University of West Hungary
Simonyi Károly Faculty of Engineering, Wood Sciences and Applied Arts

József Cziráki Doctoral School of Wood Sciences and Technologies
Infocommunication Technologies in Wood Sciences

Some Performance Analysis Applications of Stochastic Modeling

Ph.D. Dissertation
of

Ádám Horváth

Research Supervisors:
Károly Farkas, PhD.
Tien Van Do, DSc.

Sopron
2014

Alulírott *Horváth Ádám* kijelentem, hogy ezt a doktori értekezést magam készítettem, és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

I, the undersigned *Ádám Horváth* hereby declare that this Ph.D. dissertation was made by myself, and I only used the sources given at the end. Every part that was quoted word-for-word, or was taken over with the same content, I noted explicitly by giving the reference of the source.

Sopron, March 31st, 2014

. . . .

Ádám Horváth

Abstract

Models which have been used for a long time to understand and analyze processes, try to capture the essence of problems, and – as far as possible – simply describe the operation of these processes. In this dissertation, we deal with two main topics. First, we investigate the spreading of services (applications), model the spreading process, and evaluate the models. Then, we propose a model for modeling the opportunistic spectrum access and analyzing its effects.

Nowadays, mobile applications became even more popular with the proliferation of smart phones. Applications can be purchased usually through a web shop. However, there exist decentralized technologies like self-organized or ad hoc networks, in which users can download and try out applications directly from each other (secure payment in this environment is still a challenging issue today, therefore, direct application download concerns only the trial versions of the applications) [16]. The possibility of trying out an application can give additional motivation to purchase it. Moreover, there are other advantages like the community experience in case of a multi-player game. This approach can better motivate the users to purchase an application than they would have seen only some advertisements [1, 2, 3, 4, 5]. Since one of the goals in this work is to point out the relation of this area and the wood industry, we also modeled the production process of a company producing wooden windows. Using the model, the leaders of the company can find answers for questions like determining the bottleneck in the producing process [7].

Our other topic is also a novel one. However, the basic idea of the opportunistic spectrum access is not a novelty [17]. Nowadays, it is widely recognized that spectrum management reform and dynamic spectrum access can provide a solution to an existing problem (the shortage of usable radio frequencies and the under-utilization of the licensed spectrum) [18, 19, 20], the application of dynamic and opportunistic spectrum access is rarely found in practice. A lot of issues [18, 19, 20, 21] must be solved before the widespread application of opportunistic spectrum access. Amongst these issues, a question concerning the investment (on the licensed frequencies and technology) protection of the incumbent operators plays an important role regarding the acceptance of opportunistic spectrum access.

In both areas, we made an effort to reconsider the inflexible structure of the existing approaches and give alternative solutions for the main questions of the investigated topics. In the first part of the dissertation, we investigate the mobile application spreading. In the proposed model, the application spreading process is based on the direct (ad hoc) communication between the users. In case of the mobile cellular networks, we can observe also an inflexible structure: the service providers can get exclusive right to certain frequency bands on auctions held by the government. The exclusiveness involves the bad utilization of the means, although utilization becomes a key factor with the increasing demand for the bandwidth in this area. In the second part of the dissertation, we propose the introduction of the opportunistic spectrum access.

Kivonat

A folyamatok megértéséhez és elemzéséhez régóta használunk modelleket, amelyek egy probléma lényegét próbálják megfogni, s lehetőség szerint egyszerűen leírni a folyamatok működését. Jelen értekezésben két fő területet vizsgálunk meg: az egyik a szolgáltatások (alkalmazások) terjedése, a terjedés modellezése, valamint a modellek kiértékelése; a másik az opportunista spektrum hozzáférés modellezése és elemzése mobil cellás hálózatokban.

Napjainkban az okostelefonok térhódításával méginkább népszerűbbek lettek a mobil alkalmazások, melyeket hagyományosan egy webes áruházon keresztül szerezhetünk meg. Léteznek azonban már olyan decentralizált technológiák, mint az önszerveződő vagy ad hoc hálózatok, melyekben a felhasználók egymástól közvetlen módon tudnak alkalmazásokat letölteni, és kipróbálni azokat (a biztonságos vásárlás ebben a környezetben még nem megoldott, ezért itt csak az alkalmazások próba verziójáról van szó) [16]. Az alkalmazás kipróbálásának lehetősége további motivációt nyújthat a vásárláshoz, nem beszélve olyan előnyökről, mint például a közösségi élmény egy többfelhasználós játék esetében. Ez a szemlélet több motivációt jelenthet a vásárláshoz, mint ha valaki csak megnéz egy hirdetést [1, 2, 3, 4, 5]. Mivel a dolgozat céljai között az is szerepel, hogy rámutassunk a terület faiparhoz kapcsolódó pontjaira, elkészítettük egy ablakgyártó cég gyártási folyamatának modelljét is. A modell segítségével olyan kérdésekre kaphatunk választ, mint a szűk keresztmetszet meghatározása a gyártási folyamatban [7].

Másik témánk is újszerű területet ölel fel, habár az opportunista spektrum hozzáférés alapötlete nem újkeletű [17]. Ma már széles körben felismert tény, hogy a spektrum menedzsment reformja és a dinamikus spektrum hozzáférés megoldást nyújtana egy létező problémára: a felhasználható rádió spektrum hiányára és a már kiosztott spektrum alacsony kihasználtságára [18, 19, 20]. Ennek ellenére a dinamikus vagy opportunista spektrum hozzáférés nem terjedt el a gyakorlatban, mivel számos feladatot [18, 19, 20, 21] kell megoldani ahhoz, hogy az említett eljárások széles körben alkalmazhatóak legyenek. Ezen feladatok közül kiemelt szerepet játszik a kizárólagos használatba befektetett vagyon védelmének kérdése, ha el akarjuk fogadtatni az opportunista spektrum hozzáférést.

Munkánk során mindkét területen arra törekedtünk, hogy a meglévő rendszerek sokszor merev struktúráját újragondolva alternatív megoldásokat adjunk a vizsgált terület fő kérdéseire. Az alkalmazások terjedésének vizsgálatakor a központosított terjedéssel szemben egy olyan modellt állítottunk fel, amelynek alapja a felhasználók közötti közvetlen, ad hoc kommunikáció, erről szól a dolgozat első része. A mobil cellás hálózatok esetén szintén egy merev struktúrát figyelhetünk meg: a szolgáltatók a kormányzat által kiírt aukciókon szerezhetnek licenszet adott frekvencia sávok kizárólagos használatára. Ez a kizárólagosság viszont az erőforrások rossz kihasználtságát vonja maga után, pedig az egyre növekvő sávszélesség igény mellett a hatékony kihasználtság kulcsfontosságú tényező lesz ezen a területen. Javaslatunk az opportunista spektrum hozzáférés bevezetése, melyet munkánk második részében tárgyalunk.

Acknowledgments

I am thankful to my supervisors: Dr. Károly Farkas and Prof. Tien Van Do, Department of Networked Systems and Services, Budapest University of Technology and Economics, for the professional support and guidance during the preparation of this dissertation.

I would like to thank Prof. László Jereb, Institute of Informatics and Economics, University of West Hungary, for the support and encouragement provided so many years, without which this work would not have been born.

I also thank my former co-authors, especially Do Hoai Nam and Dr. András Horváth, who helped me a lot in my research.

Last but not least, I am heartily thankful to my family for supporting my studies.

This research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the last few years, the number of mobile devices equipped with wireless communication interfaces has increased significantly (the ratio of the smart phones reached 26% in Hungary at the end of 2012 [22]). Besides, the number of mobile applications and the demand for broadband spectrum access show an increasing tendency, too. The change of the typical user behavior demands innovative solutions from the researchers: alternative directions have to be offered for mobile users, which were designed according to the current trends. Accordingly, we propose a novel approach for mobile application spreading in the first part of this dissertation; and describe our opportunistic spectrum access model for mobile cellular networks in the second part.

1.1 Motivation

In spite of the above mentioned tendency, mobile application spreading is not a frequently investigated research topic today. Traditionally, mobile applications are spread via a central entity, like an Internet web shop. Users can browse the web site of the merchant (or use a mobile market application), select, purchase and download the application they like. However, decentralized technologies such as self-organized or mobile ad hoc networks allow the users to get the application software directly from each other [16]. Hence, direct application download can change the characteristics of traditional application spreading. The participants of this direct communication can even try out the applications and be motivated to purchase the ones they liked (purchasing is available only via a traditional way, because secure payment in this environment is still a challenging issue today). In this type of communication, a lot of factors can change the characteristics of the spreading process, especially from economic viewpoint, which have not taken into consideration yet (e.g., community experience when playing a multi-player game). These factors can give more motivation to the users to purchase the application than they would have seen only some advertisements [1, 2, 3, 4, 5].

Another promising topic is the spectrum sharing in mobile cellular networks, typically in Global System for Mobile Communications (GSM) networks, where the increasing demand for broadband spectrum access results in the scarcity of the available spectrum. At present, the exclusive access right to certain radio frequency bands is licensed to mobile network operators by the governments. The license of the frequency bands is guaranteed based on the result of spectrum auctions. However, it is already recognized that this exclusive access may lead to an inefficient use of the spectrum [23, 24]. Network operators can utilize spectrum renting to increase the efficiency of the spectrum usage [23, 24, 25, 26, 27, 28] and to relieve the temporary capacity shortage of a particular cell in a mobile cellular network. For example, when the number of calls increases in a specific area, a network operator could decide to rent a frequency band from another operator to keep or enhance the grade of service of calls. From the service providers' point of view, it is crucial to show that besides improving the technical parameters, the cooperation is also financially beneficial for each cooperating party.

1.2 Research problem statement

With the spreading of the smart phones, more and more opportunities are offered for the users, which have been exploited only in part yet. The engineers and the developers find innovative solutions and applications, what can open new directions in the future. However, there are still open questions, which need to be answered. In this dissertation, we deal with two problem sets.

Problem 1. *Application spreading in mobile ad hoc environments.* The actual form of application spreading is a centralized one. However, the distributed approach offers new ways of communications, which must not be ignored: exploiting the direct connections between the mobile devices could affect the application spreading process. Being aware of the characteristics of application spreading is important for the application provider not only from technical, but also from economic point of view. The application provider has to know or at least assess how much money he can earn from the purchases of a given application; how much time is needed to realize it; and which factors influence the spreading process and how [1, 2, 3, 4, 5].

Problem 2. *Bottlenecks in the manufacturing process of wooden windows.* In many manufacturing processes, the work can be divided into different disjoint phases having deterministic holding times. If a company wants to increase its production, it is advantageous to do it by the expansion of a single work phase, which is a bottleneck in the system. Therefore, it is necessary to identify the bottleneck of the manufacturing process. Moreover, it is also desirable to determine the measure of the expansion for the elimination of the main bottleneck [7].

Problem 3. *Opportunistic spectrum access in mobile cellular networks.* The current regulation of spectrum licensing guarantees exclusive access to certain frequency bands for the mobile telecommunication service providers. As some recent research papers have pointed out [23, 24], the exclusiveness has a negative effect on the efficiency. For the realization of a spectrum sharing system, several obstacles must be beaten off. Many researchers work in this area to solve the physical/technical problems, such as identifying the idle spectral ranges, interference problems, time and frequency synchronization [21]. Besides, there are still open questions if the physical difficulties are got over. Since the service providers pay license fee for using the spectrum, a basic requirement is to show that another approach can protect their investments. Moreover, even an extra profit can be achieved by using spectrum sharing. The other interesting question is the quality of the service, which is mainly the mobile subscribers' interest. Finally, the scarcity of the available spectrum will be sooner or later a great challenge for the governments.

1.3 Results

In the following, we briefly present the contributions of this dissertation. We additionally give the corresponding publications and chapters of the dissertation that describe a particular contribution in detail.

Two Performance Analysis Applications Based on Deterministic and Stochastic Petri Net Models

- We proposed a Closed Queuing Network (CQN) model which can be simply used for giving an analytical lower and upper bounds on the number of application purchases.
  Related publications: [1], [2], [3], [4], [5]
  Chapter: 3

- We demonstrated that the mean field based methodology can be applied for obtaining the transient solution of a Stochastic Petri Net (SPN), if the underlying Markov chain of the SPN is density dependent.
  Related publications: [5], [6]
  Chapter: 3

- We applied the mean field based methodology for obtaining the transient solution of our basic SPN model, and we gave an analytical approximation on the number

of application purchases in the order of seconds.
Related publications: [5], [6]
Chapter: 3

- We proposed an extended version of the basic SPN model, for which the main properties of the application spreading process can be determined by running transient simulation.
  Related publications: [5]
  Chapter: 3

- Using a Deterministic and Stochastic Petri Net (DSPN) model, we identified the bottleneck of the wooden window production process and determined the measure of the extension for eliminating the main bottleneck.
  Related publications: [7]
  Chapter: 3

Opportunistic spectrum access in mobile cellular networks

- We elaborated a spectrum sharing policy based on the idea of opportunistic spectrum access. In the model, a high level of cooperation is realized between the mobile service providers. Besides, the model considers the current technical constraints, too, which were ignored by most of the related works.
  Related publications: [8], [9], [10]
  Chapter: 4

- We demonstrated via simulations that the service quality can be improved applying the elaborated spectrum sharing policy. Moreover, we also demonstrated that the cooperating parties can realize more profit using our model than in the current environment.
  Related publications: [8]
  Chapter: 4

- We elaborated the mathematical model of the above mentioned spectrum sharing policy. We used a two-dimensional Continuous-Time Markov Chain (CTMC) to get the numerical results of the model. Since the results correspond to the simulation results, the Markovian mathematical model can be considered as a good approximation of the original model, where the channel holding times and the interarrival times are log-normally distributed.
  Related publications: [10]
  Chapter: 4

- We identified and measured the main drawback of our model, the forced termination phenomenon. In a heavily loaded system, the forced termination increases to a level that is annoying for the mobil subscribers. To handle this problem, we elaborated a method for the protection of the ongoing calls based on the Adaptive Random Early Detection (ARED) rule [30].
  Related publications: [8], [10]
  Chapter: 4

1.4 Outline

The outline of this dissertation is as follows.

Chapter 2: This chapter provides a summary of the applied research methodology in the dissertation, such as queuing networks, continuous-time Markov chains and stochastic Petri nets.

Chapter 3: In this chapter, we present two performance analysis applications based on deterministic and stochastic Petri net models. In Section 3.1, we investigate application spreading in mobile ad hoc environments. We give an overview of the related works concerning application spreading in Section 3.1.1. Section 3.1.2 presents a short description about how the application spreading process using ad hoc networks takes place. In Section 3.1.3, we describe the communication model and define the user behavior types. We present two techniques to investigate the application spreading process, thus our CQN and SPN models together with some results in Section 3.1.4 and Section 3.1.5, respectively. To point out the applicability of the mentioned modeling processes, we describe the operation model of a wooden window producing company and present some proposal to improve the effectiveness of the production in Section 3.2. Finally, we summarize the results of the chapter in Section 3.3.

Chapter 4: This chapter describes our work in the area of opportunistic spectrum access. In Section 4.1, we give an overview about the topic with the most relevant related works. In Section 4.2, we present the opportunistic spectrum access model, while we describe the numerical results of the model in Section 4.3. To alleviate the negative effect of the proposed scheme, we refined the model, in which we increased the protection of the ongoing calls against the forced termination. The refined model is discussed in Section 4.3.3. We summarize the chapter in Section 4.4.

Chapter 2

Research Methodology

This chapter provides the short overview of methodology applied to solve the problems covered in the dissertation. It is worth to emphasize that performance evaluation can be carried out by different methods [31]:

- using a simulation software;
- by mathematical analysis with numerical procedures;
- by building the system and then measure its performance.

In this dissertation, we applied the first and the second method for model evaluation. Although simulation allows us to construct more sophisticated models, mathematical analysis generally needs lower computational effort [32], and produces exact solution.

2.1 Queuing Networks

Queuing networks are used for modeling systems, which can be considered as a set of interacting services.

An Open Queuing Network (OQN) can be considered as a service, which has customers from the "outside world". Customers will be transferred also to the "outside world" after getting the service [33]. This approach is based on the fact that the arrival process can be approximately described assuming infinite customer population if the number of customers is very large. In these models, the number of customers being under service does not influence the arrival intensity of the system. Fig. 2.1 shows an example of OQNs.

A CQN has a constant number of customers circulating throughout the system. In this model, the number of idle customers determine the arrival intensity (Fig. 2.2). This approach ensures a more precise model, and appropriate even for smaller customer

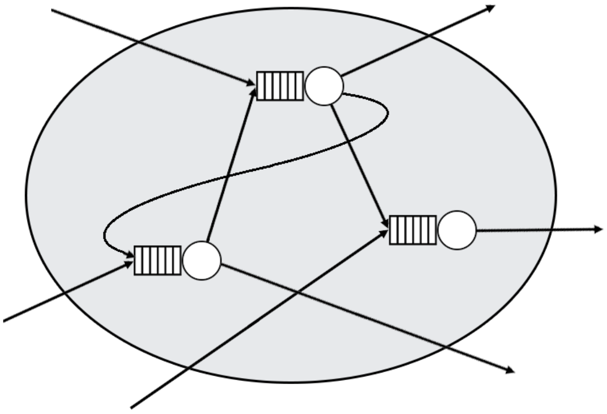


Figure 2.1: An example of open queuing networks

population. On the other hand, the analysis of these models is more challenging than in case of OQNs.

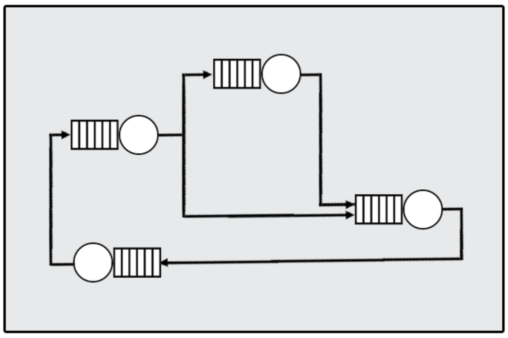


Figure 2.2: An example of closed queuing networks

In the area of queuing networks, many results were obtained in the last century [34]. These results are either exact analytical ones [35, 36] or approximations [37].

For OQNs, the Jackson networks [38, 39] have an efficient product-form solution. For CQNs, the Gordon-Newell theorem [40] and the mean value analysis [35] can be used to obtain the exact solution. However, the above mentioned processes assume that the queuing network is well-formed, i.e., the following criteria hold [33]:

- every station is reachable from any other with a non-zero probability in case of a CQN;
- in case of an OQN, add a virtual station 0 that represents the external behavior that generates external arrivals and absorbs all departing customers, so obtaining a CQN, for which the first criterion can be applied.

In our application spreading models, the user population is finite. Therefore, using CQNs to obtain the behavior of our system is an obvious idea. However, the acyclicity of our model does not allow the use of the mentioned traditional processes, since our CQN is not well-formed (if a user purchases an application, he will never lose it again). Therefore, we have to use other methods. By exploiting the special characteristics of the CQN model, we present an approximate analytical solution in this dissertation, while we validate the results via simulations.

2.2 Continuous-Time Markov Chains

In this dissertation, we often refer to CTMCs [41] as a fundamental modeling tool for stochastic processes. In Chapter 4, we use a CTMC for modeling opportunistic spectrum access in mobile cellular networks. In Section 2.3, we will show the relationship between CTMCs and SPNs, which are used in Chapter 3 for modeling application spreading in mobile ad hoc networks. In this section, we collected the main definitions of CTMCs which are important in the remaining chapters of the dissertation.

2.2.1 The Formal Definition of CTMC

A CTMC is a stochastic process $X(t) \mid t \geq 0,\ t \in \mathbb{R}$ such that for all $t_0, ..., t_{n-1}, t_n,\ t \in \mathbb{R}$, $0 \leq t_0 < ... < t_{n-1} < t_n < t$, for all $n \in \mathbb{N}$

$$\begin{gathered} P(X(t) = x \mid X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, ..., X(t_0) = x_0) = \\ P(X(t) = x \mid X(t_n) = x_n) \end{gathered} \tag{2.1}$$

2.2.2 Homogeneity

If we consider a discrete state space, and we denote

$$p_{ij}(t, s) = P(X(t+s) = j \mid X(t) = i) \tag{2.2}$$

for $s > 0$, a CTMC is called homogenous if

$$p_{ij}(t, s) = p_{ij}(s) \tag{2.3}$$

for all $t \geq 0$.

2.2.3 Transition Rates

In a homogeneous CTMC, $p_{ij}(s)$ is the probability of jumping from state i to state j during an interval time of duration s. The instantaneous transition rate from state i

to state j can be defined as

$$q_{ij} = \lim_{\Delta t \to 0} \frac{p_{ij}(\Delta t)}{\Delta t}, \tag{2.4}$$

and the exit rate from state i as $-q_{ii}$, where

$$q_{ii} = -\sum_{j \neq i} q_{ij} = \lim_{\Delta t \to 0} \frac{p_{ii}(\Delta t) - 1}{\Delta t}. \tag{2.5}$$

$Q = [q_{ij}]$ is called infinitesimal generator matrix or transition rate matrix.

2.2.4 Transient and Steady State Solution

In the following, we briefly describe how to obtain the transient and the steady state probabilities of a CTMC.

- Let denote $\pi_i(t) = P(X(t) = i)$ the distribution at time instant t, in matrix form $P(t) = [p_{ij}(t)]$. Then, $\pi(t) = \pi(u)P(t-u)$ for $u < t$.
- Substituting $u = t - \Delta t$ and subtracting $\pi(t - \Delta t)$ we get

  $$\pi(t) - \pi(t - \Delta t) = \pi(t - \Delta t)[P(\Delta t) - I], \tag{2.6}$$

  where I is the identity matrix.
- Dividing by Δt and taking the limit we get

  $$\frac{d}{dt}\pi(t) = \pi(t) \lim_{\Delta t \to 0} \frac{P(\Delta t) - I}{\Delta t}. \tag{2.7}$$

- Then, using definition of $Q = [q_{ij}]$, we obtain the Kolmogorov differential equation [42]

  $$\frac{d}{dt}\pi(t) = \pi(t)Q. \tag{2.8}$$

- The transient solution for the state probabilities $\pi(t)$ can be expressed as

  $$\pi(t) = \pi_0 e^{Qt}. \tag{2.9}$$

- Since $\pi(t)e = 1$ with $e = (1, 1, ..., 1)$, if $\lim_{t \to \infty} \pi(t)$ exists, then taking the limit of Kolmogorov differential equation we get the equation for the steady state probabilities:

  $$\begin{aligned} \pi Q &= 0 \\ \pi e &= 1 \end{aligned} \tag{2.10}$$

2.2.5 A Two-State Example for CTMCs

In the second half of the last century, CTMCs were used for modeling many known systems in telecommunications. In this section, we show an example system known as M/M/1/1 system (Fig. 2.3). In this simple two-state system, we assume

- infinite user population with arrival rate λ,
- single server with service rate μ, and
- no buffer capacity.

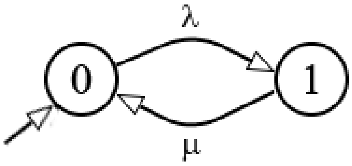


Figure 2.3: The M/M/1/1 system.

The corresponding infinitesimal generator matrix of this system is

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}, \tag{2.11}$$

while the Kolmogorov differential equation yields

$$\begin{aligned} \frac{d}{dt}\pi_0(t) &= -\lambda\pi_0(t) + \mu\pi_1(t) \\ \frac{d}{dt}\pi_1(t) &= \lambda\pi_0(t) - \mu\pi_1(t) \\ \pi_0(t) + \pi_1(t) &= 1 \end{aligned} \tag{2.12}$$

Given that $\pi_0(0) = 1$, we get the transient solution as

$$\begin{aligned} \pi_0(t) &= \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \\ \pi_1(t) &= \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \end{aligned} \tag{2.13}$$

Using equation (2.10), the steady state probabilities can be obtained as

$$\begin{aligned} \pi_0 &= \frac{\mu}{\lambda+\mu} \\ \pi_1 &= \frac{\lambda}{\lambda+\mu} \end{aligned} \tag{2.14}$$

Note that equation (2.14) can also be obtained by taking the limits as $t \to \infty$ of equation (2.13).

2.3 The Relationship Between Continuous-Time Markov Chains and Stochastic Petri Nets

The work of several authors helped the evolution of Petri net models that led to the proposal of SPNs [43, 44]. A SPN is a Petri net extended with time handling, that makes the Petri net suitable for performance analysis purposes. The firing delays in a SPN are exponentially distributed, which is a memoryless distribution.

As we will see later in this section, it is relatively easy to show that SPNs are isomorphic to the CTMCs [45]. Therefore, the standard approach for analyzing SPNs is to construct the CTMC corresponding to the underlying stochastic behavior of the SPN [41] and perform the steady state or transient analysis analytically [46] or by simulation. The association can be obtained by applying the following rules [45]:

- The CTMC state space $S = s_i$ corresponds to the reachability set $RS(M_0)$ of the PN associated with the SPN ($M_i \leftrightarrow s_i$).
- The transition rate from state s_i (corresponding to marking M_i) to state s_j (M_j) is obtained as the sum of the firing rates of the transitions that are enabled in M_i and whose firings generate marking M_j.

Based on these rules, the infinitesimal generator (or the state transition rate matrix) of the isomorphic CTMC can be automatically constructed from the description of the SPN [45]. An example of the association is illustrated in the following. In Fig. 2.4, the SPN model of shared memory system is shown, in which two processors use a common shared memory.

Omitting the detailed description, the reachability set of this system can be determined starting from the initial marking M_0. The reachability set is presented in Table 2.1.

The reachability graph of a SPN denotes how the different markings concerning to different system states can be accessed. Knowing the reachability graph and the transition intensities of the SPN, it is straightforward to obtain the state transition rate diagram of the corresponding CTMC (Fig. 2.5). Finally, Fig. 2.6 shows the infinitesimal generator matrix of the system.

However, the above mentioned approach becomes unfeasible due to the size of the state space if we consider a network composed of a large number of components. In this dissertation, we describe the mean field approach, which is a fluid approximation

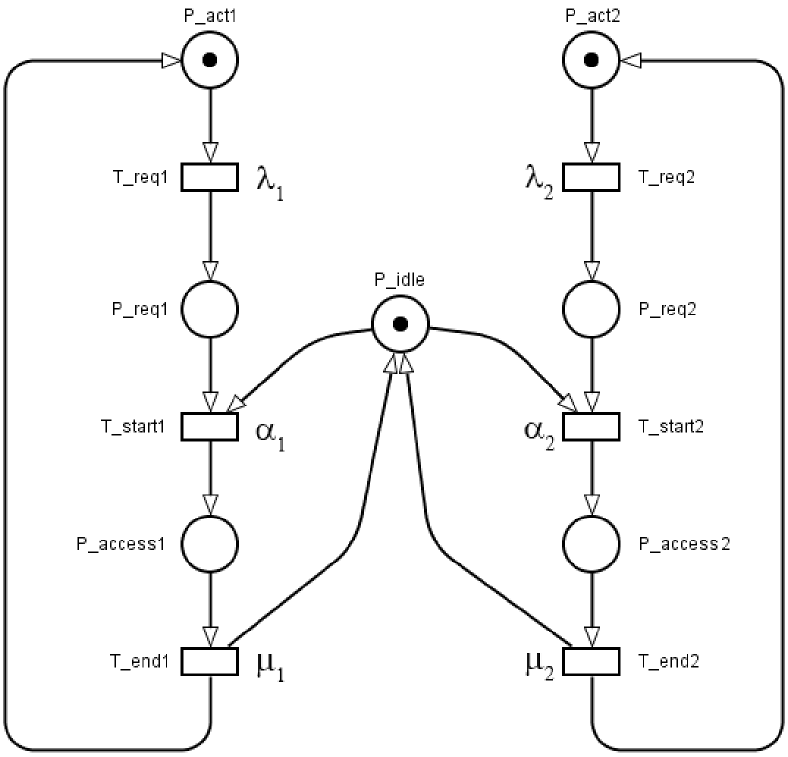


Figure 2.4: The SPN model of shared memory system.

Table 2.1: The reachability set of SPN model of Fig. 2.4

.

M_0	=	P_act1	+					P_idle	+	P_act2		
M_1	=			P_req1	+			P_idle	+	P_act2		
M_2	=					P_access1	+			P_act2		
M_3	=					P_access1	+				P_req2	
M_4	=			P_req1	+			P_idle	+		P_req2	
M_5	=	P_act1	+					P_idle	+		P_req2	
M_6	=	P_act1	+									P_access2
M_7	=			P_req1	+							P_access2

method for model evaluation. Applying this method, the analysis will terminate within a few seconds, even when the state space explodes due to the high number of tokens. This method is based on [29], while we present it in a form that is directly related to the applied definition of SPN [6]. Besides, we provide a formal relation between the CTMC and its fluid approximation in Chapter 3.

Unfortunately, if a SPN model contains inhibitor arcs, the fluid approximation method is not feasible. Instead, simulation can be used to evaluate the model. The simulation of stochastic Petri nets is supported by many known tools, like the ones presented in [47, 48, 49].

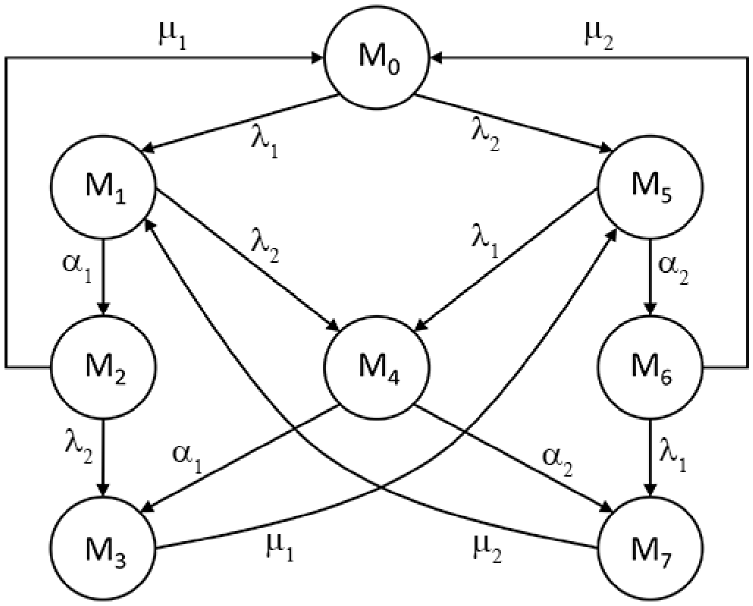


Figure 2.5: The state transition rate diagram of the CMTC associated with the SPN in Fig. 2.4.

$$Q = \begin{bmatrix} -\lambda_1-\lambda_2 & \lambda_1 & 0 & 0 & 0 & \lambda_2 & 0 & 0 \\ 0 & -\alpha_1-\lambda_2 & \alpha_1 & 0 & \lambda_2 & 0 & 0 & 0 \\ \mu_1 & 0 & -\lambda_2-\mu_1 & \lambda_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu_1 & 0 & \mu_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_1 & -\alpha_1-\alpha_2 & 0 & 0 & \alpha_2 \\ 0 & 0 & 0 & 0 & \lambda_1 & -\alpha_2-\lambda_1 & \alpha_2 & 0 \\ \mu_2 & 0 & 0 & 0 & 0 & 0 & -\lambda_1-\mu_2 & \lambda_1 \\ 0 & \mu_2 & 0 & 0 & 0 & 0 & 0 & -\mu_2 \end{bmatrix}$$

Figure 2.6: The infinitesimal generator matrix of the CTMC in Fig. 2.5.

2.4 Deterministic and Stochastic Petri Nets

In a manufacturing process, the work phases have deterministic delay. A process, in which the delays of the transitions are either exponentially, or deterministically distributed, can be appropriately described by a DSPN [50]. DSPNs are similar to SPNs, except that deterministically delayed transitions are also allowed in the Petri net model.

Although DSPNs have greater modeling strength than SPNs, there are some restrictions in the model evaluation phase. However, Lindemann and Shedler [51], and Lindemann and Thümmler [52] showed that in some special cases, the DSPNs can be handled analytically even with concurrently enabled deterministic transitions, the numerical analysis of DSPNs are generally limited to the case when there is at most one enabled deterministic transition in each marking (for further details, see Section 3.2). In these cases, simulation can be used to get the steady state or the transient solution of the DSPN. The simulation of DSPNs is supported by many known tools, e.g. by TimeNet [47].

2.5 Simulation

In this dissertation, several simulation results are presented. This section provides a short overview about the simulation of stochastic processes.

2.5.1 Discrete-Event Simulation

Discrete-event simulation is used for modeling a system which can change at only a *countable* number of points in time. Events occur in these points, where an *event* is defined as an instantaneous occurrence that may change the state of the system [53].

In each discrete-event simulation, a variable called *simulation clock* stores the current value of the simulated time. In this dissertation, we follow the next-event time-advance approach. The simulation clock is initialized to zero, and we determine the future events and their times of occurrence. Then, the simulation clock is advanced to the first future event, when the system is updated, the future events and their times of occurrence are determined again, and so on. The simulation process terminates when an investigated variable reaches a predefined value (e.g. the simulation clock reaches its predefined maximal value).

In most discrete-event simulation software, the following components can be found [53].

- *System state*: The set of state variables, which describe the system.
- *Simulation clock*: A variable for storing the current value of simulated time.
- *Event list*: A list of events containing the next time of occurrence of each event.
- *Statistical counters*: Variables for storing statistical information about the system.
- *Initialization routine*: A function for initialize the system before the simulation starts.
- *Timing routine*: A function that determines the next event in the event list and advances the simulation clock to the time of occurrence of that event.

- *Event routine*: A function for updating the system when a particular type of event occurs.
- *Library routines*: A set of functions for generating random observations from probability distributions.
- *Report generator*: A function for producing a report when the simulation terminates.
- *Main program*: The main program invokes the timing routine, checks for termination and invokes the report generator when the simulation is over.

The discrete-event simulation process is illustrated in Fig. 2.7 [53].

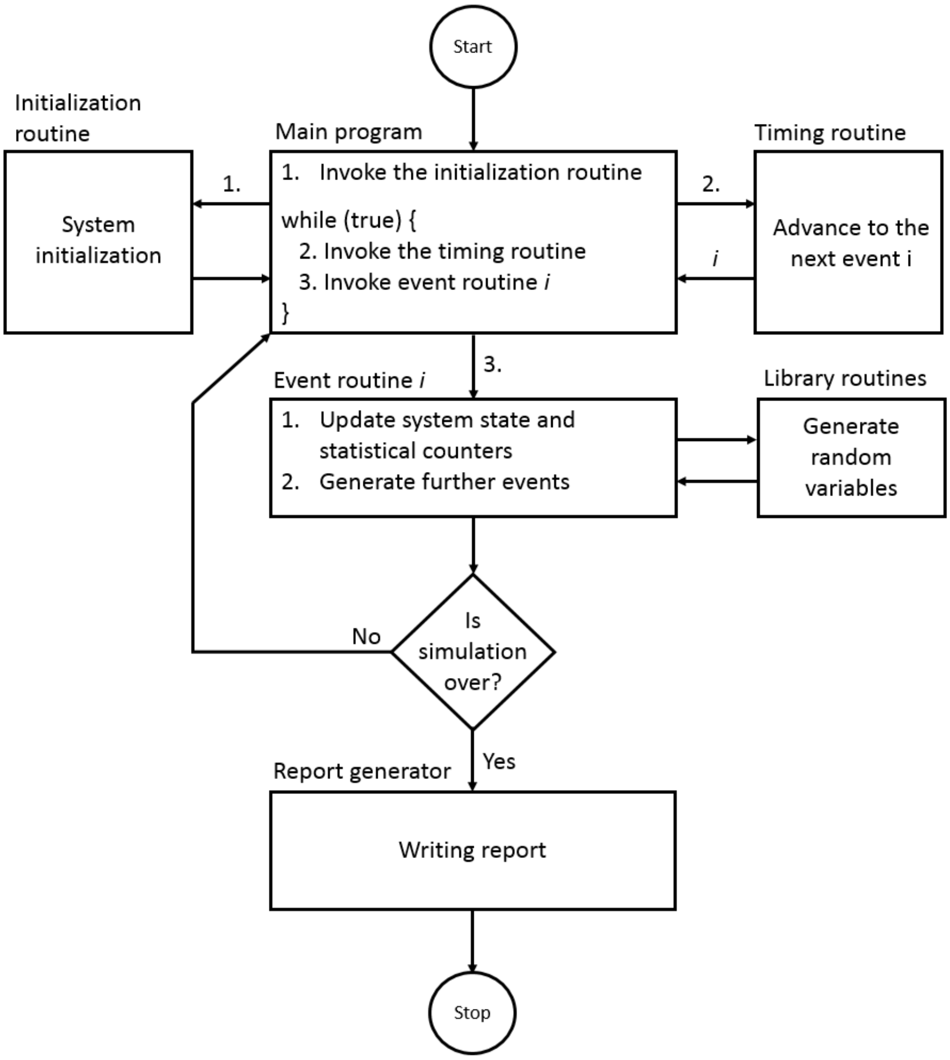


Figure 2.7: The flow chart of discrete-event simulation process.

2.5.2 Generating Random Variables

Usually, simulation processes have random aspects, which involves the demand for generating random variables from probability distributions. For this task, it is essential to have a reliable pseudo-random number generator function producing random numbers with uniform distribution in $U[0,1)$. Then, random numbers with the desired probability distribution can be generated as follows:

- Generate a pseudo-random number in $U[0,1)$.
- Let ξ a variate with probability distribution $F_\xi(x) = P(\xi < x)$.
- If $U \in [0,1)$ has uniform distribution, then the $\xi = F_\xi^{-1}(U)$ variate is a random variable with probability distribution $F_\xi(x)$.

Example: Let be $F_\xi(x) = P(\xi \leq x) = 1 - e^{-\lambda x}$. Then

$$\xi = F_\xi^{-1}(U) = \frac{ln(\frac{1}{1-U})}{\lambda} = \frac{ln(\frac{1}{U_1})}{\lambda}, \tag{2.15}$$

where $U_1 \in [0,1)$ is a variate with uniform distribution.

2.5.3 The Precision of Simulation

For determining the precision of simulation, interval estimation is a commonly used technique.

$$P(|\bar{\gamma}(n) - \mathbf{E}[\gamma]| \leq \Delta) = 1 - \alpha, \tag{2.16}$$

where

- $1 - \alpha$ is the confidence level,
- Δ is the half-width of the confidence interval, and
- $\epsilon = \Delta/\bar{\gamma}$ is the relative precision.

If n is “sufficiently large”, then

$$\Delta(n) = z_{1-\alpha/2}\hat{\sigma}_\gamma(n), \tag{2.17}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

It is also known, that if the deviation of variate γ is σ_γ, then the deviation of n experiment $\sigma_\gamma(n)$:

$$\sigma_\gamma(n) = \frac{\sigma_\gamma}{\sqrt{n}} \tag{2.18}$$

Therefore, in case of n^2 experiments,

- the deviation of the result decreases to its n^{th},
- the width of the confidence interval also decreases to its n^{th}.

Chapter 3

Two Performance Analysis Applications Based on Deterministic and Stochastic Petri Net Models

> *"But innovation comes from people meeting up in the hallways or calling each other at 10:30 at night with a new idea, or because they realized something that shoots holes in how we've been thinking about a problem."*
> (Steve Jobs)

In this chapter, we offer an alternative approach for the application providers to spread an application exploiting the direct communication between the users. Based on our proposed communication model and user types, we present different models and techniques for obtaining results regarding to the application spreading process. Besides, we show that our applied modeling techniques can be used in the wood industry, too, where we describe a wooden window manufacturing process with a DSPN model.

The outline of the chapter is as follows. Section 3.1 presents a CQN and two SPN models for application spreading in mobile ad hoc environments. In the evaluation of these models, different techniques are used depending on the complexity of the models. In Section 3.2, we model the wooden window manufacturing process with a DSPN to demonstrate that the stochastic models can be applied in the wood industry. Finally, we summarize this chapter in Section 3.3.

3.1 Application Spreading in Mobile Ad hoc Environments

In this section, we present our mobile application spreading models. We emphasize that our communication model described in Section 3.1.3 works in an environment, which *possibly* will exist, since the proliferation of ad hoc networks is not guaranteed today. Therefore, the validation of our results is also a difficult question. However, our approach will surely not exist if no one shows that it *can* work (the situation is similar to the chicken-and-egg problem). On the other hand, some innovative proposals like the multi-touch screen of the smart phones had also seemed meaningless for many experts until someone built them and met with success. In the current phase, our work described in this section can be considered as a pioneer one, which can be a base of further works in this area.

In Section 3.1.1, we give a short overview about the related works. In Section 3.1.3, we describe our communication model and present the user types, which we assume in our models. Section 3.1.4 presents the CQN model, which can be simply used for giving an analytical estimation on the number of application purchases. For obtaining more sophisticated results, we present two SPN models in Section 3.1.5. In the first SPN model, the underlying CTMC must be density dependent in order to apply a mean field based methodology, which provides an analytical approximation of the transient solution of the SPN. In the second model, we can handle a finer model, since the underlying CTMC of the SPN does not have to be density dependent. However, we have to use simulation to obtain the solution of the SPN.

3.1.1 Related Works

With the proliferation of modern communication paradigms, the investigation of application spreading using new ways becomes more and more important. However, it has not got too much attention so far. Besides our previous contributions [1, 2, 3, 4, 5, 6], only a few papers touch even the commercial use of ad hoc networks and direct communication.

On the other hand, epidemic spreading is a popular research topic today and this area is similar to the context of application spreading. In [54], the authors present a model, by which they investigate the propagation of a virus in a real network. In [55], the authors present scale-free networks for modeling the spreading of computer viruses and also give an epidemic threshold, which is an infection rate. Information spreading is also investigated by using epidemic spreading models, such as the Susceptible-Infected-Resistant (SIR) model [56], or other models based on the network topology [57, 58]. In [59], malicious software spreading over mobile ad hoc networks is investigated. The authors propose the use of the Susceptible-Infected-Susceptible (SIS) model based on the theory of Closed Queuing Networks. In [60], the authors propose the commercial use of ad hoc networks and present a radio dispatch system using mobile ad hoc communication. In the proposed system, the connectivity of the nodes is the key element of information dissemination.

In our models, we do not consider the network topology as a key element of application spreading, since no real-time information dissemination is needed between the users. For the same reason, we do not deal with mobility models such as random walk model, which are well presented in many contributions [61, 62, 63].

Although the above mentioned proposals show some similarities with our work, none of them deals with application spreading and, except [60], they do not touch the commercial benefits of direct communication. Moreover, the authors in [60] consider information dissemination as a tool, and not as a goal.

3.1.2 The Application Spreading Process using Mobile Ad Hoc Connections

Nowadays, nearly all mobile devices have got at least one wireless interface (e.g. Wi-Fi), which is appropriate for ad hoc communications. Ad hoc networks offer a good opportunity to communicate even when no central infrastructure is available. However, this advantageous property implies some difficulties, too. First, the communication range is limited due to the limited strength of the emitted signal in the Industrial, Scientific and Medical (ISM) band. Furthermore, the lack of business interest (no internet service provider is needed) does not facilitate the software development for ad hoc environments. Finally, the lack of central entities makes the network configuration more difficult.

On the other hand, the short communication range involves some advantages, too. For example, the community experience of a multi-player game is significantly increased in a Local Area Network (LAN) environment, where the players can hear and/or see each other. However, LAN games are popular even now, they are available mostly in a pre-configured environment with fixed location, typically using desktop computers. Our approach is a more general extension of LAN applications, since the location is not fixed. However, the above mentioned configuration problem still exists. To ease the network management in ad hoc networks, a service provisioning framework is proposed in [16]. Using the framework, the nodes discover each other, and have the opportunity to make services available for other nodes. Besides, the framework can alleviate the configuration difficulties.

Our approach assumes that meeting points will be formed, where people who want to use multi-user applications can find each other, and can form ad hoc networks. The network can be maintained even with low number of users, too, and is not sensitive to the relatively frequent topology changes caused by the appearing and disappearing of nodes.

3.1.3 Communication Model and User Types

In this section, we present the communication model which we use in our investigations. Moreover, we introduce three user types based on different user behaviors.

Communication Model

We refer to the individuals who are interested in the use of the application as users. The population that we investigate is composed of users only, and we do not take uninterested users into consideration, because they do not influence the spreading process. Therefore, we assume a closed user population.

We investigate the spreading of a given multi-user application having two versions, a trial and a full version. The users can be categorized into different classes depending on whether they do possess any version of the given application or do not. We named the classes after the terminology of epidemics, since our model shows similarity to the epidemic spreading models. A user is called *i*) infected, if he has got the full version of the application; *ii*) susceptible, if he possesses only the trial version of the application; and *iii*) resistant, if he has got none of them, or he has already lost the interest of using the application.

Users with their mobile devices form self-organized networks from time to time, in which direct communication takes place. The trial version of the application is free and available in these networks, so users can download it and even try it out. However, it has some restrictions (see later), so the users have to purchase the application for unrestricted usage via a traditional way of purchasing. Later, also these users can spread the trial version of the purchased application further.

Since application providers want susceptible users to be motivated in purchasing the full version of the application, some limitations must be made in using the trial version. Therefore, we apply a limit (leech[1] limit) that restricts how many nodes possessing the trial version (leech[1]) can connect to a node possessing the full version (seed[1]). In this sense, the seeds can be considered as servers, which can serve a limited number of clients. A seed is always an infected user, while a leech may be either infected or susceptible. Fig. 3.1 depicts the case when a susceptible user purchases the application.

The devices form an ad hoc network, in which the dark devices depict susceptible users, while the light ones depict infected users. In this example, the leech limit is two, so two susceptible users can peer to the only infected user, while the other two have to wait (the connection symbol represents application level peering). After one of them purchased the application, they can also use it, as shown in the right side of Fig. 3.1.

---

[1] After the terminology of BitTorrent [64]

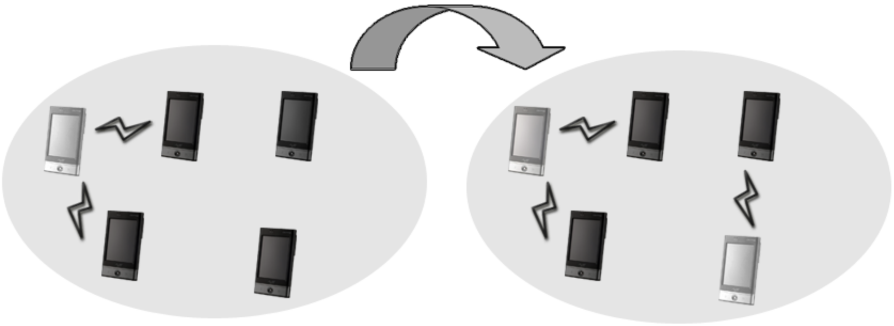


Figure 3.1: Change of application usage when a susceptible user purchases the application.

User Types

Beyond the basic communications, we distinguish three different user types based on the users' behavior.

- $Type_A$ users are interested in using the given application. Therefore, they are its potential buyers even without trying it out.
- $Type_B$ users also purchase the application very likely, but they will do it with a given intensity, only if they cannot find a seed from time to time which they can connect to.
- $Type_C$ users instead will never purchase the application but still they influence the spreading process because they decrease the probability that other users find an available seed to connect to.

In our CQN model, all the three types of users are present, and we can give a lower and an upper bound for the expected number of application purchases analytically. In our basic Petri net model, we use the mean field approach to obtain analytical results. However, our basic Petri model cannot handle the $Type_B$ users, since their behavior's description results in the violation of the density dependent property in the underlying CTMC of the SPN. Therefore, we omit $Type_B$ users from the basic model. In our extended Petri net model, we can handle the presence of $Type_B$ users, too, but only using transient simulation. Of course, additional user types can be introduced, too. However, the more user types we capture the more complex model we get, which can make the handling of the model difficult.

3.1.4 Modeling with Closed Queuing Networks

In self-organized networks, where spontaneous communication takes place, the network topology can change rapidly due to the high degree of mobility. These topology changes can be modeled by stochastic processes [65]. We can appropriately describe a stochastic process in a closed population, which is interesting from our point of view, using CQNs [34]. Moreover, ordering transition intensities to the state changes we can capture the time behavior of the application spreading process, too.

Spreading Model

For modeling the application spreading with CQNs, we propose the model depicted in Fig. 3.2. In this CQN model, the different states represent the whole user population. Each user is in one state depending on his current user class. Resistant users which possess neither the full version nor the trial version of the application are in state *Init*. We call also resistant the users, who have already lost the interest in using the application. However, they possess either its trial (state *RS*) or its full version (state *RI*). The susceptible users are in state *PS* and *AS*, depending on that they are currently using the application (Active Susceptibles, *AS*) or not (Passive Susceptibles, *PS*). Similarly, the active and passive infected users are in state *PI* and *AI*, respectively.

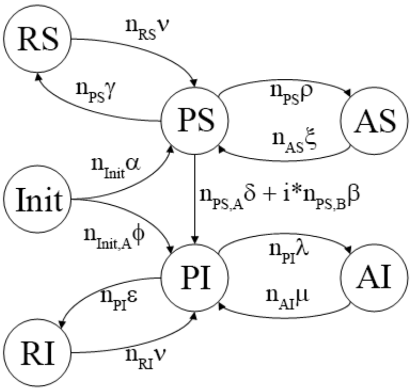


Figure 3.2: The proposed CQN model.

The Greek letters in Fig. 3.2 denote transition intensities regarding to a single user, n_x represents the number of users in state x, while $n_{x,A}$, $n_{x,B}$ and $n_{x,C}$ represent the number of $Type_A$, $Type_B$ and $Type_C$ users in state x, respectively ($n_x = n_{x,A} + n_{x,B} + n_{x,C}$). The transition intensity is a real number assigned to the state transition, which denotes how many times a state transition takes place during a given time interval. These transitions are described in Table 3.1.

Usage of the Spreading Model

We can unambiguously describe the state of the system with the user distribution (n_{Init}, n_{PS}, n_{AS}, n_{PI}, n_{AI}, n_{RS}, n_{RI}). The transition intensities (α, β, γ, δ, ϵ, ϕ, λ,

Table 3.1: The state transitions of the proposed CQN model

Transitions	Description
$PS \rightarrow AS$	A susceptible user starts to run the application and tries to connect to a seed in the network. If he cannot find one, he has to wait.
$AS \rightarrow PS$	A susceptible user stops running the application.
$PI \rightarrow AI$	An infected user starts to run the application, then either he tries to connect to an available seed, or will be a seed himself to which leeches can connect.
$AI \rightarrow PI$	An infected user stops running the application.
$Init \rightarrow PS$	A resistant user downloads the trial version of the application.
$PI \rightarrow RI$	An infected user becomes resistant losing the interest in using the application.
$PS \rightarrow RS$	A susceptible user becomes resistant losing the interest in using the application.
$Init \rightarrow PI$	A resistant user purchased the application without trying it out.
$PS \rightarrow PI$	A susceptible user becomes infected by purchasing the application. A susceptible user purchased the application. $n_{PS,A}$ and $n_{PS,B}$ denote the number of $Type_A$ and $Type_B$ users in state PS, respectively. This transition is enabled only for $Type_B$ users when there is no free seed available in the network which they can connect to. Therefore, the indicator variable i is zero if there is no free seed available, and one otherwise. $Type_C$ users never purchase the application, so this state transition is not allowed to take place for $Type_C$ users. However, $Type_C$ users can also connect to seeds, so decreasing the probability that other users find a free seed which they can connect to.
$RS \rightarrow PS$	It is possible that a resistant user, who lost the interest in using the trial version, wants to use the application again after a while. If so, his state becomes susceptible again.
$RI \rightarrow PI$	Similarly, if a resistant user possessing the full version of the application wants to use it again, his state changes to infected.

μ, ν, ρ and ξ) regarding to a single user are the system parameters, which are hard to be determined theoretically. In this dissertation, we set the system parameters based on common sense. The parameter setting can be fine-tuned experimentally, what is beyond the scope of this dissertation.

In each system state, we can generate the holding time h (the time that the system is expected to spend in a given system state) as an exponentially distributed random variable[2] in the following way:

$$h = \frac{-lnRND}{\sum\limits_{\forall state\ x} out_x}, \tag{3.1}$$

where $0 \leq RND < 1$ is a pseudo-random number, and out_x denotes the sum of the intensities for each transition with source state x. We must compute h after each state change, because the user distribution changes when a transition takes place. We generate the next system state based on the ratio of the current transition values. After we generated the transition that takes place, we move one user from its source to its destination state, then compute the holding time of the new system state, and so on.

At the beginning, each user is in state $Init$, which is the initial state of the system. The users will leave this state and change their states from time to time. After a while, a user will lose the interest in the application usage (reaches state RS or RI), but it does not mean that he cannot be interested again later on. Thus, the state transitions $RS \rightarrow PS$ and $RI \rightarrow PI$ are also enabled, however, it is allowed only with low intensity values for them. Therefore, we will reach a system state (rest state) at time instant τ, in which each user is either in state RS or RI. The rest state can be defined as follows:

Definition 1 *The system is in rest state, if for* $\forall state\ x \notin \{RS, RI\}, n_x = 0$.

The rest state is not the steady state, however, our investigation will stop here. Taking into account the asymmetry of this system (if someone purchases the application, he will never lose it), each user will be in one of state RI, PI or AI by reaching the steady state (even product-form solution exists). However, it is meaningless to consider this state in the model investigation, since after reaching the rest state, the holding times become extremely large, so the system changes very slowly. Hence, our investigations are always transient and consider only a time period which is interesting from the merchant's point of view.

We can determine how many pieces of the application were sold, upon reaching the rest state, by summing the number of $PS \rightarrow PI$ and $Init \rightarrow PI$ transitions. Running simulations and evaluating the results, we can observe the time characteristics of the spreading process, too.

Results Derived from CQN

In this section, we describe some analytical results [4] and validate them via simulations.

---

[2] Using exponentially distributed holding times is a usual modeling simplification in this context [59, 66]

Analytical solutions of CQNs can be obtained, e.g., with the well-known Mean Value Analysis (MVA) [35]. However, we investigate the transient behavior of the system, in which case this method is not feasible.

After a while, each user will leave the initial state, since we do not take the uninterested individuals into consideration. Based on the intensity value of transition $Init \rightarrow PI$ and $Init \rightarrow PS$, we can determine how many users will expectedly purchase the application without trying it out (direct purchases, $DP(\tau)$) in the following way:

$$DP(\tau) = n_{Init,A} \cdot \frac{\phi}{\alpha + \phi}, \tag{3.2}$$

where $n_{Init,A}$ denotes the initial number of $Type_A$ users in state $Init$, i.e., the total number of $Type_A$ users, and τ denotes that the number of direct purchases concerns to the time instant τ, when the system reaches its rest state. As we mentioned before, we consider only $Type_A$ users when computing $DP(\tau)$. All nodes (users) that did not purchase the application without trying it out will change their state to susceptible. Susceptible states are state PS and AS, and there are two possibilities for the users to leave these states: a user can become either (1) resistant (transition $PS \rightarrow RS$); or (2) infected (transition $PS \rightarrow PI$). Moreover, it is also allowed to return from state RS, but the intensity of transition $RS \rightarrow PS$ is very low, since losing the interest and being interested again after a while is not a typical user behavior. Therefore, we estimate the number of indirect purchases (purchase after trying out) based on the ratio of transition $PS \rightarrow PI$ and $PS \rightarrow RS$. In case of the rest of $n_A - DP(\tau)$ $Type_A$ users, $IDP_A(\tau)$ concerning to the time instant τ can be computed as

$$IDP_A(\tau) = n_{Init,A} \cdot \frac{\alpha}{\alpha + \phi} \cdot \frac{\delta}{\delta + \gamma}, \tag{3.3}$$

while we can give an upper bound on the purchases of $Type_B$ users concerning to the time instant τ as

$$IDP_B(\tau) \leq n_B \cdot \frac{\beta}{\beta + \gamma}, \tag{3.4}$$

where equality holds only if the indicator variable i is equal to one.

Summing equations (3.2) and (3.3), we can give a lower bound on the total number of purchases $TP(\tau)$, while we can give an upper bound on $TP(\tau)$ by summing equations (3.2), (3.3) and (3.4):

$$DP(\tau) + IDP_A(\tau) \leq TP(\tau) \leq DP(\tau) + IDP_A(\tau) + IDP_B(\tau), \tag{3.5}$$

where τ denotes that this result also concerns to the rest state of the system.

To validate the analytical results we ran simulations. Table 3.2 compares the analytical results to the average of 10000 individual simulation runs with the following parameters: $n_{Init,A} = n_{Init,B} = n_{Init,C} = 500$, $\alpha = 10^{-3}$, $\beta = 10^{-3}$, $\gamma = 2 \cdot 10^{-3}$, $\delta = 10^{-3}$, $\epsilon = 10^{-3}$, $\phi = 10^{-5}$, $\lambda = \rho = 4 \cdot 10^{-2}$, $\mu = \xi = 9 \cdot 10^{-1}$, $\nu = 10^{-9}$, while the leech limit was one.

Table 3.2: Comparison of the analytical and simulation results using the CQN model

Description	Analytically	By simulation
$DP(\tau)$	4.95	4.95
$IDP_A(\tau)$	165.02	164.99
$IDP_B(\tau)$	$IDP_B(\tau) \leq 166.67$	36.74
$TP(\tau)$	$169.97 \leq TP(\tau) \leq 336.64$	206.68

Further properties, such as the time behavior of the spreading process can be investigated also via simulations.

3.1.5 Modeling with Stochastic Petri Nets

In this section, we present the description and the analysis of two SPN models after introducing the fundamentals of Stochastic Petri Nets [45]. With the basic and the extended SPN model, we will be able to handle also $Type_B$ and $Type_C$ users' presence.

Since $Type_A$ users are present in all of our models (including the CQN model), we compare the models from the viewpoint of this simple user behavior, see Section 3.3.

SPN Formalism

In this section, we provide a brief introduction to SPNs, while a detailed introduction with applications can be found in [45].

SPNs are bipartite directed graphs with two types of nodes: places and transitions. An example of a SPN is the one depicted in Fig. 3.3. The places, graphically represented as circles, correspond to the state variables of the system; while the transitions, graphically represented as boxes, correspond to the events that can induce a state change. Examples of places for the SPN in Fig. 3.3 are ACT_A and ACT_B; while examples of transitions are $STOP\_A$ and $STOP\_B$. The arcs connecting places to transitions and vice versa express the relation between states and event occurrence.

Places can contain tokens drawn as black dots within places. The state of a SPN, called marking, is defined by the number of tokens in each place. In this dissertation, we use the notation M to indicate a marking in general. We will denote by $M(p)$ the number of tokens in place p in marking M. Now we recall the basic definitions that are necessary for the rest of the chapter.

Definition 2 *An SPN system is a 6-tuple*

$$(P, T, I, O, \lambda, M_0),$$

where:

- $P = \{p_i\}$ *is the set of* places *of cardinality* k*;*
- $T = \{t_i\}$ *is the set of* transitions *of cardinality* m*;*
- $I, O : T \times P \to \mathbb{N}$ *are the* input *and* output *functions that define the arcs of the net and their multiplicities;*
- $\lambda : T \to \mathbb{R}$ *is the function that assigns to each transition its firing intensity;*
- M_0 *is the* initial marking *of the net.*

A transition is "enabled" if each of its input places contains the "necessary" amount of tokens where "necessary" is defined by the input function I. Formally, transition t is enabled in marking M if for all places p of the net we have $M(p) \geq I(t, p)$. For example, transition $START\_A$ in Fig. 3.3 is enabled if the current marking contains at least one token in place $PASS\_A$ and at least one token in place $FREE\_S$. An enabled transition can fire and the firing removes tokens from the input places of the transition and puts tokens into the output places of the transition. The new marking M' after the firing of transition t is formally given $M'(p) = M(p) + O(t, p) - I(t, p)$, $\forall p \in P$.

The firing of a transition occurs after a random delay. The random delay associated with a transition has exponential distribution whose parameter depends on the firing intensity of the transition and on the actual marking. In this section, we assume that the more tokens enable a transition the faster the transition fires. This concept, called infinite server policy, is captured formally by the definition of the enabling degree.

Definition 3 *The enabling degree of transition* t *in marking* M*, denoted by* $ed(t, M)$*, is* d *iff* $\forall p \in P, M(p) \geq dI(t, p)$ *and* $\exists p \in P : M(p) < (d + 1)I(t, p)$*.*

For instance, if there are three tokens in place $PASS\_A$ and four tokens in place $FREE\_S$ then the enabling degree of transition $START\_A$ is three. The random delay associated with transition t in marking M is exponentially distributed with parameter $\lambda(t)ed(t, M)$. When a marking is entered, a random delay is chosen for all enabled transitions by sampling the associated delay distribution. The transition with the lowest delay fires and the system changes marking.

In Section 3.2, we will use another server policy called exclusive server model [45], in which the delay associated with an enabled transition t is independent from the enabling degree.

We define the effect of the firing of transition t with an integer vector $L(t)$, place indexed, defined as: $L : T \to \mathbb{N}^k$ and the ith entry of $L(t)$ is $L(t)_i = O(t, p_i) - I(t, p_i)$ for $1 \leq i \leq k$ and $\forall t \in T$. For sake of avoiding cumbersome notation we assume that $\nexists t, t' \in T : t \neq t', L(t) = L(t')$ and $\nexists t : L(t) = 0$.

Basic SPN Model

Here we describe the basic SPN model, and present a mean field based methodology for analyzing SPNs. Then using this methodology, we show the transient analysis of the basic SPN model.

Model Description The basic SPN model is depicted in Fig. 3.3. The rectangles illustrate the transitions of the SPN, while the circles represent the places. We assume the presence of 500 $Type_A$ and 500 $Type_C$ users, and initially each user is in the passive susceptible states $PASS\_A$ ($Type_A$ users) and $PASS\_C$ ($Type_C$ users). Seeds can also be passive ($PASS\_S$), however, there are no seeds in the network initially. Similarly to the CQN model, we call a user active if he is currently using the application. We keep count of the number of active $Type_A$ users (ACT_A), $Type_C$ users (ACT_C), available seeds ($FREE\_S$) and the total number of active users, including seeds (ACT_USERS). Since the leech limit is one, the available seeds show how many susceptible users can start the application in a given time. We also keep count of the number of application purchases ($PURCHASES$). After a while, users lose the interest in using the application in this model, as well. If so, they change their state to resistant ($LOST\_INT\_A$, $LOST\_INT\_C$, $LOST\_INT\_S$).

The transitions of the model and their values regarding to one user with which we analyzed the net are described in Table 3.3.

The underlying CTMC of the basic SPN model is density dependent, thus we can apply the mean field based methodology in its analysis. If we have to build a more complex model (e.g. we allow the use of inhibitor arcs resulting the violation of the density dependent property in the underlying CTMC), its analytical handling with the mean field based methodology is not possible anymore, so we must run simulations to investigate the model's behavior (see Section 3.1.5).

Mean Field Based Methodology The standard approach for analyzing SPNs is to construct the CTMC corresponding to the underlying stochastic behavior of the SPN (for details see, e.g., [41]) and perform the steady state or transient analysis analytically

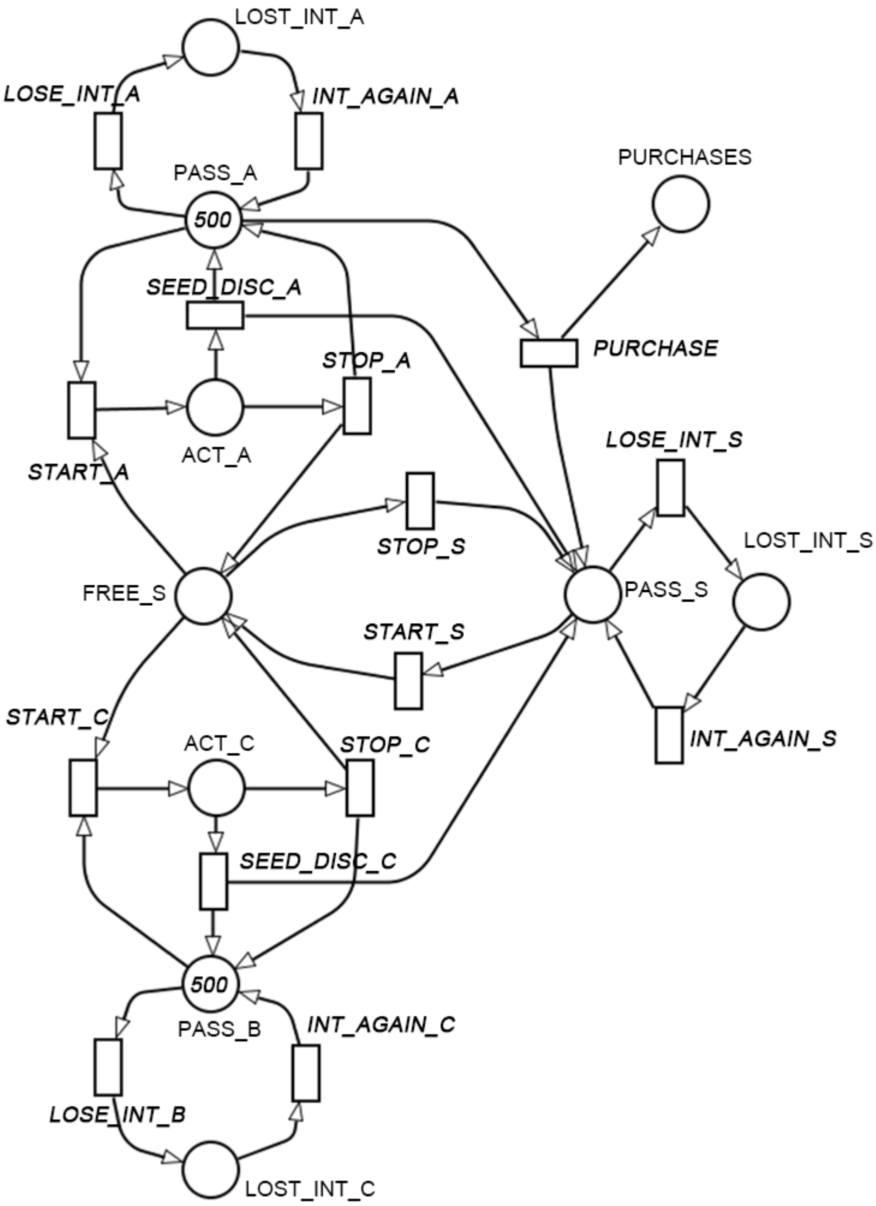


Figure 3.3: The basic SPN model.

[46] or by simulation. However, this approach becomes unfeasible due to the size of the state space if we consider a network composed of a large number of mobile components.

In the following, we describe the mean field approach, which is a fluid approximation method for model evaluation. Applying this method, the analysis will terminate within a few seconds, even when the state space explodes due to the high number of tokens. The following definition and theorem are based on [29], while Beccuti et al. presented it in [6] in a form that is directly related to the applied definition of SPN. In that paper, the authors provided a formal relation between the CTMC and its fluid approximation, too.

Table 3.3: The state transitions of the basic SPN model

Transitions	Description
$START\_A/START\_C/$ $START\_S$	A $Type_A$ user/$Type_C$ user/seed starts to run the application. They can fire only if there is at least one available seed in the network ($4 \cdot 10^{-2}/hour$ in each case).
$STOP\_A/STOP\_C/$ $STOP\_S$	A $Type_A$ user/$Type_C$ user/free seed stops running the application ($9 \cdot 10^{-1}/hour$ in each case).
$SEED\_DISC\_A/$ $SEED\_DISC\_C$	A seed to which a $Type_A/Type_C$ user had been connected stopped running the application. The state of the connected leech node becomes passive ($9 \cdot 10^{-1}/hour$ in each case).
$LOSE\_INT\_A/$ $LOSE\_INT\_C/$ $LOSE\_INT\_S$	A $Type_A$ user/$Type_C$ user/seed loses the interest in using the application ($10^{-3}/hour$ in case of a seed, $2 \cdot 10^{-3}/hour$ otherwise).
$PURCHASE$	A $Type_A$ user purchases the application ($10^{-3}/hour$).
$INT\_AGAIN\_A/$ $INT\_AGAIN\_C/$ INT_AGAIN_S	It is possible that a resistant user, who lost the interest in using the application, wants to use the application again after a while. If so, his state becomes susceptible or infected again, depending on his previous state ($10^{-5}/hour$ in each case).

Definition 4 *A parametric family of Markov chains, $X_v(t)$ with $v \in \mathbb{N}$, with state spaces $E_v \subset \mathbb{Z}^k$, is called density dependent iff there exists a continuous function $f(x,l), x \in \mathbb{R}^k, l \in \{L(t_1), ..., L(t_m)\}$, such that the non-diagonal entries of the infinitesimal generator corresponding to $X_v(t)$ can be written in the form*

$$q_{k,k+l} = vf\left(\frac{k}{v}, l\right), \ l \in \{L(t_1), ..., L(t_m)\} \tag{3.6}$$

and the initial state of the chain is $vx_0, x_0 \in \mathbb{Z}^k$, with probability 1.

Note that in definition 4 and equation (3.6), the integer vector $L(t_i)$ denotes the effect of the firing of transition $t_i \in T$, $1 \leq i \leq m$, where m is the cardinality of T (see definition 2).

Let $X(t)$ denote the solution of the Ordinary Differential Equations (ODEs)

$$\frac{dX(t)}{dt} = \sum_{l \in \{L(t_1), ..., L(t_m)\}} f(X(t), l) \tag{3.7}$$

with initial condition $X(0) = x_0$.

Theorem 1 *Under mild conditions on the function f (for details, see [29]), the following relation holds between the function $X(t)$ and a trajectory of the CTMC $X_v(t)$:*

$$\forall \delta > 0 : \lim_{v \to \infty} P \left\{ \sup_{s \leq t} \left| \frac{1}{v} X_v(s) - X(s) \right| > \delta \right\} = 0. \tag{3.8}$$

The interpretation of the above theorem is the following. Consider a CTMC modeling the interaction of k quantities whose state space hence is $\mathbb{Z}^k$. Imagine to observe a sequence of CTMCs with increasing initial state (i.e., the sequence of initial states is $x_0, 2x_0, 3x_0, ...$). If the increase of the initial states gives rise to a sequence of infinitesimal generators corresponding to the form in equation (3.6) then, as v is increased, the behavior of the CTMC converges to the solution of the ODEs given in equation (3.7). The convergence is in the sense that the probability of finding any small difference between a trajectory of the CTMC and the solution of the ODEs in a finite time horizon $(0, t)$ is zero.

It has been shown in [67] that modeling interacting populations in process algebra gives rise to a family of density dependent CTMCs and, consequently, for large population sizes the corresponding ODEs provide a good approximation of the behavior of the system. Since it is straightforward to show that the CTMCs constructed from the Petri nets as well are density dependent [67], the behavior of Petri nets with high number of tokens can be approximated by ODEs. Moreover, as we will illustrate it numerically, the average behavior of the CTMC is approximated reasonably well by the ODEs even with a lower number of tokens.

In the rest of this section we describe the ODEs which provide the fluid approximation of a SPN=$\{P, T, I, O, \lambda, M_0\}$. The state of the system is a vector of real numbers and transition $t_i, 1 \leq i \leq m$, is moving "fluid" tokens in state $x, x \in \mathbb{R}^k$, with speed

$$s(t_i, x(t)) = \lambda(t_i) \min_{j: I(t_i, p_j) \neq 0} \{x_j(t)/I(t_i, p_j)\}, \tag{3.9}$$

i.e., the speed depends on the rate of the transition, $\lambda(t_i)$, and on the quantity of tokens present in its input places. Namely, as infinite server policy is applied, the minimum amount of fluid tokens found in the input places (considering the multiplicity of the input place as well) limits the speed of the transition. The amount of tokens in the ith place is changing then according to the ODEs

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{m} s(t_j, x(t))(O(t_j, p_i) - I(t_j, p_i)), 1 \leq i \leq k, \tag{3.10}$$

i.e., if place p_i is an input (output) place of transition t_j then transition t_j is removing (adding) tokens from (to) place p_i according to the multiplicities given by the function I (O).

Given the initial state, the ODEs in equation (3.10) can be solved by numerical integration for which various reliable tools exist. There is one equation per place, so the computational complexity grows linearly with the number of places, while in case of analyzing the underlying CTMC the growth would be exponential.

Transient Analysis In the following, we illustrate the usage of the mean field approach through the transient analysis of the above mentioned basic SPN model.

The solution of the ODEs is a good approximation of the average behavior of the model. Fig. 3.4 depicts the cumulative expected value of the number of purchases after a given time, while Fig. 3.5 shows the cumulative expected value of the number of users who lost the interest in using the application as time elapses.

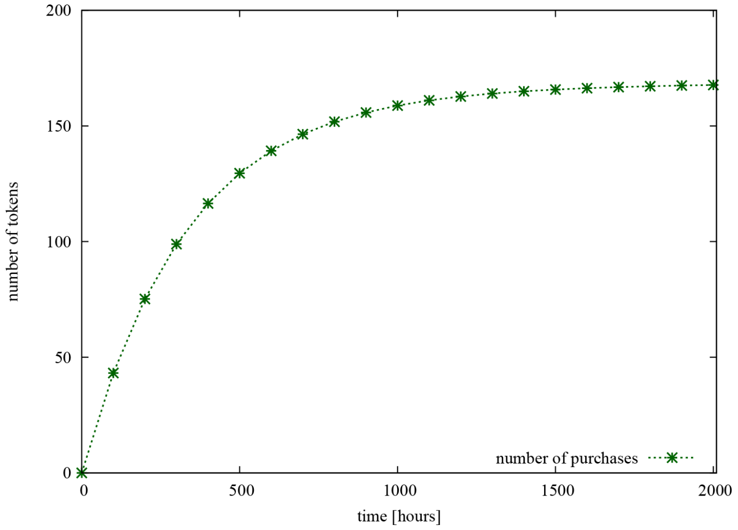


Figure 3.4: The expected value of the number of application purchases as the function of elapsed time.

In Fig. 3.4 and Fig. 3.5, we can see that applications were purchased mostly in the first 1000 hours, while the interest in using the application is approximately limited to the first 4000 hours. The latter result shows that the system really reached its rest state at $\tau \approx 4000$ hours. Using the mean field method, the run time of the evaluation process was approximately one second.

Extended SPN Model

In this section, we describe our extended SPN model and its transient simulation.

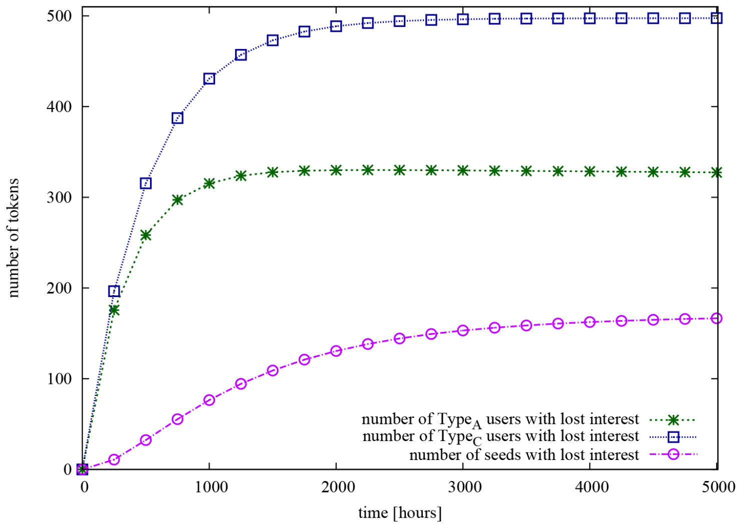


Figure 3.5: The expected value of the number of users who lost the interest in using the application.

Model Description In the extended SPN model, we can also handle more sophisticated user behaviors, like the $Type_B$ users' behavior. This requires the use of inhibitor arcs, too, violating the density dependent property in the underlying CTMC of the SPN[3]. Apart from this, the extended model is very similar to the previous one. To handle the new user type we added four places ($PASS\_B$, ACT_B, $LOST\_INT\_B$ and $PURCHASES\_B$) and six transitions ($LOSE\_INT\_B$, INT_AGAIN_B, $PURCHASE\_B$, $SEED\_DISC\_B$, $START\_B$ and $STOP\_B$) to the basic model. As we described in Section 3.1.3, $Type_B$ users purchase the application only if they cannot find an available seed which they can connect to. Therefore, the transition $PURCHASE\_B$ is enabled only if there is no token in place $FREE\_S$. This relationship is denoted in the model by an inhibitor arc with a circle on its head. Beyond that, $Type_B$ users' behavior is similar to the others'. Fig. 3.6 shows our extended SPN model. The newly added part is marked by grey background. As earlier, we set 500 users from each user type in the initial marking.

Transient Simulation Since the underlying CTMC of the extended SPN model does not satisfy the density dependent property, we cannot use the fluid approximation method to investigate the model's behavior, rather we have to run simulations. There exist many tools for modeling with SPNs, e.g., the ones presented in [47, 48, 49]. We used the transient simulation method of TimeNET [47].

---

[3] Note that the inhibitor arcs can be eliminated in most cases. However, if the Petri net not containing inhibitor arcs is equivalent with the one that contains inhibitor arcs, the underlying CTMCs will be the same. So, the density dependent property cannot be satisfied with the elimination of inhibitor arcs.

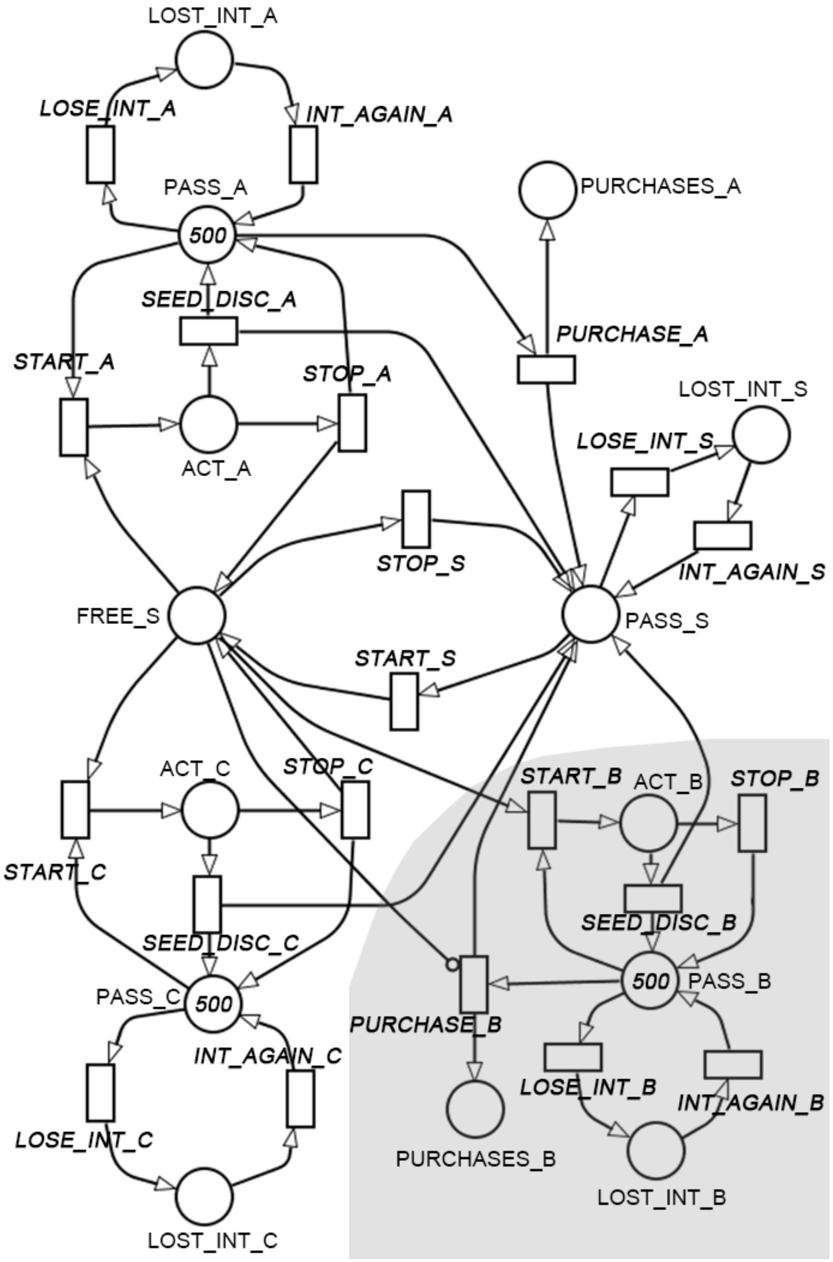


Figure 3.6: The extended SPN model.

In our investigations, we used the same transition intensities for the existing transitions as in Section 3.1.5, while the new transitions were set to the corresponding transition intensity values of the previous model. For example, transition $START\_B$ has the same intensity value as $START\_A$ and $START\_C$. The simulation results reflect 95% confidence level and 1% maximal error rating.

Fig. 3.7 shows the cumulative expected value of the number of application purchases with regard to the different user types, while Fig. 3.8 depicts the cumulative expected value of the number of users who lost the interest in using the application as time elapses.

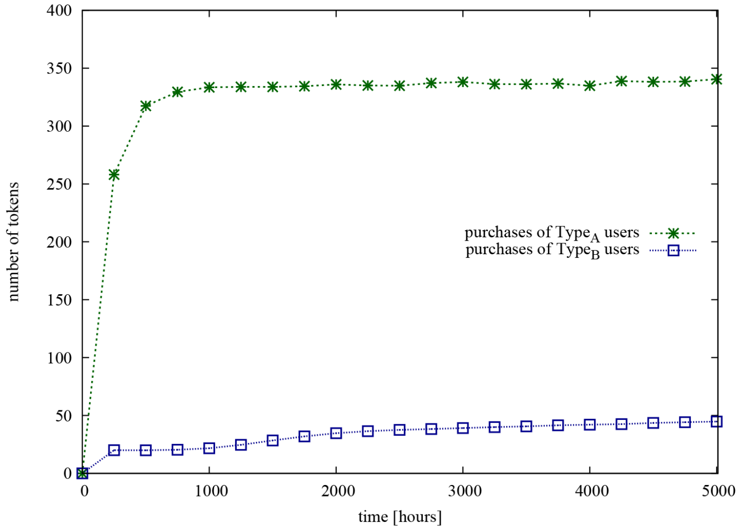


Figure 3.7: The expected value of the number of application purchases as the function of elapsed time.

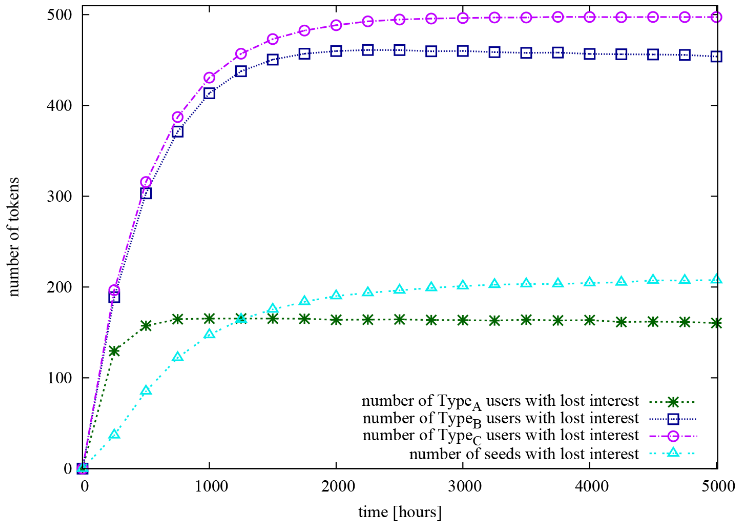


Figure 3.8: The expected value of the number of users who lost the interest in using the application.

These results show similarity to the previous investigation, namely, approximately the third of the $Type_A$ users purchased the application. On the other hand, the $Type_B$ users' purchase depends on the number of the free seeds in the network (and certainly on the initial parameter setting). According to the simulation results, 7.4% of the $Type_B$ users purchased the application. We can observe that $Type_B$ users purchased the application even when others lost the interest in using it, and therefore, the number of available seeds was low. Therefore, the application purchases were not limited to the first 2000 hours, as in the previous scenario.

3.2 Usability of Stochastic Models in the Wood Industry: a Case Study

In this section, we present a case study, in which we demonstrate the usability of the stochastic models in the wood industry. In the case study, we model the production of wooden windows of a Hungarian company (Holz-Team Ltd. [68]). Using the model, we can simply determine the bottleneck of the manufacturing process and show how to eliminate it [7].

In a manufacturing process, the work can be divided into different disjoint phases having deterministic holding times. However, there are some exponentially distributed variables even in these types of models, like the inter-arrival times of the orders and the material defects.

A process, in which the delays of the transitions are either exponentially, or deterministically distributed, can be appropriately described by a DSPN [50]. DSPNs are similar to SPNs, except that deterministically delayed transitions are also allowed in the Petri Net model. Of course, the number of waiting frames and frame members do not influence the intensity of a work phase in the model. Therefore, we do not applied the previously mentioned infinite server model, rather we use the exclusive server model.

The use of DSPNs was first proposed by Marsan and Chiola in their fundamental work [50], in which they could determine the numerical solution of the DSPN when there was at most one enabled deterministic transition in each marking. Since then, Ciardo et al. showed that DSPNs can be analyzed with concurrently enabled deterministic transitions, if the concurrently enabled deterministic transitions fire at the same time [69]. Besides, Lindemann and Shedler [51], and Lindemann and Thümmler [52] presented a novel approach, which is based on the analysis of a General State Space Markov Chain (GSSMC), whose state equations constitute a system of multidimensional Volterra differential equations [51] and Fredholm integral equations [52], respectively. However, this approach demands that the deterministic transitions which have already been enabled in the previous marking and are still enabled in the current marking, keep their remaining firing times instead of restarting the firing process.

If the above mentioned assumptions cannot be made, simulation can be used to get the steady state or the transient solution of the DSPN. The simulation of DSPNs is supported by many known tools, e.g. by TimeNet [47]. As we can see in the following, our window manufacturing model contains several concurrently enabled deterministic transitions. Moreover, the deterministic transitions restart their firing time in our case (after a material flaw is discovered, e.g. during planing, the planing of next frame member will not require less time). Therefore, the above mentioned analytical approaches cannot be used, so we will show some simulation results.

3.2.1 DSPN formalism

In this section, we give a short introduction to DSPNs. Since the formalism of SPNs is already described in Section 3.1.5, we present only the DSPN-specific parts of the formalism.

Definition 5 *A DSPN system is a tuple*

$$(P, T, I, O, H, M_0, \tau, w, e),$$

where:

- P *is the finite set of* places. *A marking* $M \in \mathbb{N}^{|P|}$ *defines the number of tokens in each place* $p \in P$*.*

- T *is the set of* transitions. T *can be partitioned into the following disjoint sets:* T^Z *is for the set of immediate transitions,* T^E *is for the set of exponential transitions, and* T^D *is for the set of deterministic transitions. Note that* $P \cap T = \emptyset$*.*

- $I, O, H : \mathbb{N}^{|P|} \to \mathbb{N}$ *are the multiplicities of the* input arc *from* p *to* t*, the* output arc *from* t *to* p*, and the* inhibitor arc *from* p *to* t*, respectively.*

- $M_0 \in \mathbb{N}^{|P|}$ *is the* initial marking *of the net.*

- $\tau : \mathbb{N}^{|P|} \to \mathbb{R}^+$ *is the mean delay for* $\forall t \in T^E \cup T^D$ *(note that* τ *may be marking-dependent).*

- $w : \mathbb{N}^{|P|} \to \mathbb{R}^+$ *is the firing weight for* $\forall t \in T^Z$ *(note that* w *may be marking-dependent).*

- $e_{t_1,t_2} : \mathbb{N}^{|P|} \to \{R, C\}$*, is the the execution policy to be used for transition* t_2 *when transition* t_1 *fires for* $\forall t_1 \in T^Z \cup T^E, t_2 \in T^D$ *(note that* e_{t_1,t_2} *may be marking-dependent).* R *and* C *denote "restart" and "continue", respectively.*

3.2.2 The Manufacturing Process of Wooden Windows

In the following, we describe the manufacturing process of wooden windows in detail.

From the company's point of view, the wooden window manufacturing process lasts from the arriving of an order to the delivery of the accomplished windows. After an order is accepted, the company starts manufacturing the sufficient number of windows. Each single window can be considered unique, since a building typically contains windows in several various sizes and types, so the quantity production cannot be carried out in this area.

In the first part of the production of a single window, the frame members are produced separately. Typically, a wooden window consists of 8 frame members: 4-4 frame members compose the frame and the casement frame, respectively. Since the manufacturing process cannot be started in the absence of raw material, the companies try to appropriately manage the quantity of the raw material in their store. The raw material typically means 6 meter long lumber.

The first work phase is cutting the raw material to size using circular saw. After cutting to size, the frame members must be planed with a planing machine. The third work phase is molding. In this work phase, a Computer Numerical Control (CNC) machine is applied to bring the frame members into their final form. CNC machines are fully automated and programmable, and their use for molding ensures high level of precision and speed.

After molding, the frame members are glued into a frame. From this point, the basic unit of the manufacturing is the frame. The next work phase is dipping, which is a surface treatment for protecting the wooden material from damages caused by insects and fungi. Then, the frames are sanded with a sanding machine to obtain a smooth surface. After sanding, priming and final coating give the final color of the frame and ensure further protection for the wooden material. The last work phase is fixing the hardware parts. In this phase, the frame and the casement frame is supplied with the required metal components and the glass. Finally, the frame and the casement frame is joined, and the window is ready for transportation.

The peculiarity of the manufacturing process is that material defects can be revealed in any work phases before the final coating. If a defect is revealed before gluing, the single frame member can easily be re-produced, while after gluing, the glued frame must be decomposed, and all work phases must be repeated with the frame, when the re-produced single frame member is available. In both cases, the re-production has priority over the other works.

3.2.3 An Operation Model for the Manufacturing Process of Wooden Windows

In this section, we model the operation of Holz-Team Ltd., a Hungarian window production company.

From modeling point of view, the events of the production can be divided into two categories, based on the distribution of the transitions modeling the given event. Accordingly, we model the orders and the occurrence of material defects with exponentially delayed transitions, while we describe the different production phases (cutting, planing, molding, gluing, dipping, sanding, priming, final coating and fixing the hardware parts) with deterministic transitions (Fig. 3.9).

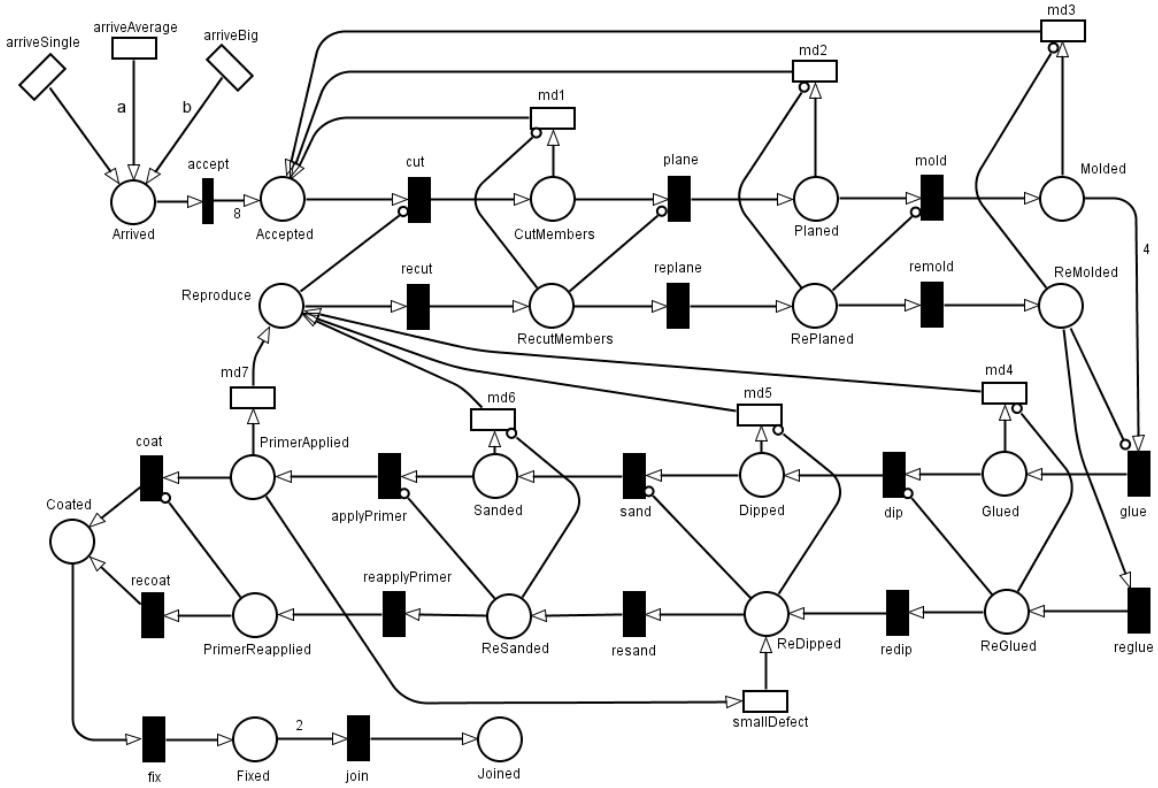


Figure 3.9: The proposed operation model for the Manufacturing Process of Wooden Windows.

In this example, there are three typical types of orders:

- for specific single window,
- for all windows of a family house, a windows on the average,
- for windows of a big building, b windows on the average.

According to the types of orders, the arrival process consists of transitions *arriveSingle*, *arriveAverage* and *arriveBig* in Fig. 3.9. In this case study, a and b are considered 18 and 65, respectively.

After the order was accepted[4], the 8 frame members of each window must be cut to size (each frame has 4 members and each casement frame has also 4 members in most cases[5]). The cut frame members must be planed, then molded (see transitions *cut*, *plane* and *mold* in Fig. 3.9). Until this point is reached, the frame members are individual entities each belonging to a specific window. Therefore, if a material defect is discovered, the given single member can be re-cut, re-planed and re-molded without causing bigger delay in the process (see transitions *md1*, *md2* and *md3* in Fig. 3.9).

As we mentioned above, the concurrently enabled deterministic transitions make the execution policy e_{t_1,t_2} crucial according to [51] and [52]. Since material defects may be revealed in any work phases in any time, we modeled these events with exponentially delayed transitions. If a material defect is revealed after a time interval $0 < t < t_D$, where t_D denotes the delay of the concurrent deterministic transitions, we cannot state that the next item in the queue needs only $t_D - t$ time in the given work phase. Therefore, the execution policy is “restart” in our case.

After gluing the 4 frame members, one token denotes a frame in the model. The next operations after gluing are the dipping of the frame, then the sanding, the priming and the final coating (see transitions *glue*, *dip*, *sand*, *applyPrimer* and *coating* in Fig. 3.9). Note that a material defect can be discovered in *any* production phase before the final coating. However, when the members compose a frame, the delay is much bigger, since the frame must be recomposed, and the single member must be re-produced (see transitions *md4*, *md5*, *md6* and *md7* in Fig. 3.9). Therefore, we assume that if a material defect is discovered after the members compose a frame, further defects will not be discovered. This assumption is based on the thorough investigation, which follows when a frame must be decomposed to prevent further delay.

In our model, there is a dedicated path for this incident: since 3 frame members wait for the re-production of the fourth one, the transitions on the dedicated path have higher priority than the concurrent ones (see transitions *recut*, *replane*, *remold*, *reglue*, *redip*, *resand*, *reapplyPrimer* and *recoat*). In the DSPN model, inhibitor arcs ensure the priority.

After priming, the smaller defects can be corrected by re-sanding, what is denoted by transition *smallDefect* in Fig. 3.9.

The last two steps of the production are fixing the hardware parts and join the casement frame and the frame (transitions *fix* and *join* in 3.9). The transitions with

---

[4] In Fig. 3.9, we took only the accepted orders into account

[5] Having the typical case is a modeling simplification.

the corresponding delays are collected in Table 3.4[6] (in case of the exponentially delayed transitions, the expected delay is shown).

Table 3.4: The transitions of the DSPN model

Transitions	Description	Delay
arriveSingle	A single wooden window is ordered.	16.64
arriveAverage	18 wooden windows are ordered.	12.07
arriveBig	65 wooden windows are ordered.	1392
$md1$	A flaw is discovered after cutting.	3.125
$md2$	A flaw is discovered after planing.	15
$md3$	A flaw is discovered after molding.	50
$md4$	A flaw is discovered after gluing.	10
$md5$	A flaw is discovered after dipping.	60
$md6$	A flaw is discovered after sanding.	33.33
$md7$	A flaw is discovered after priming.	22.22
smallDefect	A frame is sent back to sand.	33.33
cut/recut	Cutting/re-cutting of a frame member.	0.0556
plane/replane	Planing/re-planing of a frame member.	0.0104
mold/remold	Molding/re-molding of a frame member.	0.05
glue/reglue	Gluing/re-gluing of the frame members.	0.1667
dip/redip	Dipping/re-dipping of a frame.	0.0333
sand/resand	Sanding/re-sanding of a frame.	0.2
applyPrimer/reapplyPrimer	Priming/re-priming of a frame.	0.1111
coat/recoat	Coating/re-coating of a frame.	0.1111
fix	Fixing hardware parts.	0.3333
join	Join the casement frame and the frame.	0.01

3.2.4 Evaluation of the Production Process

In this section, we investigate the production process using transient simulation. So, we can determine the throughput of the system for one year, and the bottlenecks of the production process. Moreover, we can investigate the effects of the material defects.

---

[6] The values of Table 3.4 were set based on a personal discussion with the representative of the company.

However, several other aspects could be investigated, too. We used TimeNet [47] to obtain the simulation results, which reflect 99% confidence level with 1% maximal error rating.

Table 3.5 shows the expected token distribution after one year. We can observe that the company can produce more than 3000 windows a year, while the bottleneck of the system is the fixing of the hardware parts, since there are about 585 frames on the *Coated* place (as a matter of fact, the employees in this session must work overtime to accomplish the orders until their deadline).

Table 3.5: The token distribution of the DSPN model after one year

Places	Expected number of tokens
$Accepted$	189.81933
$CutMembers + RecutMembers$	0.13467
$Planed + RePlaned$	0.63745
$Molded + ReMolded$	3.25834
$Glued + ReGlued$	0.10666
$Dipped + ReDipped$	0.70611
$Sanded + ReSanded$	0.35967
$PrimerApplied + PrimerReapplied$	0.35056
$Coated$	584.55711
$Fixed$	0.52422
$Joined$	3021.378

The bottleneck (fixing the hardware parts) is a limiting factor in the production if the company wants to increase the number of accepted orders. Fig. 3.10 shows that as we increase the intensity of the *arriveAverage* transition, the production is throttled by the fixing of the hardware parts work phase. In Fig. 3.10, we also show the case when the company detects the material defects always *before* gluing (e.g. by applying quality control) to prevent the major overhead caused by these defects. Technically, we deleted the transitions *md4*, *md5*, *md6* and *md7* from the DSPN model. We can observe that eliminating the late detection of the material defects does not affect the total throughput. However, it can obviously cause a delay of certain jobs. On the other hand, the elimination of the mentioned transitions increased the number of tokens in the bottleneck place when the system was heavily loaded.

The obvious solution for the bottleneck problem is to expand the critical produc-

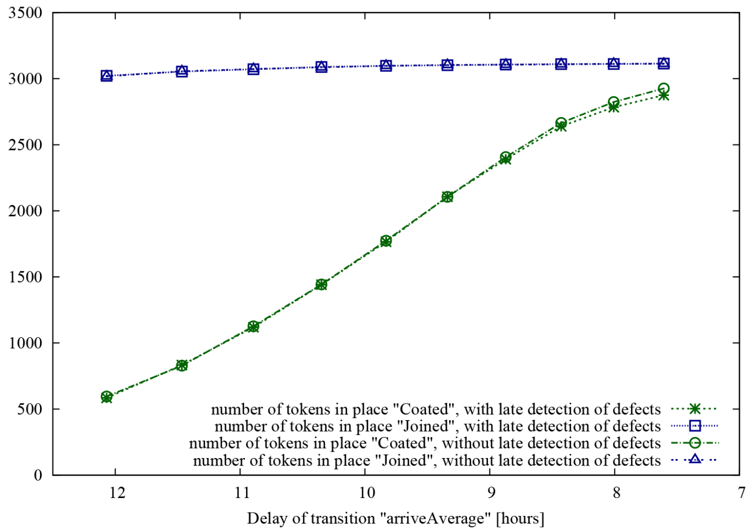


Figure 3.10: The number of tokens in places “Coated” and “Joined”.

tion contingent (i.e. by employing more workers). Therefore, we investigated the throughput of the system as a function of the delay of transition *fix*, while the delay of transition *arriveAverage* remains on the lowest value of the previous simulation (the exact value is 7.606709638). Fig. 3.11 shows that decreasing the delay of transition *fix* from 0.33 hours to 0.22 hours is enough to move the bottleneck in the system, and gives the chance to increase the throughput by about 50%. In other words, the maximal throughput can be increased by 50%, if the delay of fixing the hardware parts onto one frame could be reduced from 20 minutes to 13.2 minutes.

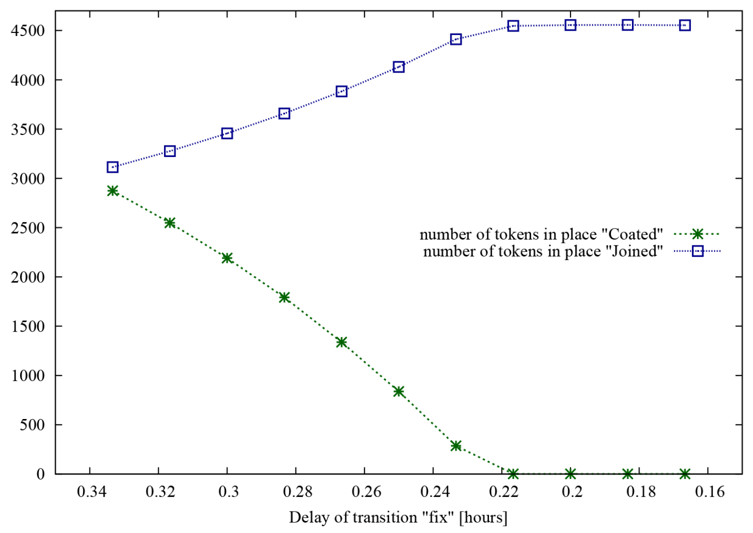


Figure 3.11: The number of tokens in places “Coated” and “Joined”.

3.3 Summary

In this chapter, we presented two techniques for modeling application spreading aided by direct communication between the users' mobile devices. Besides, we showed that the stochastic models can be used in the wood industry, too, describing the wooden window production with a DSPN model.

First, we presented a CQN model by which we can simply estimate some parameters of the application spreading process. This simple model allows a service provider to quickly calculate how much profit can be expectedly realized from application purchases.

Overcoming the limits of CQNs, we proposed two SPN models. In the basic one, we used the mean field based methodology to obtain an analytical approximation of the Petri net, which can derive results in the order of seconds. Then we presented an extended version of this basic Petri net that can accommodate an even more realistic user behavior for the price of using inhibitor arcs in the model. Unfortunately, the inhibitor arcs result in the violation of the density dependent property in the underlying CTMC of the SPN model. Therefore, the mean field approach cannot be applied in this case. So, investigated the extended Petri net model via simulations paying the fee of long runtime to produce results.

As we set the initial number of users from a given user type to the same (500) in every investigation, we can compare the results regarding to $Type_A$ users' purchases in the different models: this parameter was approximately the same in all models (169.97, 167.72 and 171.36).

In our models, with appropriate experience to set the model parameters a service provider can estimate his profit from application purchases in case of simple or even more complex scenarios. Moreover, he can learn the time behavior of the spreading process, by which additional gain can be realized, such as the refinement of the marketing strategy.

In the second part of the chapter, we presented a DSPN model of a company which produces wooden windows. Using the DSPN model, we could determine the bottleneck in the production process. By transient simulation, we showed how the main bottleneck can be eliminated.

Chapter 4

Opportunistic Spectrum Access in Mobile Cellular Networks

"It is too late to spare when you reach the dregs of the cask."
(Lucius Annaeus Seneca)

In this chapter, we propose a spectrum sharing model, in which the service providers opportunistically use each other's unutilized frequency bands. Our goal is to show that using our opportunistic spectrum access model, the service quality improves, while the service providers can realize an extra profit.

The outline of the chapter is as follows. In Section 4.1, we give a short overview about the related works in this area (Section 4.1.1). For spectrum sharing is a popular research topic nowadays, different directions are investigated. One of them is the physical realization of the opportunistic spectrum access, which we summarize in Section 4.1.2. In Section 4.2, we describe our opportunistic spectrum access model. In this section, we describe a Markovian model for the opportunistic spectrum access, and present the analytical solution of the model, too. In section 4.3, we describe some simulation results, where the holding times are log-normally distributed. In detail, we investigate the impact of our model, propose a call admission control strategy for balancing the forced blocking probability and the blocking probability of the incoming calls. Then, we present some simulation results to show the effect of our model to the average profit rate of the participant service providers. Finally, we conclude the chapter in Section 4.4.

4.1 Overview

In this section, we present the most relevant related works first, then we describe the spectrum pooling concept and its main physical obstacles.

4.1.1 Related Works

The 2002 report of the Federal Communications Commission's Spectrum Policy Task Force [18] recommended a spectrum policy reform that is based on the flexible management of spectrum and liberalized/dynamic spectrum access. The idea of dynamic spectrum access is a radical approach compared to the current centralized spectrum management that ensures the exclusive access right to certain frequency bands. Since the release of the report [18], a quest for dynamic spectrum access technique to improve spectrum efficiency has been intensively researched (see [18, 24, 25, 19, 20, 23, 21] and references therein). Buddhikot [19], Zhao and Sadler [20], and Peha [23] provided overviews and rationale concerning the efficient spectrum sharing and access.

Although it is widely recognized that spectrum management reform and dynamic spectrum access can provide a solution to an existing problem (the shortage of usable radio frequencies and the under-utilization of licensed spectrum) [18, 19, 20], the application of dynamic and opportunistic spectrum access is rarely found in practice. A lot of issues [18, 19, 20, 21] must be solved before the widespread application of opportunistic spectrum access. Amongst these issues, a question concerning the investment (on the licensed frequencies and technology) protection of the incumbent operators plays an important role regarding to the acceptance of opportunistic spectrum access.

Weiss and Jondral [21] proposed a very practical approach called "spectrum pooling" that allows the access of already licensed frequency bands without requiring any change to the licensed systems. The authors [21] described their solution for frequency-/time-division multiple access (FDMA/TDMA) based licensed systems. Furthermore, they argued [21] that spectrum pooling does not create competition to existing and upcoming 2G and 3G mobile radio standards because a capacity shortage (due to limited frequency bands and high demand) may happen in a certain area. In a such case, a specific operator can apply the spectrum pooling technique to enhance the grade of service for calls (i.e., to rent a free frequency band from another operator to serve incoming calls) [26, 27, 70, 8].

Based on the concept of spectrum pooling [21], we propose a cooperation scheme between two mobile cellular providers that hold the licenses to offer service in a specific area. We present a queuing model to analyze the performance of the proposed cooperation scheme between two operators applying opportunistic spectrum renting. It is worth emphasizing that some queuing models for spectrum renting were worked out (see [26, 27, 70]), which could not be directly applied to the present proposal. Tzeng and Huang [26], and Tzeng [27] assumed that user channels can be rented in one unit, which is not realistic. The reason is that the separate blocks of user channels are defined in each frequency band, and each block should be controlled by a single network operator. Do et al. [70] presented the first queuing model to take into account this technology aspect. Do et al. [71] analyzed the retrial phenomenon with spectrum renting. However, they [70, 71] did not consider directly the interaction of operators.

4.1.2 A Possible Realization for the "Spectrum Pooling" Concept

Nowadays, the service providers have exclusive access to the licensed spectrum. In each country, there are competitive providers accessing different frequency bands, which were assigned to them on spectrum auctions by the local government. On each frequency band, the providers can allocate 8 full rate or 16 half rate speech channels for the incoming calls of their clients. Presently, if all speech channels of a given provider are occupied, the incoming calls will be blocked even if another provider has got unused frequency bands. However, this approach must be rethought due to the increasing demand for the spectrum.

Spectrum pooling is a spectrum renting approach first proposed by Mitola [17], which enables the secondary usage of the licensed frequency bands. The basic idea of the spectrum pooling is that the different licensed frequency bands should be merged into a common pool, from which the secondary users may opportunistically rent them during the idle periods of the primary users.

For the realization of a spectrum pooling system, several obstacles must be beaten off. We collected these obstacles based on the comprehensive paper of Weiss and Jondral [21].

First, the cooperating parties should be able to identify the idle spectral ranges for the secondary usage of the spectrum. The main requirement is that detection guarantees have to be assured to license owners on the order of 99.9 percent. Otherwise, licensed users will not be motivated to share their spectrum. To achieve this, a detection algorithm was proposed in [72]. The algorithm has two basic assumptions: *i*) the higher-layer protocols of the rental system ensure that the rental users remain silent during the detection period. Therefore, spectral power is only emitted by licensed users in this time interval; *ii*) as a worst case consideration, no line of sight can be assumed between the transmitting licensed user and the detecting rental user. With these assumptions, the detection signal can be modeled as a zero mean Gaussian process. During the detection phase, the statistics of the received signal can be applied to detection algorithms based on the Neyman-Pearson criterion [73]. Moreover, the diversity approach must be applied [72] to achieve the desired detection probability: not only a central instance, but also all associated mobile terminals should do spectral measurements.

Secondly, the previously mentioned diversity approach involves an enormous signaling overhead which makes the system error-prone as interference will be occurred with the new licensed users. In [74], a boosting protocol is proposed, which moves this signaling from the MAC layer to the physical layer. The protocol consists of two phases. In the first phase, the signaling of those sub-bands takes place which were newly allocated by licensed users, while the second phase deals with the sub-bands that have become idle again.

The third obstacle is the well known mutual interference problem. Namely, the Fast Fourier Transform (FFT) and its inverse operation (Inverse Fast Fourier Transform (IFFT)) in Orthogonal Frequency-Division Multiplexing (OFDM) systems result in additional interference between the rental and the licensed system in both directions. At the worst, the mean interference power caused by the rental system can be even 5 percent of the power transmitted on one subcarrier. Fortunately, the interference level from the opposite direction remains on the order of $-20dB$ if the licensed and the rental system apply the same mean transmit powers. A possible solution of the mutual interference problem is the introduction of adaptive guard bands in the rental system. However, this method reduces the effective bandwidth of the rental system. By all means, a trade-off must be found between the interference and the bandwidth of the rental system.

Finally, the time and frequency synchronization provided by legacy WLANs cannot be applied in spectrum pooling systems [21]. In [75], the authors propose the adaptive filtering of the narrowband interferers (the licensed users). They investigated preambles based on a much higher correlation length than in the original Wireless Local Area Network (WLAN) standards. The filtering can be carried out with the combination of the allocation vector and the existing FFT/IFFT operation in the OFDM transceiver, and it should be conducted in both the transmitter's and the receiver's side. The main advantages of this method are i) the licensed users are not disturbed in their allocated sub-bands; ii) the interference caused by the licensed users can be suppressed before the preamble is fed into the correlator. The authors showed that with this method, the frame start detection ratio can exceed 95 percent, and the method is also appropriate for frequency synchronization.

4.2 An Opportunistic Spectrum Access Model

In GSM networks, each service provider divides its network into disjoint cells and tries to cover the whole area of the given country. In most cases, the mobile cells of the different providers cover each other, so more than one Service Provider (SP) is available on a given geographical place. Presently, if all speech channels of one service provider are occupied new incoming calls are blocked while there may be available channels at another operator.

However, if two arbitrary SPs (SP_1 and SP_2) would cooperate according to the scheme presented below, more clients could be served. To facilitate it, we propose the following cooperation scheme. The SPs continuously monitor each other's spectrum to obtain which frequency bands are unused (a frequency band is unused, if there is no allocated call on it). If all speech channels of SP_1 are occupied, SP_1 may opportunistically allocate the new incoming calls on SP_2's unused frequency band f. Due to the current technical constraints, the SPs cannot share a frequency band for simultaneous use [70]. Therefore, while SP_1 is using f, SP_2 cannot use it, even if only one speech channel is occupied on f.

Based on the idea of [21], we propose an opportunistic spectrum sharing scheme for two operators in a specific area as follows. Both operators mutually act as licensed users and secondary users. Their licensed spectrum bands form a pool two operators can opportunistically use. It is worth emphasizing that the tenant operator can only use a frequency band owned by the renter if the specific band is idle. Furthermore, the tenant operator should vacate all calls from the rented band if the renter needs the specific band (i.e., to place ongoing calls to its own frequency band if it is possible). Those calls that could not be reallocated will be forced to terminate. Therefore, the tenant operator should increase the utilization of its own frequency bands as much as possible to reduce the forced blocking probability.

4.2.1 Notations and Assumptions

Let n_k, $(k = 1, 2)$ be the number of licensed frequency bands of SP k in the area. The number of channels is n in each frequency band. Therefore, the number of channels owned by SP k, $(k = 1, 2)$, is $N_k = n_k n$. The total number of licensed frequencies in a specific area is $N = N_1 + N_2$.

To obtain the steady state probabilities and performance measures within the Markovian framework, we have to construct the mathematically tractable model. Therefore, we follow the classical approach frequently applied in the queuing theory for the performance evaluation of wireless cellular networks [26, 27, 76, 77, 78, 79]. Fresh calls and handover calls arrive according to Poisson processes. The inter-arrival times of new calls and handover calls are exponentially distributed with rate $\lambda_k^{(F)}$ and $\lambda_k^{(H)}$, $(k = 1, 2)$, respectively. We assume that call durations (of new admitted calls and handover admitted calls) in the cell are exponentially distributed with mean $1/\mu_k$, $(k = 1, 2)$, to obtain a mathematically tractable model.

A fresh call that requires the service of SP k, $(k = 1, 2)$, is admitted according the Fractional Guard Channel (FGC) policy [70] to ensure a protection for handover calls. Let $\beta_{k,l}$, $k = 1, 2$, denote the probability that a fresh call is accepted by SP k if the number of busy channels handled by SP k is l. Therefore, the arrival rate of calls to SP k is $\lambda_{k,l} = \beta_{k,l}\lambda_k^{(F)} + \lambda_k^{(H)}$, when the number of busy channels handled by SP k is l.

Let $I_1(t)$ and $I_2(t)$ be the number of busy channels handled by SP 1 and SP 2, respectively (see Fig. 4.1).

For the better readability, we collected the notations of the section in Table 4.1.

Table 4.1: Notations

Notations	Description
n_k, $k = 1, 2$	The number of licensed frequency bands of SP k.
n	The number of channels in each frequency band.
N_k, $k = 1, 2$	The number of channels own by SP k.
$\lambda_k^{(F)}$, $k = 1, 2$	The arrival intensity of fresh calls at SP k.
$\lambda_k^{(H)}$, $k = 1, 2$	The arrival intensity of handover calls at SP k.
$1/\mu_k$, $k = 1, 2$	The mean of call durations at SP k.
$\beta_{k,l}$, $k = 1, 2$	The probability that a fresh call is accepted by SP k if the number of busy channels handled by SP k is l.
$I_k(t)$, $k = 1, 2$	The number of busy channels handled by SP k.

4.2.2 Model Description

Based on the assumptions, the system is modeled by the CTMC $\{I_1(t), I_2(t)\}$ on the state space $S = \{(i,j) | 0 \leq i \leq N, 0 \leq j \leq N, \lceil i/n \rceil + \lceil j/n \rceil \leq n_1 + n_2\}$. The steady state probabilities are denoted by $p_{i,j} = \lim_{t\to\infty} \Pr(I_1(t) = i, I_2(t) = j)$ and $\boldsymbol{\pi}_j = (p_{0,j}, \ldots, p_{f(j),j})$, where $f(k) = N - \lceil k/n \rceil n$. The illustration of the state space and the transitions of the CTMC is depicted in Fig. 4.2 for a case where each SP owns 2 frequency bands and each frequency band has 2 channels.

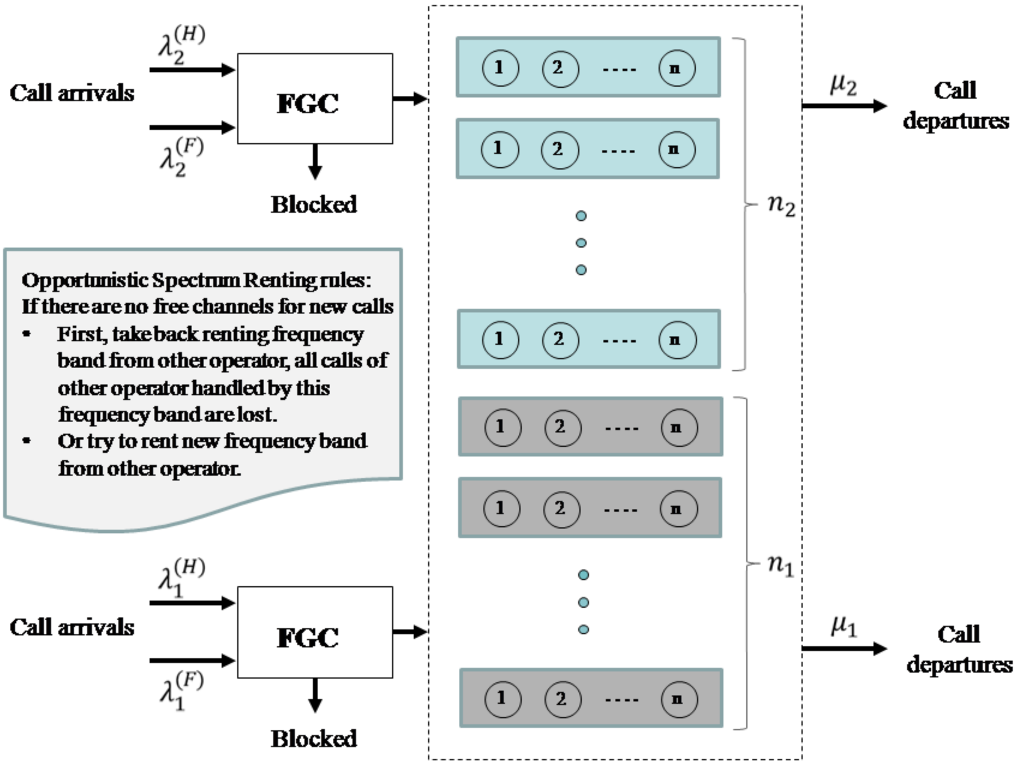
Therefore, the following types of transitions are possible between the states of the CTMC:

- $(i,j) \Rightarrow (i+1,j)$ for $0 \leq i < N$, $0 \leq j < N$ and $\lceil (i+1)/n \rceil + \lceil j/n \rceil \leq n_1 + n_2$ These transitions are due to the acceptance of calls for the second provider,

- $(i,j) \Rightarrow (i,j+1)$ for $0 \leq j < N$, $0 \leq i < N$ and $\lceil i/n \rceil + \lceil (j+1)/n \rceil \leq n_1 + n_2$. These transitions are due to the acceptance of calls for the first provider,

- $(i,j) \Rightarrow (i-1,j)$ for $i > 0$: these transitions happen when a call of the first provider departs from the system,

- $(i,j) \Rightarrow (i,j-1)$ for $j > 0$: these transitions happen when a call of the second provider departs from the system,



Opportunistic Spectrum Renting rules:
If there are no free channels for new calls
- **First, take back renting frequency band from other operator, all calls of other operator handled by this frequency band are lost.**
- **Or try to rent new frequency band from other operator.**

Figure 4.1: Resource contention and spectrum renting.

- $(i, j) \Rightarrow (i+1, (\lceil j/n \rceil - 1)n)$ for $mod(i) = 0$, $i < N_1$ and $\lceil i/n \rceil + \lceil j/n \rceil = n_1 + n_2$, where $mod(i) = i \mod n$. These transitions are initiated by the arrival of a call and one rented frequency band is taken back in case of first provider,

- $(i, j) \Rightarrow ((\lceil i/n \rceil - 1)n, j+1)$ for $mod(j) = 0$ and $j < N_2$ and $\lceil i/n \rceil + \lceil j/n \rceil = n_1 + n_2$. These transitions are initiated by the arrival of a call and one rented frequency band is taken back in case of second provider,

The matrix $(N + 1 - \lceil j/n \rceil n) \times (N + 1 - \lceil j/n \rceil n)$ A_j $(j = 0, \ldots, N)$ contains the transition rate $A_j(i, k)$ from state (i, j) to state (k, j). We can write

$$A_j(i,k) = \begin{cases} \lambda_{1,i} & \text{if } k = i+1,\ 0 \leq i < N, \\ i\mu_1 & \text{if } k = i-1,\ 0 < i \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

The $(N+1-\lceil j/n \rceil n) \times (N+1-\lceil (j+1)/n \rceil n)$ matrix B_j $(j = 0, \ldots, N-1)$ includes

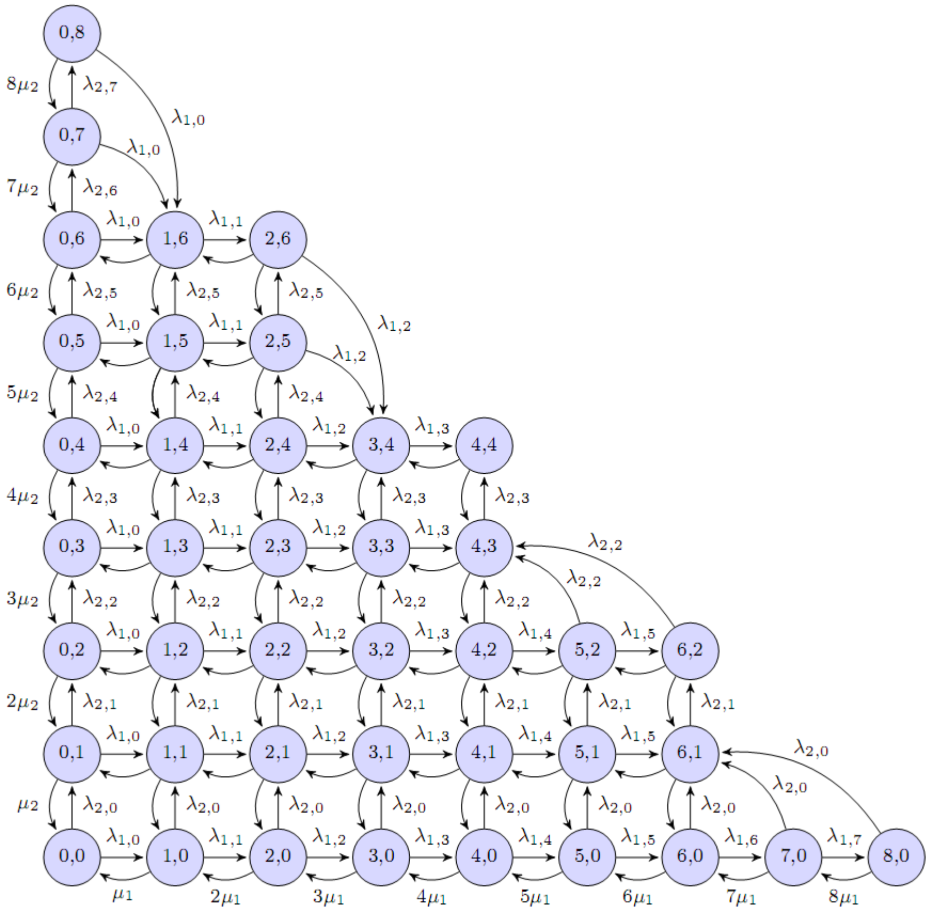


Figure 4.2: Illustration of state transition diagram.

the rate of transitions from state (i, j) to state state $(k, j + 1)$. We can obtain

$$B_j(i,k) = \begin{cases} \lambda_{2,j} & \text{if } k = i, \\ \lambda_{2,j} & \text{if } mod(j) = 0 \text{ and } j < N_2, \\ & k = (\lceil i/n \rceil - 1)n, \\ & \lceil j/n \rceil + \lceil i/n \rceil = n_1 + n_2 \\ 0 & \text{otherwise.} \end{cases}$$

The elements of the $(N+1-\lceil j/n \rceil n) \times (N+1-\lceil (j-1)/n \rceil n)$ matrix C_j $(j = 1, \ldots, N)$ are the rates of transitions from state (i, j) to state $(k, j - 1)$ due to the departures of calls. We can obtain

$$C_j(i,k) = \begin{cases} j\mu_2 & \text{if } k = i, \\ 0 & \text{otherwise.} \end{cases}$$

For $j > N_2$, the transition rate from state (i, j) to state $(k, (\lceil j/n \rceil - 1)n)$ due to the arrival of calls to the first provider is included in matrix $C_{j,(\lceil j/n \rceil - 1)n} = C_{j,*}$

of size $(N+1-\lceil j/n \rceil n) \times (N+1-(\lceil j/n \rceil - 1)n)$ with only one nonzero element $C_{j,(\lceil j/n \rceil -1)n}(i,i+1) = \lambda_{1,i}$ if $\lceil j/n \rceil + \lceil i/n \rceil = n_1 + n_2$ and $mod(i) = 0$.

4.2.3 The Computation of the Steady State Probabilities

Define

$$A_j^{(1)} = \begin{cases} A_0 - D^{A_0} - D^{B_0} \text{ if } j = 0, \\ A_j - D^{A_j} - D^{B_j} - D^{C_j} \text{ if } 0 < j \leq N_2, \\ A_j - D^{A_j} - D^{B_j} - D^{C_j} - D^{C_{j,*}} \text{ if } N_2 < j < N, \\ A_N - D^{A_N} - D^{C_N} - D^{C_{N,*}} \text{ if } j = N. \end{cases}$$

Note that D^Z $(Z = A_j, B_j, C_j)$ is a diagonal matrix whose diagonal element is the sum of all elements in the corresponding row of Z.

The balance equations, which equate the probability fluxes from and to the states of CTMC can be written as follows:

$$\pi_0 A_0^{(1)} + \pi_1 C_1 = 0. \tag{4.1}$$

For $N_2 \leq k < N$ and $mod(k) = 0$, we have

$$\pi_{k-1} B_{k-1} + \pi_k A_k^{(1)} + \pi_{k+1} C_{k+1} + \sum_{l=1}^{n} \pi_{k+l} C_{k+l,*} = 0. \tag{4.2}$$

For $0 < k < N$ and $mod(k) \neq 0$, we can write

$$\pi_{k-1} B_{k-1} + \pi_k A_k^{(1)} + \pi_{k+1} C_{k+1} = 0. \tag{4.3}$$

The last balance equation is

$$\pi_{N-1} B_{N-1} + \pi_N A_N^{(1)} = 0. \tag{4.4}$$

To compute the steady state probabilities, we shall proceed as follows. Let us define matrices R_k's such as $\pi_k = \pi_{k-1} R_k$ for $k = 1, 2, \ldots, N$. Then, based on equations (4.2-4.4), matrices R_k's can be recursively computed using

$$R_N = -B_{N-1} (A_N^{(1)})^{-1}, \tag{4.5}$$

$$R_k = -B_{k-1} [A_k^{(1)} + R_{k+1} C_{k+1}]^{-1}, \;\; (k = N-1, \ldots, 1 \text{ and } mod(k) \neq 0). \tag{4.6}$$

In case of $mod(k) = 0$, the equation (4.2) can be rewritten as follows:

$$\pi_{k-1}B_{k-1} + \pi_k A_k^{(1)} + \pi_{k+1}C'_{k+1} = 0 \tag{4.7}$$

with

$$C'_{k+1} = C_{k+1} + C_{k+1,*} + \sum_{l=2}^{n}\prod_{i=2}^{l} R_{k+i}C_{k+l,*} \tag{4.8}$$

as

$$\pi_{k+l} = \pi_{k+1}R_{k+2}\dots R_{k+l-1}R_{k+l} = \pi_{k+1}\prod_{i=2}^{l} R_{k+i}.$$

Therefore it leads to form of equation (4.6):

$$R_k = -B_{k-1}[A_k^{(1)} + R_{k+1}C'_{k+1}]^{-1}, \ \ (N_2 \le k < N \text{ and } mod(k) = 0). \tag{4.9}$$

This means, π_k $(k = 1, 2, \dots, N)$ can be expressed in π_0 and the already computed R_k's as

$$\pi_k = \pi_0 \prod_{i=0}^{k} R_k, \ (k = 0, \dots, N), \tag{4.10}$$

where R_0 is the identity matrix by definition. Equation (4.1) can be rewritten as

$$\pi_0(A_0^{(1)} + R_1C_1) = 0. \tag{4.11}$$

To compute π_0 we utilize the normalization equation $\sum_{k=0}^{N}\pi_k\mathbf{e}_k = 1$ and equation (4.11). Next, substituting π_0 into equation (4.10), we get the stationary probabilities.

4.2.4 Performance Measures

The performance measures can be computed as follows:

- the blocking probability of handover calls

$$P_{H,1} = \sum_{j=0}^{N_2} p_{f(j),j},$$

$$P_{H,2} = \sum_{i=0}^{N_1} p_{i,f(i)},$$

- the blocking probability of fresh calls

$$P_{F,1} = \sum_{j=0}^{N}\sum_{i=0}^{f(j)} p_{i,j}(1-\beta_{1,i}),$$

$$P_{F,2} = \sum_{i=0}^{N} \sum_{j=0}^{f(i)} p_{i,j}(1 - \beta_{2,j}),$$

- the ratio of interrupted calls

$$P_{L,1} = \frac{\text{Total forced termination rate of SP 1}}{\text{Total arrival rate of SP 1}}$$

$$= \frac{\sum_{l=0}^{n_2-1} \sum_{k=0}^{n-1} (n-k) \lambda_{2,n\,l}\, p_{f(n\,l)-k,n\,l}}{\lambda_1},$$

$$P_{L,2} = \frac{\text{Total forced termination rate of SP 2}}{\text{Total arrival rate of SP 2}}$$

$$= \frac{\sum_{l=0}^{n_1-1} \sum_{k=0}^{n-1} (n-k) \lambda_{1,n\,l}\, p_{n\,l,f(n\,l)-k}}{\lambda_2},$$

- the ratio of successfully completed calls on rented frequencies:

$$R_{s,1} = 1 - \frac{\text{Total forced termination rate of SP 1}}{\text{Total connection rate of SP 1 on rented frequencies}}$$

$$= 1 - \frac{\sum_{l=0}^{n_2-1} \sum_{k=0}^{n-1} (n-k) \lambda_{2,n\,l}\, p_{f(n\,l)-k,n\,l}}{\sum_{j=0}^{N_2-n} \sum_{i=N_1}^{f(j)-1} \lambda_{1,i}\, p_{i,j}},$$

$$R_{s,2} = 1 - \frac{\text{Total forced termination rate of SP 2}}{\text{Total connection rate of SP 2 on rented frequencies}}$$

$$= 1 - \frac{\sum_{l=0}^{n_1-1} \sum_{k=0}^{n-1} (n-k) \lambda_{1,n\,l}\, p_{n\,l,f(n\,l)-k}}{\sum_{i=0}^{N_1-n} \sum_{j=N_2}^{f(i)-1} \lambda_{2,j}\, p_{i,j}},$$

- the average number of realized calls

$$\overline{X_1} = \sum_{j=0}^{N} \sum_{i=0}^{f(j)} p_{i,j} * i,$$

$$\overline{X_2} = \sum_{i=0}^{N} \sum_{j=0}^{f(i)} p_{i,j} * j,$$

- the average number of rented frequency bands

$$\overline{Y_1} = \sum_{j=0}^{N_2-n} \sum_{i=N_1+1}^{f(j)} p_{i,j}(\lceil i/n \rceil - n_1),$$

$$\overline{Y_2} = \sum_{i=0}^{N_1-n} \sum_{j=N_2+1}^{f(i)} p_{i,j}(\lceil j/n \rceil - n_2).$$

- the average number of calls on rented frequencies:

$$\overline{X_{r,1}} = \sum_{j=0}^{N_2-n} \sum_{i=N_1+1}^{f(j)} p_{i,j}(i - N_1),$$

$$\overline{X_{r,2}} = \sum_{i=0}^{N_1-n} \sum_{j=N_2+1}^{f(i)} p_{i,j}(j - N_2).$$

- the average number of calls on own frequencies:

$$\overline{X_{o,1}} = \overline{X_1} - \overline{X_{r,1}} = \sum_{j=0}^{N} \sum_{i=0}^{f(j)} p_{i,j} * \min(i, N_1),$$

$$\overline{X_{o,2}} = \overline{X_2} - \overline{X_{r,2}} = \sum_{i=0}^{N} \sum_{j=0}^{f(i)} p_{i,j} * \min(j, N_2).$$

4.3 Numerical Results

In this section, we present the numerical results of the proposed opportunistic spectrum access model.

4.3.1 A Simulation Model with Log-normally Distributed Holding Times

The analysis of call traces in mobile cellular networks, performed by Jedrzycki and Leung [80] showed that the holding times of calls in cellular networks follow the log-normal distribution. Therefore, we compare the results obtained by our analytical model and a simulation model, where the channel holding times are log-normally distributed with the mean of 3.29 s and the standard deviation of 1.17 s in the simulation model. Note that these parameter values are taken from [80]. The simulation runs were performed with the confidence level of 99.9%. The confidence interval is ±0.6% of the collected data.

In Fig. 4.3, we plot performance measures vs $\rho_1 = \lambda/(N_1\mu)$ and $\rho_2 = \lambda/(N_2\mu)$ for $n_1 = 6$, $n_2 = 6$ and $n = 8$. From the curves, we can observe the excellent fitting between the analytical and the simulation results, which shows that our model with the exponential distribution of channel holding times can be used to predict the performance of a cell, where an opportunistic spectrum access is applied and the holding times of calls follow the log-normal distribution.

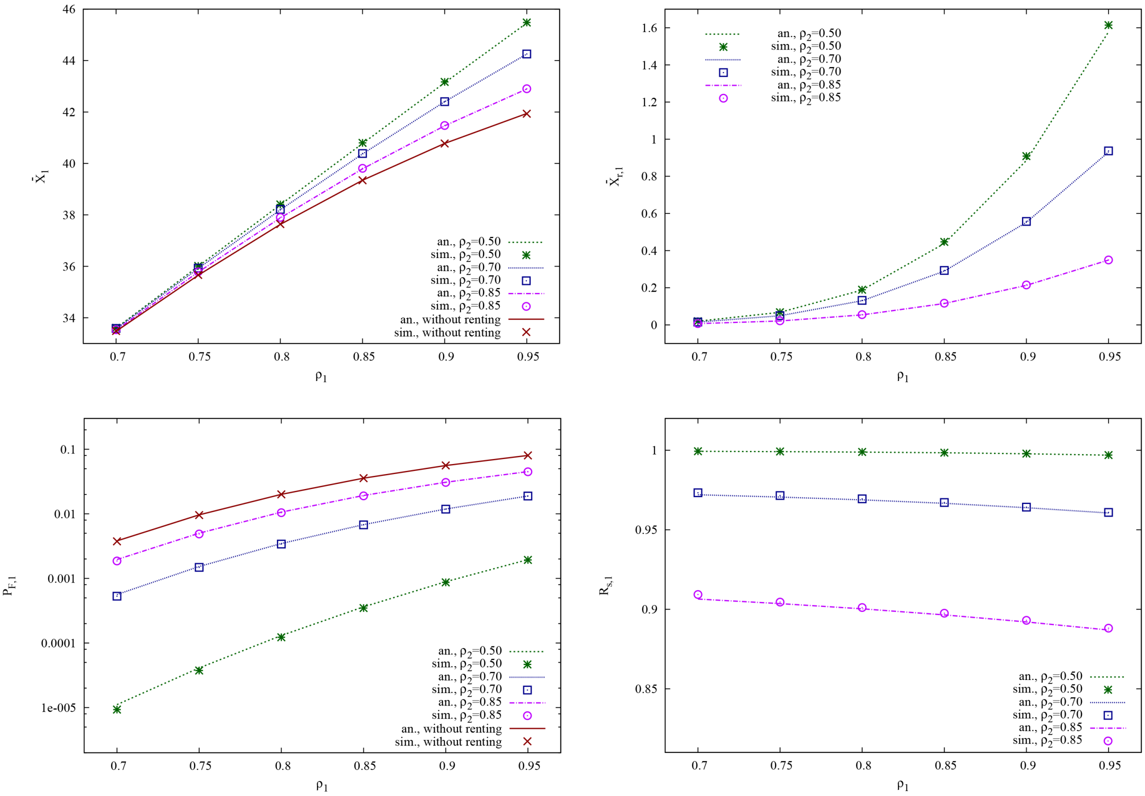


Figure 4.3: Comparison with simulation $(n_1 = n_2 = 6,\ 1/\mu_1 = 1/\mu_2 = 53.22s)$

4.3.2 Impact of Opportunistic Spectrum Access

Figures 4.4 and 4.5 clearly show the advantage of the opportunistic spectrum access. The blocking probability is decreased by one order of magnitude in a large range of offered traffic. It is worth emphasizing that *the utilization of own frequency bands* is increased due to spectrum renting (i.e., the average number of calls realized on own frequency bands is increased in comparison to the average number of established calls when no renting is performed). For example, at $\rho_1 = 0.95$ and $1/\mu_2 = 360s$, the average number of calls established on own frequency bands is approximately 43.41 and 41.94 for the renting case and the no renting case, respectively (see Fig. 4.4). The increase in the average number of calls realized on own frequencies is more than the average number of calls allocated on rented frequencies (at $\rho_1 = 0.95$ and $1/\mu_2 = 360s$, the increase is 43.41-41.94=1.47, while the average number of calls on rented frequencies is 1.02).

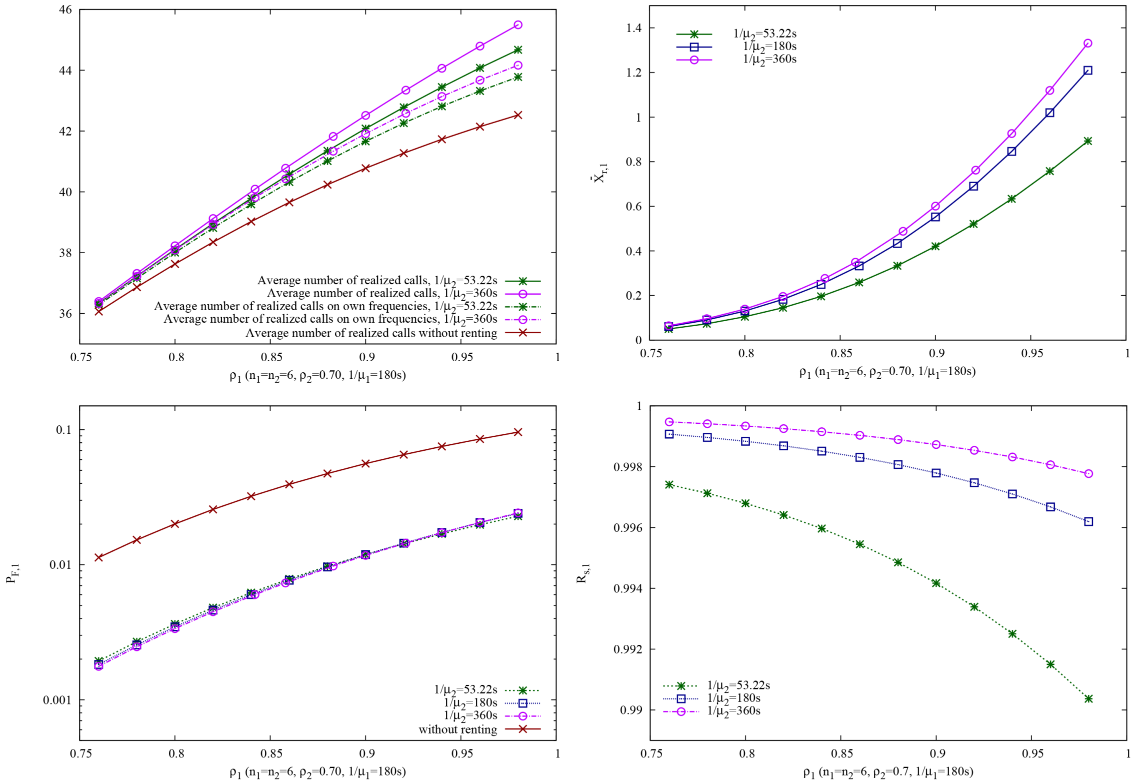


Figure 4.4: Performance measures for $n_1 = n_2 = 6$, $1/\mu_1 = 180s$ and $\rho_2 = 0.7$

On the one hand, the positive impact of spectrum renting on the blocking probability and the average number of calls realized on own frequencies is indeed a good news for the SPs. On the other hand, the price of opportunistic spectrum access is that calls may be forced to leave a rented frequency band and only a portion of subscribers who are allocated in rented frequency band can successfully complete their calls (see Figures 4.4 and 4.5). We will propose a procedure to balance the annoyance concerning the interrupted calls, when the renter takes back an opportunistically used frequency band in Section 4.3.3.

4.3.3 Balancing the Forced Blocking Probability and the Blocking Probability of Calls

To ease the comprehension, we present a procedure in the following from the aspect of SP 1, which acts as a tenant.

To increase the number of admitted calls, the opportunistic renting of free bands from SP 2 is performed by SP 1. Although SP 1 reallocates ongoing calls realized in a rented band to its own frequency band when a call allocated in an own band departures, it is still possible that ongoing calls are forced to leave the system because

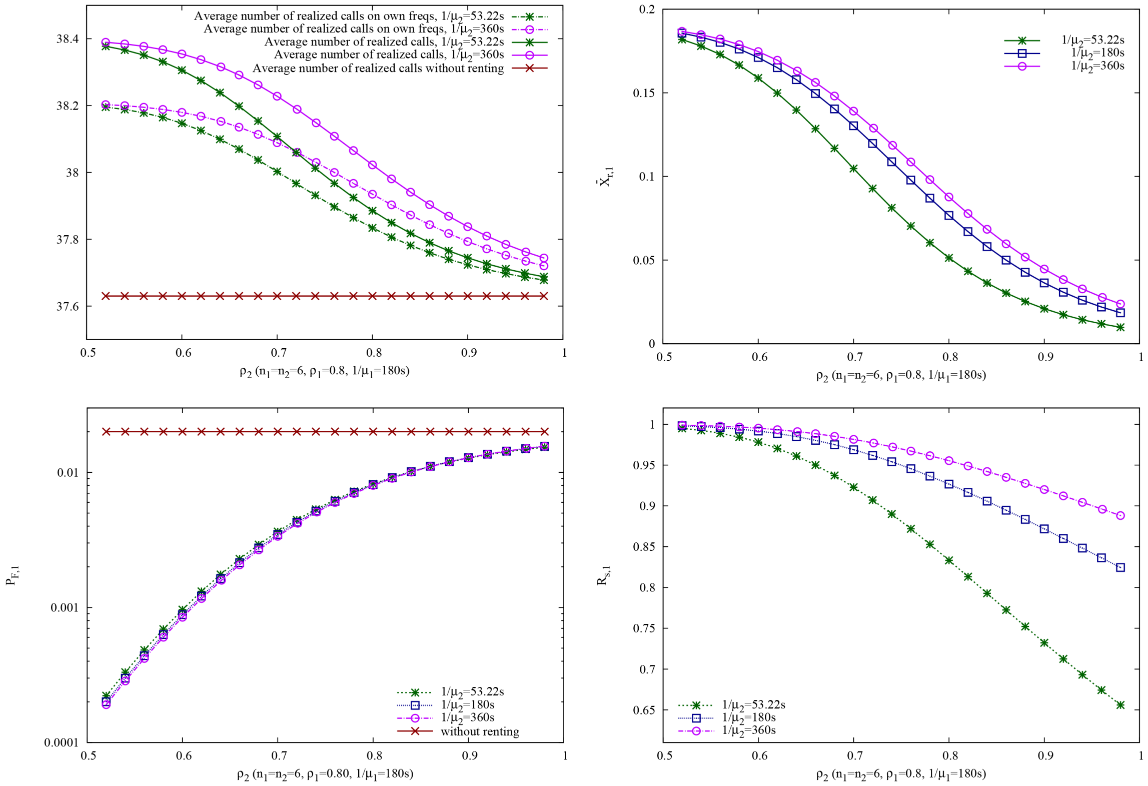


Figure 4.5: Performance measures for $n_1 = n_2 = 6$, $1/\mu_1 = 180s$ and $\rho_1 = 0.8$

SP 2 takes back a frequency band. This is the price of opportunistic spectrum renting indeed, which may cause an inconvenience for subscribers.

To minimize the number of the interrupted calls, it is reasonable to control the number of calls that are placed in rented frequency bands on the one hand, but the advantage of renting frequency bands should be preserved on the other hand. That is, SP 1 should not allocate calls in a specific rented frequency band if it can "foresee" that the specific band will be taken back by SP 2 soon. Therefore, the essential step of a proposed call admission control procedure is to continuously monitor the tendency regarding the number of free frequency bands owned by SP 2. SP 1 computes the exponential weighted moving average $\nu_2(t)$ of the number of free bands $B_2(t)$ that belong to SP 2 and are not used by any SP at time instant t_m is

$$\nu_2(t_m) \leftarrow (1 - 2/(w_2 + 1))\nu_2(t_{m-1}) + 2B_2(t_m)/(w_2 + 1), \qquad (4.12)$$

where w_2 is the window size (weight) chosen by SP 1. Note that w_2 reflects the number of the past changes is closely considered in equation (4.12).

The following Call Admission Control (CAC) decision is proposed

- if $I_1(t) < N_1$ calls are admitted, where $I_1(t)$ the number of busy channels handled by SP 1.

- if $I_1(t) \geq N_1$
  - if $(B_2(t) > 1)$ or $((B_2(t) == 1)$ and $(I_1(t)\ mod(n) \neq 0))$, an arriving call is admitted
  - else if $0 \leq \nu_2 \leq 1$, rejecting an arriving call with probability $1 - max_p \times \nu_2$.

To predict whether the specific band will be taken back by SP 2, SP 1 should incorporate a tendency regarding the change of the number of free bands owned by SP 2 (observed by SP 1). This can be done by the appropriate choice of sampling points t_m. It is normally expected that $\nu_2(t)$ is updated at every instant when $B_2(t)$ is changed. Since a CAC decision is made by SP 1 at the arrival instants of calls, it is reasonable to choose the arrival instants of calls at SP 1 as sampling instants as well (we will justify this hypothesis later).

The impact of the CAC decision is illustrated in Fig. 4.6, where results are obtained through simulation. Namely, we limited the allocation of incoming calls in rented frequency bands in order to reduce the chance that the ongoing calls are forced to leave the system. Therefore, the average number of realized calls decreased. For example, at $\rho_1 = 0.95$, the average number of realized calls is approximately 43.23 and 44.25 for the case when the CAC decision was applied and in the case of pure renting (when no CAC decision was applied), respectively. However, it is observed that limiting the number of calls allocated in rented frequency bands has only a small impact on the utilization of own frequency bands. Moreover, for the price of the decreased utilization, we could significantly improve the service quality of the calls allocated on rented frequency bands (see Fig. 4.6, $R_{s,1}$).

To "validate" the applied sampling hypothesis, we define the following sampling rules:

- t_m's include the set of instants when $B_2(t)$ is changed (Sampling Rule (SR) 1),
- t_m's include the set of arrival instants (SR 2),
- t_m's include the set of arrival instants and instants when $B_2(t)$ is changed (SR 3).

As we mentioned before, we used $SR3$ in Fig. 4.6. In Fig. 4.7, we can observe the previously depicted properties applying the different sampling rules. Moreover, we depicted the results in Fig. 4.7 without using any renting ("no renting") and in the case of pure renting. We can clearly observe in Fig. 4.7, that SR 3 is the best sampling strategy for protecting the calls allocated on rented frequency bands. Moreover, the level of the protection decreases with the increasing of the number of samples used when computing ν_2.

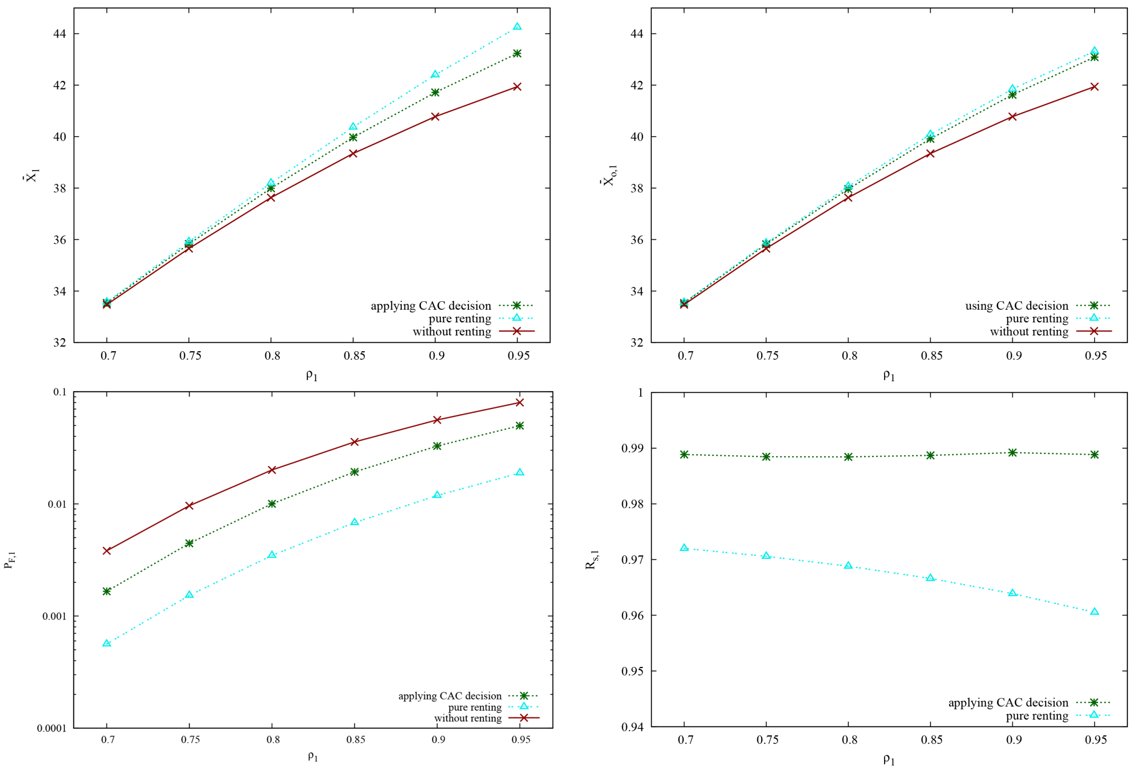


Figure 4.6: Performance measures for $n_1 = n_2 = 6$, $1/\mu_1 = 1/\mu_2 = 53.22s$ and $\rho_2 = 0.7$

To demonstrate that the operators can balance the forced blocking probability and the blocking probability, we present some other simulation results, where we investigated the effect of max_p parameter. Since the service quality of calls allocated on rented frequency bands was most preserved applying SR 3, we used this strategy in the following investigation, too. As we can see in Fig. 4.8, max_p is an additional tool in the hands of the SPs to balance the forced blocking probability and the blocking probability of calls. Varying this parameter, the SPs can fine-tune the level of protection for the ongoing calls allocated on rented frequency bands.

4.3.4 The Effect of Opportunistic Spectrum Access to the Average Profit Rate

To really demonstrate the financial benefits of the proposed cooperation scheme, we explicitly investigate a financial parameter, the Average Profit Rate (APR), which can be computed as follows.

$$APR = \alpha \times \bar{S} - \beta \times (\bar{R} - \bar{L}), \tag{4.13}$$

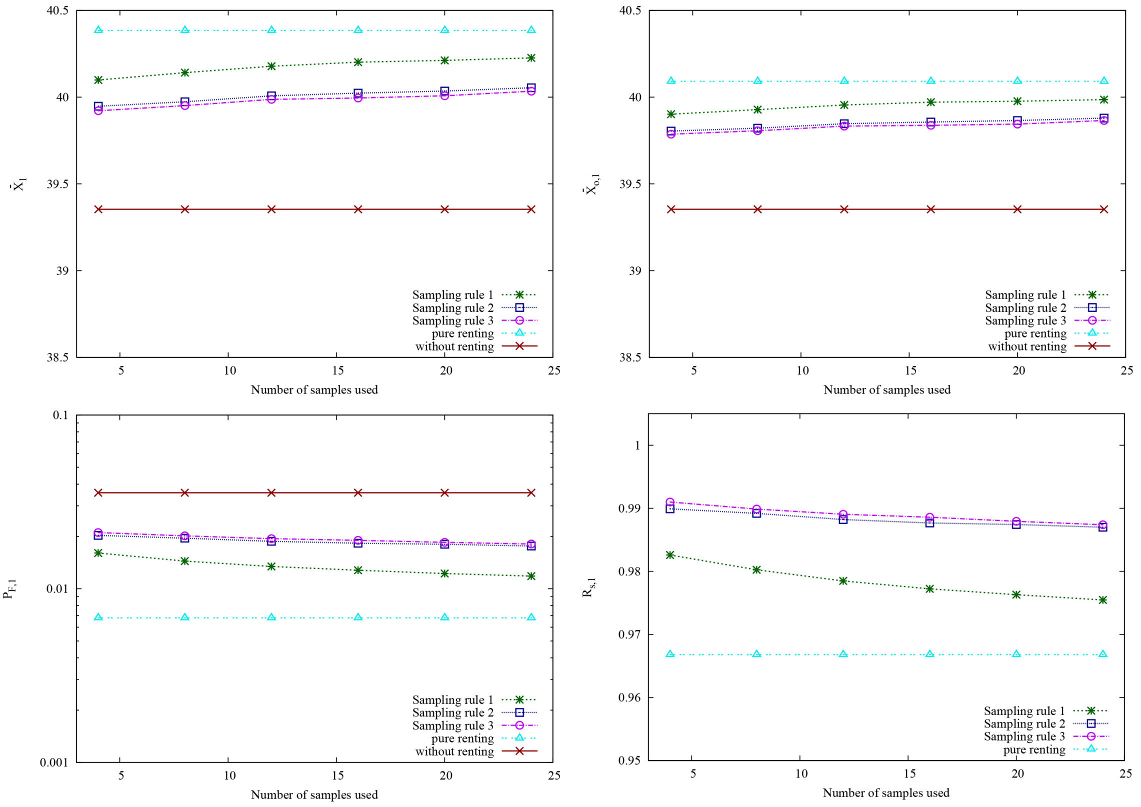


Figure 4.7: Performance measures for $n_1 = n_2 = 6$, $1/\mu_1 = 1/\mu_2 = 53.22s$, $\rho_1 = 0.85$ and $\rho_2 = 0.7$

where α and β are cost coefficients measured in cost units per time unit for the calls and the renting, respectively [10]; $\bar{S}$ is the total duration of the calls, $\bar{R}$ is the total duration of the renting, and $\bar{L}$ is the duration when the other SP was using the current SP's frequency bands. When we determine the renting fee, we must take into account that a SP will have c additional channels when renting a frequency band. In our model, we assume that at least one ongoing call is using the rented frequency band. On the other hand, the rented channel is usually not fully utilized, since there are expectedly less than 8 ongoing calls on it. Therefore, the renting fee should be less than the total income from a fully utilized frequency band (i.e. $\alpha \times c$). So, we define a discount factor and calculate the renting fee as follows.

$$\beta = \alpha \times \frac{c}{d}, \tag{4.14}$$

where $c > d \geq 1$ is the discount factor (note that $d = 1$ means no discount).

Besides, we apply a policy, according to which the SPs do not pay any renting fee to each other in the cases when the given frequency band was withdrawn. This rule can be a compensation for breaking the ongoing connections (of course, $\bar{R}$ will be not

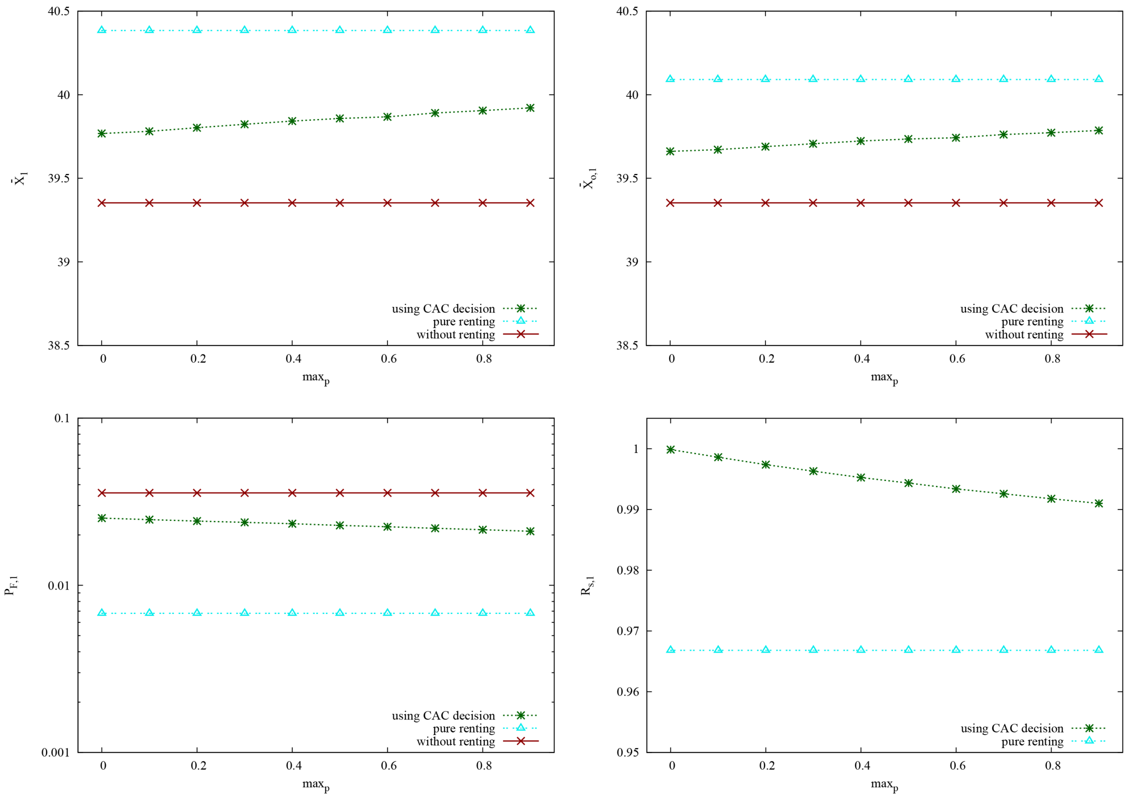


Figure 4.8: Performance measures for $n_1 = n_2 = 6$, $1/\mu_1 = 1/\mu_2 = 53.22s$, $\rho_1 = 0.85$ and $\rho_2 = 0.7$

increased with the time interval of the broken renting when computing the APR).

In the following investigations, the channel holding times are log-normally distributed with the mean of 4.0 seconds and the standard deviation of 1.17 seconds. This means that the expected channel holding time is 108.25 seconds. Moreover, Pattavina and Parini showed that the interarrival time distribution also follows the log-normal distribution [81]. Based on the parameter settings of [81], the mean interarrival time runs from 7.83 to 7.21 seconds to achieve a blocking ratio of 1-2% when the SPs did not cooperate.

The simulation results are obtained with the confidence level of 99.9% and the relative precision (i.e. the ratio of the half-width of the confidence interval and the mean of collected observations) of 0.8%.

In the first scenario, the normalized load is increasing at both SPs. In Fig. 4.9, the APR value is illustrated. We investigated the APR with different values of d, however, the results show that there was no impact of d on the APR in this case. The explanation of this is that when the traffic is equal and equally changes at both SPs, the probability that they will use each other's channel is also equal. Henceforth, we will demonstrate the impact of d in the following.

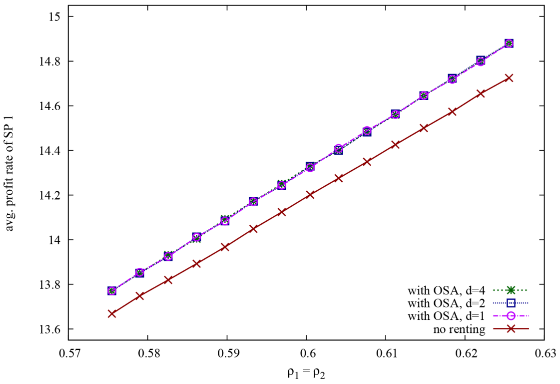


Figure 4.9: The APR for $n_1 = n_2 = 3$ and $1/\mu_1 = 1/\mu_2 = 108.25s$

In the second scenario, we investigated the APR when the load of SP 1 is increasing as in the previous simulation run, while the load of SP 2 is constant with around 1% expected blocking ratio ($\rho_2 = 0.575$). The graphs of 4.10 represent the APR values of SP 1 (left side) and SP 2 (right side). As the load of SP 1 increases, the APR of SP 1 increases, as well. Of course, the higher the discount factor is, the higher APR can be realized for the renter. Even if the discount factor is one, the APR is higher than without cooperation, which can be explained as follows. First, as we showed previously, that the utilization of the own frequency bands is increased due to spectrum renting (see Fig. 4.4). Then, SP 1 also has income from the renting fee, since SP 2 also can use the frequency bands of SP 1 when necessary. Finally, the compensation rule allows SP 1 not to pay any renting fee if the given frequency band was withdrawn.

On the other hand, SP 2 can increase its profit only due to the increase of the traffic of SP 1, when the SPs cooperate. Even if SP 2 offers a discount of $d = 2$ or $d = 4$, the gain is significant (see the right graph of Fig. 4.10).

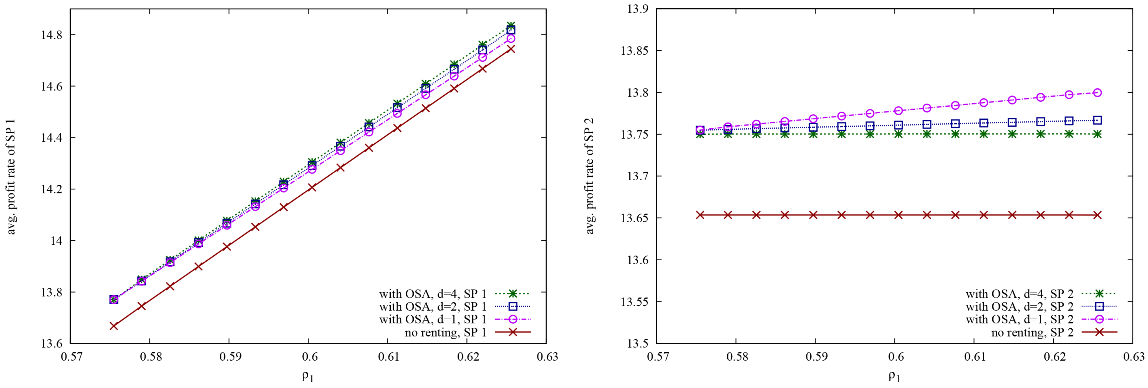


Figure 4.10: The APR for $n_1 = n_2 = 3$, $1/\mu_1 = 1/\mu_2 = 108.25s$ and $\rho_2 = 0.575$

4.4 Summary

In this chapter, we have proposed a spectrum access cooperation scheme for two service providers operating in a specific area.

We have elaborated an analytical model to evaluate the performance of an opportunistic spectrum access scenario between two operators. We have shown the positive impact of the spectrum renting on the blocking probability and the average number of calls realized in own frequency bands.

We presented a simulative study, too, in which we showed that our Markovian model using exponentially distributed holding times is a good approximation of the real operation, where the holding times follow the log-normal distribution. Besides, the simulation results showed an excellent agreement with the analytical ones, which is a good validation of the Markovian model.

Furthermore, we investigated the forced termination problem, which is a disadvantage of the opportunistic spectrum access model. Based on the investigation, we provided a solution to control the number of calls forced to leave a rented frequency band.

Finally, we investigated the effect of our opportunistic spectrum access model to the average profit rate. The results showed that the cooperation is financially beneficial for both cooperating parties.

Conclusion

Modeling a process always involves some simplifications, which make the model manageable. At this point, we always have to make trade-off in the complexity of the models to appropriately preserve the manageability, while we must not lose the essence of the modeled process.

In this dissertation, we dealt with two main problem sets. In the first one, we proposed a novel approach for modeling application spreading in mobile ad hoc environments. Based on the complexity of the models, we showed different techniques to obtain the transient solution of the models, which is more interesting in this context than the steady state solution. First, we presented a simple CQN model to give an analytical estimation on the number of application purchases. Then, we introduced our SPN models, in which we can handle more sophisticated user behavior. Namely, the first SPN model can handle two of the presented user types, and we got the transient solution of the model using a mean field based methodology, which is an analytical approximation. Together with that, we demonstrated that with some restrictions, the mean field based methodology can be applied for Petri nets, too. In the second SPN model, we could handle the third user behavior type. However, the use of inhibitor arcs resulted in the violation of the density dependent property in the underlying CTMC of the SPN, so the mean field approach could not be used, and our investigations were limited to simulation. An even more complex member of the Petri net family is the DSPN, which is useful if the transitions of a system are either exponentially, or deterministically delayed. In a production process, where the different work phases are deterministic processes, a DSPN is an appropriate modeling tool. In this dissertation, we presented the operation model of a wooden window production company. Obtaining the solution of the model, we identified the main bottleneck of the production process. Moreover, we determined the measure of improving the given work phase in order to eliminate the main bottleneck from the process. With this example, we demonstrated that the stochastic models can be applied in the wood industry.

Our second problem set was the opportunistic spectrum access in cellular networks. Based on the spectrum pooling concept, we elaborated a model for the opportunistic use of the licensed spectrum. We presented the analytical solution of the model, which we validated via simulations. We showed that the performance measures like utilization and blocking can be improved using opportunistic spectrum access, when the service providers utilize each other's idle frequency bands. We identified the main drawback of our model, the forced termination problem, which occurs when the license holder takes back an opportunistically rented frequency band for its own use. Using simulation, we showed that our concept for alleviating the negative effects of the forced termination problem is operable: we can balance the blocking probability of the calls and the forced termination applying a call admission control. Finally, we justified the financial benefits of the opportunistic spectrum access. Investigating the average profit rate, we showed that both cooperating parties can realize extra profit for sharing their licensed spectrum.

Appendix

Acronyms

APR Average Profit Rate

ARED Adaptive Random Early Detection

CAC Call Admission Control

CNC Computer Numerical Control

CQN Closed Queuing Network

CTMC Continuous-Time Markov Chain

DSPN Deterministic and Stochastic Petri Net

FFT Fast Fourier Transform

FGC Fractional Guard Channel

GSM Global System for Mobile Communications

GSSMC General State Space Markov Chain

IFFT Inverse Fast Fourier Transform

ISM Industrial, Scientific and Medical

MVA Mean Value Analysis

LAN Local Area Network

ODEs Ordinary Differential Equations

OFDM Orthogonal Frequency-Division Multiplexing

OQN Open Queuing Network

SIR Susceptible-Infected-Resistant

SIS Susceptible-Infected-Susceptible

SP Service Provider

SPN Stochastic Petri Net

SR Sampling Rule

WLAN Wireless Local Area Network

Own Publications Related to this Dissertation

[1] Á. Horváth, "Modeling opportunistic application spreading," in *Proceedings of the Second International Workshop on Mobile Opportunistic Networking*, pp. 207–208, ACM, 2010.

[2] Á. Horváth and K. Farkas, "Alkalmazások terjedésének vizsgálata mobil ad hoc hálózatokban," in *Proceedings of IKT2010, Dunaújváros, Hungary*, pp. 1–6, 2010.

[3] Á. Horváth and K. Farkas, "Modeling self-organized application spreading," in *Proc. of the Fifth International ICST Conference on Access Networks (ACCESSNETS 2010)*, pp. 71–80, Springer, 2011.

[4] Á. Horváth and K. Farkas, "Modeling application spreading using mobile ad hoc networks," in *Wireless and Mobile Networking Conference (WMNC), 2010 Third Joint IFIP*, pp. 1–6, IEEE, 2010.

[5] Á. Horváth and K. Farkas, "Techniques for modeling mobile application spreading," *Infocommunications Journal*, vol. IV, no. 1, pp. 13–20, 2012.

[6] M. Beccuti, M. De Pierro, A. Horváth, Á. Horváth, and K. Farkas, "A mean field based methodology for modeling mobility in ad hoc networks," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pp. 1–5, IEEE, 2011. **Number of independent citations: 2**.

[7] Á. Horváth, "Usability of deterministic and stochastic petri nets in the wood industry: a case study," *In Proceedings of the* 2^{nd} *International Conference on Computer Science, Applied Mathematics and Applications (ICCSAMA 2014)*, pp. 119–127, 2014.

[8] Á. Horváth, "Applying opportunistic spectrum access in mobile cellular networks," *Infocommunications Journal*, vol. V, no. 2, pp. 36–40, 2013.

[9] J. Boros and Á. Horváth, "GSM szolgáltatók közötti együttműködés vizsgálata," in *Proceedings of OGIK 2012.*, 2012.

[10] T. V. Do, N. H. Do, Á. Horváth, and J. Wang, "Modelling opportunistic spectrum renting in mobile cellular networks," *Revised version submitted to Elsevier Journal of Network and Computer Applications*, 2014.

Other Own Publications

[11] T. Bérczes and Á. Horváth, “A finite-source queuing model for spectrum renting in mobile cellular networks,” in *Accepted in Proceedings of the 10th International Conference Elektro 2014, Rajecké Teplice, Slovakia*, 2014.

[12] P. Schaffer, K. Farkas, Á. Horváth, T. Holczer, and L. Buttyán, “Secure and reliable clustering in wireless sensor networks: A critical survey,” *Computer Networks*, vol. 56, no. 11, pp. 2726–2741, 2012. **Impact factor: 1.2. Number of independent citations: 10**.

[13] J. Boros, Á. Horváth, and L. Jereb, “Szűk keresztmetszetek feltárása többrétegű architektúrákban,” in *Proceedings of IF2011, Debrecen, Hungary*, pp. 758–765, 2011.

[14] L. Bacsárdi and Á. Horváth, “Mobile ad hoc networks in the applied informatics,” *GÉP (Journal of Scientific Society of Mechanical Engineering, Hungary)*, vol. 61, no. 1-2, pp. 25–27, 2010.

[15] Á. Horváth and T. Kárász, “A konszolidáció hatása az igények rendelkezésre állására,” in *Student Conference organized by BME and HTE, Budapest, Hungary*, pp. 1–4, 2007.

References

[16] K. Farkas, L. Ruf, M. May, and B. Plattner, “Framework for service provisioning in mobile ad hoc networks,” *Proceedings of the First International Conference on Telecommunications and Computer Networks, San Sebastian, Spain*, 2004.

[17] J. Mitola III, “Cognitive radio for flexible mobile multimedia communications,” in *Mobile Multimedia Communications, 1999.(MoMuC’99) 1999 IEEE International Workshop on*, pp. 3–10, IEEE, 1999.

[18] Kolodzy, P, “Spectrum policy task force report,” *Federal Communications Commission ET Docket 02-135*, 2002.

[19] M. M. Buddhikot, “Understanding dynamic spectrum access: Models, taxonomy and challenges,” in *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pp. 649–663, April 2007.

[20] Q. Zhao and B. M. Sadler, “A survey of dynamic spectrum access,” *Signal Processing Magazine, IEEE*, vol. 24, pp. 79–89, May 2007.

[21] T. A. Weiss and F. Jondral, “Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency,” *Communications Magazine, IEEE*, vol. 42, pp. 8–14, March 2004.

[22] eNET and Telekom, “Már okostelefon-felhasználó a magyar lakosság több mint 1/4-e,” May 2013. `http://www.enet.hu`, Last visited on July 30, 2013.

[23] J. M. Peha, “Sharing spectrum through spectrum policy reform and cognitive radio,” *Proceedings of the IEEE*, vol. 97, pp. 708–719, 2009.

[24] B. Jabbari, R. Pickholtz, and M. Norton, “Dynamic spectrum access and management Dynamic Spectrum Management,” *IEEE Wireless Communications*, vol. 17, pp. 6–15, August 2010.

[25] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, “Towards real-time dynamic spectrum auctions,” *Computer Networks*, vol. 52, no. 4, pp. 879–897, 2008.

[26] S.-S. Tzeng and C.-W. Huang, “Threshold based call admission control for qos provisioning in cellular wireless networks with spectrum renting,” in *Novel Algorithms and Techniques in Telecommunications and Networking*, pp. 17–22, Springer, 2010.

[27] S.-S. Tzeng, “Call admission control policies in cellular wireless networks with spectrum renting,” *Computer Communications*, vol. 32, no. 18, pp. 1905–1913, 2009.

[28] E. Noam, “Taking the next step beyond spectrum auctions: open spectrum access,” *IEEE Communications Magazine*, vol. 33, pp. 66–73, 1995.

[29] T. G. Kurtz, “Solutions of ordinary differential equations as limits of pure jump markov processes,” *Journal of Applied Probability*, vol. 7, no. 1, pp. 49–58, 1970.

[30] S. Floyd and V. Jacobson, “Random early detection gateways for congestion avoidance,” *IEEE/ACM Trans. Netw.*, vol. 1, pp. 397–413, Aug. 1993.

[31] L. Kleinrock, “On the modeling and analysis of computer networks,” *Proceedings of the IEEE*, vol. 81, no. 8, pp. 1179–1191, 1993.

[32] T. V. Do and R. Chakka, “Simulation and analytical approaches for estimating the performability of a multicast address dynamic allocation mechanism.,” *Simulation Modelling Practice and Theory*, vol. 18, no. 7, pp. 971–983, 2010.

[33] S. Balsamo and A. Marin, “Queueing networks,” in *Formal Methods for Performance Evaluation*, pp. 34–82, Springer, 2007.

[34] T. G. Robertazzi, *Computer Networks and System: Queueing Theory and Performance Evaluation.* Springer, 2000.

[35] M. Reiser and S. S. Lavenberg, "Mean-value analysis of closed multichain queuing networks," *Journal of the ACM (JACM)*, vol. 27, no. 2, pp. 313–322, 1980.

[36] S. Balsamo and M. Clò, "A convolution algorithm for product-form queueing networks with blocking," *Annals of Operations Research*, vol. 79, pp. 97–117, 1998.

[37] S. Balsamo and A. Rainero, "Closed queueing networks with finite capacity queues: approximate analysis.," in *ESM*, pp. 593–600, Citeseer, 2000.

[38] J. R. Jackson, "Networks of waiting lines," *Operations Research*, vol. 5, no. 4, pp. 518–521, 1957.

[39] J. R. Jackson, "Jobshop-like queueing systems," *Management Science*, vol. 10, no. 1, pp. 131–142, 1963.

[40] W. J. Gordon and G. F. Newell, "Closed queuing systems with exponential servers," *Operations Research*, vol. 15, no. 2, pp. 254–265, 1967.

[41] G. Yin and Q. Zhang, *Continuous-time Markov chains and applications.* Springer New York, 1998.

[42] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, vol. 3. Springer Berlin, 1985.

[43] F. J. W. Symons, *Modeling and Analysis of Communication Protocols Using Numerical Petri Nets.* PhD thesis, University of Essex, 1978.

[44] M. K. Molloy, *On the Integration of Delay and Throughput Measures in Distributed Processing Models.* PhD thesis, UCLA, Los Angeles, CA, 1981.

[45] M. A. Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis, *Modelling with Generalized Stochastic Petri Nets.* John Wiley and Sons, 1995.

[46] W. J. Stewart, *Introduction to the numerical solution of Markov chains*, vol. 41. Princeton University Press Princeton, 1994.

[47] A. Zimmermann and M. Knoke, *TimeNET 4.0: A software tool for the performability evaluation with stochastic and colored Petri nets; user manual.* TU, Professoren der Fak. IV, 2007.

[48] S. Baarir, M. Beccuti, D. Cerotti, M. De Pierro, S. Donatelli, and G. Franceschinis, “The greatspn tool: recent enhancements,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 36, no. 4, pp. 4–9, 2009.

[49] P. Bonet, C. M. Lladó, R. Puijaner, and W. J. Knottenbelt, “Pipe v2. 5: A petri net tool for performance modelling,” in *Proc. 23rd Latin American Conference on Informatics (CLEI 2007)*, 2007.

[50] M. A. Marsan and G. Chiola, “On petri nets with deterministic and exponentially distributed firing times,” in *Advances in Petri Nets 1987*, pp. 132–145, Springer Berlin Heidelberg, 1987.

[51] C. Lindemann and G. S. Shedler, “Numerical analysis of deterministic and stochastic petri nets with concurrent deterministic transitions,” *Perform. Eval.*, vol. 27-28, pp. 565–582, Oct. 1996.

[52] C. Lindemann and A. Thümmler, “Transient analysis of deterministic and stochastic petri nets with concurrent deterministic transitions,” *Performance Evaluation*, vol. 36, pp. 35–54, 1999.

[53] W. D. Kelton and A. M. Law, *Simulation modeling and analysis.* McGraw Hill Boston, MA, 2000.

[54] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, “Epidemic spreading in real networks: An eigenvalue viewpoint,” in *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*, pp. 25–34, IEEE, 2003.

[55] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Physical review letters*, vol. 86, no. 14, p. 3200, 2001.

[56] F. Fu, L. Liu, and L. Wang, “Information propagation in a novel hierarchical network,” *arXiv preprint math/0605293*, 2006.

[57] A. Khelil, C. Becker, J. Tian, and K. Rothermel, “An epidemic model for information diffusion in manets,” in *Proceedings of the 5th ACM international workshop on Modeling analysis and simulation of wireless and mobile systems*, pp. 54–60, ACM, 2002.

[58] O. Sekkas, D. Piguet, C. Anagnostopoulos, D. Kotsakos, G. Alyfantis, C. Kassapoglou-Faist, and S. Hadjiethymiades, “Probabilistic information dissemination for manets: the ipac approach,” in *The Internet of Things*, pp. 375–385, Springer, 2010.

[59] V. Karyotis, A. Kakalis, and S. Papavassiliou, "Malware-propagative mobile ad hoc networks: Asymptotic behavior analysis," *Journal of Computer Science and Technology*, vol. 23, no. 3, pp. 389–399, 2008.

[60] E. Huang, W. Hu, J. Crowcroft, and I. Wassell, "Towards commercial mobile ad hoc network applications: A radio dispatch system," in *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pp. 355–365, ACM, 2005.

[61] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless communications and mobile computing*, vol. 2, no. 5, pp. 483–502, 2002.

[62] S. Gowrishankar, T. Basavaraju, and S. K. Sarkar, "Effect of random mobility models pattern in mobile ad hoc networks," *International Journal of Computer Science and Network Security*, vol. 7, no. 6, pp. 160–164, 2007.

[63] M. J. Kumar and R. Rajesh, "Performance analysis of manet routing protocols in different mobility models," *Proceedings of the International Journal of Computer Science and Network Security, IJCSNS*, vol. 9, no. 2, pp. 22–29, 2009.

[64] B. Cohen, "Incentives build robustness in bittorrent," in *Workshop on Economics of Peer-to-Peer systems*, vol. 6, pp. 68–72, 2003.

[65] C. Bettstetter, "Mobility modeling in wireless networks: categorization, smooth movement, and border effects," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 3, pp. 55–66, 2001.

[66] V. Karyotis, M. Grammatikou, and S. Papavassiliou, "A closed queueing network model for malware spreading over non-propagative ad hoc networks," *Proc. of MedHocNet*, pp. 129–136, 2007.

[67] M. Tribastone, *Scalable Analysis of Stochastic Process Algebra Models*. PhD thesis, School of Informatics, University of Edinburgh, 2010.

[68] H.-T. Ltd., "The web page of holz-team ltd.," 2013. `http://www.holzteam.hu/en`, Last visited on November 28, 2013.

[69] G. Ciardo, R. German, and C. Lindemann, "A characterization of the stochastic process underlying a stochastic petri net," *Software Engineering, IEEE Transactions on*, vol. 20, no. 7, pp. 506–515, 1994.

[70] T. V. Do, N. H. Do, and R. Chakka, "A new queueing model for spectrum renting in Mobile Cellular Networks," *Computer Communications*, vol. 35, pp. 1165–1171, June 2012.

[71] T. V. Do, P. Wuechner, T. Berczes, J. Sztrik, and H. De Meer, "A new finite-source queueing model for mobile cellular networks applying spectrum renting," *Asia Pacific Journal of Operational Research (APJOR)*, vol. [NA], p. [NA], 2013. accepted; to appear.

[72] T. Weiss, J. Hillenbrand, and F. Jondral, "A Diversity Approach for the Detection of Idle Spectral Resources in Spectrum Pooling Systems," in *48th Int. Scientific Colloquium, Ilmenau, Germany*, 2003.

[73] J. Neyman and E. S. Pearson, *On the problem of the most efficient tests of statistical hypotheses.* Springer, 1992.

[74] T. Weiss, J. Hillenbrand, A. Krohn, and F. K. Jondral, "Efficient signaling of spectral resources in spectrum pooling systems," in *Proc. 10th Symposium on Communications and Vehicular Technology (SCVT)*, 2003.

[75] T. Weiss, A. Krohn, and F. Jondral, "Synchronization Algorithms and Preamble Concepts in Spectrum Pooling Systems," in *1st Mobile and Wireless Telecommun. Summit, Aveiro, Portugal*, 2003.

[76] P. Tran-Gia and M. Mandjes, "Modeling of customer retrial phenomenon in cellular mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 8, pp. 1406–1414, 1997.

[77] J. R. Artalejo and M. J. Lopez-Herrero, "Cellular mobile networks with repeated calls operating in random environment," *Computers & operations research*, vol. 37, no. 7, pp. 1158–1166, 2010.

[78] T. V. Do, "A new computational algorithm for retrial queues to cellular mobile systems with guard channels," *Computers & Industrial Engineering*, vol. 59, no. 4, pp. 865 – 872, 2010.

[79] T. V. Do, "Solution for a retrial queueing problem in cellular networks with the fractional guard channel policy," *Mathematical and Computer Modelling*, vol. 53, no. 11–12, pp. 2058–2065, 2011.

[80] C. Jedrzycki and V. C. M. Leung, "Probability distribution of channel holding time in cellular telephony systems," in *IEEE VTC'96*, (Atlanta, GA), pp. 247–251, 1996.

[81] A. Pattavina and A. Parini, “Modelling voice call inter-arrival and holding time distributions in mobile networks,” *Performan. Challenges Efficient Next Generat. Networks, Proc. 19th Int. Teletraf. Congr*, pp. 729–738, 2005.