Ergodicity and Throughput Bounds of Petri Nets with Unique Consistent Firing Count Vector

Javier Campos* Giovanni Chiola† Manuel Silva*

Abstract

This paper addresses ergodicity and throughput bounds characterizations for a subclass of timed and stochastic Petri nets, interleaving qualitative and quantitative theories. The considered nets represent an extension of the well known subclass of marked graphs, defined as having a unique consistent firing count vector, independently of the stochastic interpretation of the net model. In particular, persistent and mono-T-semiflow nets subclasses are considered. Upper and lower throughput bounds are computed using linear programming problems defined on the incidence matrix of the underlying net. The bounds proposed here depend on the initial marking and the mean values of the delays but not on the probability distributions (thus including both the deterministic and the stochastic cases). From a different perspective, the considered subclasses of stochastic nets can be viewed as special classes of synchronized queueing networks, thus the proposed bounds can be applied to these networks.

Index terms: Ergodicity, linear programming, Petri nets, structural analysis, synchronized queueing networks, throughput, upper and lower bounds.

1 Introduction

In this paper, which is an improved version of [1],[1] we study the possibility of obtaining *upper and lower bounds* on the steady-state performance of two Petri net (PN, for short) subclasses that are characterized by having a unique consistent firing count vector. Although restricted, the two net subclasses represent two different kind of generalizations of marked graphs (MGs, in what follows), thus some basic results from [2] are briefly recalled. In particular in this work we study the throughput of transitions, defined as the average number of firings per time unit. We derive results that depend only on the mean values and neither on the higher moments nor on the form of the probability distributions of the random variables that describe the timing of the system. Both *deterministic* and *stochastic* timings are covered by our bounds. In some sense this independence of the probability distribution is a useful generalization of the results, since higher moments of the delays and forms of the probability distributions are usually unknown for real cases, and difficult to estimate and assess. Another extension that becomes possible taking the bounding approach instead of the exact computation, is that we can derive bounds also in the case of *marking non-ergodic* systems.

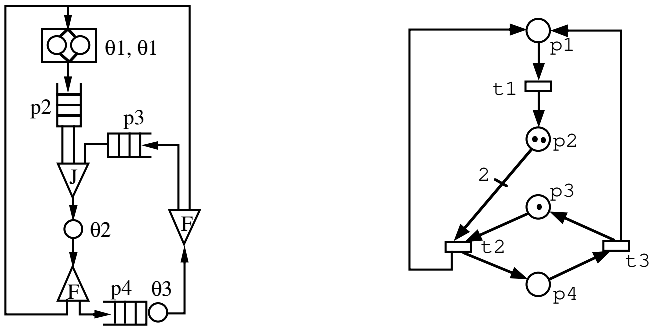
We assume the reader is familiar with the structure, firing rules, and basic properties of net models (see [3] for a nice recent survey). Let us recall some notation here: $\mathcal{N} = \langle P, T, Pre, Post \rangle$ is a net with $n = |P|$ places and $m = |T|$ transitions. If Pre and $Post$ incidence functions take values in $\{0, 1\}$, $\mathcal{N}$ is said ordinary. PRE, $POST$, and $C = POST - PRE$ are $n \times m$ matrices representing the Pre, $Post$, and global incidence functions. Vectors $Y \geq 0$, $Y^T \cdot C = 0$ ($X \geq 0$, $C \cdot X = 0$) represent P-semiflows, also called conservative components (T-semiflows, also called consistent components). M (M_0) is a marking (initial marking). Finally, σ represents a fireable sequence, while $\vec{\sigma}$ is the firing count vector associated to σ. If M is reachable from M_0 (i.e. $\exists \sigma$ s.t. $M_0[\sigma\rangle M$), then $M = M_0 + C \cdot \vec{\sigma} \geq 0$ and $\vec{\sigma} \geq 0$.

In a PN there is an obvious relation between the concepts of steady-state behaviour and that of repeatable firing sequences: sequences of transitions that are repeatable only a finite number of times cannot contribute to the steady-state performance of the model. Here we consider live bounded connected nets (thus strongly connected) which are either decision-free or such that the decision policy at effective conflicts does not change the *relative firing frequencies* of transitions in steady-state. A characteristic of these bounded nets is the existence of a unique consistent firing count vector $\vec{\sigma}_R$ associated with all marking *repetitive sequences*, i.e. $\forall M$ s.t. $M[\sigma\rangle M$: $\vec{\sigma} = k\vec{\sigma}_R, k \in \mathbb{N}$ (thus $M = M + C \cdot \vec{\sigma} \Rightarrow C \cdot \vec{\sigma} = 0$, i.e. $\vec{\sigma} = k\vec{\sigma}_R$ is a T-semiflow).

*J. Campos and M. Silva are with the Departamento de Ingeniería Eléctrica e Informática, Universidad de Zaragoza, calle María de Luna 3, E-50015 Zaragoza, Spain. This work has been supported by project PA86-0028 of the Spanish Comisión Interministerial de Ciencia y Tecnología (CICYT).

†G. Chiola is with the Dipartimento di Informatica, Università di Torino, corso Svizzera 185, I-10149 Torino, Italy. This work was performed while G. Chiola was visiting the University of Zaragoza, with the financial support of the Caja de Ahorros de la Inmaculada de Zaragoza under Programa Europa of CAI-DGA.

[1] Unless otherwise explicitly stated, all theorem proofs can be found or easily derived from results in [1].



(a) Extended QN (F: Fork; J: Join) (b) Structurally persistent net (pre(p2, t2)=2)

Figure 1: A synchronized queueing network and its PN representation (transitions = stations).

The main property of nets with unique consistent firing count vector is that their relative firing frequency vector does not depend on the stochastic interpretation (even if there exist conflicts!). Two overlapping subclasses of nets with unique consistent firing count vector are identified in this paper: persistent (behaviourally defined) and mono-T-semiflow (structurally defined) nets. Both persistent and mono-T-semiflow nets allow certain interplay between choice and concurrency. Bounded structurally persistent nets are decision-free and belong to the intersection between bounded persistent and mono-T-semiflow nets. MGs are structurally persistent nets.

From a different perspective the obtained results can be applied to the analysis of some monoclass queueing networks extended with some synchronization schemes and batch movements (e.g., see the synchronized queueing network in Figure 1.a and its PN representation).

The paper is organized as follows. In Section 2 we discuss the stochastic interpretation of nets. The liveness bound concept is also introduced. Ergodicity for the firing process must be assured, otherwise the throughput computation problem makes no sense. Firing and marking ergodicities of nets with unique consistent firing count vector are considered in Section 3. In particular weak and strong ergodicity are differentiated. After that, in Section 4 we identify two net subclasses having a unique consistent firing count vector: persistent nets (see, e.g., [4]) and the mono-T-semiflow nets subclass, which is introduced in this work. Some of their qualitative properties that can be exploited to characterize different ergodicities and to derive performance bounds are presented. Upper bounds for the throughput of nets are derived in Section 5. In the particular case of MGs (discussed in detail in [2]) the upper bounds, computed by means of proper linear programming problems, are reachable (Section 5.1). Later in Section 5.2 we propose a method to construct cyclic processes (MGs) starting from non-safe ordinary persistent models, that is an extension of a technique originally used by Ramchandani for safe persistent nets [5]. Using this extended method we derive reachable upper bounds for the case of bounded persistent nets. Section 5.3 is devoted to upper bounds for mono-T-semiflow nets. Lower bounds for the throughput of transitions are considered in Section 6. In the case of MGs, the obtained bounds are reachable. Conclusions are summarized in Section 7.

2 Stochastic interpretation of nets

In the original definition, PNs did not include the notion of time. Nevertheless the introduction of timing specification is essential if we want to use this class of models for performance evaluation of distributed systems.

2.1 Timing and firing process

Historically there have been two ways of introducing the concept of time in PN models, namely, associating a time interpretation with either *places* or *transitions*; in the latter case transitions have been defined to fire either *atomically* or in *three phases*. A more detailed discussion of the timing and firing process can be found in [6]. Since we are trying to use qualitative results derived from untimed net descriptions, we cannot change the firing mechanism at the level of the net interpretation. Hence we exclude the three-phase firing interpretation, which does change the firing mechanism when the concept of time is introduced. We simply speak of marked nets where we mean timed nets with single-phase timed transitions.

2.2 Single versus multiple server semantics: enabling and liveness bounds

Another possible source of confusion in the definition of the timed interpretation of a net model is the concept of "degree of enabling" of a transition (or re-entrance). From the queueing theory point of view [7], this can be interpreted as the number of *servers* at each station (transition). Of course an infinite server transition can always be constrained to a "k-server" behaviour by reducing its enabling bound to k, just introducing a place self-loop with k tokens around the transition. Therefore the infinite server semantics appears to be the most general one, and for this reason it is adopted in this work.

The performance of a model with infinite server semantics depends on the maximal degree of enabling of the transitions. For this reason we introduce here the concept of enabling bound:

Definition 2.1 *Enabling bound:* $\forall t \in T$, $E(t) \stackrel{\text{def}}{=} \max\{k \mid \exists M \in R(\mathcal{N}, M_0)$ *such that* $M \geq kPRE[t]\}$. $\diamond$

In particular, the steady-state performance does depend on the maximal degree of enabling of transitions in steady-state, which can be different from the maximal degree of enabling of a transition during all its evolution from the initial marking. Therefore, we introduce the concept of liveness bound $L(t)$, which allows us to generalize the classical concept of liveness of a transition:

Definition 2.2 *Liveness bound:* $\forall t \in T$,
$L(t) \stackrel{\text{def}}{=} \max\{k \mid \forall M_1 \in R(\mathcal{N}, M_0), \exists M \in R(\mathcal{N}, M_1)$ *such that* $M \geq kPRE[t]\}$. $\diamond$

Property 2.1 *Let* $\langle \mathcal{N}, M_0 \rangle$ *be a marked net, then* $\forall t \in T$: $E(t) \geq L(t)$. $\diamond$

A net is said to be *reversible* if the initial marking can always be recovered. In other words, in a reversible net the reachability graph is strongly connected (i.e. there exists no transient marking). Thus the following can be stated:

Property 2.2 *Let* $\langle \mathcal{N}, M_0 \rangle$ *be a reversible net, then* $\forall t \in T$: $E(t) = L(t)$. $\diamond$

Live MGs are reversible nets, so that enabling and liveness bounds are equal. On the other hand, this is not true for the more general cases that we consider in this paper; indeed the decision-free non ordinary net in Figure 3.b with an initial marking of two tokens in p_1 and the other two places empty gives an example of a live and bounded net in which $E(t_1) = 2 > L(t_1) = 1$.

A case of strict inequality in Property 2.1 can be interpreted as a generalization of the concept of non-liveness: there exist transitions that "contain potential servers" that are never used in the steady-state; these additional servers might only be used in a transient phase (i.e. they eventually die during the evolution of the model). On the other hand, the condition $L(t) > 0$ is equivalent to the usual liveness condition for transition t.

The two definitions above refer to behavioural properties. Since we are looking for computational techniques at the structural level, we can also introduce a structural counterpart of one of these concepts.

Definition 2.3 *Structural enabling bound:* $\forall t \in T$,
$SE(t) \stackrel{\text{def}}{=} \max\{k \mid M = M_0 + C \cdot \vec{\sigma} \geq kPRE[t],\ \vec{\sigma} \geq 0\}$. (LPP1) $\diamond$

Note that, the definition of structural enabling bound reduces to the formulation of a linear programming problem (LPP, in what follows) [8] using matrix C (the incidence matrix of the net). From the classical implication $M \in R(\mathcal{N}, M_0) \implies M = M_0 + C\vec{\sigma} \wedge \vec{\sigma} \geq 0$ [3], one can easily show that:

Property 2.3 *Let* $\langle \mathcal{N}, M_0 \rangle$ *be a marked net. Then* $\forall t \in T$: $SE(t) \geq E(t)$. $\diamond$

3 Ergodicity and measurability

In order to compute the steady-state performance of a system we have to assume that some kind of "average behaviour" can be estimated on the long run of the system we are studying. The usual assumption in this case is that the system model must be *ergodic*, meaning that at the limit when the observation period tends to infinity, the estimates of average values tend (almost surely) to the theoretical expected values of the (usually unknown) probability distribution functions (PDFs) that characterize the performance indexes of interest.

This assumption is very strong and difficult to verify in general; moreover, it creates problems when we want to include the deterministic case as a special case of a stochastic model, since the existence of the theoretical limiting expected value can be hampered by the periodicity of the model [1]. Thus we introduce the concept of *weak ergodicity* that allows the estimation of long run performance also in the case of deterministic models.

Definition 3.1 *Ergodicity:*

1) *A (not necessarily stochastic) process $X(t)$ is said to be weakly ergodic (or measurable in long run) iff the following limit exists:* $\overline{X} \stackrel{\text{def}}{=} \lim_{t\to\infty} \frac{1}{t}\int_0^t X(\tau)\,d\tau < \infty$.

2) *A stochastic process $X(t)$ is said to be strongly ergodic iff the following condition holds:* $\lim_{t\to\infty} \frac{1}{t}\int_0^t X(\tau)\,d\tau = \lim_{t\to\infty} E[X(t)] < \infty$ *(a.s.).* ◊

For stochastic PNs, weak ergodicity of the marking and the firing processes can be defined in the following terms:

Definition 3.2 *The marking process of a stochastic marked net is weakly ergodic iff the following limit exists:* $\overline{M} \stackrel{\text{def}}{=} \lim_{t\to\infty} \frac{1}{t}\int_0^t M(\tau)d\tau < \vec{\infty}$.
The firing process of a stochastic marked net is weakly ergodic iff the following limit exists: $\vec{\sigma}^* \stackrel{\text{def}}{=} \lim_{t\to\infty} \frac{\vec{\sigma}(t)}{t} < \vec{\infty}$.
The usual (i.e. strong) ergodicity concepts (see, e.g., [9]*) are defined in the obvious way taking into consideration Definition 3.1.2.* ◊

Since in this paper we are interested on the computation of bounds for the steady-state throughput of transitions, only weak ergodicity of the firing process must be assured. This is stated in the next result, for PNs with a unique consistent firing count vector.

Theorem 3.1 *If a marked PN has a unique consistent firing count vector, then its firing process is weakly ergodic.* ◊

Ergodicity of the marking and of the firing processes are, in general, unrelated properties. Let us consider an $M|M|1$ queue modelled by means of a place with one input and another output single server exponentially distributed timed transitions. If time-dependent rates λt and μt, with $\lambda < \mu$ are considered for arrival and service distributions, the marking process is strongly ergodic while the firing process is non (even weakly) ergodic. In what follows, we do not consider any more time-dependent distributions. On the other hand, if the arrival and the service rates are λ and μ, respectively, with $\lambda > \mu$, then the firing process is strongly ergodic but the marking process is non (even weakly) ergodic, because the marking of the place tends to infinity, almost surely. For bounded nets, ergodicity of the firing process does not imply marking ergodicity. This can be the case if after an initial transient phase, the model can reach different closed subsets of the state space. Even in those cases in which there does not exist a "true" mean marking (i.e. the limit marking for $t \to \infty$ is not unique), it makes sense to compute upper and lower bounds on transition throughputs.

Related to marking ergodicity, a sufficient condition for the weak ergodicity of the marking process of bounded nets is the existence of a *home state*. Home states are markings which can be reached from any other reachable marking. If a bounded PN has a home state, its associated state space has a unique closed subset of markings, and marking weak ergodicity is assured. This result provides an interesting example of possible interleaving between qualitative (home state concept) and quantitative (ergodicity concept) analysis for stochastic PNs.

Theorem 3.2 *If a bounded marked net has a home state then its marking process is weakly ergodic.* ◊

Semi-Markovian nets [9] are stochastic PNs such that their related marking process is a semi-Markov process. An important particular case of semi-Markovian PNs is that using *Coxian* distributions (i.e. characterized by having rational Laplace transform) for transition firing times. The interest of this family of distributions is that any distribution function can be approximated with a Coxian, preserving mean and higher moments [7]. For semi-Markovian bounded nets with home state, even strong marking ergodicity is assured:

Theorem 3.3 *If a semi-Markovian bounded marked net has a home state then its marking process is strongly ergodic.* ◊

The conditions of this theorem cannot be relaxed. An unbounded net can have home states but non ergodic marking process if the mean marking of a place tends to infinity. On the other hand, nets can have bounded marking mean values and be non ergodic because of the presence of more than one closed subset in the state space. However, marking ergodicity does not imply the existence of a home state: for the net in Figure 2, which is live structurally bounded mono-T-semiflow and it has not home state, an exponential distribution timing can be associated with transitions (for instance, all rates equal to 1) such that the related marking process is ergodic anyway.

In the next section, some interesting qualitative and quantitative properties for subclasses of nets with a unique consistent firing count vector are grouped.

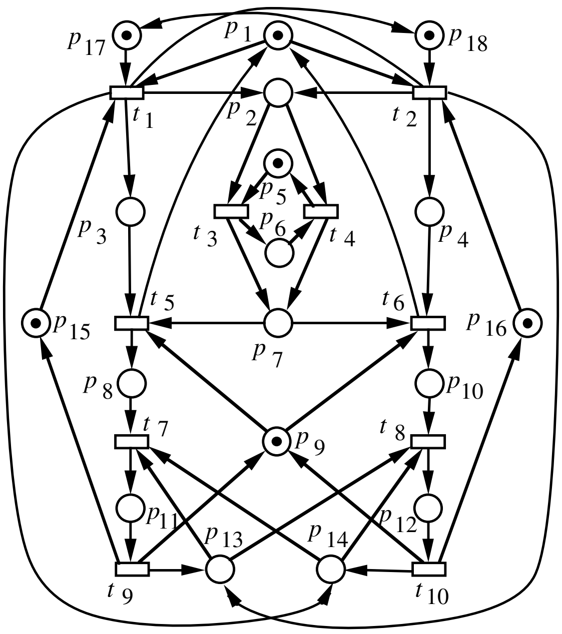


Figure 2: A live and bounded mono-T-semiflow net without home states.

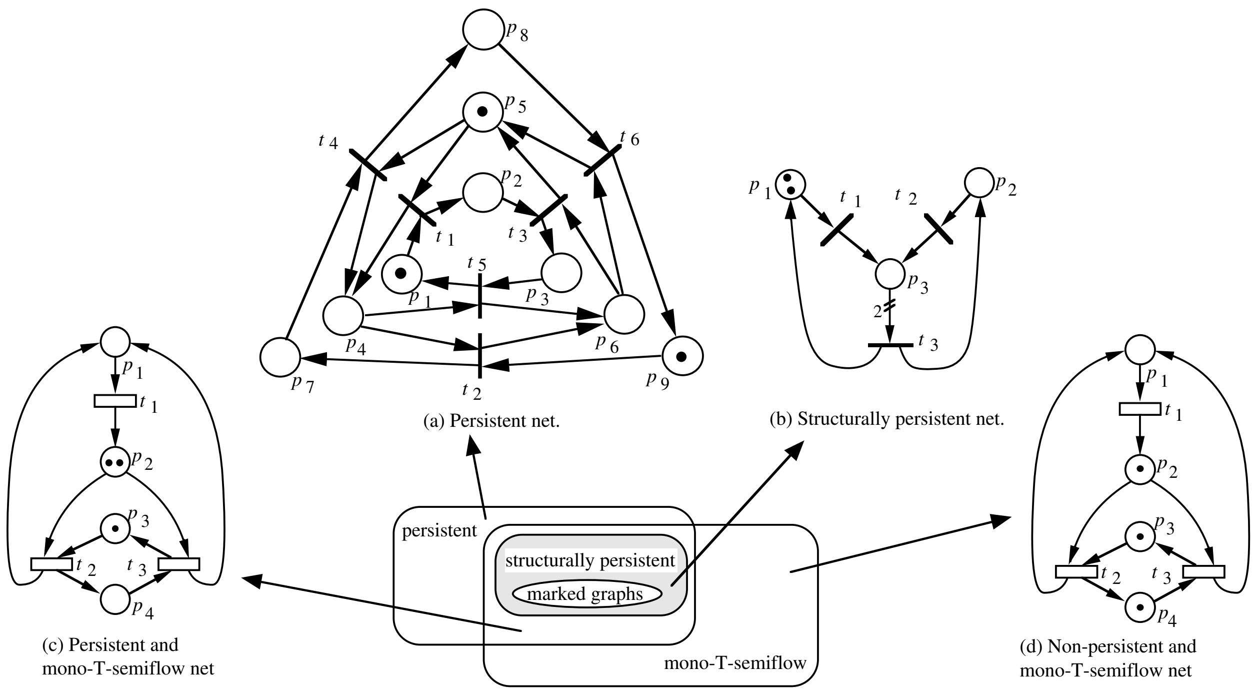


Figure 3: Bounded net subclasses having a unique consistent firing count vector.

4 Live and bounded nets with a unique consistent firing count vector

Persistent and *mono-T-semiflow* are two non disjoint subclasses of nets with a unique consistent firing count vector. Persistent nets are behaviourally defined, while mono-T-semiflow are structurally characterized. As a particular case, *structurally persistent nets* belong to the intersection of these two classes, thus possessing the good properties of both. *Marked graphs* are structurally persistent nets. This section is devoted to the introduction of these nets subclasses and to the presentation of some of their basic properties relevant from the performance point of view. Figure 3 provides an overall picture of the relations among the net subclasses considered in this section. These subclasses do not cover all PNs with unique consistent firing count vector. A neither persistent nor mono-T-semiflow net having a unique consistent firing count vector can be constructed from several persistent and mono-T-semiflow components. Nevertheless, for simplicity we restrict the discussion in this section to persistent and mono-T-semiflow subclasses.

4.1 Persistent nets

Definition 4.1 *A marked net $\langle \mathcal{N}, M_0 \rangle$ is said to be persistent iff for all reachable marking M and for all different transitions, t_1 and t_2, enabled in M, the sequence $t_1 t_2$ is fireable from M.* $\diamond$

Persistent nets are effectively conflict-free (structural conflicts could exist among transitions, but persistency implies that no actual decision is ever made). As an example look at the net in Figure 3.a. This net has structural conflicts (e.g. p_4 has two output transitions, t_2 and t_5) but for the initial marking $M_0 = (1,0,0,0,1,0,0,0,1)^T$ no state can be reached in which a decision must be taken. Persistency is a behavioural property, i.e. the same net structure with a different initial marking can give non persistent behaviour. For example, the net in Figure 3.c is persistent, but that in Figure 3.d (which has the same structure) is not. In both cases the marked nets are live and bounded, and both their unique consistent firing count vector are $X = (2,1,1)^T$.

Let us recall a property and two results that will lead to the conclusion that *the firing process associated with a bounded persistent net is weakly ergodic*. The property is that of *directedness*; this means that any two reachable markings have at least one common successor marking.

Lemma 4.1 [4] *All persistent marked nets have the directedness property.* $\diamond$

Lemma 4.2 [10] *For bounded marked nets, directedness and the existence of a home state are equivalent properties.* $\diamond$

Examples can be found [1] showing that Lemma 4.2 does not hold for unbounded persistent nets.

A place is said to be *implicit* [11] if its marking never is the only one that prevents the firing of any output transition. Using the previous lemmas, the following statement can be derived:

Theorem 4.1 *Live bounded persistent connected nets without implicit places have a unique consistent firing count vector (i.e. $\exists \vec{\sigma}_R$ such that $\forall M \in R(\mathcal{N}, M_0)$ if $M[\sigma\rangle M$ then $\vec{\sigma} = k\vec{\sigma}_R$ with $k \in \mathbb{N}$).* $\diamond$

Now, from Theorems 3.1, 3.2, 3.3, and 4.1, the next result can be stated:

Corollary 4.1 *Let $\langle \mathcal{N}, M_0 \rangle$ be a live bounded connected marked net without implicit places. If $\langle \mathcal{N}, M_0 \rangle$ is persistent then:*

1) Both the marking and the firing processes are weakly ergodic.

2) If $\langle \mathcal{N}, M_0 \rangle$ is semi-Markovian, its marking process is strongly ergodic. $\diamond$

In order to study the steady-state performance of a stochastic net, only *recurrent* markings are relevant (i.e. *transient* markings do not affect the computation). Even if bounded persistent nets are ergodic this does not mean that there exist no transient markings. The net in Figure 3.b is structurally persistent, live, and 2–bounded for $M_0 = (2,0,0)^T$, but M_0 is a transient state (i.e. it is not a home state), so that the net is not reversible.

4.2 Mono-T-semiflow nets

Let us introduce now a structurally characterized class of nets with a unique consistent firing count vector.

Definition 4.2 *A structurally bounded net $\mathcal{N}$ is called mono-T-semiflow iff there exists a unique minimal T-semiflow that contains all transitions.* $\diamond$

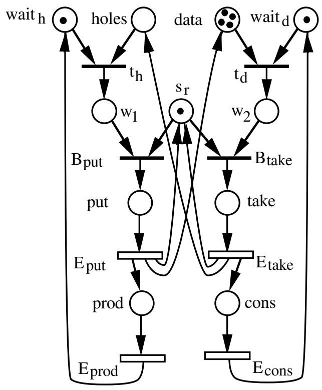


Figure 4: A producer-consumer system with mutual exclusion.

In a mono-T-semiflow net conflicts may be reached, so that different behaviours can occur. However, from the steady-state performance point of view, these decisions yield a unique consistent firing count vector, provided that the net is live (all different behaviours let the same set of transitions, characterized by the only T-semiflow of the net, fire, perhaps in a different order). For example, $\sigma_a = t_2 t_1 t_3$ and $\sigma_b = t_3 t_1 t_2$ are possible sequences in the net of Figure 3.d, both fireable from M_0. Even if the performance can be equal for any conflict resolution policy, from the functional point of view the results can be different (imagine t_2 and t_3 be two non commutative operations).

As an example, let us consider the problem of modelling a *producer-consumer system* composed by two processes, and a buffer storage with limited capacity. The process P_1 produces data that are placed at the buffer. The process P_2 takes data from the buffer for processing them. Processes P_1 and P_2 cannot operate simultaneously with the buffer, which is a shared resource. Thus the system cannot be modelled by means of MGs. The control system for the production and consumption of data is depicted in Figure 4 by means of a PN. Mutual exclusion is modelled with place s_r.

Obviously, the net in Figure 4 is not persistent. Transitions B_{put} and B_{take} can be in an effective conflict. The net is mono-T-semiflow and the unique minimal T-semiflow is the vector with all components equal 1.

Structural boundedness (i.e. $\exists Y > 0$ s.t. $Y^T \cdot C \leq 0$) [3] can be decided in polynomial time. Thus mono-T-semiflow nets can be polynomially characterized.

Property 4.1 *Let C be the incidence matrix of a mono-T-semiflow net. Then $rank(C) = m-1$.* $\diamond$

The following result is a particularization of theorem 3.1 in [12] that takes into account the existence of a single T-semiflow.

Property 4.2 *Let $\langle \mathcal{N}, M_0 \rangle$ be a live mono-T-semiflow net. For any firing sequence σ applicable in $\langle \mathcal{N}, M_0 \rangle$ we can write: $\vec{\sigma} = rX + b$, where $X > 0$ is the minimal T-semiflow, $r \in I\!N$, and $b \geq 0$ is a bounded vector such that $C \cdot b \gneq 0$ is impossible.* $\diamond$

Now, from our Theorem 3.1 and the above property, the next result follows:

Corollary 4.2 *Mono-T-semiflow nets have weakly ergodic firing process.* $\diamond$

Related to marking ergodicity, the following negative result for mono-T-semiflow nets it is shown in [1] by means of an example:

Property 4.3 *There exist live mono-T-semiflow nets without home states, so that weak ergodicity of their marking processes is not guaranteed.* $\diamond$

4.3 Structurally persistent nets and marked graphs

As it was pointed out in Section 4.1, persistency is a behavioural property. Let us introduce a subclass of persistent nets such that the persistency is inherent to the structure.

Definition 4.3 *A net $\mathcal{N}$ is said to be structurally persistent iff $\langle \mathcal{N}, M_0 \rangle$ is persistent for all finite initial marking M_0.* $\diamond$

Property 4.4 *A net is structurally persistent iff it does not exist any structural conflict in it (i.e. $\forall p \in P, |p^\bullet| \leq 1$).* $\diamond$

Structurally persistent nets are structurally decision-free. Moreover the live and bounded net in Figure 3.b shows that they may have transient states. Now it is interesting to point out a very well known subclass of these nets for which there exists no transient marking, provided the liveness for M_0.

Property 4.5 *Marked graphs (MGs) are structurally persistent nets. The reverse is not true (for example the net in Figure 3.b is not an MG).* $\diamondsuit$

Property 4.6 *Let $\mathcal{N}$ be a structurally persistent net.*

1) If $\langle \mathcal{N}, M_0 \rangle$ is live and bounded for some M_0, then $\mathcal{N}$ is mono-T-semiflow. The reverse is not true (Figure 3.c).

2) If $\mathcal{N}$ is an MG, then it is consistent, and its unique minimal T-semiflow is $\vec{1}$. $\diamondsuit$

By Properties 4.6 and 4.1, since MGs are consistent nets, the rank of their incidence matrix is $m - 1$.

Theorem 4.2 [3] *Let $\langle \mathcal{N}, M_0 \rangle$ be a live (possibly unbounded) MG. The two following statements are equivalent:*

i) $M \in R(\mathcal{N}, M_0)$, i.e. M is reachable from M_0.

ii) $B_f \cdot M = B_f \cdot M_0$, with B_f the fundamental circuit matrix of the graph (i.e. its row vectors are a basis of the left annullers of C), and $M \geq 0$. $\diamondsuit$

According to the above theorem $M \in R(\mathcal{N}, M_0) \iff M_0 \in R(\mathcal{N}, M)$. In other words:

Property 4.7 *Live MGs are reversible.* $\diamondsuit$

Property 4.8 *Let $\mathcal{N}$ be an MG.*

1) $\mathcal{N}$ is structurally bounded (i.e. $\langle \mathcal{N}, M_0 \rangle$ is bounded $\forall M_0$) iff it is strongly connected.

2) Let $\langle \mathcal{N}, M_0 \rangle$ be live. Then $\langle \mathcal{N}, M_0 \rangle$ is bounded iff $\mathcal{N}$ is structurally bounded. $\diamondsuit$

According to the above properties, strong connectivity and boundedness are equivalent properties for live MGs. Taking now into account Theorems 3.2 and 3.3, the following can be stated:

Corollary 4.3 *If $\langle \mathcal{N}, M_0 \rangle$ is a strongly connected live MG then:*

1) $\langle \mathcal{N}, M_0 \rangle$ has weakly ergodic marking process.

2) If $\langle \mathcal{N}, M_0 \rangle$ is semi-Markovian, its marking process is strongly ergodic. $\diamondsuit$

Finally, an interesting property of live MGs, that allows an efficient computation of liveness bounds, is the following:

Property 4.9 *Let $\langle \mathcal{N}, M_0 \rangle$ be a live MG, then $\forall t \in T$: $L(t) = E(t) = SE(t)$ (i.e. $L(t)$ can be computed in polynomial time by solving* (LPP1)*; see Definition 2.3).* $\diamondsuit$

5 Upper bounds for the steady-state throughput

In this section, upper bounds are presented for the steady-state transition throughputs of nets with a unique consistent firing count vector. First we derive some general structural results, and then we particularize them to bounded persistent nets and mono-T-semiflow nets.

5.1 General approach

Let us account only for the first moment of the PDFs associated with transitions. In the following, let θ_i be the mean value of the random variable associated with the firing of transition t_i (*service time* of t_i, with queueing networks terminology), and D the diagonal matrix with elements θ_i, $i = 1, \ldots, m$.

The limit firing count vector per time unit (under weak ergodicity assumption) is $\vec{\sigma}^* = \lim_{t \to \infty} \vec{\sigma}(t)/t$, and the mean time between two consecutive firings of a selected transition t_i (*mean cycle time* of t_i), $\Gamma_i = 1/\vec{\sigma}_i^*$. The *relative firing frequency vector* (or vector of *visiting ratios*, from the queueing networks point of view), denoted by $F_i = \Gamma_i \vec{\sigma}^*$, is the limit firing flow vector $\vec{\sigma}^*$ normalized for having the i^{th} component equal 1 (note that this makes sense only if $\vec{\sigma}_i^* \neq 0$). Then the components of $PRE \cdot D \cdot F_i$ represent the product of the number of tokens needed for firing the transitions and the mean length of time that these tokens reside in each place between two consecutive firings of t_i.

For live mono-T-semiflow nets $F_i = F_i(\mathcal{N}) = X$ is the minimal T-semiflow, thus independent of the initial marking (provided that liveness is guaranteed) and of the timing interpretation. For persistent nets $F_i = F_i(\mathcal{N}, M_0)$ (Figure 3.a) is a T-semiflow (possibly non minimal), and it is independent of the timing interpretation.

Let $\overline{M}$ be the limit vector of the average number of tokens in each place (i.e. $\overline{M} = \lim_{t\to\infty} \frac{1}{t}\int_0^t M(s)ds$). Then, provided that this limit exists, $\Gamma_i \overline{M}$ is the vector of products of the mean number of tokens and the length of one cycle and we have:

$$\Gamma_i \overline{M} \geq PRE \cdot D \cdot F_i \tag{1}$$

From this inequality, a lower bound for the mean cycle time associated with transition t_i, Γ_i^{lb}, can be derived.

$$\Gamma_i^{lb} = \max_{p_j \in P} \frac{PRE[p_j] \cdot D \cdot F_i}{\overline{M}(p_j)} \tag{2}$$

Since the vector $\overline{M}$ is unknown, (2) cannot be solved. However, the following marking invariant can be written using P-semiflow Y:

$$Y^T \cdot M_0 = Y^T \cdot M = Y^T \cdot \overline{M} \tag{3}$$

Now, from (1) and (3): $Y^T \cdot \Gamma_i M_0 \geq Y^T \cdot PRE \cdot D \cdot F_i$, and a lower bound for the mean cycle time in steady-state is:

$$\Gamma_i^{lb} = \max_{Y \in \{P-semiflow\}<} \frac{Y^T \cdot PRE \cdot D \cdot F_i}{Y^T \cdot M_0} \tag{4}$$

Of course, an upper bound for the throughput of t_i is $1/\Gamma_i^{lb}$.

The previous lower bound for the mean cycle time can be formulated in terms of a particular class of optimization problems called *fractional programming problems* [8], and later transformed into a linear programming problem.

Theorem 5.1 *For any marked net with unique consistent firing count vector, a lower bound for the mean cycle time of transition t_i can be computed by the following linear programming problem:*
$\Gamma_i^{lb} = \max\{Y^T \cdot PRE \cdot D \cdot F_i \mid Y^T \cdot C = 0,\ Y^T \cdot M_0 = 1,\ Y \geq 0\}$ (LPP2) $\diamond$

The previous theorem shows that the problem of finding an upper bound for the steady-state throughput in a persistent or mono-T-semiflow net can be solved looking at the cycle times associated with P-semiflows, considered in isolation. These cycle times can be computed making the summation of the average firing times of all the transitions involved in the P-semiflow, and dividing by the number of tokens present in it. From a different perspective, from (LPP2) it can be stated that Γ_i^{lb} is finite iff all P-semiflows are marked (a necessary condition for liveness).

For the particular case of strongly connected MGs, $F_i = \vec{1}$ (unique T-semiflow), and the above bound is the same that has been obtained for the deterministic case by other authors (see, e.g., [5, 13]). For deterministic timed MGs, the reachability of this bound has been shown. Since deterministic timing is just a particular case of stochastic timing, the reachability of the bound is assured for our proposes. Even more, this bound cannot be improved only on the base of the knowledge of the coefficients of variation for the transition firing times [2].

5.2 Upper bounds for bounded persistent nets

For bounded persistent nets, weak ergodicity of the firing process is assured (Corollary 4.1.1). Thus for these nets a unique limit firing behaviour exists, and bounds can be computed for the steady-state throughput.

As remarked in Section 4.1, persistent nets are behaviourally defined. This means that a behavioural analysis must be made before computing performance bounds in order to check for the persistency of the net. Few results are known in the literature related to bounds for the performance of bounded persistent nets. A partial result was presented in [5] for safe and persistent nets with deterministic timing. For these nets a behaviourally equivalent safe MG (*behaviour graph*) can be built. The method consists in drawing the initially marked places and enabled transitions. After that, firing all transitions and drawing the output places and repeating the procedure until a marking in the process is re-found (see Figure 5). Then, the method explained in Section 5.1 can be applied for computing the bounds for this MG and so for the steady-state performances of the initial persistent net. Unfortunately, this analysis is not possible for bounded (non safe) nets when non deterministic timing is considered.

Let us now introduce some general results useful for computing bounds for the performance of bounded persistent nets. Later we shall improve some of these results.

Let us consider live bounded persistent nets without implicit places. According to Theorem 4.1 consistent firing count vectors are proportional to $\vec{\sigma}_R$. Thus the problem (LPP2) can be applied for computing a lower bound of the steady-state cycle time of a selected transition t_i taking into account that $F_i = k\vec{\sigma}_R$ is a T-semiflow (non minimal if there exist more than one; see Figure 3.a) with $F_i(t_i) = 1$.

The optimal value of the previous problem is a non reachable bound in general (i.e. there exist net models such that no stochastic interpretation allows to reach the computed bound, Γ_i^{lb}). To see it, let us consider for

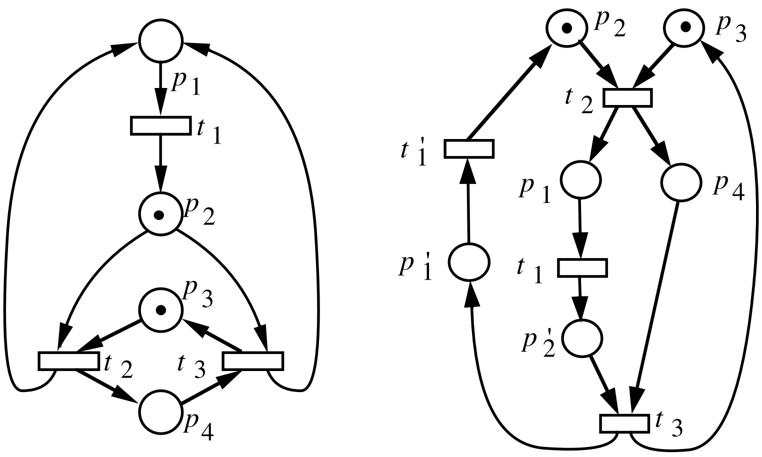


Figure 5: Equivalent MG for safe persistent net.

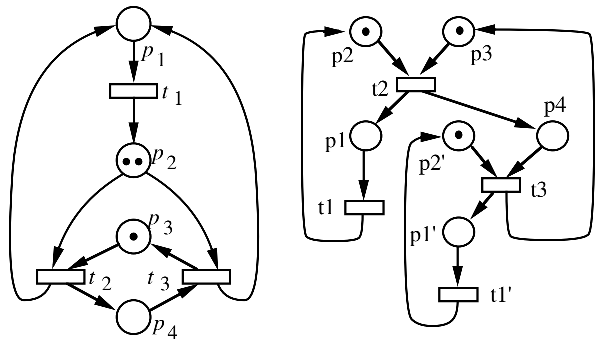


Figure 6: Equivalent MG for deterministic timing.

example the net in Figure 3.c. Selecting transition t_2, the vector F_i for (LPP2) is $F_i = (2, 1, 1)^T$ and the obtained bound is $\Gamma_2^{lb} = \max\{\theta_2 + \theta_3, (2\theta_1 + \theta_2 + \theta_3)/2\}$. Now, considering deterministic timing for all transitions with $\theta_1 = 2, \theta_2 = 60, \theta_3 = 1$, the obtained bound is $\Gamma_2^{lb} = 61$ while the actual cycle time for transition t_2 is bigger because of the sequence $t_1 t_2$ which takes 62 units of time. Nevertheless, for safe (and ordinary, in order to be live) persistent nets, the bound given by (LPP2) can be always reached: it would be obtained by deriving the equivalent MG (according to [5]) and computing the bound, using (LPP2).

Even though the cycle time bound obtained from (LPP2) can be non reachable for non-safe persistent nets, it can be pointed out that the bound is finite if and only if the actual cycle time is finite, and this trivially characterizes the liveness of the model.

Theorem 5.2 *Let $\langle \mathcal{N}, M_0 \rangle$ be a bounded persistent net. The following three statements are equivalent:*

i) The optimal value Γ_i^{lb} of (LPP2) *for $\langle \mathcal{N}, M_0 \rangle$ is finite.*

ii) The actual cycle time of $\langle \mathcal{N}, M_0 \rangle$ is finite.

iii) $\langle \mathcal{N}, M_0 \rangle$ is live. $\diamond$

The above result is not true for other net classes. In Section 5.3 it is proven that for mono-T-semiflow nets the actual cycle time can be infinite (so that the net is non live) while the lower bound obtained from (LPP2) is finite.

Improving the upper bound for ordinary nets. Let us now describe a method to compute a reachable upper bound for the steady-state throughput of bounded ordinary persistent nets. Deterministic timing yields the best performance for a given persistent (and conflict-free, see Section 4.1) net and average firing delay associated to transitions. If we consider only deterministic timing, a behaviourally equivalent MG can be derived in an analogous way to that proposed in [5] (see the example in Figure 6): (1) Split the places into instances in such a way that their safe markings represent conditions for the enabling of transitions. (2) Develop the behaviour graph of the net (under deterministic timing assumption) from the initial marking. Since the original net is live, the behaviour graph can be indefinitely extended. (3) Identify those instances of places that must be superposed, in such a way that the relative firing frequency of transitions is preserved: an MG has been derived.

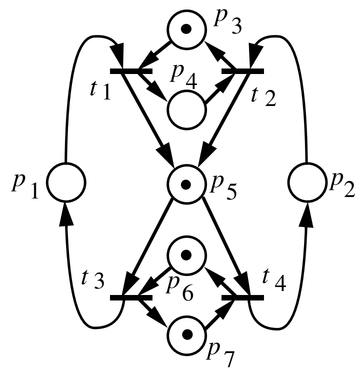


Figure 7: Non-live mono-T-semiflow net, even if all P-semiflows are marked and thus Γ_i^{lb} in (LPP2) is finite.

Considering general timing distributions, the original net and the derived MG are not behaviourally equivalent. In fact, the steady-state throughput for the MG is greater than or equal to the one of the original net. Nevertheless for deterministic timing the equality holds and this provides the following method for computing a reachable bound for the throughput: (1) consider a bounded ordinary persistent net with general distribution timing; (2) develop the cyclic process for the deterministic case (i.e. the behaviourally equivalent MG for the deterministic timing); and (3) compute the upper bound for the steady-state throughput of the MG using (LPP2).

The above computed value is a reachable bound for the steady-state throughput of the original net, because in the deterministic case the maximal throughput for the net is always obtained and under this condition the throughputs of the cyclic process and of the original net are equal.

5.3 Upper bounds for mono-T-semiflow nets

Let us now consider mono-T-semiflow nets (a polynomial time characterizable net subclass, Section 4.2) and give upper bounds for their steady-state throughput.

Unfortunately, the existence of a unique consistent minimal T-semiflow, X, does not guarantee the ergodicity of the marking process (Property 4.3). However, even in the case in which a net is marking non-ergodic, the computation of the throughput bounds makes sense. The values that we compute in this section are bounds for all possible steady-state marking behaviours of the net.

The problem (LPP2) defining an upper bound for the steady-state throughput of a net with a unique consistent firing count vector (Section 5.1) can be used taking $F_i = kX$ where $kX(t_i) = 1$. Nevertheless, in general this bound is not reachable. Moreover, a mono-T-semiflow net can be non live and the obtained lower bound for the cycle time be finite (see Figure 7). In other words:

Property 5.1 *For mono-T-semiflow nets, liveness is not characterized by the finiteness of the lower bound of the cycle time computed by means of* (LPP2). ◇

Comming back to the producer-consumer system of Figure 4, let us suppose that transitions t_h, t_d, B_{put}, and B_{take} are immediate (i.e. they fire in zero time) and that the mean values of random variables associated with the rest of transitions are: $\theta(E_{put}) = 2$; $\theta(E_{take}) = 3$; $\theta(E_{prod}) = 4$; and $\theta(E_{cons}) = 2$.

The problem (LPP2) gives a lower bound for the mean cycle time of the net: $\Gamma^{lb} = \max\{\theta(E_{put}) + \theta(E_{take}), (\theta(E_{put})+\theta(E_{take}))/4, \theta(E_{put})+\theta(E_{prod}), \theta(E_{take})+\theta(E_{cons})\} = 6$. As it is remarked in Property 5.1, this bound is non-reachable, in general. However, in this case, the lower bound is equal to the actual cycle time for deterministic timing. In fact, if deterministic timing is considered, the buffer capacity (initial number of tokens at place *data*) can be reduced to 1, without modifying the actual cycle time. This is because, in this particular case, there exist two different P-semiflows Y_1 and Y_2 (with $||Y_1|| = \{s_r, put, take\}$ and $||Y_2|| = \{holes, data, w_1, w_2, s_r, put, take\}$) involving the same set of timed transitions (E_{put} and E_{take}). Since $Y_1^T \cdot M_0 = 1$, then the number of tokens at place *data* can be reduced to 1, and the same optimum value in problem (LPP2) is preserved.

6 Lower bounds for the steady-state throughput

A trivial lower bound in steady-state performance for a live net with a unique consistent firing count vector is of course given by the sum of the firing times of all the transitions weighted by the firing flow vector itself. Since the net is live all transition must be fireable, and the sum of all firing times multiplied by the number of occurrences of each transition in the (unique) average cycle of the model corresponds to any *complete sequentialization* of all the activities represented in the model. This *pesimistic* behaviour can be reached in some particular cases (e.g., for live and safe MGs) if random variables with arbitrarily large coefficient of variation are conveniently selected for transitions firing times.

Property 6.1 *For any live net with unique consistent firing count vector, an upper bound for the mean cycle time of transition t_i is:*

$$\Gamma_i^{ub} = \sum_{j=1}^{m} F_i(t_j)\theta_j$$

where θ_j is the mean firing time of transition t_j and F_i is the relative firing frequency vector of the net for transition t_i (i.e. the unique consistent firing count vector, normalized for having $F_i(t_i) = 1$). $\diamond$

In order to improve the previous bound, an intuitive idea could be to take into account that some work can be done in parallel at each transition, since infinite server semantics is assumed. From a queueing theory perspective and considering the steady-state behaviour, the number of servers at each *station* (transition) is equal to the corresponding enabling bound in steady-state (i.e. liveness bound), and the contribution of each transition to the duration of the complete sequentialization of all activities can be divided by its liveness bound. Thus we could conjecture the following upper bound for the mean cycle time:

$$\Gamma_i \overset{?}{\leq} \sum_{j=1}^{m} \frac{F_i(t_j)\theta_j}{L(t_j)} \tag{5}$$

The same value would be obtained taking the algorithm used for the computation of the lower bound (Section 5.1), substituting in it the "max" operator with the sum of the firing times of all transitions involved, and making some manipulation to avoid counting more than once the contribution of the same transition.

The conjecture (5) has been shown to be true for strongly connected MGs in [2]. In fact, for this subclass of nets the upper bound for the mean cycle time given by (5) has been shown to be reachable for any net topology, for any specification of the mean firing delays, and for some assignement of PDFs to the firing delays of transitions. Moreover, taking into account Property 4.9, the liveness bound can be efficiently computed for strongly connected MGs by means of the problem (LPP1, Section 2.2).

Concerning structurally persistent nets the conjecture (5) is false (thus it is false for persistent and for mono-T-semiflow nets). This can be shown considering, for example, the structurally persistent net in Figure 1.b with mean firing times $\theta_1, \theta_2, \theta_3$ for transitions. For this net, the relative firing frequency vector is $F_2 = (2, 1, 1)^T$, and the liveness bounds of transitions are given by $L = (2, 1, 1)^T$. Thus, the conjecture (5) would give the value $\theta_1 + \theta_2 + \theta_3$ as upper bound for Γ_2. If exponentially distributed random variables (with means $\theta_1, \theta_2, \theta_3$; $\theta_1 \neq \theta_3$) are associated with transitions, the steady-state cycle time for transition t_2 is $\Gamma_2 = \theta_1 + \theta_2 + \theta_3 + \theta_1^2/2(\theta_1 + \theta_3)$, which is greater than the value obtained applying (5), thus the conjecture is false.

Unfortunately, the trivial bound given by Property 6.1 is non-reachable in general, and in some cases its value can be too pesimistic. An improving of this bound would probably require more information about the PDFs of firing times than their mean values, and this approach is not within the scope of this work.

7 Conclusions

We have addressed the problem of computing upper and lower bounds for the throughput of transitions in Petri net models (or the corresponding synchronized queueing networks) having a unique consistent firing count vector. The results presented here represent, in some sense, an extension of those described in [2] for the case of bounded marked graphs. The upper bound in case of persistent nets is a generalization of that obtained for marked graphs, and has been shown to be reachable for ordinary nets. The technique proposed for the derivation of behaviour graphs for non-safe ordinary persistent nets is an extension of a method originally proposed by Ramchandani for safe persistent nets, that was not directly applicable.

For what concerns the lower bound on throughput, only the trivial bound computed as the inverse of the sum of all transition firing delays weighted by the firing count vector can be easily borrowed from the marked graph case. However this can be too pessimistic in case of persistent or live mono-T-semiflow nets.

In any case both the upper and lower bounds are independent of any assumption on the probability distribution of the delays associated with transitions, and their value can be computed based on the knowledge of the averages. This represents a generalization with respect to the usual assumptions needed for the exact performance evaluation of a Petri net model. A second generalization, implicit in the choice of computing throughput bounds instead of actual values, is that the analysis of marking non-ergodic models still makes sense.

Besides the results on the computation of bounds, this paper identifies two subclasses of bounded nets having a unique consistent firing count vector, and contains a discussion of their ergodicity conditions. In particular, the concept of liveness bound for transitions is a new behavioural property, that comes directly from considerations related to the timing semantics of a timed Petri net model. It generalizes the usual concept of liveness for a transition, and provides an example of possible interleaving between qualitative and quantitative analysis for timed and stochastic Petri nets. Another example has been provided by stablishing the strong connection between marking ergodicity and home state concepts.

Alternative extensions of the results concerning marked graphs are already being considered. In particular, live and bounded free choice nets are studied in [14]. In this case, the idea is that several consistent firing count vectors can be reproduced in steady-state. However the free choice property leads to the fact that selections are completely governed by the structure and the stochastic (routing) interpretation of the net. Thus the "relative firing frequency vector" can be computed independently of the marking of the net, provided that liveness is guaranteed.

References

[1] J. Campos, G. Chiola, and M. Silva. Properties and steady-state performance bounds for Petri nets with unique repetitive firing count vector. In *Proceedings of the* 3^{rd} *International Workshop on Petri Nets and Performance Models*, IEEE-CS Press, pp. 210–220, Kyoto, Japan, December 1989.

[2] J. Campos, G. Chiola, J. M. Colom, and M. Silva. Tight polynomial bounds for steady-state performance of marked graphs. In *Proceedings of the* 3^{rd} *International Workshop on Petri Nets and Performance Models*, IEEE-CS Press, pp. 200–209, Kyoto, Japan, December 1989.

[3] T. Murata. Petri nets: properties, analysis, and applications. *Proceedings of the IEEE*, 77(4):541–580, April 1989.

[4] G. W. Brams. *Réseaux de Petri: Théorie et Pratique. T.1. théorie et analyse.* Masson, Paris, 1983.

[5] C. Ramchandani. *Analysis of Asynchronous Concurrent Systems by Petri Nets.* PhD thesis, M.I.T. Cambridge, Mass. February 1974.

[6] M. Ajmone Marsan, G. Balbo, A. Bobbio, G. Chiola, G. Conte, and A. Cumani. The effect of execution policies on the semantics and analysis of stochastic Petri nets. *IEEE Transactions on Software Engineering*, 15(7):832–846, July 1989.

[7] E. Gelenbe and G. Pujolle. *Introduction to Queuing Networks.* John Wiley, 1987.

[8] K. G. Murty. *Linear Programming.* John Wiley & Sons, 1983.

[9] G. Florin and S. Natkin. Les réseaux de Petri stochastiques. *Technique et Science Informatiques*, 4(1):143–160, 1985.

[10] E. Best and K. Voss. Free-choice systems have home states. *Acta Informatica*, (21):89–100, 1984.

[11] M. Silva and J. M. Colom. On the computation of structural synchronic invariants in P/T nets. In G. Rozenberg, editor, *Advances in Petri Nets 1988*, pp. 386–417, LNCS 340, Springer-Verlag, Berlin, 1988.

[12] M. Silva. Towards a synchrony theory for P/T nets. In K. Voss, H. Genrich, and G. Rozenberg, editors, *Concurrency and Nets*, pp. 435–460, Springer-Verlag, Berlin, 1987.

[13] C. V. Ramamoorthy and G. S. Ho. Performance evaluation of asynchronous concurrent systems using Petri nets. *IEEE Transactions on Software Engineering*, 6(5):440–449, September 1980.

[14] J. Campos, G. Chiola, and M. Silva. Properties and performance bounds for closed free choice synchronized monoclass queueing networks. *Research Report GISI-RR-90-2, Dpto. Ingeniería Eléctrica e Informática, Universidad de Zaragoza*, Spain, January 1990.