta^{1,1}(\ell-1,\ell-1)\Big),\\\frac{dX_\theta^{M,0}}{dt}&\overset{(d)}{=}\lambda\Big(-X_\theta^{M,0}\pi_\theta^{0,1}(M,0)+X_\theta^{M-1,0}\pi_\theta^{1,0}(M-1,0)\\&\qquad+X_\theta^{M,1}\pi_\theta^{0,-1}(M,1)\Big),\\\frac{dX_\theta^{M,M}}{dt}&\overset{(e)}{=}\lambda\Big(-X_\theta^{M,M}\pi_\theta^{0,-1}(M,M)+X_\theta^{M,M-1}\pi_\theta^{0,1}(M,M-1)\\&\qquad+X_\theta^{M-1,M-1}\pi_\theta^{1,1}(M-1,M-1)\Big),\end{cases}\quad(28)$$

for any $\ell=1\dots M-1$, with the initial conditions $X_\theta^{0,0}(0)=1$ and $X_\theta^{\ell,k}(0)=0$, $\forall\ell,k\neq0$. Note that the state equations (26-28) are nonlinear, since each $\pi_\theta^{\delta{\rm m},\delta{\rm d}}$ depends on $X_\theta^{\ell,k}$, see (6) and (9).

VI. Analysis of the DTN state equations

In what follows, the asymptotic behavior of the DTN state equations (26-28) is characterized.

A. Evolution of $X_\theta^{\ell,k}$ with $\ell<M$

In this section, one focuses on the evolution of $X_\theta^{\ell,k}(t)$ when $\ell<M$. The aim is to determine conditions under which the DTN state reaches the permanent regime.

Proposition 1. *For any $\ell<M$ and $k\leqslant\ell$, $\lim_{t\to\infty}X_\theta^{\ell,k}(t)=0$.*

Proof: To prove Proposition 1, one needs Lemma 1 and Lemma 2, stated in the following. The proof is by induction. Starting with (28.a), one has

$$X_\theta^{0,0}(t)=\exp\left(-\lambda\int_0^t\left(p_0p^{00}(\tau)+p_1p^{10}(\tau)\right)d\tau\right)\quad(29)$$

As a consequence, if (30) is satisfied, for any $\xi>0$, there exists $t_{00}>0$ such that $t>t_{00}$ implies $X_\theta^{0,0}(t)<\xi$ and $\lim_{t\to\infty}X_\theta^{0,0}(t)=0$.

Then, assume that for any $\ell\leqslant M-1$, and $\xi>0$, there exists $t_{(\ell-1)0}>\dots>t_{00}$ such that $t>t_{(\ell-1)0}$ implies $X_\theta^{j,0}(t)<\xi$ for $j=0,\dots,\ell-1$. One has to show now that there exists $t_{\ell0}>t_{(\ell-1)0}$ such that $X_\theta^{\ell,0}(t)<\xi$ for all $t>t_{\ell0}$.

Define $Z_\theta^{\ell,0}(t)=\sum_{j=0}^{\ell}X_\theta^{j,0}(t)$. From (28$a$) and (28$b$), one has

$$\frac{dZ_\theta^{\ell,0}}{dt}=-\lambda\left(v(t)Z_\theta^{\ell-1,0}(t)+\left(p_0p^{00}(t)+p_1p^{10}(t)\right)X_\theta^{\ell,0}(t)\right),$$

where $v(t) = \pi_\theta^{1,1}(t,\ell,k)$, since $\pi_\theta^{1,0}$ and $\pi_\theta^{1,1}$ do not depend on ℓ and k when $\ell < M$. Using (31) one has $dZ_\theta^{\ell,0}/dt < 0$ for any $X_\theta^{\ell 0} > 0$. As a consequence, $Z_\theta^{\ell,0}(t)$ decreases until $X_\theta^{\ell,0}(t)$ reaches 0. Hence, for any $\xi > 0$, there exists $t_{\ell,0} > t_{(\ell-1)0}$, such that $X_\theta^{\ell,0} < \xi$ and $\lim_{t\to\infty} X_\theta^{\ell,0}(t) = 0$.

In the same way, using (28c) and the previous results that $X_\theta^{\ell k}(t) \to 0$ with $k = 1,\ldots,M-2$ and $\ell = k,\ldots,M-2$, one can prove that for any $k = 1,\ldots,M-1$, $X_\theta^{\ell',(k+1)}(t)$ tends to zero as $t \to \infty$, with any $\ell' = k+1,\ldots,M-1$. ■

Lemma 1. *If*

$$\lim_{t\to\infty} \int_0^t \left(p_0 p^{00}(\tau) + p_1 p^{10}(\tau)\right) d\tau = \infty \tag{30}$$

then $p_0 p^{00}(t) + p_1 p^{10}(t) > 0$ *for all* $t \in \mathbb{R}^+$.

Proof: Since $p_0 > 0$, $p_1 > 0$, $p^{00} \geqslant 0$, and $p^{10} \geqslant 0$, one has to prove that

$$p^{00}(t) + p^{10}(t) \neq 0 \qquad \forall t > 0. \tag{31}$$

Assume that there exists a time instant $t^* > 0$, such that $p^{00}(t^*) + p^{10}(t^*) = 0$. From the definition of p^{00} and p^{10} introduced in (6) and (9), one deduces that at t^* all nodes in the network believe themselves as carrying defective sensors. As a consequence, no node will transmit its data to its neighbors. No LODTs is performed and the state of nodes will no longer change. Therefore, if $p^{00}(t^*) + p^{10}(t^*) = 0$ for some t^*, then $p^{00}(t) + p^{10}(t) = 0$ for all $t > t^*$. Then, there exists some constant C, such that $\lim_{t\to\infty} \int_0^t \left(p^{00}(\tau) + p^{10}(\tau)\right) d\tau < C$, which contradicts (30). ■

Lemma 2. *The property* (30) *is always satisfied.*

Proof: Assume that (30) does not hold, *i.e.*, there exists C^* such that $\lim_{t\to\infty} \int_0^t \left(p_0 p^{00}(\tau) + p_1 p^{10}(\tau)\right) d\tau \geq C^*$ for all $t > 0$. Then

$$\begin{aligned} X_\theta^{0,0}(t) &= \exp\left(-\lambda \int_0^t \left(p_0 p^{00}(\tau) + p_1 p^{10}(\tau)\right) d\tau\right) \\ &> \exp\left(-\lambda C^*\right) > 0. \end{aligned} \tag{32}$$

However, from (6), one has $p^{00}(\tau) \geqslant X_\theta^{0,0}(\tau)$, leading to

$$\int_0^t \left(p^{00}(\tau) + p^{10}(\tau)\right) d\tau \geqslant \int_0^t X_\theta^{0,0}(\tau)\, d\tau > \exp\left(-\lambda C^*\right) t$$

and $\exp\left(-\lambda C^*\right) t \to \infty$ as $t \to \infty$, which violates (32). Hence, (30) is always satisfied. ■

B. Equilibrium point of $X_\theta^{\ell,k}$

The algorithm may reach an equilibrium $\overline{X}_\theta^{M,k}$ at which the proportions of nodes in different states $X_\theta^{M,k}$ do not vary any more. In Section VI-A, one shows that $\lim_{t\to\infty} X_\theta^{\ell,k}(t) = 0$, $\forall \ell < M$. As a consequence, the only possible values at equilibrium of $X_\theta^{\ell,k}(t)$ with $\ell < M$ is 0. To evaluate $\overline{X}_\theta^{M,k}$, one thus considers the following simplified dynamics derived from (28)

$$\begin{cases} \frac{dX_\theta^{M,0}}{dt} = \lambda\Big(-X_\theta^{M,0}\pi_\theta^{0,1}(M,0) + X_\theta^{M,1}\pi_\theta^{0,-1}(M,1)\Big), \\ \frac{dX_\theta^{M,M}}{dt} = \lambda\Big(-X_\theta^{M,M}\pi_\theta^{0,-1}(M,M) + X_\theta^{M,M-1}\pi_\theta^{0,1}(M,M-1)\Big), \\ \frac{dX_\theta^{M,k}}{dt} = \lambda\Big(-X_\theta^{M,k}\left(\pi_\theta^{0,-1}(M,k) + \pi_\theta^{0,1}(M,k)\right) \\ \qquad + X_\theta^{M,k+1}\pi_\theta^{0,-1}(M,k+1) + X_\theta^{M,k-1}\pi_\theta^{0,1}(M,k-1)\Big). \end{cases} \tag{33}$$

At equilibrium, one has $dX_\theta^{M,k}(t)/dt = 0$ for all $k \leqslant M$. Moreover, the transition probabilities will not vary any more. Let $\overline{\boldsymbol{X}}_\theta^M = \left[\overline{X}_\theta^{M,1},\ldots,\overline{X}_\theta^{M,M}\right]^{\mathrm{T}}$ be the state values at equilibrium, $a_\theta(k) = \pi_\theta^{0,1}(M,k)$, and $b_\theta(k) = \pi_\theta^{0,-1}(M,k)$. At equilibrium, from 33 one should have $\boldsymbol{\Psi}\cdot\overline{\boldsymbol{X}}_\theta^M = \mathbf{0}$ where

$$\boldsymbol{\Psi} = \begin{pmatrix} -a_\theta(0) & b_\theta(1) & & & \\ a_\theta(0) & -a_\theta(1) - b_\theta(1) & b_\theta(2) & & \\ & \ddots & \ddots & \ddots & \\ & & a_\theta(M-1) & -b_\theta(M) \end{pmatrix}.$$

Summing Lines 1 to $k+1$, for all $k = 0,\ldots,M-1$, one obtains $b_\theta(k+1)\,\overline{X}_\theta^{M,k+1} = a_\theta(k)\,\overline{X}_\theta^{M,k}$, which leads to

$$\overline{X}_\theta^{M,k} = \overline{X}_\theta^{M,0} \prod_{j=0}^{k-1} \frac{a_0(j)}{b_0(j+1)}. \tag{34}$$

One evaluates

$$\frac{a_\theta(j)}{b_\theta(j+1)} = \frac{\pi_\theta^{0,1}(M,j)}{\pi_\theta^{0,-1}(M,j+1)} = \eta_\theta \frac{M-j}{j+1}, \tag{35}$$

where η_θ is obtained using (23) and (24),

$$\begin{cases} \eta_0 = \frac{p_0 q_{\mathrm{FA}}(2)\overline{p}^{00} + p_1 q_{\mathrm{D}}(1,1)\overline{p}^{10}}{p_0(1-q_{\mathrm{FA}}(2))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(1,1))\overline{p}^{10}} \\ \eta_1 = \frac{p_0 q_{\mathrm{D}}(1,1)\overline{p}^{00} + p_1 q_{\mathrm{D}}(0,2)\overline{p}^{10}}{p_0(1-q_{\mathrm{D}}(1,1))\overline{p}^{00} + p_1(1-q_{\mathrm{D}}(0,2))\overline{p}^{10}} \end{cases} \tag{36}$$

with

$$\overline{p}^{\theta 0} = \sum_{k:k/M<\nu} \overline{X}_\theta^{M,k}. \tag{37}$$

From (34) and (35), one has

$$\overline{X}_\theta^{M,k} = \overline{X}_\theta^{M,0} \prod_{j=0}^{k-1} \left(\eta_\theta \frac{M-j}{j+1}\right) = \binom{M}{k} \eta_\theta^k \overline{X}_\theta^{M,0}. \tag{38}$$

Since $\sum_{k=0}^M \overline{X}_\theta^{M,k} = 1$, one obtains

$$\overline{X}_\theta^{M,k} = \binom{M}{k} \left(\frac{\eta_\theta}{\eta_\theta+1}\right)^k \left(\frac{1}{\eta_\theta+1}\right)^{M-k}, \tag{39}$$

$\forall k = 0,\ldots,M$. Introduce (39) into (37), one obtains (40) (at the top of the next page), which have to be solved to determine $\overline{p}^{00}$ and $\overline{p}^{10}$. Then $\overline{X}_\theta^{M,k}$ can be deduced from $\overline{p}^{\theta 0}$ directly, using (39).

$$\begin{cases} \overline{p}^{00} = F_0\left(\overline{p}^{00}, \overline{p}^{10}\right) = \sum_{k:k/M<\nu} \binom{M}{k} \left(\frac{p_0 q_{\text{FA}}(2)\overline{p}^{00} + p_1 q_{\text{D}}(1,1)\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^k \left(\frac{p_0(1-q_{\text{FA}}(2))\overline{p}^{00} + p_1(1-q_{\text{D}}(1,1))\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^{M-k}, \\ \overline{p}^{10} = F_1\left(\overline{p}^{00}, \overline{p}^{10}\right) = \sum_{k:k/M<\nu} \binom{M}{k} \left(\frac{p_0 q_{\text{D}}(1,1)\overline{p}^{00} + p_1 q_{\text{D}}(0,2)\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^k \left(\frac{p_0(1-q_{\text{D}}(1,1))\overline{p}^{00} + p_1(1-q_{\text{D}}(0,2))\overline{p}^{10}}{p_0\overline{p}^{00} + p_1\overline{p}^{10}}\right)^{M-k}. \end{cases} \quad (40)$$

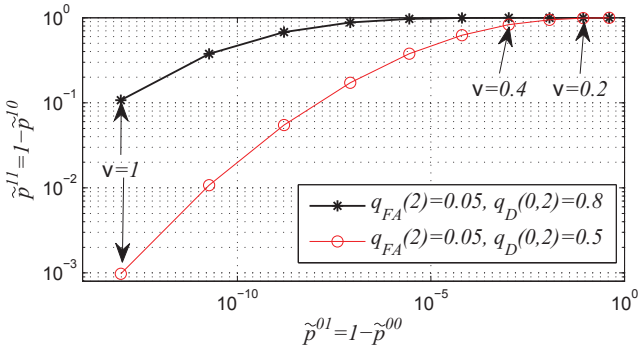


Fig. 3. Approximate $\overline{p}^{11}$ as a function of approximate $\overline{p}^{01}$, for various ν and fixed $M = 10$.

VII. Approximations of the Equilibrium

Closed form expressions for $\overline{p}^{00}$ and $\overline{p}^{10}$ are difficult to obtain from (40). An approximation of (40) is thus proposed, to help choosing ν. Since $\overline{p}^{10}$ represents the proportion of nodes with defective sensors that have not detected their status, the value of $\overline{p}^{10}$ should be small. One has $\lim_{\overline{p}^{10}\to 0} \eta_0 = q_{\text{FA}}(2) / (1 - q_{\text{FA}}(2))$ and $\lim_{\overline{p}^{10}\to 0} \eta_1 = q_{\text{D}}(1,1) / (1 - q_{\text{D}}(1,1))$, thus one may consider the following approximations

$$\eta_0 \approx \widetilde{\eta}_0 = \frac{q_{\text{FA}}(2)}{1 - q_{\text{FA}}(2)}, \quad \eta_1 \approx \widetilde{\eta}_1 = \frac{q_{\text{D}}(1,1)}{1 - q_{\text{D}}(1,1)}. \quad (41)$$

Therefore, (40) may be rewritten as

$$\begin{cases} \widetilde{p}^{00} = \sum_{k:k/M<\nu} \binom{M}{k} \left(q_{\text{FA}}(2)\right)^k \left(1 - q_{\text{FA}}(2)\right)^{M-k}, \\ \widetilde{p}^{10} = \sum_{k:k/M<\nu} \binom{M}{k} \left(q_{\text{D}}(1,1)\right)^k \left(1 - q_{\text{D}}(1,1)\right)^{M-k}. \end{cases} \quad (42)$$

from which one deduces approximate values of $\widetilde{X}_0^{M,k}$ at equilibrium

$$\begin{cases} \widetilde{X}_0^{M,k} = \binom{M}{k} \left(q_{\text{FA}}(2)\right)^k \left(1 - q_{\text{FA}}(2)\right)^{M-k}, \\ \widetilde{X}_1^{M,k} = \binom{M}{k} \left(q_{\text{D}}(1,1)\right)^k \left(1 - q_{\text{D}}(1,1)\right)^{M-k}. \end{cases} \quad (43)$$

For any fixed M, $q_{\text{FA}}(2)$, and $q_{\text{D}}(1,1)$, the values of the *detection rate* (DR) and of the *false alarm rate* (FAR) at equilibrium can be predicted by using (42), where FAR $= \overline{p}^{01} = 1 - \overline{p}^{00}$ and DR $= \overline{p}^{11} = 1 - \overline{p}^{10}$. For a fixed $M = 10$, Figure 3 presents the approximate evolution of DR as a function of FAR for different values of ν. This figure provides guidelines to choose the value of ν to meet different performance requirements.

VIII. Numerical results

A. Numerical verification of theoretical results

This section presents first the solution of the state equation (26-28) describing the evolution of the proportion of nodes in various states. Algorithm 1 is verified considering a random displacement of nodes without constraint on their speed. Consider a LODT where $q_{\text{FA}}(0,2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, and $q_{\text{D}}(0,2) = 0.9$. Besides, $p_0 = 0.9$, $p_1 = 0.1$, $M = 4$, and $\nu = 0.4$. Figure 4 presents the evolution of the proportion of nodes with good sensors (top part) and defective sensors (bottom part) in different states, obtained by solving (26-28) using $\lambda = 1$. One observes that the proportion of each state becomes constant as $t/\Delta t \approx 15$, which means that the system reaches fast its equilibrium point. For the good nodes, only two states are such that $X_0^{\ell,k} > 0.05$, *i.e.*, $(0,4,0)$ and $(0,4,1)$, while the others are very close to 0. For the defective nodes, only $X_1^{4,4}$, $X_1^{4,3}$, and $X_1^{4,2}$ are relatively large. Observe that since there does not exist a common (ℓ,k) such that both $X_0^{\ell,k}$ and $X_1^{\ell,k}$ have unnegligeble values, the accuracy of the algorithm is good. With $\nu = 0.4$, one has $\overline{p}^{00} = 0.985$ and $\overline{p}^{10} = 0.027$, which means that FAR is 1.5% and the *non-detection rate* (NDR) is NDR $= 1 -$ DR $= \overline{p}^{10} = 2.7\%$.

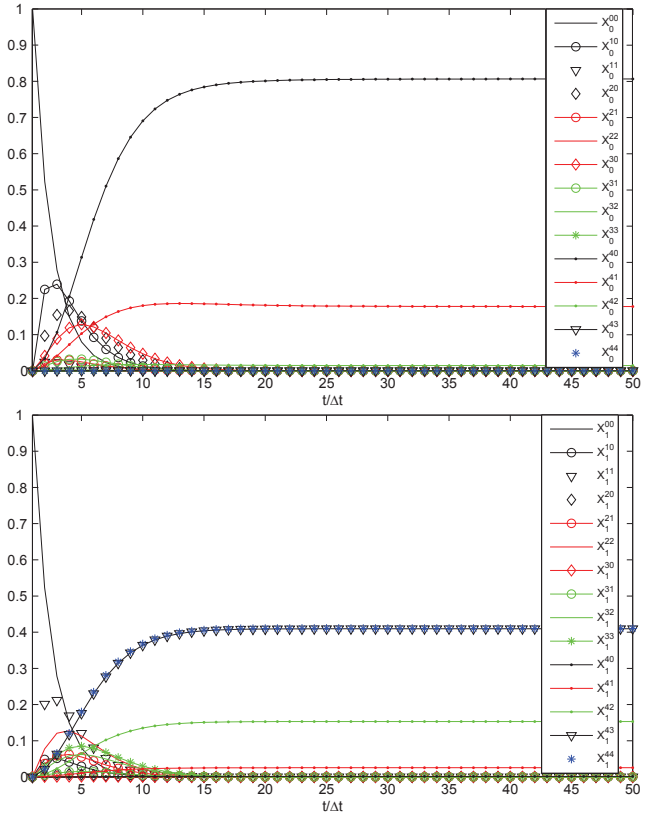


Fig. 4. Evolution of $X_\theta^{\ell,k}(t)$ obtained from (26-28), when $q_{\text{FA}}(0,2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, $q_{\text{D}}(0,2) = 0.9$, $M = 4$, $\nu = 0.4$, and $\lambda = 1$.

Consider now a set $\mathcal{S}$ of $N_S = 1000$ moving nodes uniformly distributed over a square of unit area. One considers

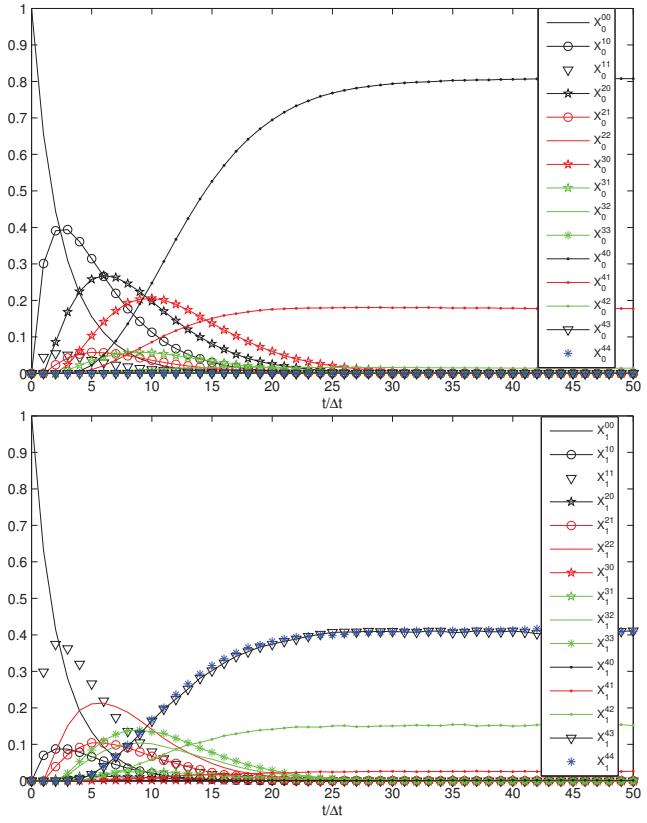


Fig. 5. Evolution of $X_\theta^{\ell,k}(t)$ by simulations with the ideal displacement model, when $q_{\text{FA}}(0,2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, $q_{\text{D}}(0,2) = 0.9$, $M = 4$, $\nu = 0.4$, and $\lambda\Delta t = 0.33$.

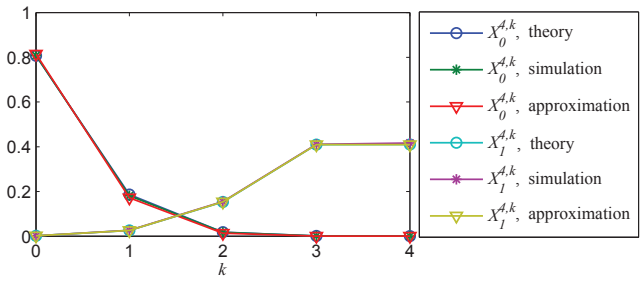


Fig. 6. Comparison of $\overline{X}_\theta^{4,k}$ at the equilibrium.

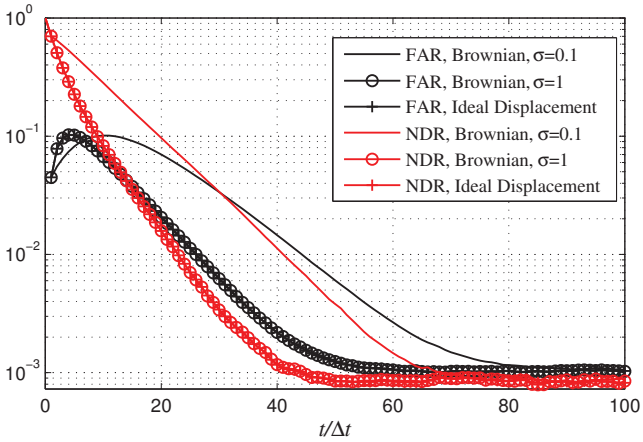


Fig. 7. Evolution of FAR and NDR for the two moving models, with $\sigma \in \{0.1, 1\}$, $q_{\text{FA}}(2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, $q_{\text{D}}(0,2) = 0.9$, $M = 10$ and $\nu = 0.4$.

a motion model: Node i randomly chooses its location at time instant $(k+1)\Delta t$, independently from its previous location at time $k\Delta t$. Two nodes communicate only at discrete time $k\Delta t$ when their distance is less than r_0. Node i has its neighbors in the set $\mathcal{N}_i = \{j \in \mathcal{S} : 0 < R_{i,j} \leqslant r_0\}$, where $R_{i,j}$ is the distance between i and j. Then, we assume that Node i can contact its neighbor if and only if $|\mathcal{N}_i| = 1$. Denote $\rho = \pi r_0^2 N_S$ as the average value of $|\mathcal{N}_i|$. The cardinality of $\mathcal{N}_i$ approximately follows a Poisson distribution as N_S is large enough, the inter-contact probability during Δt is thus

$$\lambda\Delta t = \mathbb{P}\{|\mathcal{N}_i| = 1\} = \rho \exp(-\rho).$$

In the Monte-Carlo simulations, we used $r_0 = 0.014$, so that $\rho \approx 0.6$ and $\lambda\Delta t \approx 0.33$. The simulation results are shown in Figure 5. Comparing Figure 4 and Figure 5, the state evolution in the transient phase has similar shape but with different convergence speed, which depends mainly on λ. Figure 6 shows a good match between theory and simulation for the proportions of states at equilibrium. The approximation of $\overline{X}_\theta^{4,k}$ using (43) is also presented in Figure 6, which is very close to its actual value.

B. Simulation with Brownian motion model

Consider now a Brownian motion model: each node is moving with a random speed, and it simply changes its orientation when it reaches the boundary of the unit square. Define $\boldsymbol{O}_i = \left(o_x^i, o_y^i\right)$ as the location of Node i. Consider a second order mobility model, *i.e.*,$d^2o_x/dt^2 = v_x$ and d^2o_x/dt^2 where $v_x, v_y \sim \mathcal{N}\left(0, (\sigma r_0)^2\right)$.

Consider $\sigma \in \{0.1, 1\}$, $M = 10$, $q_{\text{FA}}(2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, and $q_{\text{D}}(0,2) = 0.9$. Figure 7 presents the evolution of FAR and NDR as functions of time for both models, with the fixed $\rho \approx 0.6$. At the equilibrium, the performance obtained in the two moving models is quite close. However, the convergence speed depend on the inter-contact rate λ. When $\sigma = 0.1$, the algorithm converges slowly in the Brownian model. When $\sigma = 1$, which results to a larger value of λ, the evolution of FAR and NDR with the Brownian model are close to the ideal displacement model.

Now we characterize the impact of the parameters p_1 and M on the performance of the algorithm. To this purpose we assume $q_{\text{FA}}(2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, $q_{\text{D}}(0,2) = 0.9$ and consider the Brownian model with $\sigma = 0.1$. The evolution of NDR (or $\overline{p}^{10}$) and FAR (or $\overline{p}^{01}$) in case of various $p_1 \in \{0.1, 0.5\}$ and $M \in \{4, 10, 20\}$ are shown in Figure 8. Observe that a larger M leads to a better performance but requires longer convergence time. The proportion of the nodes with defective sensors, *i.e.*, p_1, also has an impact on the convergence speed of the algorithm since a larger p_1 implies longer time to achieve the same performance.

Consider now a fixed $q_{\text{D}}(0,2) = 0.9$ and $M = 10$. For

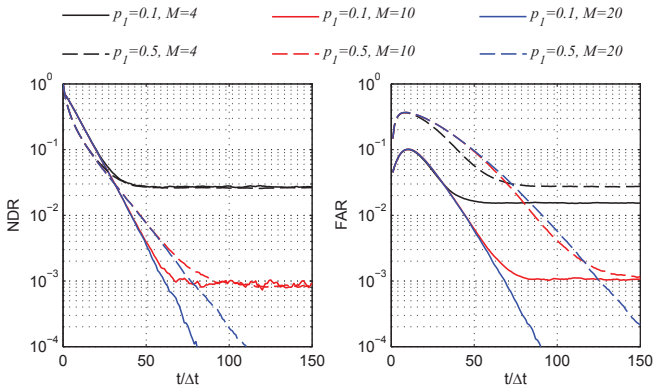


Fig. 8. Evolution of NDR and FAR for various $M \in \{4, 10, 20\}$ and $p_1 \in \{0.1, 0.5\}$, with $q_{\text{FA}}(2) = 0.05$, $q_{\text{D}}(1,1) = 0.8$, $q_{\text{D}}(0,2) = 0.9$.

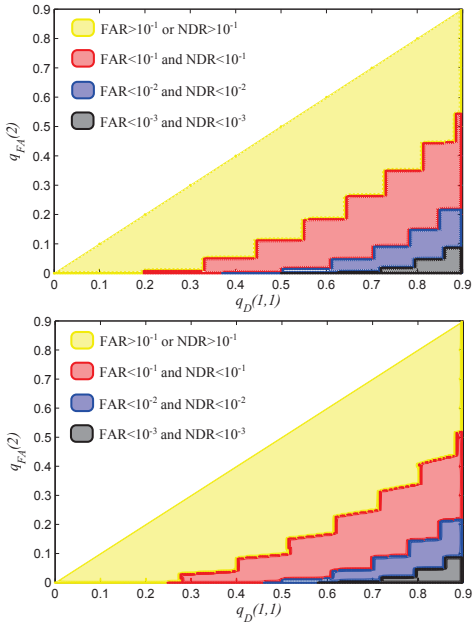


Fig. 9. FAR and NDR that can be obtained for different values of the pair $(q_{\text{D}}(1,1), q_{\text{FA}}(2))$ when $p_1 = 0.1$ (top) and for $p_1 = 0.5$ (bottom).

$p_1 \in \{0.1, 0.5\}$, we investigate the achievable NDR and FAR for $0 \leqslant q_{\text{FA}}(2) < q_{\text{D}}(0,2)$ and $q_{\text{FA}}(2) < q_{\text{D}}(1,1) \leqslant q_{\text{D}}(0,2)$. The performance level is obtained solving (42) for the best value of ν. Four performance levels were considered: i) Level 0: either NDR or FAR cannot be less than 10^{-1}; ii) Level 1: both NDR and FAR can be less than 10^{-1}; iii) Level 2: both NDR and FAR can be less than 10^{-2}; iv) Level 3: both NDR and FAR can be less than 10^{-3}. Figure 9 shows the performance levels that can be reached for different values of the pair $(q_{\text{D}}(1,1), q_{\text{FA}}(2))$. Observe that for large values of $q_{\text{D}}(1,1)$ and small values of $q_{\text{FA}}(2)$, good performance can be obtained. The value of p_1 does not have a significant impact on the performance at the equilibrium, even when $p_1 = 0.5$.

IX. Conclusions

This paper presents a fully distributed algorithm allowing each node of a DTN to estimate the status of its own sensors using LODT performed during the meeting of nodes. The DFD algorithm is analyzed considering a Markov model of the evolution of the proportion of nodes with a given belief in their status. Interestingly, the proportions at the equilibrium follow approximately a binomial distribution. Approximations of the various proportions of nodes at equilibrium provide a guideline to properly choose the decision parameter ν of the DFD algorithm. The simulations show a good match with the theoretical part. The convergence speed of the DFD algorithm depends on the inter-contact rate and on the proportion of nodes with defective sensors. The latter has no have significant effect the NDR and FAR at equilibrium. Further research will be dedicated to an adaptation of ν with time to increase the convergence speed of the proposed DFD algorithm. This may be particularly important in variants of the considered problem, such as malware detection.

References

[1] M. J. Khabbaz, C. M. Assi, and W. F. Fawaz, "Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 607–640, 2012.

[2] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay-tolerant networks," *IEEE Trans. on Mobile Computing*, vol. 10, no. 11, pp. 1576–1589, 2011.

[3] V. N. Soares, J. J. Rodrigues, and F. Farahmand, "Geospray: A geographic routing protocol for vehicular delay-tolerant networks," *Information Fusion*, vol. 15, pp. 102–113, 2014.

[4] H. Zhu, S. Du, Z. Gao, M. Dong, and Z. Cao, "A probabilistic misbehavior detection scheme toward efficient trust establishment in delay-tolerant networks," *IEEE Trans. on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 22–32, 2014.

[5] L. Galluccio, B. Lorenzo, and S. Glisic, "Sociality-aided new adaptive infection recovery schemes for multicast dtns," *IEEE Trans. on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[6] W. Li, F. Bassi, D. Dardari, M. Kieffer, and G. Pasolini", "Defective sensor identification for WSNs involving generic local outlier detection tests," *IEEE Trans. on Signal and Information Processing over Networks*, vol. 2, no. 1, pp. 29–48, 2016.

[7] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[8] A. Mahapatro and P. M. Khilar, "Fault diagnosis in wireless sensor networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2000–2026, 2013.

[9] H. Dong, Z. Wang, S. X. Ding, and H. Gao, "A survey on distributed filtering and fault detection for sensor networks," *Mathematical Problems in Engineering*, 2014.

[10] J. Chen, S. Kher, and A. Somani, "Distributed fault detection of wireless sensor networks," in *Proc Workshop DIWANS*, New York, NY, 2006, pp. 65 – 72.

[11] J.-L. Gao, Y.-J. Xu, and X.-W. Li, "Weighted-median based distributed fault detection for wireless sensor networks," *Journal of Software*, vol. 18, no. 5, pp. 1208 – 1217, 2007.

[12] S. Ji, S.-F. Yuan, T.-H. Ma, and C. Tan, "Distributed fault detection for wireless sensor based on weighted average," in *Proc NSWCTC*, Wuhan, China, 2010, pp. 57 – 60.

[13] M. Panda and P. Khilar, "Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test," *Ad Hoc Networks*, vol. 25, pp. 170–184, 2015.

[14] B. Zhu, W. Zhang, W. Feng, and L. Zhang, "Distributed faulty node detection and isolation in delay-tolerant vehicular sensor networks," in *Proc. PIMRC*, Sept 2012, pp. 1497–1502.

[15] W. Peng, F. Li, X. Zou, and J. Wu, "Behavioral malware detection in delay tolerant networks," *IEEE Trans. on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 53–63, 2014.

[16] H. Zhu, L. Fu, G. Xue, Y. Zhu, M. Li, and L. Ni, "Recognizing exponential inter-contact time in vanets," in *Proc. INFOCOM*, March 2010, pp. 1–5.