Prediction of Timing Constraint Violation for Real-Time Embedded Systems with Known Transient Hardware Failure Distribution Model

Yue Yu, Shangping Ren, Ophir Frieder
Department of Computer Science
Illinois Institute of Technology
{yyu8, ren, frieder}@iit.edu

Abstract

We apply interval-based timing constraint satisfaction probability results to predict timing constraint violations in real-time embedded system with a known hardware transient failure model. A previous study indicated that hardware transient failures follow a Poisson distribution with an average failure arrival rate λ. Under this model, the distribution of time intervals between successive failures follows an exponential distribution with the same parameter λ. Our goal is to use the statistical transient failure models to calculate the earliest time at which we can predict, with a determined level of confidence, that a given timing constraint may be violated. This earlier prediction provides time-critical systems with valuable time before the deadline is reached to adapt themselves, and hence, to minimize possible negative impacts caused by timing constraint violations.

1. Introduction

Hardware units inevitably encounter transient failures. Studies [18, 19] have shown that a Poisson distribution with an average arrival rate of λ can often model transient hardware failures. While the hardware unit is in a transient failure state, it will not generate expected events or the generated events could be faulty and should be ignored. Simply, valid events can happen only in the interval between two successive transient failures. Under the Poisson distribution failure model, the time between two successive failures follows an exponential distribution with λ equaling the average failure arrival rate. Furthermore, given a transient hardware failure distribution, an expected event occurrence distribution can be obtained to statistically predict if a timing constraint related to the expected events will be violated.

Monitoring the timing behavior of time-critical systems is as important as monitoring its functional correctness. In practice, it is crucial to monitor the minimum/maximum separation between a pair of failures. For example, a distributed real-time system with a primary machine M_1 and a backup machine M_2 should monitor if failures occurring on M_1 and M_2 are separated by at least the time needed by the primary machine to recover.

Researchers developed efficient algorithms that detect constraint violations or satisfactions of timed events [13]. Later, they extended their work to deal with situations where the time point at which an event occurs is unknown [12]. They proposed a new type of timing constraints based on an interval time event model where timestamps for the events are given by time intervals. Recent work [10] further analyzed timing constraints based on time interval in which an event occurrence is *uniformly* distributed over the interval. They introduced the concept of *earliest expiration time* (*EET*) and were able to quantitatively define the relationship between *satisfaction probability* (*SP*) and the *EET*. With the earliest expiration time, one can obtain an early warning about possible timing constraint violations before the actual deadlines, and hence, provide real-time systems with opportunities to adjust their behaviors.

We use the interval-based timing constraint model to analyze the earliest time at which we can predict with a determined level of confidence that a given timing constraint may be violated due to different statistical failure models. Given the observed Poisson hardware failure model, our key contribution is the ability to predict timing constraint violations at earlier time instances with the more realistic model of exponential distributed failures than the simplified assumption of uniformly distributed failures.

2. Related Work

Failures occur due to hardware/software errors and the effect of cosmic ray radiation. Moreover, transient failures occur much more frequently than permanent failures [2, 7, 8]. In the literature, transient failures are modeled as following a Poisson distribution with an average arrival rate of λ [18, 19]. The reason is that Poisson distributions can be used to describe various phenomena of discrete nature whenever the probability of the phenomenon happening is constant in time or space. Zhu [19] investigates the dynamic energy management problem in real-time embedded systems. There, the transient failure is modeled as a Poisson model and voltage scaling affects the parameter λ associated with the distribution model. Hence this approach is more effective as it better reflects realistic situations.

Another approach investigates the problem of monitoring by studying timing constraints of events. Chodrow et al. [3] presents a constraint-graph-based algorithm for detecting violations of timing constraints. To check the satisfiability for a set of constraints, a constraint graph is instantiated from the current event histories with an all-pair shortest path algorithm run on the instantiated graph. A negative cycle indicates unsatisfiability of the constraint set.

In [9] and [14], the authors extend the timing constraint specification and violation detection algorithm to distributed real-time systems. They indicate that the derivation of implicit constraints is essential for catching timing violations at an earlier time since it is possible that an implicit constraint is violated before an explicit delay or deadline becomes unsatisfiable at run-time. They also prove that for constraint violation detection, the problem of minimizing the amount of information to be exchanged between processors is NP-hard.

In [13], Mok and Liu provide a more expressive specification language based on Real Time Logic to define timing constraints. To reduce time complexity of the event-monitoring algorithm at run-time, they resolve most of the shortest path information of the instantiated constraint graph from the uninstantiated constraint graph at compilation time. Thus, only small modifications of the graph are needed during run-time.

Two new timing constraints based on time intervals, *certain* and *possible*, are proposed in [12] to specify the desired degree of certainty whether a timing violation has occurred. The authors indicate that this interval-based timing constraint: $I_1 + d \geq I_2 U$, where I_1 and I_2 denote the time intervals in which the corresponding events may occur, is satisfied either possibly (P) or certainly (C). The authors further extend the monitoring algorithm in [13] to monitor interval-based timing constraints with probabilities.

Lee et al. [10] propose interval-based timing constraints with a confidence threshold model. The concept of *Earliest Expiration Time* (EET) is also introduced. The knowledge of EET enables the event monitor to announce the violation of timing constraints even before the actual deadlines.

Yu et al. [17] extend Lee's work by considering more general cases of the interval-based timing constraint model where events are exponentially or normally distributed. This paper uses the results from [17] to analyze and predict timing constraint violations caused by different transient failure models.

Examples of other stochastic approaches in real-time applications can be found in [1, 4, 5, 6, 11, 15].

3. Basic Assumptions, Definitions and Problem Descriptions

We now state our assumptions and definitions and then present a theoretical analysis based on them.

3.1. Failure Model

As shown in [18, 19], during the execution of an application, failures are modeled to follow a Poisson distribution with an average arrival rate λ. That is, given the average arrival rate λ in a time interval, the probability that exact k failures will occur in this period is:

$$f(k;\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Under this model, the time between successive arrivals of failures follows an exponential distribution with the same parameter λ. That is, suppose X is a random variable representing the time between successive failure arrivals then the probability density function of X is:

$$f_X(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad x < 0 \end{cases}$$

Example 1 Consider a specific kind of failure that occurs at random times with a mean rate of 6 distinct times per hour. The probability that ten or less minutes will elapse between two failures is:

$$P[X \leq 10] = \int_0^{10} \frac{1}{10} e^{-\frac{x}{10}} = 1 - e^{-1} \approx 63\%$$

Thus, if one failure occurs at 10:00 am, the probability that the second failure will happen before 10:10 am is 63%. Since we do not know the exact occurrence time of the second failure, we assume it follows a probability distribution on $[t_1, +\infty)$ (in this case, exponential distribution), where t_1 is the time point at which the previous failure occurs. In realistic analysis,

we can use arbitrarily large numbers to replace the infinity. In Fig 1, we depict the intuition in which t_1+40 is considered as $+\infty$ in this context. The dark region indicates that the failure is more likely to occur in the interval of $[t_1, t_2]$.

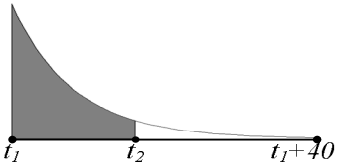


Fig 1. The occurrence of second failure can be modeled as following exponential distribution over the interval [t_1, t_1+40].

3.2. Interval-based Timing Constraint Model

Generally speaking, real-time constraints can be categorized into two classes, namely, deadline constraints or delay constraints. More specifically, a deadline constraint between two events with timestamps σ and γ is modeled as $\sigma + d \geq \gamma$, while a delay constraint is modeled as $\sigma + d < \gamma$, where $d \geq 0$ is a constant representing a deadline or a delay.

For self-completeness, we quote the related definitions (Definition 1 through 3) from [10] in the following.

Definition 1 (Timestamp) A timestamp I consists of a pair of time points: [*min_time, max_time*] where *min_time* and *max_time* are the earliest and latest time points at which an event may occur, respectively. Moreover, given a timestamp I, we assume the probability density function of X representing the time point at which the event may occur is $f(x)$.

□

Definition 2 (Function *min, max,* and *len*) Given a timestamp I = (*min_time, max_time*), the *min*, *max* and *len* functions of a timestamp I are defined as following:

$$min(I) = min\_time$$
$$max(I) = max\_time$$
$$len(I) = max(I) - min(I)$$

For brevity, we use min_k, max_k, and len_k to denote $min(I_k)$, $max(I_k)$, and $len(I_k)$, respectively, where I_k is a timestamp.

□

Definition 3 (Interval-based timing constraint) An interval-based deadline constraint with a confidence threshold is given by:

$$c^+ : I_1 + d \geq I_2 with\ P$$

and an interval-based delay constraint with a confidence threshold is given by:

$$c^- : I_1 + d < I_2 with\ P$$

where I_1 and I_2 are timestamps, $d \geq 0$ is a constant representing a deadline or a delay, and P is a confidence threshold ranging from 0% to 100%.

□

Definition 4 (Run-time timing constraint violation) A violation of a deadline timing constraint c^+: $I_1 + d \geq I_2$ *with* P [1] is said to occur at run-time when the event corresponding to I_1 occurs at time point t and the event corresponding to I_2 does not occur by s, where $s \in [min_2, t+d)$, and the remaining satisfaction probability of the timing constraint is less than the specified confidence threshold P.

□

Definition 4 states that although it is possible that the event corresponding to I_2 could occur during the interval $(s, t+d]$ that satisfies $I_1+d \geq I_2$, the specified confidence threshold P is violated.

Definition 5 (Earliest prediction time, EPT) Given a timing constraint c^+: $I_1 + d \geq I_2$ *with* P where $d \geq 0$, the earliest prediction time of deadline constraint violation from a time point t, $EPT(t)|_{c^+}$, where t is the time point at which the event corresponding to I_1 occurs, is defined as:

$$EPT(t)|_{c^+} = t' - t$$

where t' is the earliest time when an event monitor can safely claim that c^+ is *violated* in case the event corresponding to I_2 does not occur by t'.

□

Note that, our definition of the *run-time timing constraint violation* and thus the *earliest prediction time* differs from timing constraint violation and earliest expiration time introduced in [10]: In our definition, when a deadline timer is set at a time point t, we assume the event corresponding to I_1 has occurred. In other words, only when the event corresponding to I_1 occurs, do we start the timer and try to monitor the occurrence of event corresponding to I_2 for the possible satisfaction or violation of the constraint. In contrast, in [10], both events corresponding to I_1 and I_2 are still random events even after the deadline timer is set at t. In our definition, t has an intuitive and physical interpretation. Example 2 in subsection 3.3 will make the distinction more clear.

3.3 Problem Description

Before we formalize the problem, we present an intuitive example.

[1] The timing constraint violation and the earliest prediction time are defined only on a deadline constraint for the reason that delay and deadline are symmetric. Hence, it is thus sufficient to focus our presentation only on deadline constraints.

Example 2 Consider a distributed embedded system. For failure tolerance, the system has a primary machine M_1 and a backup machine M_2. During execution, transient failures may occur independently on these two machines, both of which follow a Poisson distribution with average failure arrival rates λ_1 and λ_2, respectively. To ensure the system availability, it is important that the time interval between two failures on these two machines must be at least d time units apart (where d is the machine recovery time). Since we cannot get the exact occurrence times of failures on M_1 and M_2, we are only able to give the satisfaction probability of the timing constraint. Moreover, if a failure occurs on M_1 at time t, as we cannot get the exact occurrence time of a failure on M_2, we may only know with a certain level of confidence whether the timing constraint is violated at run-time. The following figure illustrates the concept.

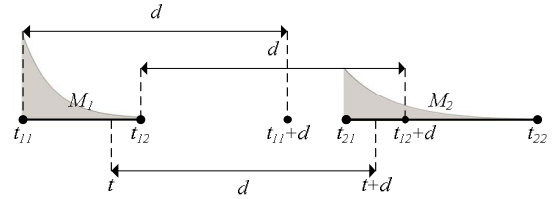


Fig 2. Independent failure occurrences on a primary machine M_1 and a backup machine M_2.

In Fig 2, failures on M_1 and M_2 follow exponential distribution on intervals $[t_{11}, t_{12}]$ and $[t_{21}, t_{22}]$, respectively. If a failure on M_1 occurs at t, the timing constraint can be satisfied with a certain probability since the failure on M_2 could occur prior to or after $t+d$. The satisfaction probability can be calculated given the failure distribution model on M_2. Similarly, prior to t_{11} when both failures have not occurred, the satisfaction probability can be calculated given the failure models on both machines.

We consider the problem of monitoring the timing constraint between two independent failures in real-time systems. The timing constraint can either be a deadline constraint or a delay constraint. We consider the more realistic situation where failures are exponentially distributed, as well as a much simpler case where we assume that the failures are uniformly distributed.

Specifically, we consider the problem of: given a transient failure timing constraint of the form c^+: $I_1+d \geq I_2$ *with* P or c^-: $I_1+d < I_2$ *with* P, determine if the constraint is satisfiable by comparing the satisfaction probability (*SP*) of the constraint with the confidence threshold P. Moreover, we look into the problem of predicting at run-time the earliest time (*EPT*) when we can claim a violation of a constraint. If the specified timing constraint cannot be satisfied with the required degree of confidence, it means that some system properties such as the required degree of failure tolerance cannot be guaranteed.

4. General Theorems for Interval-based Timing Constraints

4.1. Satisfaction Probability (SP)

Given the definitions in Section 3.2, we present Theorem 1 that calculates the satisfaction probability of a deadline constraint and under arbitrary probability density functions. If the calculated satisfaction probability is less than the confidence threshold, we know at compile time that the specified constraint is not satisfiable.

Theorem 1 Given a deadline constraint c^+: $I_1 + d \geq I_2$, where $d \geq 0$, and $f(x)$, $g(y)$ are the probability density functions of independent event occurrences on interval I_1 and I_2 , respectively, the satisfaction probability of c^+, $SP\big|_{c^+}$ is given by the expression:

$$SP\big|_{c^+} = \frac{\int_{x=MAX(\min_1,\min_2-d)}^{\max_1} f(x)\int_{y=\min_2}^{MIN(x+d,\max_2)} g(y)dydx}{\int_{\min_1}^{\max_1} f(x)dx \cdot \int_{\min_2}^{\max_2} g(y)dy} \quad (1)$$

Proof:

Let $X \in I_1$, $Y \in I_2$ be two continuous random variables with density functions $f(x)$ and $g(y)$. Since the two random variables are mutually independent, the joint density function $z(x, y)$ is simply the product of their individual density functions, as shown in Fig 3. Furthermore, the joint cumulative distribution over $I_1 = [min_1, max_1]$ and $I_2 = [min_2, max_2]$, denoted as V, is:

$$V = \int_{x=\min_1}^{\max_1}\int_{y=\min_2}^{\max_2} z(x,y)dydx = \int_{\min_1}^{\max_1} f(x)dx \cdot \int_{\min_2}^{\max_2} g(y)dy$$

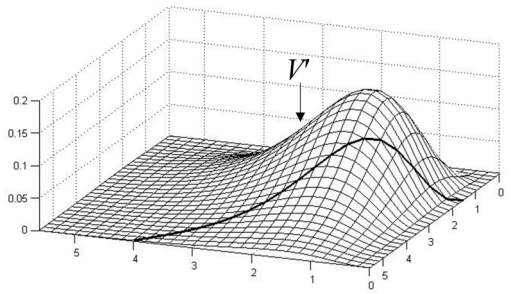


Fig 3. Joint density function of independent event occurrences on two intervals.

The satisfiable region, denoted as V', is the intersection between the region $y \leq x+d$ and V, as

shown in Fig 3. The bold line represents the intersection between the joint density function and the plane $y = x+d$.

The satisfaction probability is thus the ratio between the satisfiable region V' and the joint cumulative distribution V.

To calculate V', we project the plane $y = x+d$ and the surface $z(x, y)$ onto the X-Y plane and consider the relationship between the line $y = x+d$ and the four points (min_1, min_2), (min_1, max_2), (max_1, min_2), and (max_1, max_2). There are only six possible relationships that correspond to the six permissible configurations in [10]. Two of the six cases are trivial:

- (min_1, max_2) is below the line $y = x+d$, that is, $max_2 \leq min_1+d$, which implies a 100% satisfaction probability;
- (max_1, min_2) is above the line $y = x+d$, that is, $min_2 > max_1+d$, which implies a 0% satisfaction probability.

The four non-trivial configurations are:

- $\alpha\beta$ configuration, where $min_1+d \leq min_2 \wedge min_2 < max_1+d \leq max_2$
- $\alpha\gamma$ configuration, where $min_1+d \leq min_2 \wedge max_2 < max_1+d$
- $\beta\beta$ configuration, where $min_2 < min_1+d \leq max_2 \wedge min_2 < max_1+d \leq max_2$
- $\beta\gamma$ configuration, where $min_2 < min_1+d \leq max_2 \wedge max_2 < max_1+d$

Fig 4 gives graphical view for the four configurations.

The satisfiable region V' in each of them is as following:

- $\alpha\beta$ configuration
$$V' = \int_{x=\min_2-d}^{\max_1} f(x) \int_{y=\min_2}^{x+d} g(y)dydx$$
- $\alpha\gamma$ configuration
$$V' = \int_{x=\max_2-d}^{\max_1} f(x) \int_{y=\min_2}^{\max_2} g(y)dydx + \int_{x=\min_2-d}^{\max_2-d} f(x) \int_{y=\min_2}^{x+d} g(y)dydx$$
- $\beta\beta$ configuration
$$V' = \int_{x=\min_1}^{\max_1} f(x) \int_{y=\min_2}^{x+d} g(y)dydx$$
- $\beta\gamma$ configuration
$$V' = \int_{x=\max_2-d}^{\max_1} f(x) \int_{y=\min_2}^{\max_2} g(y)dydx + \int_{x=\min_1}^{\max_2-d} f(x) \int_{y=\min_2}^{x+d} g(y)dydx$$

Therefore, the satisfaction probability of c^+ is:

$$SP\big|_{c^+} = \frac{V'}{V} = \frac{\int_{x=MAX(\min_1,\min_2-d)}^{\max_1} f(x) \int_{y=\min_2}^{MIN(x+d,\max_2)} g(y)dydx}{\int_{\min_1}^{\max_1} f(x)dx \cdot \int_{\min_2}^{\max_2} g(y)dy}$$

□

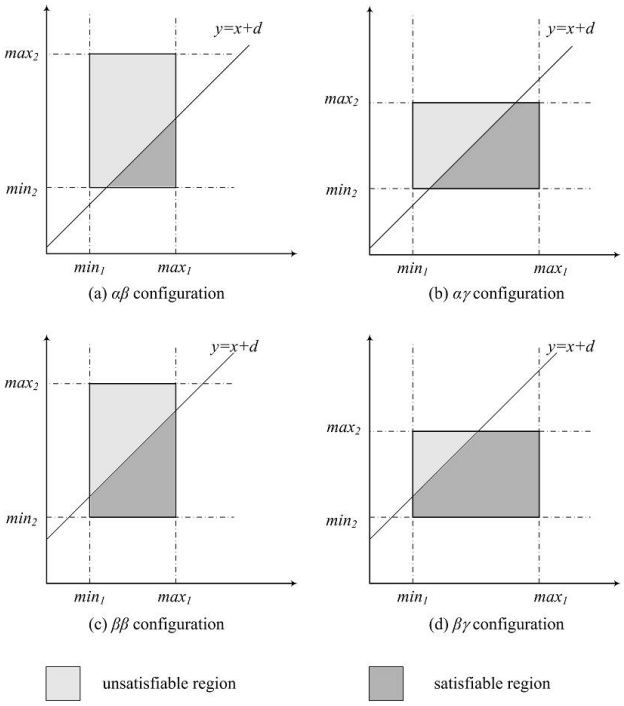


Fig 4. Four non-trivial configurations.

4.2. Earliest Prediction Time (EPT)

Theorem 2 Given a timing constraint c^+: $I_1+d \geq I_2$ *with P* where $d \geq 0$, and $f(x)$, $g(y)$ are the probability density functions of independent event occurrences on interval I_1 and I_2, respectively. The earliest prediction time of a constraint violation from a time point t for c^+, $EPT(t)\big|_{c^+}$, where t is time point at which the event corresponding to I_1 occurs, is given by the expression:

$$EPT(t)\big|_{c^+} = t' - t \tag{2}$$

where

$$t' = \inf_{s \in I_2} \left\{ s \,\middle|\, \int_{\min_2 \leq s \leq MIN(t+d,\max_2)}^{MIN(t+d,\max_2)} g(y)dy \Big/ \int_{\min_2}^{\max_2} g(y)dy \leq P \right\}$$

Proof:

Since $X \in I_1$ and $Y \in I_2$ are independent, the conditional distribution function of $Y \in I_2$ when $X \in I_1$ occurs is $g(y)$.

Based on Definition 4, a violation occurs when the event corresponding to I_2 does not occur by $s < t+d$ and the remaining satisfaction probability of the timing constraint is less than the specified confidence threshold. This is shown in the following figure:

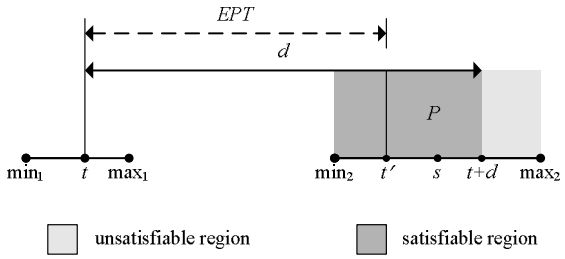


Fig 5. EPT of a deadline constraint violation.

As illustrated in the figure, although the event corresponding to I_2 could occur during the interval (s, $t+d$] that satisfies $I_1+d \geq I_2$, the specified confidence threshold P is violated. That is:

$$\exists s \in [min_2, MIN(t+d, max_2)]$$

$$s.t.$$

$$\int_s^{MIN(t+d,\max_2)} g(y)dy \Big/ \int_{\min_2}^{\max_2} g(y)dy \leq P$$

Thus, the inferior t' of such a set of s is the earliest time by which we can safely claim that c^+ is violated in case the event corresponding to I_2 does not occur.

□

It is worth pointing out that in Fig 5, t' is drawn at the point where the integration *equals* P. However, for many probability density functions, especially discrete ones, we may not find the point that *equals* P. We draw t' at the point because all the probability density functions we are dealing with in the following sections are contiguous.

5. Poisson Failure Model

As shown in Section 3.1, the random variable X representing the time between successive arrivals of transient failures (which can be viewed as events) follows an exponential distribution with parameter λ. Moreover, as shown in Section 3.3 and the following example, monitoring timing constraints between failures can actually be mapped to monitoring interval-based timing constraints discussed in the previous section.

Example 2 Revisited To map the failure monitoring problem to the interval-based timing constraint problem, we do the following reductions: since transient failures on M_1 and M_2 follow exponential distribution on intervals [t_{11}, t_{12}] and [t_{21}, t_{22}], respectively, and failure occurrences on M_1 and M_2 are independent, we can model failures on the two machine as independent event occurrences over two intervals. Moreover, the required delay between the two failures on the two machines can be mapped to monitoring if no failure occurs before a given deadline, that is, whenever a deadline constraint monitor says "yes", a delay constraint monitor says "no" and vice versa.[2] From Section 3.1, we know that when a failure follows Poisson distribution with an average arrival rate λ, the time between successive failure arrivals follows exponential distribution. Therefore, given a timing constraint of failures c^+: $I_1+d \geq I_2$ *with* P, we assume the probability density functions of failure occurrence in I_1 and I_2 are:

$$f(x) = \lambda_1 e^{-\lambda_1(x-\min_1)}, \lambda_1 > 0 \; and \; g(y) = \lambda_2 e^{-\lambda_2(y-\min_2)}, \lambda_2 > 0$$

respectively, where $min_1 = t_{11}$ and $min_2 = t_{21}$. Furthermore, without loss of generality, we assume that the individual cumulative distributions of failures over the intervals are the same, i.e.,

$$\int_{\min_1}^{\max_1} \lambda_1 e^{-\lambda_1(x-\min_1)} dx = \int_{\min_2}^{\max_2} \lambda_1 e^{-\lambda_1(x-\min_2)} dy$$

The joint density function of failures over I_1 and I_2 is illustrated in Fig 6 which is derived from Fig 2 and Fig 3.

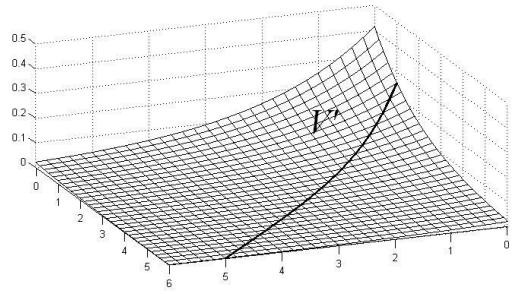


Fig 6. Joint density function of independent failure occurrences.

Therefore, the problem of monitoring timing constraint of transient failure is reduced to the problem of monitoring interval-based timing constraints. In Section 4, we presented analytical results for interval-based timing constraint satisfaction probabilities and earliest prediction time under *arbitrary* probability density functions. It may seem that the failure monitoring problem is nothing more than a special case of general interval base timing constrain monitoring problem. However, as shown in the following subsections, Poisson failure model offers many good properties which help us to improve the effectiveness of monitoring.

5.1. Satisfaction Probability (SP)

[2] Note that although it is more convenient to directly use delay constraint to model timing constraint between failures, we stick to deadline constraint instead in order to be consistent with the discussions in [10].

Substitute $f(x)$ and $g(y)$ in (1) with $\lambda_1 e^{-\lambda_1(x-\min_1)}$ and $\lambda_2 e^{-\lambda_2(y-\min_2)}$, respectively, we get the expression for the satisfaction probability under Poisson failure model:

$$\frac{\int_{x=MAX(\min_1,\min_2-d)}^{\max_1}\lambda_1 e^{-\lambda_1(x-\min_1)}\int_{y=\min_2}^{MIN(x+d,\max_2)}\lambda_2 e^{-\lambda_2(y-\min_2)}dydx}{(1-e^{-\lambda_1 len_1})(1-e^{-\lambda_2 len_2})} \quad (3)$$

Given a specific deadline timing constraint c^+: $I_1+d \geq I_2$, although it is possible for us to directly calculate the satisfaction probability using (3), sometimes, for the following reasons, this is not done:

1. For certain probability density functions (such as normal distribution) which do not have elementary antiderivatives, we may not be able to obtain an analytical representation of the integration as we do in (4) below. In this case, we can only use numerical integrations to calculate the satisfaction probability instead. However, this is time consuming and not always computationally permissible in a real-time environment.
2. Even in the case where an analytical representation of the integration can be obtained as in (4), calculating the specific value of the satisfaction probability consumes CPU cycles and introduces additional overhead to the system.

Therefore, it is desirable to derive some bounds on the satisfaction probabilities under certain configurations. For example, given a deadline timing constraint c^+: *(0, 3)+5 ≥ (6, 10) with 60%*, where the probability density functions of transient failure occurrences in I_1 and I_2 are exponentially distributed, by Theorem 3 below, we can safely claim that this constraint specification is not satisfiable without the need to even actually calculate the satisfaction probability.

Let us consider the four non-trivial cases for constraint c^+: $I_1+d \geq I_2$ *with P* in detail.

5.1.1. αβ Configuration. As defined in the previous section, under $\alpha\beta$ configuration, min_1+d and max_1+d are bounded by $min_1+d \leq min_2$ and $min_2 < max_1+d \leq max_2$, respectively, as shown in Fig 7.

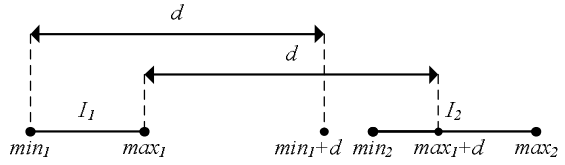


Fig 7. configuration

Therefore, the satisfaction probability under $\alpha\beta$ configuration can be simplified as:

$$SP|_{c^+} = \frac{\int_{x=\min_2-d}^{\max_1}\lambda_1 e^{-\lambda_1(x-\min_1)}\int_{y=\min_2}^{x+d}\lambda_2 e^{-\lambda_2(y-\min_2)}dydx}{(1-e^{-\lambda_1 len_1})(1-e^{-\lambda_2 len_2})} \quad (4)$$

$$= \frac{e^{-\lambda_1 len_1}[\lambda_2(e^{\lambda_1(\max_1+d-\min_2)}-1)+\lambda_1(e^{-\lambda_2(\max_1+d-\min_2)}-1)]}{(\lambda_1+\lambda_2)(1-e^{-\lambda_1 len_1})(1-e^{-\lambda_2(\max_2-\min_2)})}$$

With this analytical representation of the satisfaction probability, we can study how the satisfaction probability changes when max_2 or min_2 changes under $\alpha\beta$ configuration. This is given by the following lemma.

Lemma 1 If a timing constraint c^+: $I_1+d \geq I_2$ *with P* is in $\alpha\beta$ configuration, where transient failure occurrences on I_1 and I_2 are exponentially distributed, the satisfaction probability of the constraint increases when either max_2 or min_2 decreases.

Proof:

The first part of the lemma trivially holds since max_2 only appears on the denominator of the fraction.

To prove the second part, we can view the expression for $SP|_{c^+}$ under $\alpha\beta$ configuration as a function for min_2 and compute the partial derivative of the function on min_2:

$$\frac{\partial SP|_{c^+}}{\partial \min_2}$$

$$= \frac{\partial}{\partial \min_2}\left(\frac{e^{-\lambda_1(\max_1-\min_1)}[\lambda_2(e^{\lambda_1(\max_1+d-\min_2)}-1)+\lambda_1(e^{-\lambda_2(\max_1+d-\min_2)}-1)]}{(\lambda_1+\lambda_2)(1-e^{-\lambda_1(\max_1-\min_1)})(1-e^{-\lambda_2(\max_2-\min_2)})}\right)$$

$$= \frac{\lambda_2(\lambda_1+\lambda_2)(1-e^{-\lambda_1 len_1})e^{-\lambda_1 len_1}e^{-\lambda_2 len_2}}{[(\lambda_1+\lambda_2)(1-e^{-\lambda_1 len_1})(1-e^{-\lambda_2 len_2})]^2}\cdot$$

$$\left[(\lambda_1+\lambda_2)[e^{\lambda_1(\max_1+d-\min_2)}-1]-\lambda_1[e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)}-1]\left(\frac{e^{\max_2-\min_2}}{e^{\max_1+d-\min_2}}\right)^{\lambda_2}\right]$$

Since

$$\frac{\lambda_2(\lambda_1+\lambda_2)(1-e^{-\lambda_1 len_1})e^{-\lambda_1 len_1}e^{-\lambda_2 len_2}}{[(\lambda_1+\lambda_2)(1-e^{-\lambda_1 len_1})(1-e^{-\lambda_2 len_2})]^2} > 0 \quad (5)$$

It suffices to prove that

$$(\lambda_1+\lambda_2)[e^{\lambda_1(\max_1+d-\min_2)}-1]-\lambda_1[e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)}-1]\left(\frac{e^{\max_2-\min_2}}{e^{\max_1+d-\min_2}}\right)^{\lambda_2} < 0 \quad (6)$$

Note that under $\alpha\beta$ configuration,

$$min_2 < max_1 + d \leq max_2 \Rightarrow \left(\frac{e^{\max_2-\min_2}}{e^{\max_1+d-\min_2}}\right)^{\lambda_2} \geq 1 \quad (7)$$

Given that

$$\frac{\partial\left[(\lambda_1+\lambda_2)[e^{\lambda_1(\max_1+d-\min_2)}-1]-\lambda_1[e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)}-1]\right]}{\partial \min_2} \quad (8)$$

$$= -\lambda_1(\lambda_1+\lambda_2)e^{\lambda_1(\max_1+d-\min_2)}+\lambda_1(\lambda_1+\lambda_2)e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)} > 0$$

and

$$\left[(\lambda_1+\lambda_2)[e^{\lambda_1(\max_1+d-\min_2)}-1]-\lambda_1[e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)}-1]\right]\Big|_{\min_2=\max_1+d} = 0 \quad (9)$$

It follows from (8) and (9) that

$$\left[(\lambda_1+\lambda_2)[1-e^{-\lambda_1(\max_1+d-\min_2)}]-\lambda_1[1-e^{-(\lambda_1+\lambda_2)(\max_1+d-\min_2)}]\right]\Big|_{\min_2<\max_1+d} < 0 \quad (10)$$

Therefore, from (7) and (10), we have

$$(\lambda_1+\lambda_2)[e^{\lambda_1(\max_1+d-\min_2)}-1]-\lambda_1[e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)}-1]\left(\frac{e^{\max_2-\min_2}}{e^{\max_1+d-\min_2}}\right)^{\lambda_2} \quad (11)$$

$$\le(\lambda_1+\lambda_2)[e^{\lambda_1(\max_1+d-\min_2)}-1]-\lambda_1[e^{(\lambda_1+\lambda_2)(\max_1+d-\min_2)}-1]<0$$

where $min_2 < max_1+d \le max_2$, which is a necessary condition for $\alpha\beta$ configuration. Equation (11), together with (5), implies that:

$$\frac{\partial\, SP|_{c^+}}{\partial \min_2}<0 \quad (min_2 < max_1 + d \le max_2)$$

Thus, under $\alpha\beta$ configuration, the satisfaction probability of the constraint increases when min_2 decreases.

□

Theorem 3 If a deadline timing constraint c^+: $I_1+d \ge I_2$ *with* P is in $\alpha\beta$ configuration, where transient failure occurrences on I_1 and I_2 are exponentially distributed, the satisfaction probability reaches its maximum when $min_2 = min_1+d \wedge max_2 = max_1+d$. Under the assumption that the individual cumulative distributions over the two intervals are the same, this maximum is 50%.

Proof:

The first part of this theorem immediately follows from Lemma 1 and the necessary condition for $\alpha\beta$ configuration, i.e., $min_1+d \le min_2 \wedge min_2 < max_1+d \le max_2$.

To compute the corresponding satisfaction probability, take $min_2 = min_1+d$, $max_2 = max_1+d$, and $len_2 = max_2-min_2 = len_1$ in (4):

$$MAX(SP|_{c^+})=\frac{[\lambda_1 e^{-(\lambda_1+\lambda_2)len_1}-(\lambda_1+\lambda_2)e^{-\lambda_1 len_1}+\lambda_2]}{(\lambda_1+\lambda_2)(1-e^{-\lambda_1 len_1})(1-e^{-\lambda_2 len_1})} \quad (12)$$

Given that the individual cumulative distributions over the intervals are the same, we have:

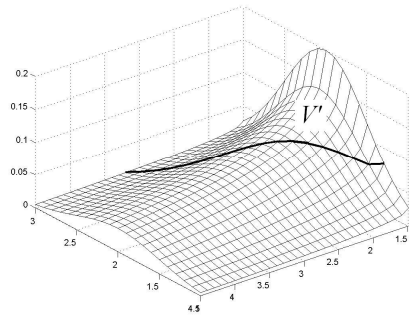
$$1-e^{-\lambda_1 len_1}=1-e^{-\lambda_2 len_2}\Rightarrow\lambda_1=\lambda_2$$

As a result, (12) can be simplified as:

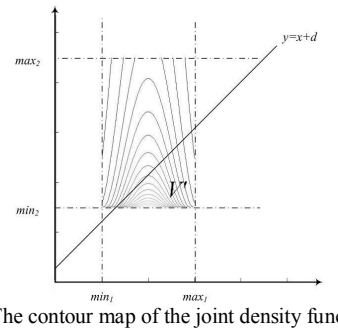
$$MAX(SP|_{c^+})=\frac{(\lambda_1 e^{-2\lambda_1 len_1}-2\lambda_1 e^{-\lambda_1 len_1}+\lambda_1)}{2\lambda_1(1-e^{-\lambda_1 len_1})^2}=\frac{1}{2}$$

□

Therefore, the satisfaction probability of timing constraint of transient failures c^+ under $\alpha\beta$ configuration is less than or equal to 50%. It is worth noting that this theorem may not hold for arbitrary probability density functions. Consider a scenario: the event occurrence on I_1 is normally distributed while the event occurrence on I_2 is exponentially distributed. Fig 8 (a) shows the joint density function of event occurrences with the bold line indicating its intersection with the plane $y = x+d$. It is clear from the contour map in (b) that the upper bound of the satisfaction probability under $\alpha\beta$ configuration could be much larger than 50% since the contour lines are much denser in the region $y \le x+d$ than in $y > x+d$.



(a) Joint density function



(b) The contour map of the joint density function.

Fig 8. Two events with normal distribution on I_1 and exponential distribution on I_2.

Theorem 3 is particularly useful when given a deadline timing constraint c^+: $I_1+d \ge I_2$ *with* P, where both failure occurrences are of exponential distribution, if I_1 and I_2 are in $\alpha\beta$ configuration (constant time decidable), and $P > 50\%$, we can safely claim that this constraint specification is not satisfiable with the required confidence and no run-time monitoring is even needed. Therefore, if we are given a timing constraint of transient failures of the form c^+: *(0, 3)+5* $\ge$ *(6, 10) with 60%*, since this constraint is intrinsically unsatisfiable, it means that the required degree of failure tolerance cannot be guaranteed.

5.1.2. $\beta\gamma$ configuration. Similar lemma and theorem hold for βγ configuration:

Lemma 2 If a timing constraint c^+: $I_1+d \ge I_2$ *with* P is in $\beta\gamma$ configuration, where transient failure occurrences on I_1 and I_2 are exponentially distributed, the satisfaction probability of the constraint decreases when either max_2 or min_2 decreases.

□

Theorem 4 If a timing constraint c^+: $I_1+d \ge I_2$ *with* P is in $\beta\gamma$ configuration, with failure occurrences on I_1 and I_2 exponentially distributed, the satisfaction probability approaches its minimum when $min_2 \rightarrow min_1+d \wedge max_2 \rightarrow max_1+d$. Under the assumption

that the individual cumulative distributions over the two intervals are the same, this minimum is 50%.

□

For remaining two non-trivial configurations, i.e., $\alpha\gamma$ and $\beta\beta$ configurations, the following observations can be made:

1. From Fig 4(b), it is easy to see that under the $\alpha\gamma$ configuration, the satisfaction probability approaches 100% when $max_2 \rightarrow min_2 \wedge min_2 \rightarrow min_1+d$ and moves toward 0% when $max_2 \rightarrow max_1+d \wedge min_2 \rightarrow max_2$.
2. Likewise, it is easy to see from Fig 4(c) that under the $\beta\beta$ configuration, the satisfaction probability approaches 100% when $max_2 \rightarrow max_1+d \wedge min_2 \rightarrow -\infty$ and moves toward 0% when $max_2 \rightarrow +\infty \wedge min_2 \rightarrow min_1+d$.

Therefore, for $\alpha\gamma$ and $\beta\beta$ configurations, no general upper bounds or lower bounds can be proven.

5.2. Earliest Prediction Time (EPT)

As the probability distribution for failure occurrence is exponentially distributed within time intervals, it is clear that the cumulative distribution on the shorter interval [*min*, *min* +(*max*−*min*)/*n*] which is $1-e^{-\lambda(max-min)/n}$ does not differ much from the cumulative distribution on the entire interval [*min, max*] which is $1-e^{-\lambda(max-min)}$, since

$$\frac{1-\varepsilon^{1/n}}{1-\varepsilon} = \frac{1}{1+\varepsilon^{1/n}+\cdots+\varepsilon^{(n-1)/n}} >> \frac{1}{n} \text{ where } \varepsilon = e^{-\lambda(\max-\min)} << 1$$

In other words, failures will be more likely to occur in the first *1/n* subinterval than in the last (*n*−*1*)/*n*. This observation provides a basis for shortening the earliest prediction time and thus improving the effectiveness of run-time monitoring.

Since the probability density function of exponential distribution is contiguous,

$$t' = \inf_{s \in I_2}\left\{ s \middle| \int_{\min_2 \le s \le MIN(t+d,\max_2)}^{MIN(t+d,\max_2)} g(y)dy \Big/ \int_{\min_2}^{\max_2} g(y)dy \le P \right\}$$

in (2) is the solution of

$$\frac{\int_{s}^{MIN(t+d,\max_2)} g(y)dy}{\int_{\min_2}^{\max_2} g(y)dy} = P \quad when \frac{\int_{\min_2}^{MIN(t+d,\max_2)} g(y)dy}{\int_{\min_2}^{\max_2} g(y)dy} > P$$

and is

$$min_2 \quad when \int_{\min_2}^{MIN(t+d,\max_2)} g(y)dy \Big/ \int_{\min_2}^{\max_2} g(y)dy \le P$$

Therefore, according to Theorem 2, the earliest prediction time of a deadline timer from a time point *t* for c^+: $I_1+d \ge I_2$ *with P*, where failure occurrences on I_1 and I_2 are exponentially distributed, is given as:

$$EPT(t)\big|_{c^+} = MAX(t', min_2) - t$$

where

$$t' = \begin{cases} min_2 + \ln\dfrac{1}{P+(1-P)e^{-\lambda_2 len_2}} & , t+d \in (max_2, +\infty) \\ min_2 + \dfrac{1}{\lambda_2}\ln\dfrac{1}{e^{-\lambda_2(t+d-\min_2)} + P(1-e^{-\lambda_2 len_2})} & , t+d \in (min_2, max_2] \end{cases} \quad (13)$$

Example 3 Consider a deadline timing constraint between failures: c^+: (*0, 3*)+*8* ≥ (*4, 9*) with 40%. Suppose a failure occurs on machine M_1 at time point 2, that is, the deadline timer is set at the time. The earliest prediction times in the cases of uniform distribution and exponential distribution are illustrated in Fig 9:

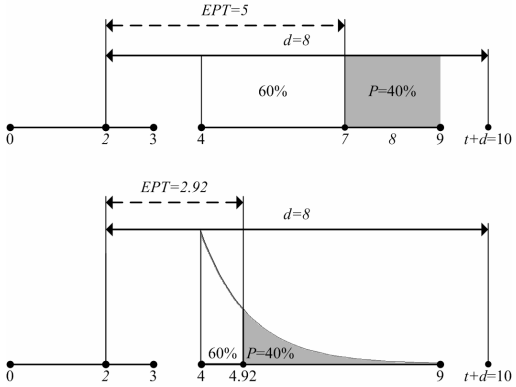


Fig 9. Earliest prediction times in the cases of uniform and exponential distribution.

1. If the failure occurrence on machine M_2 is uniformly distributed, the earliest time we can safely claim that the constraint is violated is 7 since (9−7)/(9−4) = 40%, and thus *EPT* is 7−2 = 5.
2. If the failure occurrence on machine M_2 is exponentially distributed with $\lambda_2 = 1$. According to (13), the earliest time we can safely claim that the deadline constraint is violated is $t' = 4 + \ln\dfrac{1}{0.4+(1-0.4)\times e^{-5}} \approx 4.92$ and thus *EPT* is 4.92−2 = 2.92.

□

This example shows that if transient failure occurrences are exponentially distributed, the earliest prediction time is earlier than in uniformly distributed failure occurrences in certain cases. It is thus possible for us to give earlier warnings for potential constraint violations. In Section 7, we will give analytical comparisons of the two cases.

6. Uniform Failure Model

To simplify modeling, uniform distributions are also

used to model event occurrence sometimes. To facilitate comparisons of interval-based timing constraints with different failure models in the next section, we give the results on the satisfaction probability and earliest prediction time for timing constraints based on intervals with uniformly distributed failure occurrences. This corresponds to the case where failures are equally likely to happen within a given period. Similar to the previous section, the monitoring of failures under the uniform assumption is also a special case of the theorems presented in Section 4. Our discussion is based on the following assumptions:

Given a timing constraint c^+: $I_1+d \geq I_2$ *with* P, we assume the probability density functions of event occurrence in I_1 and I_2 are:

$$f(x) = \frac{1}{max_1 - min_1} = \frac{1}{len_1} \text{ and } g(y) = \frac{1}{max_2 - min_2} = \frac{1}{len_2}$$

respectively.

6.1. Satisfaction Probability (SP)

Substitute $f(x)$ and $g(y)$ in Theorem 1 with $1/len_1$ and $1/len_2$, we get the expression for the satisfaction probability under uniform distribution:

$$\begin{aligned} SP|_{c^+} &= \int_{x=MAX(\min_1,\min_2-d)}^{\max_1} \frac{1}{len_1} \int_{y=\min_2}^{MIN(x+d,\max_2)} \frac{1}{len_2} dy dx \\ &= \frac{1}{len_1 \cdot len_2} \int_{x=MAX(\min_1,\min_2-d)}^{\max_1} MIN(x+d-min_2, len_2) dx \\ &= \frac{1}{len_1 \cdot len_2} \int_{x=\min_1}^{\max_1} MIN(MAX(x+d-min_2),0), len_2) dx \end{aligned}$$

which is Lee's Theorem 1 given in [10].

It is not hard to prove that if a timing constraint c^+: $I_1+d \geq I_2$ *with* P is in $\alpha\beta$ configuration, where failure occurrences on I_1 and I_2 are uniformly distributed, the satisfaction probability of the constraint increases when either max_2 or min_2 decreases and thus the maximum satisfaction probability under $\alpha\beta$ configuration is 50%. Similar arguments hold for $\beta\gamma$ configuration.

6.2. Earliest Prediction Time (EPT)

According to Theorem 2, the earliest prediction time of a deadline timer from a time point t for c^+: $I_1+d \geq I_2$ *with* P, where failure occurrences on I_1 and I_2 are uniformly distributed, is given as:

$$EPT(t)|_{c^+} = MAX(t', min_2) - t \qquad (14)$$

where

$$t' = \begin{cases} min_2 + (1-P)len_2 & ,t+d \in (max_2, +\infty) \\ t+d-P \cdot len_2 & ,t+d \in (min_2, max_2] \end{cases}$$

7. Comparisons on Different Failure Models

Equation (13) and (14) give the analytical representation of the earliest prediction times of exponentially distributed and uniformly distributed failure occurrences, respectively, and thus provide a basis for us to compare the *EPT*'s for the two cases. The one with a smaller *EPT* gives earlier constraint violation predictions.

Lemma 3 Given a deadline timing constraint c^+: $I_1+d \geq I_2$ *with* P where $d \geq 0$, and $g(y)$ is the probability density functions of independent failure occurrences on interval I_2. The earliest prediction time of failure occurrence on interval I_2 with exponential distribution approaches that of uniform distribution when

$$\int_{\min_2}^{\max_2} g_{\exp}(y) dy = 1 - e^{-\lambda_2 len_2} \to 0 \cdot$$

Proof:

To compare the earliest prediction time for the two failure occurrences with exponential distribution and uniform distribution, we only need to compare the corresponding t' in (13) and (14), that is:

$$t'_{exp} = \begin{cases} min_2 + \frac{1}{\lambda_2} \ln \frac{1}{P + (1-P)e^{-\lambda_2 len_2}} & ,t+d \in (max_2, +\infty) \\ min_2 + \frac{1}{\lambda_2} \ln \frac{1}{e^{-\lambda_2 (t+d-min_2)} + P(1-e^{-\lambda_2 len_2})} & ,t+d \in (min_2, max_2] \end{cases}$$

and

$$t'_{uniform} = \begin{cases} min_2 + (1-P)len_2 & ,t+d \in (max_2, +\infty) \\ t+d-P \cdot len_2 & ,t+d \in (min_2, max_2] \end{cases}$$

in the cases of exponential distribution and uniform distribution, respectively.

When $t+d \in (max_2, +\infty)$, note that,

$$\begin{aligned} &\lim_{\lambda_2 len_2 \to 0} \frac{t'_{exp} - min_2}{t'_{uniform} - min_2} \\ &= \lim_{\lambda_2 len_2 \to 0} \frac{1}{\lambda_2 (1-P) len_2} \ln \frac{1}{P + (1-P)e^{-\lambda_2 len_2}} \quad \left(\lim_{x \to 0} e^{-x} = 1 - x\right) \\ &= \lim_{\lambda_2 len_2 \to 0} \frac{1}{(1-P)\lambda_2 len_2} \ln \frac{1}{1 - \lambda_2 len_2 + P\lambda_2 len_2} \\ &= \lim_{\lambda_2 len_2 \to 0} \ln [1 + (P-1)\lambda_2 len_2]^{\frac{1}{(P-1)\lambda_2 len_2}} \quad \left(\lim_{x \to \infty} (1 + 1/x)^x = e\right) \\ &= \ln e = 1 \end{aligned}$$

Similar proof can be obtained when $t+d \in (min_2, max_2]$. Therefore, the earliest prediction time of failure occurrence with exponential distribution equals that of failure occurrence with uniform distribution when $\lambda_2 len_2 \to 0$, that is, $1 - e^{-\lambda_2 len_2} \to 0$.

□

As can be seen from the above lemma, the case of uniform distribution only gives a prediction of violation as early as exponential distribution when the cumulative distribution of the exponential distribution over the second interval approaches 0. This is a very

rare and even impossible case. In general, an earlier prediction of violation can be obtained in applications with exponential failure occurrences under certain conditions:

Theorem 5 The earliest prediction time of failure occurrence with exponential distribution is always smaller than that of failure occurrence with uniform distribution when $t+d \in (T, +\infty)$, where

$$T = min_2 + \frac{1}{\lambda_2} \ln \frac{e^{P \cdot \lambda_2 len_2} - 1}{P(1 - e^{-\lambda_2 len_2})} \in [min_2, max_2]$$

Proof:

Case 1: $t+d \in (max_2, +\infty)$

$$t'_{exp} - t'_{uniform} = \frac{1}{\lambda_2} \ln \frac{1}{P + (1-P)e^{-\lambda_2 len_2}} - (1-P) len_2$$

To prove $t'_{exp} < t'_{uniform}$, it suffices to prove:

$$\ln \frac{1}{P + (1-P)e^{-\lambda_2 len_2}} - (1-P)\lambda_2 len_2 < 0$$

Suppose $x = \lambda_2 len_2 \in (0, +\infty)$, and we have:

$$F(x) = \ln \frac{1}{P + (1-P)e^{-x}} - (1-P)x$$

Compute the derivative of $F(x)$, we obtain:

$$F'(x) = \frac{(1-P)e^{-x}}{P + (1-P)e^{-x}} - (1-P) = \left[\frac{e^{-x}}{e^{-x} + P(1-e^{-x})} - 1 \right](1-P) < 0$$

Thus, $F(x)$ is monotonically decreasing. Moreover, from Lemma 3 we have $F(x) \to 0$ when $x = \lambda_2 len_2 \to 0$. Thus, we have:

$$F(x) < 0 \quad \Rightarrow \quad t'_{exp} < t'_{uniform}$$

Case 2: $t+d \in (min_2, max_2]$

$$t'_{\exp} - t'_{uniform} = min_2 - \frac{1}{\lambda_2} \ln \frac{1}{e^{-\lambda_2 (t+d-\min_2)} + P(1 - e^{-\lambda_2 len_2})} - t - d + (1-P) len_2 \quad (15)$$

Suppose $x = t+d-min_2$, $x \in (0, len_2]$, and (15) becomes

$$F(x) = \frac{1}{\lambda_2} \ln \frac{1}{e^{-\lambda_2 x} + P(1 - e^{-\lambda_2 len_2})} - x + (1-P) len_2$$

Note that $F(x) = 0$ when

$$x = min_2 + \frac{1}{\lambda_2} \ln \frac{e^{P \cdot \lambda_2 len_2} - 1}{P(1 - e^{-\lambda_2 len_2})} = T - min_2 \quad (16)$$

Compute the derivative of $F(x)$, we obtain

$$F'(x) = \frac{e^{-\lambda_2 x}}{e^{-\lambda_2 x} + P(1 - e^{-\lambda_2 len_2})} - 1 < 0$$

Thus, $F(x)$ is monotonically decreasing and from (16), we have:

$$\begin{cases} F(x) < 0 & when\ x \in (T - min_2, len_2] \\ F(x) \geq 0 & when\ x \in (0, T - min_2] \end{cases}$$

Therefore,

$$\begin{cases} t'_{\exp} < t'_{uniform} & when\ t + d \in (T, max_2] \\ t'_{\exp} \geq t'_{uniform} & when\ t + d \in (min_2, T] \end{cases}$$

Combining the two cases, we have that

$$\begin{cases} t'_{\exp} < t'_{uniform} & when\ t + d \in (T, +\infty) \\ t'_{\exp} \geq t'_{uniform} & when\ t + d \in (min_2, T] \end{cases}$$

where

$$T = min_2 + \frac{1}{\lambda_2} \ln \frac{e^{P \cdot \lambda_2 len_2} - 1}{P(1 - e^{-\lambda_2 len_2})}$$

□

Note that T is a constant independent of $t+d$ when the timing constraint is given, that is, when P, λ_2, and len_2 are set.

8. Conclusion

In distributed embedded systems, it is crucial to monitor the timing relationship between a pair of independent failures occur on different machines. We focus on the problem of determining the satisfaction probability (SP) of transient failure timing constraints under a Poisson failure model. Moreover, we look into the problem of predicting at run-time the earliest time (EPT) we can claim a violation of a constraint. We first give general case satisfaction analysis for interval-based timing constraints where event occurrences over intervals have arbitrary probability density functions. We further present a detailed study on transient failure timing constraints with a Poisson failure model. Our analysis shows that there are upper bounds and lower bounds on the satisfaction probabilities under certain constraint configurations in this failure model. Furthermore, we present results and compare the earliest prediction times for different failure models: the EPT is smaller for exponentially distributed failures than that of uniformly distributed failures under some circumstances.

More specifically, we are able to use the general theories of interval-based timing constraints to show in this paper the following important properties regarding transient failure timing constraints:

1. Given a timing constraint between transient failure c^+: $I_1 + d \geq I_2$ *with* P, if it is in $\alpha\beta$ configuration and failure occurrences on intervals I_1 and I_2 are exponentially distributed, the satisfaction probability (SP) of the constraint increases when either min_2 or max_2, the minimum or maximum possible time the second failure could occur, decreases. Moreover, the satisfaction probability of the timing constraint in $\alpha\beta$ configuration can be no larger than 50%.
2. The earliest prediction time (EPT) of failure occurrence with exponential distribution is smaller than that of failure occurrence with uniform distribution when $t + d \in (T, +\infty)$, and is larger when $t + d \in (min_2, T]$, where

$$T = min_2 + \frac{1}{\lambda_2} \ln \frac{e^{P \cdot \lambda_2 len_2} - 1}{P(1 - e^{-\lambda_2 len_2})} \in [min_2, max_2]$$

It should be noted that in hard timing constraint settings, a real-time system should be designed taking worst-case scenarios into consideration. However, under soft timing constraints settings, when certain hard timing constraints may be intrinsically infeasible to satisfy, the interval base timing constraints with confidence threshold smaller than 100% can be applied as a relaxation. Moreover, while this paper considers two event occurrences to be monitored, this approach can be extended to a distributed system with multiple events in two ways:

1. In [10], the all-pairs shortest-path algorithm is extended to facilitate the derivation of implicit constraints from a set of timing constraints among multiple events. And the earliest prediction times derived in this paper can be well fit in to the algorithm.
2. As argued in [16], in some distributed embedded applications (e.g., distributed voting in sensor networks), the probabilistic timing behavior of a group of events can be derived from those of individual events. This allows us to study the interval based timing constraints between multiple groups of events which raises the granularity of analyses and reduces the complexity to a great extent.

9. References

[1] L. Abeni and G. Buttazzo, "Qos guarantee using probabilistic deadlines", in *Proc. of the 11th Euromicro Conference on Real-Time Systems*, 1999, pp. 242–249.

[2] X. Castillo, S. McConnel, and D. Siewiorek, "Derivation and caliberation of a transient error reliability model", *IEEE Transactions on Computer*, 31(7):658-671, 1982.

[3] S. Chodrow, F. Jahanian, and M. Donner, "Run-time monitoring of real time systems", in *Proc. of the 12th IEEE Real-Time Systems Symposium*, 1991, pp. 74–83.

[4] L. David and I. Puaut, "Static determination of probabilistic execution times", in *Proc. of the 16th Euromicro Conference on Real-Time Systems*, 2004, pp. 223–230.

[5] J. L. D´ıaz, D. F. Garc´ıa, K. Kim, C.-G. Lee, L. L. Bello, J. M. L´opez, S. L. Min, and O. Mirabella, "Stochastic analysis of periodic real-time systems", in *Proc. of the 23rd IEEE Real-Time Systems Symposium*, 2002, pp. 289–300.

[6] X. S. Hu, T. Zhou, and E. H.-M. Sha, "Estimating probabilistic timing performance for real-time embedded systems", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 833–853, 2001.

[7] R.K. Iyer and D.j.Rossetti, "A measurement-based model for workload dependence of CPU errors", *IEEE Transactions on Computers*, 33:518-528, 1984.

[8] R.K. Iyer, D.j. Rossetti, and M.C. Hsueh, "Measurement and modeling of computer reliability as affected by system activity", *ACM Transactions on Computer Systems*, 4(3):214-237, Aug. 1986.

[9] F. Jahanian, R. Rajkumar, and S. Raju, "Run-time monitoring of timing constraints in distributed real-time systems", University of Michigan, Technical Report CSE-TR 212–94, 1994.

[10] C.-G. Lee, A. K. Mok, and P. Konana, "Monitoring of timing constraints with confidence threshold requirements", in *Proc. of the 24th IEEE Real Time Systems Symposium*, 2003, pp. 178–187.

[11] S. Manolache, P. Eles, and Z. Peng, "Optimization of soft real-time systems with deadline miss ratio constraints", in *Proc. of the 10th IEEE Real-Time and Embedded Technology and Applications Symposium*, 2004, pp. 562–570.

[12] A. K. Mok, C.-G. Lee, H. Woo, and P. Konana, "The monitoring of timing constraints on time intervals", in *Proc. of the 23rd IEEE Real Time Systems Symposium*, 2002, pp. 191–200.

[13] A. K. Mok and G. Liu, "Efficient run-time monitoring of timing constraint", in *Proc. of the 3rd IEEE Real Time Technology and Applications Symposium*, 1997, pp. 252–262.

[14] S. Raju and R. Rajkumar, "Monitoring timing constraints in distributed real time systems", in *Proc. of the 13th IEEE Real-Time Systems Symposium*, 1992, pp. 57–67.

[15] S. Wang, J. R. Merrick, and K. G. Shin, "Component allocation with multiple resource constraints for large embedded real-time software design", in *Proc. of the 10th IEEE Real-Time and Embedded Technology and Applications Symposium*, 2004, pp. 219–226.

[16] Y. Yu and S. Ren, "Expected Time for Obtaining Dependable Data in Real-Time Environment", in *Proc. 22nd ACM Symposium on Applied Computing*, 2007, submitted.

[17] Y. Yu, W. Guan, S. Ren, and O. Frieder, "Satisfaction Probabilities of Interval-based Timing Constraints", *IEEE Transactions on Computers*, submitted.

[18] Y. Zhang and K. Chakrabarty, "Energy-aware adaptive checkpointing in embedded real-time systems", in *Proc. of IEEE/ACM Design, Automation and Test in Europe Conference (DATE)*, 2003.

[19] D. Zhu, "Reliability-Aware Dynamic Energy Management in Dependable Embedded Real-Time Systems", in *Proc. of the 12th IEEE Real-Time and Embedded Technology and Applications Symposium*, Apr. 2006.