Research Article

Detection of Temporal Anomalies for Partially Observed Timed PNs

Dimitri Lefebvre

Normandie Université, UNIHAVRE, GREAH, 76600 Le Havre, France

Correspondence should be addressed to Dimitri Lefebvre; dimitri.lefebvre@univ-lehavre.fr

Received 13 October 2016; Revised 1 March 2017; Accepted 16 March 2017; Published 12 April 2017

1. Introduction

The prevention of faults is a critical issue in numerous systems to preserve the safety of both equipment and human operators. These issues have been addressed in numerous studies with fault detection and diagnosis (FDD) methods. The aim of fault detection is to create an alarm each time a fault occurs, and the aim of diagnosis is to isolate the fault within a group of candidates [1]. In the domain of discrete event systems (DESs), FDD has been often formulated with automata, Petri nets (PNs) [2], in particular labeled PNs (LPNs) [3] or partially observed Petri nets (POPNs) [4]. The main reason for developing FDD tests with PN extensions is that such models include graphical representations that can be disseminated widely in numerous application domains. They also offer mathematical supports that are consistent with standard tools. The proposed methods are useful for a large variety of technological systems, ranging from computer or chemical engineering to manufacturing and intelligent transportation systems.

In numerous contributions, the faults that are considered are unexpected events that may occur in event sequences and that cannot be directly measured. Various approaches have been proposed with PN extensions to detect and isolate such unexpected events. These approaches are based either on the analysis of the PN reachability graph [5–9], on the direct properties of the PNs [10, 11], or on PN unfolding [12, 13]. A few results also concern the introduction of temporal information in the diagnosis process. At first, dates of events have been introduced in usual extensions of untimed PNs. Such dates lead to a more accurate estimation of the past and future fault occurrence probabilities [14] and are also useful to propose an evaluation of the unknown fault dates [15]. The design and identification of models that include temporal faults have been also considered [16, 17]. Then, fuzzy Petri nets have been used to model and check temporal constraints between event occurrences [18]. Partial orders with unfolding and (max, +)-linear inequalities have been used with timed PN models [19, 20]. Monotonic monitoring and stratification have been introduced, when the monitoring is fragmented because of the uncertain temporal observation [21]. Finally, indirect monitoring has been used by comparing the actual cycle periods with the expected one in order to detect faults [22].

This paper takes place in the context where both transitions and places are assumed to be partially observed and...
consider only temporal faults. For that purpose, temporal constraints are defined by tolerance intervals that are associated with the transitions and that represent the normal durations of the system operations. The aim of the diagnosis system is to generate alarms when the temporal constraints are no longer satisfied. For that purpose, timed POPNs (POTPNs) are introduced. POTPNs take into consideration some measurable events that correspond to dated and labeled transition firings and also to partial measurements of the marking vector that is dated. This formalism, fully described in [23], is useful to represent incomplete history of the marking vector that is dated. This formalism, fully detailed throughout the paper. Section 4 concludes the paper. In Section 3, the main results are detailed. In Section 2, temporal constraints and POTPNs are introduced. In Section 3, the main results are detailed. Examples are detailed throughout the paper. Section 4 concludes the paper.

2. Context and Notations

2.1. PNs with Temporal Specifications. A PN structure is defined as \( G = (P,T,W_{PP},W_{PC}) \), where \( P = \{p_1, \ldots, p_n\} \) is a set of \( n \) places and \( T = \{t_1, \ldots, t_q\} \) is a set of \( q \) transitions, \( W_{PC} \in (N)^{n \times q} \) and \( W_{PP} \in (N)^{n \times n} \) are the post- and preinci-dence matrices \( (N \) is the set of nonnegative integer numbers), and \( W = W_{PO} - W_{PR} \) is the incidence matrix. A PN is choice-free if \( |(P_i)'| \leq 1 \) (the postset of \( P_i \) contains at most a single transition). \((G,M)\) is a PN system with initial marking \( M \) and \( M \in (N)^n \) represents the PN marking vector. A PN system is 1-bounded if and only if (iff) \( M \leq 1_n \) where \( 1_n = (1 \cdots 1)^T \) (inequality \( M \leq 1_n \) is considered component wise). A transition \( T_j \) is enabled at marking \( M \) iff \( M \geq W_{PP}(\cdot, j) \), where \( W_{PP}(\cdot,j) \) is the column \( j \) of preincidence matrix; this is denoted as \( M(T_j) \). When \( T_j \) is enabled, it may fire, and when \( T_j \) fires once, the marking varies according to \( \Delta M = M' - M = W(\cdot,j) \). This is denoted as \( M(T_j)M' \). A sequence of size \( H = |\sigma| \) fired at marking \( M \) is a sequence of \( H \) transitions \( \sigma = T(1)T(2)\cdots T(H) \), with \( T(j) \in T \), \( j = 1, \ldots, H \) that consecutively fire from marking \( M \) to marking \( M' \). This is denoted as \( M(\sigma)M' \). The integer \( x_j(\sigma) \) is the number of occurrences of transition \( T_j \) in \( \sigma \), and \( X(\sigma) = (x_j(\sigma)) \in (N)^q \) is the firing count vector for \( \sigma \). A sequence \( \sigma \) fired at \( M \) leads to an untimed trajectory \((\sigma,M)\) detailed in

\[
(\sigma, M) = M(0) [T(1)] M(1) \cdots [T(H)] M(H),
\]

with \( M(0) = M \). A marking \( M \) is said to be reachable from initial marking \( M_j \) if there exists a firing sequence \( \sigma \) such that \( M_j(\sigma)M \). The set of all reachable markings from initial marking \( M_j \) is \( R(G,M_j) \).

Timed Petri nets are PNs whose behavior is driven by time. Time is measured with time units (TU). The time can be associated with the firing of the transitions or with the sojourn of the tokens in the places. In this paper the time is associated with the transitions and the firing of each transition \( T \) occurs after a firing duration \( d \) that can eventually be zero. In that case the firing is immediate; in the other cases it is delayed. In this last case, the duration \( d \) can be deterministic \((d \text{ is a constant})\) or stochastic \((d \text{ is a random variable (RV)})\) with a probability density function (pdf) \( f(d) \). In this article, stochastic durations are considered at first but the results are also applicable to deterministic durations. No particular assumption is made on the pdfs of the firing durations but the pdf of each transition is assumed to be known. The set of pdfs for all transitions is referred to as PDF. Two classes of pdfs are of particular interest for this work: bounded uniform (Figure 1(a)) and symmetrical triangular pdfs (Figure 1(b)) defined, respectively, with equations (2) and (3).

Bounded uniform pdf is as follows:

\[
f(d) = \begin{cases} 
1/(b-a) & \text{if } d \in [a,b], \\
0 & \text{otherwise} 
\end{cases}
\]

Symmetrical triangular pdf is as follows:

\[
f(d) = \frac{4}{(b-a)^2} \cdot (d-a) \quad \text{if } d \in \left[a, \frac{(a+b)}{2}\right],
\]
where $\sigma = T(j_1, t_1)T(j_2, t_2)\cdots T(j_H, t_H)$ where $j_1, \ldots, j_H$ are the labels of the transitions and $t_1, \ldots, t_H$ represent the dates of the firings that satisfy $t_0 \leq t_1 \leq t_2 \cdots \leq t_H \leq \tau_{\text{end}}$. This leads to the timed trajectory $(\sigma, M)$ detailed in the following with $M(\tau_0) = M$:

\[
(\sigma, M)
= M(\tau_0) \left[ T(t_1) \right] M(t_1) \cdots \left[ T(t_H) \right] M(\tau_{\text{end}}).
\]

Note that we refer to timed and untimed firing sequences with the same notation $\sigma$ as long as the notation is not confusing; otherwise we use $\sigma_{j, \tau}$ to refer to untimed firing sequence and $\sigma$ to refer to timed ones.

### 2.2. Partially Observed Timed Petri Nets

Partially observed Petri nets are considered to represent the system sensor. $L : T \rightarrow E \cup \{\varepsilon\}$ is a labeling function that assigns a label to each transition where $E = \{e_1, \ldots, e_p\}$ is the set of labels that are assigned to observable transitions and $\varepsilon$ is the null label that is assigned to the silent ones. The concatenation of labels obviously satisfies the following: $\varepsilon \varepsilon = \varepsilon$ and $e_k \varepsilon = e_k$.

For simplicity, each label $e_k$ is represented by the elementary vector $e_k$ of dimension $p$ such that $e_k = (e_{kj})$ with $e_{kj} = 1$ for $j \neq k$ and $e_{kk} = 1$. The null label is represented by the zero vector $\varepsilon = 0_p$ of dimension $p$. The labeling function is linear and defined by the matrix $L = (l_{kj}) \in (N)^{p \times q}$ such that $l_{kj} = 1$ if $L \cdot X(T_j) = e_k$; otherwise $l_{kj} = 0$ ("*" stands for the product operator). The marking sensor matrix $H \in (R)^{n \times n}$ ($R$ is the set of real numbers) defines the projection of the marking vector $M$ over $n_0$ subsets of places. The observable part of the marking is denoted as $M_0 = H \cdot M$.

Thus, partially observed timed PNs with temporal constraints (POTPN) are defined as $(G, L, H, PDF, TC, M_0)$, where PDF is the set of pdfs, TC is the set of temporal constraints, $G$ is a Petri net structure, $L$ is an event sensor matrix, and $H$ is a marking sensor matrix. The matrices $L$ and $H$ define the sensor configuration.

Measurements are collected over the time interval $[\tau_0, \tau_{\text{end}}]$ When the POTPN marking varies with the firing of a single transition $T$ at date $t \in [\tau_0, \tau_{\text{end}}]$, the measurement function $\Gamma(T, \tau_0, \tau_{\text{end}})$ is defined by

\[
\Gamma(T, \tau_0, \tau_{\text{end}}) = \left( H \cdot M(\tau_0) \right) \left[ L \cdot X(T(t)) \right] \left( H \cdot M(\tau_{\text{end}}) \right)
\]

if 

\[
(H \cdot M(\tau_0) \neq H \cdot M(\tau_{\text{end}})) \lor (L \cdot X(T(t)) \neq \varepsilon),
\]

(5)

\[
\Gamma(T, \tau_0, \tau_{\text{end}}) = (H \cdot M(\tau_{\text{end}}))
\]

if 

\[
(H \cdot M(\tau_0) = H \cdot M(\tau_{\text{end}})) \land (L \cdot X(T(t)) = \varepsilon).
\]
Roughly speaking, the measurement function $\Gamma$ collects a new label each time a transition fires that is not silent or that changes the measurement of the marking. The measurement function $\Gamma$ is then extended to timed trajectories of the form (4) measured over the time interval $[\tau_0, \tau_{\text{end}}]$:

$$
\Gamma((\sigma T, M), \tau_0, \tau_{\text{end}}) = \Gamma((\sigma, M), \tau_0, \tau_{\text{end}}) L \cdot X(T(t)) (H \cdot M(\tau_{\text{end}}))
$$

if $(H \cdot M(t) \neq H \cdot M(\tau_{\text{end}})) \lor (L \cdot X(T(t)) \neq \varepsilon)$,

$$
\Gamma((\sigma, M), \tau_0, \tau_{\text{end}}) = M_{O}(\tau_0) e_O(\tau_1) M_{O}(\tau_1) e_O(\tau_2) \cdots M_{O}(\tau_{K-1}) e_O(\tau_K) M_{O}(\tau_{\text{end}}),
$$

where $M_{O}(\tau_0) = M$, $K$ is the length of the sequence that satisfies $K \leq h$, and $\tau = \{\tau_k, k = 1,\ldots,K\}$ refers to the set of measurement dates. Note that $M_{O}(\tau_0)$ does not necessarily correspond to the measurement of initial marking $M_I$. A timed trajectory $(\sigma, M)$ is said to be consistent with a given timed measured trajectory $TR_O$ in time interval $[\tau_0, \tau_{\text{end}}]$ if it satisfies $\Gamma((\sigma, M), \tau_0, \tau_{\text{end}}) = TR_O$. In the next, it is assumed that the time interval starts at time 0 (i.e., $\tau_0 = 0$) and that it ends at the last measurement date (i.e., $\tau_{\text{end}} = \tau_K$) (Assumption B).

The objective of the present work is to estimate the probability that any given timed measured trajectory satisfies the temporal constraints. An immediate application of the proposed estimation is to provide an algorithm that generates alarms when this probability goes down a specific threshold $\gamma$. For the best of our knowledge, it is the first time this problem is considered with PNs. Note that POTPNs cannot be encoded as a Hidden Markov Model (HMM) [26] because in a HMM each state successively reached by the system delivers an observation that is not certain and depends on the emission probabilities. On the contrary, in a POTPN model, the states and the transitions deliver certain but partial observations and in some cases the states do not deliver any observation at all.

3. Temporal Specifications Checking

The proposed diagnosis systems operate with three stages.

(i) For any timed measured trajectory $TR_O$ with $K$ measurements, the set of timed trajectories that are consistent with $TR_O$ are first computed with an integer linear programming approach developed in our previous work [23, 27].

(ii) For each possible trajectory, the probability that this trajectory is consistent with the temporal constraints is estimated.

(iii) The probability that a timed measured trajectory is consistent with the temporal constraints is obtained as a consequence by computing the probability of each consistent trajectory [15].

3.1. Untimed Trajectories Consistent with $TR_O$. In this section, the set $U^{-1}(TR_O)$ of all untimed trajectories $(\sigma_U, M)$ that are consistent with a given measured trajectory $TR_O$ is computed. Note that this problem cannot be solved using standard algorithms (as the Viterbi algorithm) [28] issued from dynamic programming because such algorithms aim to find only the trajectory of maximum probability from the measured trajectory, but not all trajectories. For diagnosis issues it is however required to consider all trajectories. It is assumed that the set of the reachable states and also the reachability graph of the net are known. Let us define $A_{\varepsilon}$ as the matrix of the reachability graph of the unobservable part of the PN where all transitions that are observable or whose firing changes the measured part of the marking have been removed and assume that this graph is acyclic (Assumption C). In that case, the maximal number of consecutive silent events is upper bounded by $h_{\text{max}} - 1$ [23] with

$$
h_{\text{max}} = \min\{h \geq 0 \text{ such that } (A_{\varepsilon})^h = 0\}. \tag{8}
$$

Let us consider a timed measured trajectory $TR_O$ of the form (7) with $K$ measurements in time interval $[0, \tau_{\text{end}}]$. An untimed trajectory $(\sigma_U, M)$ with $\sigma_U = T(1)T(2)\cdots T(H)$ is consistent with $TR_O$ iff the following conditions are satisfied [27]:

(1) $H \cdot M = M_O(0)$;

(2) There exists $h_0 = 0, h_1, \ldots, h_K$ such that $h_k - h_{k-1} \leq h_{\text{max}}$, $k = 1,\ldots,K$, and the untimed firing sequence $\sigma_U$ is rewritten as $\sigma_U = T(1)\cdots T(h_1)\cdots T(h_K)$ and satisfies the following:
\[
\left( -I_q \ 0 \ \cdots \ 0 \ 0 \\
0 \ -I_q \ \cdots \ 0 \\
\vdots \ \ddots \ \cdots \ 0 \\
0 \ \cdots \ 0 \ -I_q \\
(1_q)^T \ 0 \ \cdots \ 0 \\
0 \ (1_q)^T \ \cdots \ 0 \\
\vdots \ \ddots \ \cdots \ 0 \\
0 \ \cdots \ 0 \ (1_q)^T \\
W_{PR} \ 0 \ \cdots \ 0 \ -I_n \\
-W \ W_{PR} \ \cdots \ -I_n \\
\vdots \ \ddots \ \cdots \ 0 \\
-W \ \cdots \ -W \ W_{PR} \ -I_n \\
\right) \\
\left( X(T(1)) \\
\vdots \\
X(T(h_1)) \\
X(T(h_1 + 1)) \\
\vdots \\
X(T(h_2)) \\
\vdots \\
X(T(h_K)) \\
M \\
\vdots \\
0_n \\
\right) \leq \\
\left( 0_q \\
\vdots \\
0_q \\
1 \\
\vdots \\
1 \\
0_n \\
\vdots \\
0_n \\
\right) \\
(9)
\]

where all inequalities are taken component-wise. Roughly speaking inequality (9) means that the firing count vector of each transition in \( \sigma \) is positive, unitary, and feasible (i.e., leading to a positive marking). Equality (10) means that for \( k = 1, \ldots, K \), \( h_k - 1 \) first transitions are silent and only the last one may provide a label \( e_O(\tau_k) \). Similarly it ensures that \( h_k - 1 \) first marking measurement does not provide any information and only the last one may provide marking changes \( M_O(\tau_k) - M_O(\tau_{k-1}) \). The combined use of (9) and (10) leads to the exhaustive set of untimed trajectories that are consistent with \( TR_O \) [23, 27]. Note that \( \Gamma^1_U(\text{TR}_O) \) does not include the silent closure of the trajectories (i.e., the continuations of the trajectories that provide no event nor marking measurement) because the time interval \([0, \tau_K]\) ends with the last measurement (Assumption B) and the no immediate firing occurs (Assumption A5). If required, the silent closure can easily be added to \( \Gamma^1_U(\text{TR}_O) \) by considering the following equation in addition to (9) and (10):

\[
\left( L \ 0 \ \cdots \ 0 \\
0 \ \cdots \ 0 \\
0 \ \cdots \ L \\
H \cdot W \ 0 \ \cdots \ 0 \\
0 \ \cdots \ 0 \\
0 \ \cdots \ 0 \\
H \cdot W \ \cdots \ H \cdot W \ H \cdot W \\
\right) \\
\left( X(T(h_{k-1} + 1)) \\
\vdots \\
X(T(h_k - 1)) \\
X(T(h_k)) \\
\right) \\
= \\
\left( \varepsilon \\
\vdots \\
\varepsilon \\
e_O(\tau_k) \\
0_n \\
\vdots \\
0_n \\
M_O(\tau_k) - M_O(\tau_{k-1}) \\
\right) \\
(10)
\]

for \( k = 1, \ldots, K \).

(9) and (10) leads to the exhaustive set of untimed trajectories that are consistent with \( TR_O \) [23, 27]. Note that \( \Gamma^1_U(\text{TR}_O) \) does not include the silent closure of the trajectories (i.e., the continuations of the trajectories that provide no event nor marking measurement) because the time interval \([0, \tau_K]\) ends with the last measurement (Assumption B) and the no immediate firing occurs (Assumption A5). If required, the silent closure can easily be added to \( \Gamma^1_U(\text{TR}_O) \) by considering the following equation in addition to (9) and (10):

\[
\left( L \ 0 \ 0 \\
0 \ \cdots \ 0 \\
0 \ \cdots \ L \\
H \cdot W \ 0 \ \cdots \ 0 \\
0 \ \cdots \ 0 \\
0 \ \cdots \ 0 \\
H \cdot W \ \cdots \ H \cdot W \ H \cdot W \\
\right) \\
\left( X(T(h_K + 1)) \\
\vdots \\
X(T(h_{K+1})) \\
\right) \\
= \\
\left( \varepsilon \\
\vdots \\
\varepsilon \\
0_n \\
\vdots \\
0_n \\
\right) \\
(11)
\]

where \( T(h_{K+1}) \cdots T(h_{K+1}) \) stands for the silent closure.
Let us consider the marked POTPNI of Figure 2 with 

\[ M_I = (1 \ 0 \ 0)^T, \] 
a single observable transition \( T_2, \) and no observable place (unobservable places and transitions are highlighted in grey). The set of labels is \( E = \{e\} \) and the matrices \( L = (0 \ 1) \) and \( H = (0 \ 0 \ 0) \) define the sensor configuration. Measurements are collected over the time interval \( [0, \tau] \). Assume that the measured trajectory \( TR_0 = 0(0)\ 0(\tau) \) is measured according to \( L \) and \( H \). Note at first that the given example satisfies Assumptions A to C and that (8) leads to \( h_{\text{max}} = 2 \). Thus, untimed trajectories \( (\sigma_U, M) \) are searched with \( M \in \{(1 \ 0 \ 0)^T; (0 \ 1 \ 0)^T; (0 \ 0 \ 1)^T\} \) and \( \sigma_t = T(1)T(2) = T(2) \). Two untimed trajectories \( (T_1T_2, (1 \ 0 \ 0)^T) \) and \( (T_2, (0 \ 1 \ 0)^T) \) are consistent with \( TR_0 \) in this particular example. If we assume that the first measured marking is \( M_I \), then the single trajectory \( (T_1T_2, (1 \ 0 \ 0)^T) \) remains consistent with \( TR_0 \). 

3.2. Probability of a Timed Trajectory with a Set of Given Firing Dates. A timed trajectory \( (\sigma, M) \) with \( \sigma = T(t_1) \cdots T(t_2) \cdots T(t_H) \) is consistent with \( TR_0 \) iff the corresponding untimed trajectory \( (\sigma_U, M) \) obtained by making abstraction of time satisfies the previous conditions (conditions 1 and 2 in Section 3.1) and if the date \( t_h, h = 1, \ldots, H \) satisfies the following conditions [23, 27]:

\[ \text{(3)} \ t_1 \leq t_2 \leq \cdots \leq t_H \ \text{(i.e., the chronological order of the events results from} \sigma_U). \]

\[ \text{(4)} \ \text{There exists } K \text{ dates } t_{h_1}, \ldots, t_{h_K} \text{ such that } t_{h_c} = \tau_c. \]

\[ \text{(5)} \ \text{The probability that each transition } T(t_{h_i}) \in \sigma, \tau_h \not\in \{t_{h_i}, \ldots, t_{h_K}\} \text{ fires within a small interval } [t_{h_i}, t_{h_i} + dt] \text{ of width } dt \text{ is nonzero.} \]

For each transition \( T(t_{h_i}) \in \sigma \) that fires at date \( t_{h_i} \), let us consider the firing duration \( d_t \) of transition \( T(t_{h_i}) \) and introduce \( t_{h_i} \) and \( M(t_{h_i}) \) as the date and marking from which transition \( T(t_{h_i}) \) remains enabled. Thus \( T(t_{h_i}) \) is enabled from date \( t_{h_i} \) and fires at date \( t_{h_i} \).

\[ (\sigma, M) = M(0) \left[ T(t_1) \right] M(t_1) \cdots \left[ T(t_H) \right] M(t_H) \cdots \left[ T(t_H) \right] M(t_H), \]

\[ \text{-----------------------------} \]

\[ d_t = t_{h_i} - t_{h_i}. \]

3.3. Probability That a Timed Trajectory Satisfies the Temporal Constraints. The probability that any timed trajectory \( (\sigma, M) \), obtained from untimed trajectory \( (\sigma_U, M) \) and consistent with \( TR_0 \), satisfies the set of temporal constraints \( TC \) results from the extension and integration of (13) with respect to the 3 types of situations described in Table 1. This leads to the following:

\[ \text{prob} ((\sigma, M) \text{ satisfies TC} | (\sigma_U, M) \text{ and } \tau) \]

\[ = \frac{S_1(\sigma_U, M, TC, \tau)}{S_2(\sigma_U, M, \tau)} \]

with

\[ S_1(\sigma_U, M, TC, \tau) = \int_0^{\tau} g_1(t_1 - \tau_0) \cdots \left( \int_0^{\tau} g_h(t_h - t_{h_i}) \cdots \left( \int_0^{\tau} g_H(t_H - t_{H_i}) \cdots d_{t_k} \right) \cdots d_{t_i} \right) \cdots d_{t_1}, \]

\[ S_2(\sigma_U, M, \tau) = \int_0^{\tau} g_1(t_1 - \tau_0) \cdots \left( \int_0^{\tau} g_h(t_h - t_{h_i}) \cdots \left( \int_0^{\tau} g_H(t_H - t_{H_i}) \cdots d_{t_k} \right) \cdots d_{t_i} \right) \cdots d_{t_1}, \]
Table 1: Deterministic and random variables introduced to evaluate \( \text{prob}(\{ t_{\mathcal{U}}, h \in [1: \mathcal{H}] \} | (\sigma_{\mathcal{U}}, M) \) and \( \tau )

<table>
<thead>
<tr>
<th>( t_h / \tau_h )</th>
<th>Measured: ( t_h = \tau_h )</th>
<th>Unmeasured: ( t_h ) is a RV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured: ( t_h = \tau_h )</td>
<td>Type 1: ( d_h ) is a deterministic variable ( d_h = \tau_h - t_h ) measured</td>
<td>Type 3: ( d_h ) is a new RV ( d_h = t_h - \tau_h )</td>
</tr>
<tr>
<td>Unmeasured: ( t_h ) is a RV</td>
<td>Type 2: ( d_h ) is a RV that depends on a RV of type 3 ( d_h' ) ( d_h = \tau_h - t_h )</td>
<td>Type 3: ( d_h ) is a new RV ( d_h = t_h - \tau_h )</td>
</tr>
</tbody>
</table>

Table 2: Bounds and functions for (15).

<table>
<thead>
<tr>
<th>( d_h )</th>
<th>Type 2: ( t_h ) is a measured date</th>
<th>Type 3: ( t_h ) is an unknown date</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g_h(t_h) )</td>
<td>( f_h(t_h' - t_h) ) ( \Gamma_{\text{Class}(h')}(f_h(t_h - t_h')) )</td>
<td>( f_h(t_h - \tau_h) )</td>
</tr>
<tr>
<td>( b_h )</td>
<td>( \max(t_{h-1}, t_h + \delta_h', \max \Gamma_{\text{Class}(h')}(f_h(t_h - \Delta_h'))) ) *</td>
<td>( \max(t_{h-1}, t_h' + \delta_h) )</td>
</tr>
<tr>
<td>( B_h )</td>
<td>( \min(\tau_h, t_h' + \Delta_h', \min \Gamma_{\text{Class}(h')}(f_h(t_h - \delta_h'))) ) *</td>
<td>( \min(\tau_h, t_h' + \Delta_h) )</td>
</tr>
</tbody>
</table>

where dates \( b_h \) and \( B_h \) and functions \( g_h \) are defined in Table 2 for \( h \in [1: \mathcal{H}] \). Note that if \( d_h \) is of type 1 (i.e., \( d_h \) is a deterministic variable), then \( \text{prob}(t_h | (\sigma_{\mathcal{U}}, M) \) and \( \tau ) = 1 \) if \( d_h \in [\delta_h, \Delta_h] \); otherwise \( \text{prob}(t_h | (\sigma_{\mathcal{U}}, M) \) and \( \tau ) = 0 \). Thus situations of type 1 are no longer considered and (15) does not necessarily include all variables \( d_h \). To simplify the notation let us also divide the transitions of type 2 in several classes referred to as \( \text{Class}(h') \) formally defined by the following:

\[
\text{Class}(h') = \{ h \text{ such that } T(t_h) \\
\text{is of type 2 and } T(t_h) \\
\text{remains enabled from } M(t_h') \}
\]

In other words, \( \text{Class}(h') \) is the set of transitions of type 2 (their firing date is measured) that are enabled at the same date \( t_h' \).

Roughly speaking, \( S_1(\sigma_{\mathcal{U}}, M, \mathcal{T}, \tau) \) is a multi-integral that evaluates the sum of the duration variables \( d_h \) over their possible range of variation constrained by the measurement time interval \( [0, \tau_c] \), the dates of measurements \( \tau_c \), and the temporal specifications \( [\delta_h, \Delta_h] \). Similarly, \( S_2(\sigma_{\mathcal{U}}, M, \tau) \) evaluates the sum of the duration variables \( d_h \) over their possible range of variation constrained only by the measurement time interval \( [0, \tau_c] \) and the dates of measurements \( \tau_c \) (the temporal specifications \( [\delta_h, \Delta_h] \) are not considered). The ratio of both evaluations leads to the probability that any timed trajectory \((\sigma, M)\), resulting from untimed trajectory \((\sigma_{\mathcal{U}}, M)\) and consistent with \( \tau \), satisfies the set of temporal constraints \( \mathcal{T} \). From a numerical point of view the calculation of \( S_1(\sigma_{\mathcal{U}}, M, \mathcal{T}, \tau) \) and \( S_2(\sigma_{\mathcal{U}}, M, \tau) \) is obtained with a recursive algorithm.

3.4. Detection of the Temporal Faults. In the general case, several untimed trajectories \((\sigma_{\mathcal{U}}, M)\) may be consistent with a given timed measured trajectory \( \mathcal{T}_\mathcal{C} \) in time interval \( [0, \tau_c] \) (i.e., \( \Gamma_{\mathcal{C}}(\mathcal{T}_\mathcal{C}) \) contains more than one trajectory). This situation is due to two different reasons: (1) several markings \( M \) may be consistent with the first marking measurement \( M_0(0) \); (2) from a given marking \( M \), several untimed sequences may be consistent with the measured trajectory \( \mathcal{T}_\mathcal{C} \). In order to deal with (1), let us consider \( M_0(\mathcal{T}_\mathcal{C}) \) as the set of markings consistent with \( M_0(0) \) and \( \pi_0(M) \) as the probability that \( M \) is the current marking at date 0 such that \( \sum_{M \in M_0(\mathcal{T}_\mathcal{C})} \pi_0(M) = 1 \). The set \( M_0(\mathcal{T}_\mathcal{C}) \) and the probability \( \pi_0(M) \) for each \( M \in M_0(\mathcal{T}_\mathcal{C}) \) are assumed to be known (Assumption D). Note that Assumption D can be relaxed for PNs without absorbing subsets of markings by considering the steady state probability of \( M \) as \( \pi_0(M) \). In order to deal with (2), the probability of each sequence issued from the same marking \( M \) and consistent with \( \mathcal{T}_\mathcal{C} \) is evaluated with the following:

\[
\text{prob}(\{ (\sigma_{\mathcal{U}}, M) | \mathcal{T}_\mathcal{C} \}
\]

\[
= \frac{S_2(\sigma_{\mathcal{U}}, M, \tau)}{\sum_{(\sigma_{\mathcal{U}'}, M) \in \Gamma_{\mathcal{C}}(\mathcal{T}_\mathcal{C})} S_2(\sigma_{\mathcal{U}'}, M, \tau)}.
\]

Finally the probability of each trajectory \((\sigma_{\mathcal{U}}, M) \) obtained with

\[
\text{prob}(\{ (\sigma_{\mathcal{U}}, M) | \mathcal{T}_\mathcal{C} \}
\]

\[
= \pi_0(M) \cdot \left( \frac{S_2(\sigma_{\mathcal{U}}, M, \tau)}{\sum_{(\sigma_{\mathcal{U}'}, M) \in \Gamma_{\mathcal{C}}(\mathcal{T}_\mathcal{C})} S_2(\sigma_{\mathcal{U}'}, M, \tau)} \right)
\]

and \( \text{prob}(\mathcal{T}_\mathcal{C} \) satisfies \( \mathcal{T} \) \) \( \text{prob}(\{ (\sigma_{\mathcal{U}}, M) | \mathcal{T}_\mathcal{C} \} \) \) \( \text{prob}(\sigma_{\mathcal{U}}, M \) satisfies \( \mathcal{T} \) \( \text{prob}(\sigma_{\mathcal{U}}, M \) and \( \tau) \). This last equation leads to

\[
\text{prob}(\mathcal{T}_\mathcal{C} \) satisfies \( \mathcal{T} \) \) \( \text{prob}(\{ (\sigma_{\mathcal{U}}, M) | \mathcal{T}_\mathcal{C} \} \) \) \( \text{prob}(\sigma_{\mathcal{U}}, M \) and \( \tau) \). This last equation leads to

\[
\text{prob}(\mathcal{T}_\mathcal{C}) \) satisfies \( \mathcal{T} \) \) \( \text{prob}(\{ (\sigma_{\mathcal{U}}, M) | \mathcal{T}_\mathcal{C} \} \) \) \( \text{prob}(\sigma_{\mathcal{U}}, M \) and \( \tau) \). This last equation leads to

\[
\text{prob}(\mathcal{T}_\mathcal{C} \) satisfies \( \mathcal{T} \) \) \( \text{prob}(\sigma_{\mathcal{U}}, M \) and \( \tau) \). This last equation leads to

\[
\text{prob}(\mathcal{T}_\mathcal{C} \) satisfies \( \mathcal{T} \) \) \( \text{prob}(\sigma_{\mathcal{U}}, M \) and \( \tau) \). This last equation leads to
new measurements that are collected during time interval 
\[ (k - 1) \cdot \Delta t, k \cdot \Delta t \]. Formally, if \( TR_O(k) \) refers to the measured trajectory collected during time interval \( [0, k \cdot \Delta t] \),
the probability \( \text{prob}(TR_O(k) \text{ satisfies } TC) \) will be compared to a given threshold \( \gamma \) and an alarm is generated each time \( \text{prob}(TR_O(k) \text{ satisfies } TC) < \gamma \).

Let us consider again POTPN1 of Figure 2 introduced in Section 3.1. If the first measured marking is assumed to coincide with \( M_1 \) then \( M_0(TR_O) = \{(1 0 0)^T, (0 1 0)^T \} \) with \( \pi_0(1 0 0)^T = 1 \) and \( \pi_0(0 1 0)^T = 0 \). The timed trajectories \((\sigma, M)\) consistent with \( TR_O = 0(0) e(\tau_1) 0(\tau_1) \) satisfy also \( \sigma = T(t_1)T(t_2) \) with \( T(t_1) = T_1, T(t_2) = T_2, t_2 = \tau_1, \) and \( t_1 \leq \tau_1 \). Two successive cases are considered to illustrate the computation of \( \text{prob}(TR_O \text{ satisfies } TC) \) with (19).

In case A, the pdfs of the transition durations are assumed to be bounded uniform with the same support \([0, 10]\) for \( T_1 \) and \( T_2 \). The temporal constraints are arbitrarily defined by \( TC_1 = [4, 6] \) for \( T_1 \) and \( TC_2 = [3, 7] \) for \( T_2 \). Prob\( (TR_O \text{ satisfies } TC) \) computed with (19) is reported in Figure 3(a) (full line) in function of the date \( \tau_1 \). For the considered example, this equation can be rewritten as follows:

\[
\text{prob} (TR_O \text{ satisfies } TC) = \int_{\min(\Delta t, \tau_1 - \Delta t)}^{\max(\Delta t, \tau_1 - \Delta t)} f_1(t) \cdot f_2(\tau_1 - t) \cdot dt
\]

\[ = \int_{0}^{\tau_1} f_1(t) \cdot f_2(\tau_1 - t) \cdot dt \]  

(20)

Note that this probability is zero for \( \tau_1 < 7 \text{ TU} \) (one lower bound at least is not satisfied for the temporal constraints) and also for \( \tau_1 > 13 \text{ TU} \) (one upper bound at least is not satisfied for the temporal constraints). This computation is confirmed with a series of 1000 Monte Carlo simulations that coincide with \( TR_O \) (dashed line). Depending on the choice of the threshold \( \gamma \), an alarm may be generated. For example, for \( \gamma = 0.2 \), an alarm is generated if \( \tau_1 \in [8.5, 11.3] \).

In case B, the pdfs of the transition durations are assumed to be symmetrical triangular with supports [1, 4] for \( T_1 \) and [2, 7] for \( T_2 \). The temporal constraints are, respectively, \( TC_1 = [2, 3] \) for \( T_1 \) and \( TC_2 = [3, 6] \) for \( T_2 \). Prob\( (TR_O \text{ satisfies } TC) \) is computed with (19) in Figure 3(b) (full line) in function of the date \( \tau_1 \). This computation is also validated with a series of 1000 Monte Carlo simulations that coincide with \( TR_O \) (dashed line).

3.5. Numerical Complexity. The numerical complexity of the whole diagnosis schema is due (a) to the computation of \( \Gamma_U^{-1}(TR_O) \); (b) to the numerical evaluation of (19).

(a) The complexity to compute \( \Gamma_U^{-1}(TR_O) \) is related to the resolution of (9) and (10) that include \( h_{max} \cdot K \cdot (n + q + 1) \) inequalities and \( h_{max} \cdot K \cdot (n_0 + p) \) equalities with \( h_{max} \cdot K \cdot q + n \) unknown integer variables. Basically, the complexity is exponential with respect to the number \( N \) of reachable markings in \( R(G, M_1) \) and to the length \( K \) of the measured trajectories \((h_{max} \) is a constant parameter). Branch and bound algorithms can be used to solve (9)-(10) as an integer linear programming problem (LPP) [24]. These algorithms have a general nonpolynomial complexity but limit the computational effort in many practical situations.

An algorithm of linear complexity has been also developed in our previous work that limits the length \( K \) of the timed measured trajectories under test. It considers measured trajectories within a sliding window of constant size \( K_0 \) instead of increasing size \( K \) [25] and leads to an algorithm of linear complexity with respect to \( K \). Note also that the complexity with respect to \( N \) is no longer exponential if the set \( M_0(TR_O) \) is known (Assumption D).

(b) The numerical evaluation of (19) is obtained according to a recursive scheme with a deep equal to the number of transitions of type 2 or 3 in the considered sequence. Consequently the computation effort increases rapidly in time and in space with respect to the sequence length. To limit the computational complexity, the trajectory \((\sigma, M)\) is divided into \( S \) subtrajectories that are considered successively and independently: \( (\sigma, M) = (\sigma_1, M_1) \cdots (\sigma_S, M_S) \).
with $M_1 = M$. Each subtrajectory $(\sigma_j, M_j)$ is of minimal length such that (i) $(\sigma_j, M_j)$ ends with a measurement $M(\tau_k)$; (ii) all transitions in $\sigma_j$ are enabled from a marking that belongs to the same subtrajectory $(\sigma_j, M_j)$ and not to a previous one. For this reason each subtrajectory can include several measurements. For example, the trajectory in (21) is divided into two subtrajectories $(\sigma_1, M_1)$ and $(\sigma_2, M_2)$.

Trajectory decomposition is as follows:

\[
\begin{align*}
(\sigma, M) &= (\sigma_1, M_1) \\
(\sigma, M) &= M(0) \left[ T(t_1) \right] M(t_1) \cdots [T(t_{\nu'})] \\
\hspace{1cm} d_{h''} &= t_{h''} - t_1 \\
(\sigma, M) &= M(t_{\nu'}) \left[ T(t_{h'}) \right] M(t_{h'}) \cdots [T(t_H)] M(t_H) \\
\hspace{1cm} d_{h} &= t_{h} - t_{h'} \\
\hspace{1cm} d_{H} &= t_{H} - t_{H''}.
\end{align*}
\]

(21)

3.6. Example. Let us consider the marked POTPN2 of Figure 4 (unobservable places and transitions are highlighted in grey) that represents a cycle of tasks. The state of the system is not measured and only two events are observed. The set of labels is $E = \{e_1, e_2\}$. The matrices $L = \begin{bmatrix} (1 & 0 & 0 & 0 & 1) \end{bmatrix}^T$ and $H = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$ define the sensor configuration. Measurements are collected over the time interval $[0, 20]$ and the measured trajectory $T_{E_2} = (00) e_1(3.7) 0(3.7) e_2(20) 0(20)$ is considered. For this example $h_{max} = 4$ and a single untimed trajectory $h_{max} = 4$ and a single untimed trajectory $(\sigma_{U}, M_{U})$ with $\sigma_{U} = T_1(1) T_2(2) T_3(3) T_4(4) T_5(5)$ is consistent with $T_{E_2}$. The timed trajectories $(\sigma_j, M_j)$ consistent with $T_{E_2}$ satisfy $\sigma = T(t_1) T(t_2) T(t_3) T(t_4) T(t_5)$ with $T(t_1) = T_1(1), T(t_2) = T_2(2), T(t_5) = T_5(5), t_1 = 3.7 \text{ TU}, t_5 = 20 \text{ TU}$, and $t_1 \leq t_2 \leq t_3 \leq t_4 \leq t_5$. The pdfs of the transition durations are assumed to be bounded uniform with the same support $[0, 10]$ for $T_j$, $j = 1:5$. The temporal constraints are also assumed to be identical $T_{C_j} = [1, 9]$ for $j = 1:5$.

$\text{Prob}(T_{E_2} \text{ satisfies } \mathbf{T_C}) = S_1/S_2 = 0.46$ is obtained by (19) with

\[
\begin{align*}
S_1 &= \int_{t_1}^{t_1 + \Delta_1} f_2(t_2 - t_1) \cdot \left( \int_{t_2}^{t_2 + \Delta_2} f_3(t_3 - t_2) \cdot \left( \int_{t_3}^{\max(t_3 + \Delta_3, t_5 - \Delta_3)} f_4(t_4 - t_3) \cdot f_5(t_5 - t_4) \, dt_4 \right) \, dt_3 \right) \, dt_2 = 0.024, \\
S_2 &= \int_{0}^{t_5} f_2(t_2 - t_1) \cdot \left( \int_{t_2}^{t_5} f_3(t_3 - t_2) \cdot \left( \int_{\min(t_3, t_5)}^{t_5} f_4(t_4 - t_3) \cdot f_5(t_5 - t_4) \, dt_4 \right) \, dt_3 \right) \, dt_2 = 0.054.
\end{align*}
\]

This computation is validated with a series of simulations that leads to a probability of 0.47. The evaluation of $\text{Prob}(T_{E_2} \text{ satisfies } \mathbf{T_C})$ with (19) saves time compared to the numerical evaluation based on simulation.

4. Conclusion

This article has proposed a diagnosis system that checks if the heterogeneous measurements obtained from a stochastic timed discrete event system with an uncomplete sensor configuration are consistent or not with a set of temporal constraints that specify tolerance intervals for the system operations. For this purpose, the set of trajectories consistent with a given timed measured trajectory are first characterized. Then the consistency of each trajectory with the temporal constraints is estimated as a probability. Finally the probability of each trajectory is also evaluated and the global probability that the temporal constraints are satisfied results from the previous steps. The diagnosis system returns an alarm each time this probability goes below a given threshold. The contribution is validated with simulation results.

In the future, we will consider the isolation of the temporal constraints that are unsatisfied. We will relax some assumptions, in particular Assumption B, in order to consider the silent closure of the trajectories. We will also study the problem from a structural point of view by providing some results to decide whether a set of sensors is suitable or not to check that a set of temporal constraints is satisfied. Finally we aim to apply the proposed approach to larger systems.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.
Acknowledgments

This study was supported by the Region Normandie and the European Union (Project MRT MADNESS 2016–2019).

References


