Timed pattern diagnosis in timed workflows: a model checking approach
Yannick Pencolé, Audine Subias

To cite this version:
Yannick Pencolé, Audine Subias. Timed pattern diagnosis in timed workflows: a model checking approach. 14th Workshop on Discrete Event Systems (WODES 2018), May 2018, Sorrento Coast, Italy. hal-01827362

HAL Id: hal-01827362
https://hal.laas.fr/hal-01827362
Submitted on 2 Jul 2018

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.
Timed pattern diagnosis in timed workflows: a model checking approach

Yannick Pencolé * Audine Subias. *

* LAAS-CNRS, Université de Toulouse, CNRS, INSA, Toulouse, France
(y pencole@laas.fr, subias@laas.fr).

Abstract
In this article we introduce the problem of timed pattern diagnosis in timed system and propose a method to solve a subclass of these problems: the problem of timed pattern diagnosis of workflows. It consists in searching for all the possible evolutions of the workflow that are consistent with a timed sequence of observations and determining whether these evolutions match a pattern of timed events instead of a single fault event. The formal characterization of the diagnosis problem is based on the notion of pattern matching. Defined as a reachability problem, the diagnosis problem is then solved by model checking.

Keywords: Diagnosis, Time Petri Net, Timed Pattern, Model checking

1. INTRODUCTION

The problem of fault diagnosis in DESs was introduced in Sampath et al. (1995). Later, extensions of the problem handling time have been proposed using timed automata (Tripakis (2002)) and more recently Petri nets with several interpretations of time and methods, for instance: Chatain and Jard (2005), Jiroveanu and Boel (2006), Basile et al. (2007), Bonhomme (2015), Boel and Jiroveanu (2013), Basile et al. (2015), Liu (2014), Wang et al. (2015). These contributions describe some methods for the same problem that is the diagnosis of single faults in a timed system. In this article we propose to extend the diagnosis problem to more complex behaviors than single events called timed patterns and to apply it to timed workflows. Timed patterns extend with time the classical notion of pattern (Jéron et al. (2006); Gougam et al. (2017)). Beyond faulty behaviors, classical patterns can represent any untimed normal behavior of interest or any behavior associated with a system specification as in Jiang and Kumar (2004). In Bozzano et al. (2015), such patterns are called diagnosis conditions and are represented with LTL (Linear Time Logic) properties with past operators.

As classical patterns, the proposed timed pattern models more complex behaviors than a single event (sequences, concurrent events,...) but it also add the possibility to model a time interval between the occurrence of two events in the patterns. Pattern model can then naturally model unexpected/acceptable/faulty time delays between two unobservable events. In this article we firstly propose to formally define the general problem of timed pattern diagnosis in timed system. We then propose a method based on model-checking techniques that solves a subclass of this new problem that is the timed pattern diagnosis in timed workflows. Given the observable part of the run of a workflow, the diagnosis problem consists in determining whether the considered timed pattern has occurred or not during the run of the workflow. The proposed method first converts the problem into a reachability problem and then use model checking tool to solve it. Section 2 formally defines the general problem. The proposed method is presented in Section 3. An illustrative example about a service workflow is described in Section 4.

2. GENERAL PROBLEM STATEMENT

The objective of the approach is to diagnose the occurrence of a timed behavior (i.e. a timed pattern) in a timed system. It supposes the implementation of a diagnosis function that takes as input a sequence of observations and returns one of the following three results:

(1) The pattern’s behavior has certainly occurred.
(2) The pattern’s behavior is certainly absent.
(3) The pattern’s behavior may have occurred, ambiguity.

This section defines the different parts of the problem that is: how to model timed systems/timed patterns, how to formally define the occurrence of a timed pattern, what is finally the diagnoser function.

2.1 Preliminaries

This paper deals with timed sequences. Let Σ be an alphabet that is a finite set of symbols, each symbol represents an event. A timed sequence over Σ is an element of $T(Σ) = (N^+ × (Σ ∪ \lambda))^*$. Suppose for instance that $Σ = \{a, b, c\}$, $ρ = (a, 3), (λ, 1), (c, 1), (c, 1), (b, 2), (λ, 5)$ also denoted $3a1λ1c1c2b5λ$ is a timed sequence, $ρ' = 1b1b40λ$ is another one. Symbol $λ$ defines a time point (basically $1λ2λ$ is the same as $3λ$). The empty sequence of $T(Σ)$ is given by $0λ$.

For any timed sequence it is possible to define a canonical form $δ_0e_1...δ_ne_nλ, e_i ∈ Σ$ that is obtained by removing useless time points. For instance the canonical form of $ρ = 3a1λ1c1c2b5λ$ is $3a2c1c2b5λ$. In this article and without loss of generality only canonical forms are considered. The projection of a timed sequence $ρ$ over a sub-set $Σ_ρ$ of $Σ$ (denoted $P_{Σ_ρ}(ρ)$) is the timed sequence obtained by removing any symbols that does not belong to $Σ_ρ$ and by summing the durations between two symbols that are still present after the removal. Let us consider the canonical form of $ρ$, it corresponds
to an event \( a \) that occurs at date 3 after the origin of time, an event \( c \) that occurs 2 units of time after \( a \) (so \( c \) occurs at date 5), followed by an event \( b \), 1 unit of time later (so at date 6), and so on. Given \( \Sigma_p = \{a, b\} \) then \( P_{\Sigma_p}^\rho(\rho) = 3a5b5\lambda \).

Any sub-set (canonical sub-set) of \( T(\Sigma) \) is a timed language over \( \Sigma \). Let \( \text{time}(\rho) \) denote the sum of all durations in \( \rho \), then \( \text{time}(3a + 2b) = 5 + 2 + 1 + 2 + 4 = 14 \). \( \rho' \) is a prefix of \( \rho \) if \( \text{time}(\rho') \leq \text{time}(\rho) \) and there exists \( \rho'' \) such that \( \rho' \rho'' \) and \( \rho \) have the same canonical form.

### 2.2 Matching

Like untimed discrete event system, a timed discrete event system generates a finite set of events \( \Sigma \). The difference is that in the timed case, the possible time durations between the occurrence of two successive events are defined, whereas in the untimed case, the events are only connected by a precedence relationship. Of course in both cases (timed and untimed), two events may not be related i.e even the order of their occurrence is not known. An evolution of a timed discrete event system is a timed sequence on its alphabet. The set of possible evolutions of a timed system is thus a set of timed sequences, that is, a timed language.

**Definition 2.1.** (Timed system). A timed discrete event system whose set of events is \( \Sigma \), is a timed language \( L \subseteq T(\Sigma) \) such that if \( \rho \in L \) then any of its prefixes is in \( L \).

The set of events \( \Sigma \) generated by the system is partitioned into two subsets: observable events \( \Sigma_o \) and unobservable events \( \Sigma_u \). Two timed sequences can be linked by the notion of matching that states whether a sequence is present inside another.

**Definition 2.2.** (Sequence matching). A sequence \( \rho \in T(\Sigma) \) matches a sequence \( \sigma \in T(\Sigma) \), denoted \( \rho \triangleright \sigma \), if:

- \( \sigma = \delta\lambda \) and \( \text{time}(\rho) \geq \delta \), with \( \delta \in \mathbb{N} \); or
- \( \sigma = \delta_0 e_0 \sigma_1 \), such that: \( \delta_0 \in \mathbb{N} \), \( e_0 \in \Sigma \), \( \sigma_1 \in T(\Sigma) \) and there exist two timed sequences \( \rho_0, \rho_1 \in T(\Sigma) \) and a duration \( \delta \in \mathbb{N} \) such that:
  1. \( \rho = \rho_0 \delta e_0 \rho_1 \);
  2. \( \delta_0 = \text{time}(\rho_0) + \delta \);
  3. \( \rho_1 \triangleright \sigma_1 \).

In other words, \( \rho \) matches \( \sigma \) if \( \sigma \) is a sub-word of \( \rho \). Figure 1 illustrates how the sequence \( \rho = 2b3a25a1c3 \lambda \) matches the sequence \( \sigma = 7a7c6\lambda \). The first event of \( \sigma \) is \( a \) and occurs at date 7. In \( \rho \), there is also an event \( a \) that occurs at date 7 = 2 + 3 + 2. The second event of \( \sigma \) is \( c \) that occurs 7 time units later than \( \rho \). This event is matched with the third event \( c \) of \( \rho \) that occurs 5 + 1 + 7 = 13 time units after \( \rho \). Finally, \( \sigma \) ends instantaneously after event \( c \) while \( \rho \) lasts 3 time units more so \( \rho \) matches \( \sigma \).

We define a timed pattern as a set \( \mathcal{R} \) of timed sequences which define a behavior of interest. For example, it may be the occurrence of a single event in a given time period, the occurrence of several events that occur in any order, but whose occurrence dates have constraints or the occurrence of several events that occur in a fixed order with specific time constraints. As in the original diagnosis problem that considers unobservable fault event, we consider that the behavior of the pattern is not observable i.e no event of the pattern is observable in the system.

**Definition 2.3.** (Timed pattern). A timed pattern over a set of events \( \Sigma_{\mathcal{R}} \subseteq \Sigma_u \) is a timed language \( \mathcal{R} \subseteq T(\Sigma_{\mathcal{R}}) \) such that if \( \mathcal{R} \triangleright \sigma \), \( \mathcal{R} \) is an \( \mathcal{R} \)-diagnoser.

![Figure 1. How \( \rho = 2b3a25a1c3\lambda \) matches \( \sigma = 7a7c6\lambda \):](image)
\[ \Delta_{\mathcal{R}}(\sigma) = \mathcal{R} - \text{absent} \] if for any sequence \( \rho \in \mathcal{L} \) such that \( \mathcal{P}_{\mathcal{R}}(\rho) = \sigma \), \( \rho \) is not \( \mathcal{R} \)-matching;

\[ \Delta_{\mathcal{R}}(\sigma) = \mathcal{R} - \text{ambiguous otherwise.} \]

Considering a set of patterns \( \mathcal{R}_1, \ldots, \mathcal{R}_n \), the diagnoser function is defined by:

\[ \Delta: T(\Sigma_\omega) \rightarrow \prod_{i=1}^{n} (\mathcal{R}_i - \text{certain}, \mathcal{R}_i - \text{absent}, \mathcal{R}_i - \text{ambiguous}) \]

such that \( \Delta(\sigma) = (\Delta_{\mathcal{R}_1}(\sigma), \ldots, \Delta_{\mathcal{R}_n}(\sigma)) \).

### 3. Diagnosis by Model Checking

The proposed method defines the problem of timed pattern diagnosis as a problem of reachability and uses model checking techniques to solve it. This resolution relies on the TINA toolbox (Time Petri Net Analyzer) (Berthomieu et al. (2004)). Given a system and a pattern, the method proceeds as follows.

1. The timed languages of the system and of the pattern are modeled by L-type Labeled Time Petri nets respectively denoted \( \Theta \) and \( \Omega \).
2. From the two previous nets, a new net is built (\( \Theta_{\Omega} \)) that captures the matching relation between the system and the pattern.
3. The input observation sequence is transformed into an L-type Labeled Time Petri net (denoted \( \text{Obs} \)).
4. The result (\( \Theta_{\Omega} \)) is then synchronised with the Petri net model of the observation sequence (\( \text{Obs} \)) to obtain the net (\( \Theta_{\Omega}[\text{Obs}] \)).
5. The model checker of TINA is used to verify the matching property between the system and the pattern according to the input observation sequence on the net (\( \Theta_{\Omega}[\text{Obs}] \)).
6. The diagnosis result is thus given according to the answer of the model checker.

All the steps are detailed hereafter.

#### 3.1 Timed languages modeling

The approach relies on the following class of Time Petri nets (see Berthomieu et al. (2006) and Peterson (1977)).

**Definition 3.1.** (LLTPNPr). An L-type Labeled Time Petri Net with Priorities (LLTPNPr for short) is a tuple \( \mathcal{N} = (P, T, \succ, \ell, \lambda, I_\omega, I_s, Q, M_0) \) with:

- \( P \): a finite set of nodes called places;
- \( T \): a finite set of nodes called transitions and \( P \cap T = \emptyset \);
- \( A \subseteq (P \times T) \cup (T \times P) \) is the set of arcs between nodes.
- \( \succ \subseteq T \times T \): the priority relation, not reflexive, not symmetric and not transitive; \((t_1 \succ t_2)\) means that if \( t_1 \) and \( t_2 \) are both enabled then only \( t_1 \) is firable.
- \( \ell : P \cup T \rightarrow L \cup \Sigma \cup \{ \lambda \} \): the deterministic labelling function where \( L \) (resp. \( \Sigma \)) is the set of place labels (resp. transition labels); For \( r = \delta_1 t_1 \delta_2 t_2 \ldots \in (\mathcal{N} \times T)^* \), \( \ell(r) = \delta_1 \ell(t_1) \ell(t_2) \ldots \).
- \( I_\omega : T \rightarrow I^+ \): the static time interval function, where \( I^+ \) is the non empty set of time intervals with non negative rational bounds.
- \( Q \): the set of final markings and \( M_0 \) the initial marking.

A marking \( M \) is a function from \( P \) to \( \mathcal{N} \) which maps each place with a number of tokens. For an LLTPNPr \( \mathcal{N} \), the functions \( \text{pre}(t) = \{ p \in P : (p, t) \in A \} \) and \( \text{post}(t) = \{ p \in P : (t, p) \in A \} \) are defined. A state of \( \mathcal{N} \) is a couple \( (M, I) \), with: \( M \) a marking and \( I \) the firing interval of the transition.
model represents the synchronized evolutions of the system and the pattern but also the evolutions of the system alone as a pattern evolution corresponds potentially (in case of matching) to a sub-word of the system. The combination of both nets is obtained by performing the union of the system places and the pattern places. The transitions of the system that do not share any event with the pattern are kept with their initial time intervals. The other ones must be synchronized with the pattern’s transitions. As in Time Petri nets compositionality is kept only when synchronizations involve transitions with an interval $[0, +\infty)$ (Berthomieu et al. (2006)), we need to proceed to a time decomposition of the transitions to synchronize (see Figure 3). Transition $t_1$ (on the left) is labeled by $a$ with a finite time interval $[a, \beta]$. $t_1$ is replaced by a sub-net (on the right) composed by a silent transition (labeled by $\lambda$ not indicated on the figure) of interval $[\alpha, \alpha]$ whose firing enables both a transition labeled by $a$ ($t_2$) with the interval $[0, +\infty]$ and a silent transition ($t_2$) of interval $[\beta - \alpha + 1, \beta - \alpha + 1]$. At $\beta + 1$ (i.e. $\beta$ times unit + one time unit) as this transition has no priority, the silent transition ($t_2$) is fired leading to an unacceptable behavior represented by the marking of a specific place unAcc. This transformation applies to transitions of the system as well as transitions of the pattern. Therefore in the system the specific place is named sys_unAcc whereas in the pattern it is patt_unAcc.

Finally Figure 4 illustrates the complete combination in $\Theta_{\Omega}$ of a system transition labeled by $b$ and assigned with an interval $[1, 4]$ with a pattern transition labeled by $b$ with an interval $[3, 6]$. Two transitions labeled by $b$ appear: the one of the right results from the synchronization of the system and the pattern transitions and the left one represents the standalone evolution of the system. Priorities (dashed lines) ensure that the synchronized transition is triggered just in case of matching. For each possible run of the system, the net $\Theta_{\Omega}$ represents whether this run matches the pattern or not. Finally, to solve the diagnosis problem, this model must now be confronted with a sequence of observations.

3.4 Integration of the observation sequence

The observation sequence is first modeled by an L-Type Labeled Time Petri net denoted $Obs$. For instance for an observation sequence $1\omega_13\omega_2$, $Obs$ has one transition $t_1$ labeled $\omega_1$ assigned with a static time interval $[1, 1]$. The output place of this transition enables a second transition $t_2$ labeled by $\omega_2$ with an interval $[3, 3]$. The output place of $t_2$ is denoted $p_{Obs}$. Afterwards, the combination system-pattern (i.e the net $\Theta_{\Omega}$) is synchronized (Hack (1975)) with the Petri model of the observation sequence ($Obs$) to obtain the net ($\Theta_{\Omega}||Obs$). To perform the synchronization the transitions of the observation sequence net ($Obs$) and the transitions of the combination ($\Theta_{\Omega}$) are transformed so that the synchronizations involve only transitions with $[b, +\infty]$ intervals as explained in Section 3.3. The verification of the matching property between the system and the pattern according to the input observation sequence on $\Theta_{\Omega}||Obs$ is then performed by the model checker. Then the diagnosis function ($R$-diagnoser) generates its results according to the answers of the model checker as presented now.

3.5 Implementation of the $R$-diagnoser function by Model checking

Back to the description of the method in Section 3, the overall computational complexity of the method lies in step 5 and we propose to use an off-the-shelf model checking tool to solve this part: the TINA toolkit (Berthomieu et al. (2004)). TINA generates enriched labeled Kripke structures from which it is possible to check properties written in SE-LTL (State/Event Linear Temporal Logic) which extends LTL in the way presented hereafter. A formula $\psi$ is a SE-LTL formula if it is a universally quantified formula: $\psi := \forall \phi$ such that:

\footnote{This property is true because the systems that are considered here are workflows (see Definition 3.4).}
The proposed method is illustrated on an example denoted as the _foodshop_ example. This example is inspired from the 'Web services' systems that were studied in the European project WSDIAMOND (Di Nitto et al., 2009). Figure 5 presents the workflow of a foodshop service. In this type of application, fault patterns usually represent misbehaviors not compatible with the service specification instead of physical faults on hardware. In this workflow, a foodshop first receives the order of a client (_sh_reqv_). Within the next 3 hours, the shop must start two activities (_sh_st[1,2]) in parallel. The first activity dispatches the order into two categories: the storable products (not perishable) and the non-storable products (perishable ones). This dispatch activity ends with the event _sh_end_spl_. The second activity selects a warehouse. For storable products, the shop queries the warehouse to check the availability of the products (_sh_reg_whcheck_). The warehouse then checks (_wh_check_) and provides a feedback (available _wh_send_av_, unavailable _wh_send_unav_). For non-storable products, the shop sends directly a request to a supplier (_sh_reg_suppcheck_) that also provides feedbacks (_sp_send_av_, _sp_end_deliv_). The shop requires that the feedbacks are given within the next 4 hours. After the reception of the feedbacks, the shop decides to implement or not the order based on the availabilities (_sh_decide_). If one product is missing the shop cancels the order (_sh_cancel_), if not the shop requests the preparation for delivery to the warehouse and the supplier (_sh_reg_deliv_). When preparations are finished (_wh_end_prep_deliv_ and _sp_end_prep_deliv_), it is finally delivered. In Figure 5 (on the left), observable events are in bold. They are events from the shop that manages the whole delivery process. To illustrate the method, we consider two patterns.

4. APPLICATION TO THE DIAGNOSIS OF A TIMED WORKFLOW OF A FOODSHOP

4.1 Occurrence of _sp_send_unav_

This pattern is the occurrence of _sp_send_unav_. This pattern models the fact that the supplier sends as feedback that the goods are not available. This simple pattern aims at illustrating that our method extends the classical problem where fault are modeled by single events. This pattern is composed of two places _pp0_, _pp1_ and one transition. Initial marking is _M0(pp0) = 1_ and _M0(pp1) = 0_ and the only accepting marking _M ∈ Qf_ is such that _M(pp0) = 0_ and _M(pp1) = 1_. Consider one run of the timed workflow and suppose that it produces the observable sequence _σ1_ whose conversion to Petri net is shown in Figure 5 (on the right). To solve the problem, we first build the product between the system and the pattern that is then synchronized with the Petri net of _σ1_. On the resulting Petri net, we first query to check whether any run of the system that can produce _σ1_ badge the pattern with the following query _ϕ1:

\[
\{\text{sys\_unAcc = 0} \lor p9 = 1\} \Rightarrow (p1 = 1)
\]

TINA’s response is negative: there exists at least one run of the system that produces _σ1_ and for which _sp_send_unav_ has not happened. Then we check whether these runs do not match (i.e. contain) _sp_send_unav_ with the formula _ϕ2:

\[
\{\text{sys\_unAcc = 0} \lor p9 = 1\} \Rightarrow (p1 = 1)
\]

This is a positive answer. So, the diagnosis of _σ1_ is that _sp_send_unav_ did not happen. Let us consider now the observable sequence _σ2_ composed of the first five events of _σ1_ followed by _sh_send_av_. In this case _ϕ1_ is true: any run generating _σ2_ contains _sp_send_unav_. Querying TINA with _ϕ2_ is useless: the diagnosis is certain.

4.2 Occurrence of a sequence pattern

Consider _wh_send_av[0, +∞[ \rightarrow wh_end_prep_deliv[70, 200]]. This second pattern illustrates a misbehaviour: if the warehouse sends as a feedback that the goods are available then the preparation ends within 70 and 200 time units which is considered as an unacceptable duration (quality of services is impacted). Let us consider the following observed sequence: _σ3_ = _sh_regv_ _sh_st_ _sh_end_spl_ _sh_reg_whcheck_ _sh_reg_suppcheck_ _sh_decide_ _sh_reg_deliv_ _sh_end_deliv_. TINA concludes that the pattern has not happened. This is due to the fact that the delivery date is time 50 after the order request. Now if we modify _σ3_ by increasing the date 50 to 80, the diagnosis becomes certain. Only the observation of time is discriminating here.

5. CONCLUSIONS ET PERSPECTIVES

This work extends the original problem of the diagnosis of simple faults to the diagnosis of timed patterns in timed systems. The diagnoser function is based on the notion of matching between the system evolutions that are consistent with an input sequence of observations and the pattern. A resolution of the problem based model checking techniques is proposed to solve the problem on timed workflows. This work must be completed by a diagnosability analysis inspired from Gougam et al. (2017). We are also planning to investigate other subclasses of Time Petri nets where the problem is decidable.

REFERENCES

Basile, F., Chiaccio, P., and De Tommasi, G. (2007). Improving on-line fault diagnosis for discrete event systems using...


