Improving throughput upper bounds for net based models of manufacturing systems

J. Campos, J. M. Colom, and M. Silva

Dpto. de Ingeniería Eléctrica e Informática, Centro Politécnico Superior, Universidad de Zaragoza, María de Luna 3, 50015 Zaragoza, Spain

Abstract

This paper addresses the improvement of throughput upper bounds for live and bounded stochastic Petri nets, presented by the authors in previous works. The introduction of a greater amount of structural information, traps and implicit places, allows to improve the bounds using linear programming problems defined on the net structure, on the routing probabilities, and on the mean service time of transitions. The obtained bounds can be applied for the analysis of manufacturing systems modelled by means of stochastic Petri nets. An example is presented and evaluated using the introduced techniques.

1 INTRODUCTION

Timed and stochastic Petri nets constitute an adequate model for the evaluation of performance measures of concurrent and distributed systems (see, e.g., [1, 2, 3, 4]). Nevertheless, one of the main problems in the actual use of these models for the evaluation of large systems is the explosion of the computational complexity of the analysis algorithms.

The computation of upper bounds for the *throughput of transitions*, defined as the average number of firings per time unit, is considered in this paper. Previous results [5, 6, 7], based only on the net structure, on the routing probabilities, and on the mean values of the service time of transitions, are newly interpreted and improved, with the introduction of more information from the net structure: *traps* and *implicit places*.

We assume the reader is familiar with the structure, firing rule, and basic properties of net models (see [8] for a 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 the Pre- and $Post$-incidence functions take values in $\{0, 1\}$, $\mathcal{N}$ is said to be 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). The support of P-semiflows (T-semiflows) is defined by $||Y|| = \{p \in P | Y(p) > 0\}$ ($||X|| = \{t \in T | X(t) > 0\}$). A (P- or T-) semiflow I has minimal support iff there exist no other semiflow I' such that $||I'| \subset ||I||$. A (P- or T-) semiflow is canonical iff the greatest common divisor of its components is 1. A (P- or T-) semiflow is elementary iff it is canonical and has minimal support. M (M_0) is a marking (initial marking). With σ we represent a firing 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 $\vec{\sigma} \geq 0$. The reachability set $R(\mathcal{N}, M_0)$ is the set of all markings reachable from the initial marking. $L(\mathcal{N}, M_0)$ is the set of all firing sequences and their suffixes in $\langle \mathcal{N}, M_0 \rangle$: $L(\mathcal{N}, M_0) = \{\sigma | M[\sigma\rangle \text{ with } M \in R(\mathcal{N}, M_0)\}$.

State machines are ordinary nets such that $\forall t \in T : |^\bullet t| = |t^\bullet| = 1$. Marked graphs are ordinary nets such that $\forall p \in P : |^\bullet p| = |p^\bullet| = 1$. Free choice nets are ordinary nets such that $\forall p \in P : |p^\bullet| > 1 \Rightarrow {}^\bullet(p^\bullet) = \{p\}$. Simple nets are ordinary nets such that each transition has at most one shared input place, i.e., $\forall t \in T, |\{p \in {}^\bullet t : |p^\bullet| > 1\}| \leq 1$. The previous net subclasses are characterized by local structural properties. The following is a net subclass characterized by a global structural property: mono-T-semiflow nets are nets having only one minimal T-semiflow.

The introduction of a timing specification is essential in order to use Petri net models for performance evaluation of distributed systems. We consider nets with *deterministically or stochastically* timed transitions with *one phase* firing rule, i.e., a timed enabling (called the *service time* of the transition) followed by an atomic firing. The service times of transitions are supposed to be mutually independent and time independent.

In order to avoid the coupling between resolution of conflicts and duration of activities, we suppose that transitions in conflict are *immediate* (they fire in zero time). Decisions at these conflicts are taken according to *routing rates* $\mathcal{R}$ associated with immediate transitions (*generalized stochastic Petri nets* [9]). In other words, each subset of transitions $\{t_1, \ldots, t_k\} \subset T$ that are in conflict in one or several reachable markings are considered immediate, and the constants $r_1, \ldots, r_k \in \mathbb{N}^+$ are explicitly defined in the net interpretation in such a way that when $t_1, \ldots, t_k$ are enabled, transition t_i ($i = 1, \ldots, k$) fires with probability (or with long run rate, in the case of deterministic conflicts resolution policy) $r_i / (\sum_{j=1}^{k} r_j)$. Note that the routing rates are assumed to be strictly positive, i.e., all possible outcomes of any conflict have a non-null probability of firing. This fact guarantees a *fair* behaviour for the non-autonomous Petri nets that we consider.

Throughout the paper we consider a simplified version of the following manufacturing example in order to illustrate the computation of performance measures and their corresponding improvements introduced here. Let us consider a flexible manufacturing system consisting of two production cells. Both of them manufacture two same kinds of parts, O_1 and O_2, but in different sites and with different processes P_1 and P_2. Process P_1 manufactures a proportion of q of the overall production and P_2 $(1-q)$. The last step in the FMS consists of an assembly operation of one part O_1 and one part O_2 available in the assembly station input stores. The duration of assembly operation is s_5 time units. Process P_1 starts with a loading operation of raw materials that are conveyed to the inputs of machines M_1 and M_2, respectively. The machine M_1 produces, after s_1 time units in average, a part of type O_1 that is conveyed to the input of the assembly station. In a similar way, machine M_2 consumes s_2 time units in average for producing a part of type O_2 that later is put in the input of the assembly station. Process P_2 works in an analogous way but in this case the production of parts O_1 and O_2 is done on the machines M_3 and M_4. The average service times of these machines are s_3 and s_4, respectively. The transport of raw materials to the inputs of machines of a process is provided by a guided vehicle. The time consumed by the transport is negligible with respect to the manufacturing times.

The finished parts are put by the producer machines in two stores, one for each part class. These stores are the inputs for the assembly station. The consumed time to store

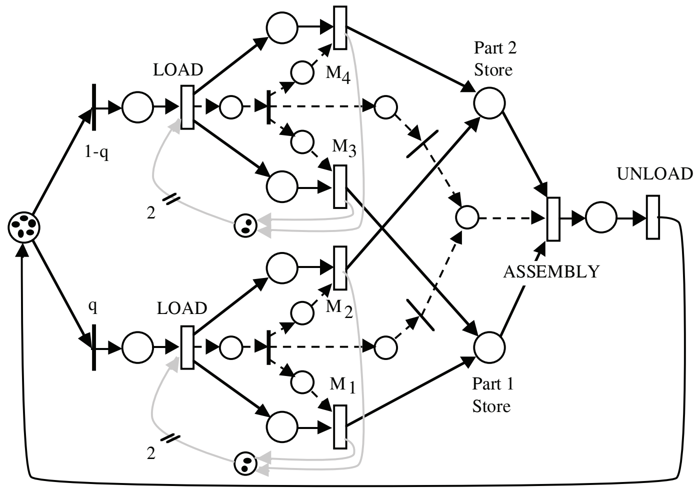

Figure 1: A flexible manufacturing system.

the parts is also negligible with respect to the manufacturing times. Finally, the same conveyor that transports the raw materials is used for conveying the assembled product to an unloading station. The Petri net that models the above described FMS is depicted in figure 1. The dashed lines correspond to the transport of raw materials to the inputs of machines and the movement of vehicles towards the assembly station. The shaded arcs limit the capacity of machines to 1. Because of the lack of space, in the paper we consider a simplified version of the net in figure 1, that is presented in figure 2.a. The simplifications arise from considering negligibles the duration times of operations $LOAD$, M_1, M_4, and transport, and from considering a safe (1–bounded) net system.

The paper is organized as follows. In section 2, an inequality derived from the Little's law is presented as a tool to derive throughput upper bounds of transitions of (live and bounded) stochastic Petri nets. In order to do that, linear marking relations are used. A structural family (i.e., computable from the net structure) of linear marking relations is presented in section 2.1, and two of this kind of relations, derived from P-semiflows and traps, are used in sections 2.2 and 2.3 for the computation of bounds. These bounds are valid for all distribution functions defining the service time of transitions, and they are based on the knowledge of the mean values of those distributions. In section 3, the bounds obtained using traps are reinterpreted using implicit places. Moreover, their introduction increases the number of P-semiflows, thus performance bounds improvements can be expected. Finally, it is shown that the introduction of implicit places improves also the traps-based bounds. Conclusions are summarized in section 4.

sc marked graphs	$\vec{v}^{(i)} = \mathbb{1}$ (constant)
mono-T-semiflow nets	$\vec{v}^{(i)} = \vec{v}^{(i)}(\mathcal{N})$
LB free choice nets	$\vec{v}^{(i)} = \vec{v}^{(i)}(\mathcal{N}, \mathcal{R})$
simple nets	$\vec{v}^{(i)} = \vec{v}^{(i)}(\mathcal{N}, M_0, \mathcal{R}, \vec{s})$

Table 1: Relative firing frequency vector and net subclasses.

2 LITTLE'S LAW AND STRUCTURAL LINEAR MARKING RELATIONS FOR STOCHASTIC PETRI NETS

In this work, we consider stochastic (or deterministic timed, as a particular case) live and bounded Petri nets as synchronized queueing networks [7, 10] and apply Little's result [11] to each place of the net. Let us denote $\overline{M}(p_i)$ the limit average number of tokens at place p_i, $\vec{\sigma}^*$ the limit vector of transition throughputs, and $R(p_i)$ the average time spent by a token within the place p_i (response time at place p_i). Then the above mentioned relationship can be stated as follows:

$$\overline{M}(p_i) = (PRE[p_i] \cdot \vec{\sigma}^*) R(p_i) \tag{1}$$

where $PRE[p_i]$ is the i^{th} row of PRE, thus $PRE[p_i] \cdot \vec{\sigma}^*$ is the output rate of place p_i.

The conditions under which equation (1) holds are very weak, thus this equation is widely applicable. In this work we consider the steady-state behaviour of stochastic Petri nets, under *weak ergodicity* assumption for the firing and the marking processes [5, 6]. Weak ergodicity of a process is considered instead of the usual *strong ergodicity* concept [12] because this one creates problems when we want to include the deterministic case as a special case of a stochastic model. In particular, the marking and the firing processes of a stochastic net are said to be weakly ergodic if the following limits exist:

$$\overline{M} = \lim_{t\to\infty} \frac{1}{t} \int_0^t M(\tau) d\tau < \vec{\infty}; \quad \vec{\sigma}^* = \lim_{t\to\infty} \frac{\vec{\sigma}(t)}{t} < \vec{\infty} \tag{2}$$

$\overline{M}$ and $\vec{\sigma}^*$ are called the limit average marking and the limit throughput vector, respectively.

In the study of distributed systems, Little's law is frequently used when two of the related quantities are known and the third one is needed. This is not exactly the case here. In this case, $R(p_i)$ and $\overline{M}(p_i)$ are unknown, while information about $\vec{\sigma}^*$ can be easily computed only for interesting net subclasses. In effect, here we consider those stochastic nets whose relative firing frequency vector $\vec{v}^{(i)} = \Gamma_i \vec{\sigma}^*$ (or vector of visit ratios to transitions, provided its normalization for a t_i, i.e., $\Gamma_i = 1/\vec{\sigma}_i^*$) can be computed in an efficient way from the net structure $\mathcal{N}$ and from the relative frequency of conflict resolutions $\mathcal{R}$ (i.e., the routing probabilities associated with decisions). This is the case, for instance, for strongly connected marked graphs, live and bounded mono-T-semiflow nets, and live and bounded free choice nets (see table 1). Unfortunately, for simple nets, the relative firing frequency vector depends also on the initial marking M_0 and on the service times $\vec{s}$ of transitions [7].

The *response* times at places $R(p_i)$ are unknown. In fact, they can be expressed as sums of the *waiting* times due to the synchronization schemes and the *service* times associated with transitions, and only service times are known: s_i, $i = 1, \ldots, m$. Thus the response times can be *lowerly bounded* from the knowledge of the average service times, and the following system of inequalities can be derived from (1):

$$\Gamma_i \overline{M} \geq PRE \cdot \vec{D}^{(i)} \tag{3}$$

where $\vec{D}^{(i)}$ is the vector of *average service demands* of transitions, with components $\vec{D}^{(i)}(t_j) = s_j \vec{v}^{(i)}(t_j)$, and Γ_i is called the *mean cycle time for transition* t_i (i.e., the inverse of its throughput).

A goal of this paper is the computation of lower bounds for the mean cycle time Γ_i of transitions t_i (i.e., average time between two consecutive firings of t_i), based on the inequalities (3). Since the limit average marking $\overline{M}$ is unknown, linear marking relations derived from the underlying net will be considered to achieve this goal:

$$Z^T \cdot M \leq k, \quad \forall M \in R(\mathcal{N}, M_0), \text{ with } Z \gneqq 0 \tag{4}$$

Linearity is required in the above relation because, taking into account the definition of the limit average marking, a similar inequality can be derived for $\overline{M}$:

$$Z^T \cdot \overline{M} = \lim_{t \to \infty} \frac{1}{t} \int_0^t Z^T \cdot M(\tau) d\tau \leq \lim_{t \to \infty} \frac{1}{t} \int_0^t k = k \tag{5}$$

In this case, the (unknown) vector $\overline{M}$ can be substituted in (3), premultiplied by Z, obtaining:

$$\Gamma_i \geq (Z^T \cdot PRE \cdot \vec{D}^{(i)})/k \tag{6}$$

(if $k > 0$) and so Γ_i is a lower bound for the mean cycle time of t_i.

An additional advantage can be taken of the use of linear relations, since this linearity will lead, in most cases, to *polynomial complexity* calculations, based on *linear algebra* and *linear programming* techniques.

2.1 Structural linear marking relations

A first family of linear marking relations is obtained considering those being *structurally characterized*. These are stronger conditions than those expressed by inequality (4) (behaviourally defined), but they provide easier and more efficient techniques for their manipulation. Structural linear marking relations can be expressed using the incidence matrix C of the net:

$$a) \quad Y^T \cdot C = 0, Y \geq 0 \Longrightarrow \forall M_0, \forall M \in R(\mathcal{N}, M_0) : Y^T \cdot M = Y^T \cdot M_0 \tag{7}$$

$$b) \quad Z^T \cdot C \lneqq 0, Z \geq 0 \Longrightarrow \forall M_0, \forall M \in R(\mathcal{N}, M_0) : Z^T \cdot M \leq Z^T \cdot M_0 \tag{8}$$

$$c) \quad Z^T \cdot C \gneqq 0, Z \geq 0 \Longrightarrow \forall M_0, \forall M \in R(\mathcal{N}, M_0) : Z^T \cdot M \geq Z^T \cdot M_0 \tag{9}$$

Let us consider firstly the equation (7). Vectors $Y \geq 0$ verifying this equation are called *P-semiflows*, and they have been used for the computation of throughput upper bounds by premultiplying the inequality (3) (see [5, 6, 7]). The obtained results using P-semiflows as well as their limitations for the computation of reachable (i.e., tight) bounds for some nets subclasses are summarized in the next section.

Regarding structural linear inequality relations for the reachable markings of a marked Petri net, vectors $Z \geq 0$ verifying (8) could be considered. Premultiplying the state equation of the net by such vectors, the following sequence of inequalities is obtained for each sequence of successor markings, and for all initial marking M_0:

$$Z^T \cdot M_0 \geq \cdots \geq Z^T \cdot M_{i-1} \geq Z^T \cdot M_i \geq Z^T \cdot M_{i+1} \geq \cdots \tag{10}$$

Moreover, $Z^T \cdot C \neq 0$ implies that there exists (at least) a transition t_j such that $Z^T \cdot C[t_j] < 0$, and if $M_i[t_j\rangle M_{i+1}$ then $Z^T \cdot M_i > Z^T \cdot M_{i+1}$ in the sequence of inequalities (10) (i.e., strict inequality). But in this case the net cannot be live (because if it was live then transition t_j could be fired an infinite number of times, an infinite number of strict inequalities would appear in (10), and this is impossible if the initial marking is finite). Thus linear inequalities of the form $Z^T \cdot C \lneq 0$ are not usefull for us, because no steady-state is possible firing (at least) the transition t_j.

Non-negative vectors satisfying the inequality (9): $Z^T \cdot C \gneq 0$ cannot be used directly for the substitution of $\overline{M}$ in (3) (because they give inequalities in the other direction). Nevertheless, P-semiflows Y could be considered such that $Y - Z \geq 0$, and this is always possible for structurally live and structurally bounded nets, since they are conservative (see, e.g., [8]):

$$(Y - Z)^T \cdot C = \underbrace{Y^T \cdot C}_{0} - Z^T \cdot C \lneq 0 \tag{11}$$

Then the trick of premultiplying inequality (3) by $Y - Z$ could be used, obtaining:

$$\Gamma_i \geq \frac{(Y - Z)^T \cdot PRE \cdot \vec{D}^{(i)}}{(Y - Z)^T \cdot M_0} \tag{12}$$

Unfortunately, the existence of such vectors Z for live nets would imply unboundedness of the marking of places (see, e.g., [8]) and we are not considering this case.

Alternatively, other linear marking inequalities of the form $Y_\Theta^T \cdot M \geq 1$, for all (non-transient) marking M can be derived considering vectors $Y_\Theta \geq 0$ having a *trap* Θ as support. Traps are sets of places which remain marked once they have gained at least one token. This structural concept can be used to improve the throughput upper bound computed by means of Little's law and P-semiflows, as it is explained in section 2.3.

2.2 Little's law and P-semiflows

Considering P-semiflows Y and using relation (6), the following lower bound for the mean cycle time of a given transition t_i can be derived:

$$\Gamma_i \geq \max_{Y \in \{P-semiflow\}} \frac{Y^T \cdot PRE \cdot \vec{D}^{(i)}}{Y^T \cdot M_0} \tag{13}$$

The previous lower bound can be formulated in terms of a fractional programming problem and later, after some considerations, transformed into a linear programming problem (LPP, in the sequel):

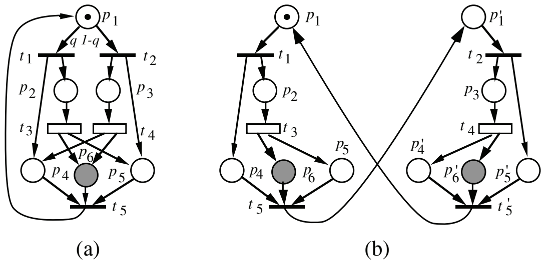

Figure 2: (a) A live and safe free choice net with an implicit place p_6 and (b) its behaviourally equivalent safe marked graph for deterministic resolution of conflict and $q = 1/2$.

Theorem 2.1 [6] *For any net, a lower bound for the mean cycle time* Γ_i *of transition* t_i *can be computed by the following LPP:*

$$\begin{array}{lll} \Gamma_i \geq & \textit{maximize} & Y^T \cdot PRE \cdot \vec{D}^{(i)} \\ & \textit{subject to} & Y^T \cdot C = 0; Y^T \cdot M_0 = 1; Y \geq 0 \end{array} \qquad (14)$$

If the solution of the above problem is unbounded, since it is a lower bound for the mean cycle time of transition t_i, the non-liveness can be assured (infinite cycle time). If the visit ratios of all transitions are non-null, the unboundedness of the above problem implies that a total deadlock is reached by the net. This result has the following interpretation: if the problem (14) is unbounded then there exists an unmarked P-semiflow, and the net is non-live.

The advantage of theorem 2.1 lies on the fact that the *simplex method* for the solution of LPP's has almost linear complexity in practice, even if it has exponential worst case complexity. In any case, algorithms of *polynomial* worst case complexity can be found in [13].

For strongly connected marked graphs, the bound derived from theorem 2.1 has been shown to be reachable for arbitrary mean values and coefficients of variation associated with transition service times [5]. Unfortunately, this is not the case for more general net subclasses. Let us consider, for instance, the live and safe free choice net in figure 2.a without place p_6. Let s_3 and s_4 be the mean service times associated with t_3 and t_4, respectively. Let t_1, t_2, and t_5 be *immediate* transitions (i.e., they fire in zero time). Let $q, 1-q \in (0,1)$ be the probabilities defining the resolution of conflict at place p_1. The relative (to t_5) firing frequency vector is $\vec{v}^{(5)} = (q, 1-q, q, 1-q, 1)^T$. The elementary P-semiflows are $Y_1 = (1,1,0,0,1)^T$ and $Y_2 = (1,0,1,1,0)^T$. Then, applying the problem (14) to this net, the following lower bound for the mean cycle time of transition t_5 is obtained: $\Gamma_5 \geq \max\{qs_3, (1-q)s_4\}$, while the actual cycle time for this transition is $\Gamma_5 = qs_3 + (1-q)s_4$, independently of the higher moments of the probability distribution functions associated with transitions t_3 and t_4. Therefore the bound given by theorem 2.1 is non-reachable for the net in figure 2.a.

2.3 Little's law and traps

A trap in a Petri net $\mathcal{N}$ is a subset of places $\Theta \subseteq P$ such that $\Theta^\bullet \subseteq {}^\bullet\Theta$. A well-known property of these structural elements is recalled now.

Theorem 2.2 [14] *Let $\langle \mathcal{N}, M_0 \rangle$ be a marked Petri net and $\Theta \subseteq P$ a trap. If Θ is initially marked, then Θ is marked throughout the net's evolution.*

This property can be expressed in algebraic terms considering the vector Y_Θ associated with a given trap Θ, and defined as $Y_\Theta(p) = \chi_\Theta(p)$ (we denote by χ_Θ the *characteristic function of the set* Θ, i.e., $\chi_\Theta(p) = 1$ if $p \in \Theta$, and $\chi_\Theta(p) = 0$ otherwise). The next inductive invariant is true: if $Y_\Theta^T \cdot M_0 \geq 1$ then $Y_\Theta^T \cdot M \geq 1$ for all reachable marking M.

Let us consider the vector Y_Θ associated with a trap Θ of a net, and a P-semiflow Y such that $Y - Y_\Theta \geq 0$ (it always exists for conservative nets). The following linear relations can be derived for all reachable marking M and for $\overline{M}$:

$$(Y - Y_\Theta)^T \cdot M \leq Y^T \cdot M_0 - 1 \tag{15}$$

$$(Y - Y_\Theta)^T \cdot \overline{M} \leq Y^T \cdot M_0 - 1 \tag{16}$$

Premultiplying inequality (3) by $Y - Y_\Theta$, a lower bound for Γ_i is derived:

Theorem 2.3 *For any net $\mathcal{N}$ and for any trap Θ of $\mathcal{N}$, a lower bound for the mean cycle time Γ_i of transition t_i is given by:*

$$\begin{array}{rll} \Gamma_i \geq & \textit{maximize} & \dfrac{(Y - Y_\Theta)^T \cdot PRE \cdot \vec{D}^{(i)}}{Y^T \cdot M_0 - 1} \\ & \textit{subject to} & Y^T \cdot C = 0; Y - Y_\Theta \geq 0; Y_\Theta(p) = \chi_\Theta(p) \end{array} \tag{17}$$

Going back to the net in figure 2.a, the unique minimal trap different from the P-semiflows is $\Theta = \{p_1, p_4, p_5\}$. Considering the P-semiflow $Y = (2, 1, 1, 1, 1)^T$, we have $Y \geq Y_\Theta = (1, 0, 0, 1, 1)^T$, and theorem 2.3 can be applied: $\Gamma_5 \geq q s_3 + (1 - q) s_4$. Therefore the bound obtained in section 2.2 using only P-semiflows has been improved (in fact the bound computed now coincides with the actual cycle time).

In order to explain in an intuitive way (with the example) the reason of the previous improvement, let us derive a behaviourally equivalent safe marked graph [15] (figure 2.b without places p_6 and p'_6) for the safe free choice net of figure 2.a without place p_6, assuming for the sake of simplicity that the resolution of conflict at place p_1 is deterministic with $q = 1/2$ (i.e., transitions t_1 and t_2 fire once each one, alternatively). The lower bound for the mean cycle time of this marked graph based on theorem 2.1 (i.e., using the P-semiflows) is $\Gamma_{\mathrm{MG}} \geq s_3 + s_4$ (in fact it is reached) and corresponds to the circuit $\langle p_1, p_2, p_5, p'_1, p_3, p'_4 \rangle$. Since transition t_5 appears instantiated twice in the marked graph, the obtained bound for the cycle time of this transition is $\Gamma_5 \geq (s_3 + s_4)/2$. In the original free choice net there does not exist any minimal P-semiflow including both p_2 and p_3 in its support, thus the previous bound is not obtained.

In the next section we recall the concept of *implicit* place. It allows us to reinterpret the improvement of the upper bound on throughput using the P-semiflows of a behaviourally equivalent net, obtained by adding implicit places, instead of the trap structures in the original one. Therefore, the application of (14) to the transformed net can improve our bounds.

3 A NEW PERSPECTIVE: IMPLICIT PLACES

Let $\mathcal{N}$ be a net and $\mathcal{N}^p$ be the net resulting from adding a place p to $\mathcal{N}$. If M_0 is an initial marking of $\mathcal{N}$, M_0^p denotes the initial marking of $\mathcal{N}^p$, i.e., $M_0^p(p') = M_0(p')$ for all $p' \neq p$. Given a net $\langle \mathcal{N}^p, M_0^p \rangle$, the place p is *implicit* [16, 17, 18] iff $L(\mathcal{N}^p, M_0^p) = L(\mathcal{N}, M_0)$ (i.e., its elimination preserves the firing sequences).

Implicit places are defined according with *interleaving semantics* of concurrent systems. In order to preserve the behaviour of timed net systems, *true concurrency* must be preserved (at least the *step semantics*). We consider a *step* of $\langle \mathcal{N}, M_0 \rangle$ as a multiset of transitions such that there exists a reachable marking M that allows to fire simultaneously all transitions in the multiset.

Proposition 3.1 [17] *If p is a selfloop-free place then p is implicit iff its elimination preserves the steps of the net system.*

In the following we only consider selfloop-free implicit places. As an example let us consider once more the net in figure 2.a. Place p_6 is implicit since its elimination does not change the firing sequences of the net.

Implicit places are behaviourally defined. The structural counterpart requires that for all M_0 of $\mathcal{N}$ the place under consideration is implicit for at least one finite initial marking (for this place). These places will be called structurally implicit (SIP).

In [18] linear programming techniques are used for deriving necessary and sufficient conditions for a place to be SIP as well as for computing an upper bound of the minimum initial marking that makes it implicit.

Theorem 3.1 [18]

1. *A place p is SIP in $\mathcal{N}^p$ iff $\exists Y \geq 0$ such that $Y^T \cdot C \leq l_p$, where C is the incidence matrix of $\mathcal{N}$ and l_p is the incidence vector of place p in $\mathcal{N}^p$.*

2. *Let p be a SIP. An upper bound of the minimum initial marking $M_0^p(p)$ for which p is implicit is $M_0^p(p) = \max\{0, v\}$, where*

$$
\begin{array}{rll}
v = & \textit{minimize} & Y^T \cdot M_0 + \mu \\
 & \textit{subject to} & Y^T \cdot C \leq l_p \\
 & & Y^T \cdot PRE[t_k] + \mu \geq pre(p, t_k), \forall t_k \in p^\bullet \\
 & & Y \geq 0
\end{array}
\tag{18}
$$

As it is remarked in [18], theorem 3.1.2 allows to detect SIP's (p is SIP iff (18) has a feasible solution) and implicit places (if the initial marking of p is greater than or equal to that computed by (18)) in polynomial time.

Now, we reinterpretate the linear marking relations derived from traps using implicit places. Let us consider again the net in figure 2.a without place p_6 and its behaviourally equivalent (for $q = 1/2$) marked graph depicted in figure 2.b without places p_6 and p_6'. The elementary circuits (P-semiflows) of this marked graph are $c_1 = \langle p_1, p_2, p_5, p_1', p_5' \rangle$, $c_2 = \langle p_1, p_4, p_1', p_3, p_4' \rangle$, $c_3 = \langle p_1, p_2, p_5, p_1', p_3, p_4' \rangle$, and $c_4 = \langle p_1, p_4, p_1', p_5' \rangle$. The circuits c_1 and c_2 correspond to the elementary P-semiflows of the original net Y_1 and Y_2, respectively.

Thus, these circuits cannot contribute to the improvement of the bound computed for the original net based on the P-semiflows. This is not the case for the circuits c_3 and c_4. These circuits add linear information which is not reflected by P-semiflows in the original net. The circuit c_4 does not include any timed transition and must not be considered. On the other hand, the circuit c_3 reflects the sequentialization of transitions t_3 and t_4, and it gives the actual cycle time of the net system.

A given elementary circuit of the derived marked graph does not correspond to any elementary P-semiflow of the original free choice net when it includes several instances of a unique transition and each instance has as input (or output) places which are instances of different original places. This is the case, for example, for the circuit c_3 of the marked graph of figure 2.b without p_6 and p'_6. It includes instances t_5 and t'_5 of a unique transition, and the input places of these transitions in circuit c_3 are p_5 and p'_4, respectively, which are instances of different original places.

Now, let us increment the number of circuits of the marked graph of figure 2.b, by adding the places p_6 and p'_6. Places p_6 and p'_6 are replicas of places p_5 and p'_4, respectively (thus they are implicit), and can be supposed to be different instances of a new (implicit) place in the original net (place p_6 of the net in figure 2.a). The addition of this place generates a new elementary P-semiflow $Y_3 = (1,1,1,0,0,1)^T$. With this P-semiflow, the lower bound for the cycle time computed with problem (14) is $\Gamma_5 \geq qs_3 + (1-q)s_4$, which is the same obtained in section 2.3 using relations derived from trap structures.

Let us remark that the relation between the implicit place p_6 of the net in figure 2.a and the trap $\Theta = \{p_1, p_4, p_5\}$ considered in section 2.3 is straightforward: $l_{p_6} = Y_\Theta^T \cdot C$, that is, the incidence vector of p_6 is the sum of those of places p_1, p_4, and p_5.

The following linear relation can be derived from the trap Θ (and the P-semiflow $Y = Y_1 + Y_2$):

$$(Y - Y_\Theta)^T \cdot M = M(p_1) + M(p_2) + M(p_3) \leq 1, \forall M \in R(\mathcal{N}, M_0) \tag{19}$$

While this one follows from the new P-semiflow Y_3 that includes the implicit place p_6:

$$Y_3^T \cdot M = M(p_1) + M(p_2) + M(p_3) + M(p_6) = 1, \forall M \in R(\mathcal{N}, M_0) \tag{20}$$

It can be pointed out that the information given by relation (19) is included in that given by the new P-semiflow (equation (20)), because $M(p_6) \geq 0$.

Now, technical details related with the addition of implicit places which improve the throughput upper bound computed by means of P-semiflows and traps are considered.

Let us consider an initially marked trap Θ of a given net $\mathcal{N}$, and its associated vector Y_Θ defined as in previous sections. The following result, which follows from theorem 3.1.1, assures that a structurally implicit place p_Θ associated with Θ, can be added to $\mathcal{N}$.

Corollary 3.1 *Let Θ be an initially marked trap of $\mathcal{N}$, $Y_\Theta(p) = \chi_\Theta(p)$, $Y_\Theta^T \cdot M_0 \geq 1$, and $\mathcal{N}^{p_\Theta}$ the net resulting from the addition of place p_Θ with incidence vector $l_{p_\Theta} = Y_\Theta^T \cdot C$ to $\mathcal{N}$. Then p_Θ is structurally implicit in $\mathcal{N}^{p_\Theta}$.*

The importance of the previous structural implicit place lies on the fact that, if a marking makes it implicit (e.g., the marking given by theorem 3.1.2), then the lower bound for the mean cycle time of a transition computed using P-semiflows of the augmented net can improve the bounds based on P-semiflows of the original net (theorem 2.1) and on the trap Θ (theorem 2.3). Before to state this result we firstly present a technical lemma.

Lemma 3.1 *Let Θ be an initially marked trap of $\mathcal{N}$, $Y_\Theta(p) = \chi_\Theta(p)$. Let p_Θ be a place defined as $l_{p_\Theta} = Y_\Theta^T \cdot C$. Then Y_Θ, $\mu_\Theta = -1$ is a feasible solution of the problem (18) and $M_0^{p_\Theta}(p_\Theta) \leq Y_\Theta^T \cdot M_0 + \mu_\Theta$ (place p_Θ is assumed to be pure, i.e., selfloop-free).*

Proof. $Y_\Theta^T \cdot C = l_{p_\Theta} \Rightarrow \forall t \in {}^\bullet p_\Theta : Y_\Theta^T \cdot POST[t] - Y_\Theta^T \cdot PRE[t] = -pre(p_\Theta, t)$. Taking into account that Y_Θ is the characteristic function of a trap we have in the last equality that $Y_\Theta^T \cdot PRE[t] > 0$ if and only if $Y_\Theta^T \cdot POST[t] > 0$. Therefore, $\forall t \in {}^\bullet p_\Theta$: $Y_\Theta^T \cdot PRE[t] > pre(p_\Theta, t)$ and from the problem (18) we conclude that Y_Θ and $\mu_\Theta = -1$ is a feasible solution. From (18) we also conclude directly that $M_0^{p_\Theta}(p_\Theta) \leq Y_\Theta^T \cdot M_0 + \mu_\Theta$ because $Y_\Theta^T \cdot M_0 \geq 1$. ∎

Theorem 3.2 *Let $\langle \mathcal{N}, M_0 \rangle$ be a marked net, Θ an initially marked trap of $\mathcal{N}$, $Y_\Theta(p) = \chi_\Theta(p)$, and $\langle \mathcal{N}^{p_\Theta}, M_0^{p_\Theta} \rangle$ the marked net resulting from the addition to the original net of the structural implicit place p_Θ with incidence vector $l_{p_\Theta} = Y_\Theta^T \cdot C$ and with $M_0^{p_\Theta}(p_\Theta)$ given by theorem 3.1.2.*

1. *A lower bound $\Gamma_i^{p_\Theta}$ for the mean cycle time Γ_i of transition t_i in $\langle \mathcal{N}, M_0 \rangle$ can be computed applying theorem 2.1 to the net $\langle \mathcal{N}^{p_\Theta}, M_0^{p_\Theta} \rangle$.*

2. *If Γ_i^{PS} and Γ_i^{Θ} are the lower bounds of Γ_i derived from the direct application of theorems 2.1 and 2.3, respectively, to the original net, then $\Gamma_i^{p_\Theta} \geq \Gamma_i^{PS}$ and $\Gamma_i^{p_\Theta} \geq \Gamma_i^{\Theta}$.*

Proof. $\Gamma_i^{p_\Theta}$ is a lower bound for the mean cycle time of t_i in $\langle \mathcal{N}^{p_\Theta}, M_0^{p_\Theta} \rangle$ by theorem 2.1. Since p_Θ is implicit, t_i has the same mean cycle time in $\langle \mathcal{N}, M_0 \rangle$ and in $\langle \mathcal{N}^{p_\Theta}, M_0^{p_\Theta} \rangle$. Then, $\Gamma_i^{p_\Theta}$ is a lower bound for the mean cycle time of t_i in $\langle \mathcal{N}, M_0 \rangle$.

$\Gamma_i^{p_\Theta} \geq \Gamma_i^{PS}$ because if Y is a P-semiflow of $\mathcal{N}$, then $Z = [Y^T | 0]^T$ is a P-semiflow of $\mathcal{N}^{p_\Theta}$.

Finally, we prove that $\Gamma_i^{p_\Theta} \geq \Gamma_i^{\Theta}$. Let Y be a P-semiflow of $\mathcal{N}$ such that $Y - Y_\Theta \geq 0$. Then $Z = [(Y - Y_\Theta)^T | 1]^T$ is a P-semiflow of $\mathcal{N}^{p_\Theta}$. Now, applying equation (13) for $\Gamma_i^{p_\Theta}$:

$$
\begin{aligned}
\Gamma_i^{p_\Theta} &\geq \frac{[(Y - Y_\Theta)^T | 1] \cdot PRE^{p_\Theta} \cdot \vec{D}^{(i)}}{Y^T \cdot M_0 - Y_\Theta^T \cdot M_0 + M_0^{p_\Theta}(p_\Theta)} = \\
&= \frac{(Y - Y_\Theta)^T \cdot PRE \cdot \vec{D}^{(i)}}{Y^T \cdot M_0 - Y_\Theta^T \cdot M_0 + M_0^{p_\Theta}(p_\Theta)} + \frac{pre(p_\Theta) \cdot \vec{D}^{(i)}}{Y^T \cdot M_0 - Y_\Theta^T \cdot M_0 + M_0^{p_\Theta}(p_\Theta)}
\end{aligned} \tag{21}
$$

And this value is greater than or equal to that given by equation (17) in theorem 2.3 because the second term of the above sum is non negative and the first term in the above sum is greater than or equal to that given by equation (17) in theorem 2.3 (take into account that $M_0^{p_\Theta}(p_\Theta) \leq Y_\Theta^T \cdot M_0 - 1$, by lemma 3.1, and that the denominator is less than or equal to $Y^T \cdot M_0 - 1$). ∎

Theorem 3.2.2 tells that the addition of implicit places allows to obtain better bounds than those computed using traps or the P-semiflows of the original net. The problems that remain are how to add implicit places (i.e., which ones allow to improve the bounds) and when no more improvements are possible with the technique.

In the previous section, the net of figure 2.a without p_6 has been considered as an example in which the bound computed using the trap $\Theta = \{p_1, p_4, p_5\}$ is tight because it reaches the actual value of the mean cycle time. It is also shown that the same value can be derived, after the addition of the associated implicit place p_6, considering the new P-semiflow $(1,1,1,0,0,1)^T$.

Let us consider the same net of figure 2.a without p_6, but assuming now that transition t_5 is not immediate but timed, with mean service time equal to s_5. If problem (14) is applied to the net, the following bound is obtained: $\Gamma_5^{PS} = \max\{qs_3 + s_5, (1-q)s_4 + s_5\}$. If trap $\Theta = \{p_1, p_4, p_5\}$ and P-semiflow $Y = (2,1,1,1,1)^T$ are considered, inequality (17) gives the bound: $\Gamma_5^{\Theta} = qs_3 + (1-q)s_4$. If the implicit place associated with Θ is added to the net, theorem 3.2 gives the bound: $\Gamma_5^{p_\Theta} = qs_3 + (1-q)s_4 + s_5$ (for the P-semiflow $(1,1,1,0,0,1)^T$), which improves both Γ_i^{PS} and Γ_i^{Θ}, and, in fact, it gives the actual cycle time of transition t_5 (i.e., it is tight). Note that, in this case, the improvement is due to the non-null second term of the expresion (21).

Finally, for the original net of figure 1, if $s_l, s_1, s_2, s_3, s_4, s_{ass}, s_{ul}$ are the average service times of transitions $LOAD$, M_1, M_2, M_3, M_4, $ASSEMBLY$, $UNLOAD$, the application of problem (14) (based on P-semiflows of the original net) gives the bound:

$$\begin{aligned}\Gamma_{ul} \geq \max\{ \ & (2qs_l + qs_1 + qs_2)/2, (2(1-q)s_l + (1-q)s_3 + (1-q)s_4)/2, \\ & (qs_l + (1-q)s_l + qs_1 + (1-q)s_3 + s_{ass} + s_{ul})/5, \\ & (qs_l + (1-q)s_l + qs_2 + (1-q)s_4 + s_{ass} + s_{ul})/5\}\end{aligned}$$

If implicit places p_1 and p_2 are added, with ${}^\bullet p_1 = \{M_1, M_4\}$, $p_1^\bullet = \{ASSEMBLY\}$, ${}^\bullet p_2 = \{M_2, M_3\}$, and $p_2^\bullet = \{ASSEMBLY\}$, the application of theorem 3.2 gives the value:

$$\begin{aligned}\Gamma_{ul} \geq \max\{ \ & (2qs_l + qs_1 + qs_2)/2, (2(1-q)s_l + (1-q)s_3 + (1-q)s_4)/2, \\ & (qs_l + (1-q)s_l + qs_1 + (1-q)s_3 + s_{ass} + s_{ul})/5, \\ & (qs_l + (1-q)s_l + qs_2 + (1-q)s_4 + s_{ass} + s_{ul})/5, \\ & (qs_l + (1-q)s_l + qs_1 + (1-q)s_4 + s_{ass} + s_{ul})/5, \\ & (qs_l + (1-q)s_l + qs_2 + (1-q)s_3 + s_{ass} + s_{ul})/5\}\end{aligned}$$

And this bound improves the previous one, because two new P-semiflows have been considered.

4 CONCLUSIONS

We have addressed the problem of computing upper bounds for the throughput of transitions in Petri net models (or the corresponding synchronized queueing networks), using only the net structure, the routing probabilities, and the mean service time of transitions.

Until now, only P-semiflows had been used from the net structure in order to compute such bounds. In this paper, other structural linear marking relations have been considered and, in particular, traps have been used for the computation of better bounds. Parallelling the improvements for the analysis of qualitative properties in [18], implicit places have been shown to be an interesting tool for the improvement of structure-based performance bounds.

5 ACKNOWLEDGEMENTS

This work was partially supported by project PA86-0028 of the Comisión Interministerial de Ciencia y Tecnología, the DEMON Esprit Basic Research Action 3148, and the Plan Nacional de Investigación TIC-0358/89.

6 REFERENCES

[1] *Proceedings of the International Workshop on Timed Petri Nets*, Torino, Italy, July 1985. IEEE-Computer Society Press.

[2] *Proceedings of the International Workshop on Petri Nets and Performance Models*, Madison, WI, USA, August 1987. IEEE-Computer Society Press.

[3] *Proceedings of the* 3^{rd} *International Workshop on Petri Nets and Performance Models*, Kyoto, Japan, December 1989. IEEE-Computer Society Press.

[4] *Proceedings of the* 4^{rd} *International Workshop on Petri Nets and Performance Models*, Melbourne, Australia, December 1991. IEEE-Computer Society Press.

[5] 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*, pages 200–209, Kyoto, Japan, December 1989. IEEE-Computer Society Press.

[6] J. Campos, G. Chiola, and M. Silva. Ergodicity and throughput bounds of Petri nets with unique consistent firing count vector. *IEEE Transactions on Software Engineering*, 17(2):117–125, February 1991.

[7] J. Campos, G. Chiola, and M. Silva. Properties and performance bounds for closed free choice synchronized monoclass queueing networks. *IEEE Transactions on Automatic Control*, 36(12):1368–1382, December 1991.

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

[9] M. Ajmone Marsan, G. Balbo, and G. Conte. A class of generalized stochastic Petri nets for the performance evaluation of multiprocessor systems. *ACM Transactions on Computer Systems*, 2(2):93–122, May 1984.

[10] J. Campos. *Performance Bounds for Synchronized Queueing Networks.* PhD thesis, Departamento de Ingeniería Eléctrica e Informática, Universidad de Zaragoza, Spain, October 1990. Research Report GISI-RR-90-20.

[11] J. D. C. Little. A proof of the queueing formula $L = \lambda W$. *Operations Research*, 9:383–387, 1961.

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

[13] G. L. Nemhauser, A. H. G. Rinnooy Kan, and M. J. Todd, editors. *Optimization*, volume 1 of *Handbooks in Operations Research and Management Science*. North-Holland, Amsterdam, The Netherlands, 1989.

[14] M. H. T. Hack. Analysis of production schemata by Petri nets. M. S. Thesis, TR-94, M.I.T., Boston, USA, 1972.

[15] J. Campos and J. M. Colom. A reachable throughput upper bound for live and safe free choice nets. In *Proceedings of the Twelfth International Conference on Application and Theory of Petri Nets*, pages 237–256, Gjern, Denmark, June 1991.

[16] M. Silva. *Las Redes de Petri en la Automática y la Informática.* Editorial AC, Madrid, 1985. In Spanish.

[17] J. M. Colom. *Análisis Estructural de Redes de Petri, Programación Lineal y Geometría Convexa.* PhD thesis, Departamento de Ingeniería Eléctrica e Informática, Universidad de Zaragoza, Spain, June 1989. Research Report GISI-RR-89-11. In Spanish.

[18] J. M. Colom and M. Silva. Improving the linearly based characterization of P/T nets. In G. Rozenberg, editor, *Advances in Petri Nets 1990*, volume 483 of *LNCS*, pages 113–145. Springer-Verlag, Berlin, 1991.