BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
DEPARTMENT OF TELECOMMUNICATIONS

Unified Analysis of Cyclic Polling Models with BMAP

Ph.D. Thesis summary

by

Zsolt Saffer

Advisor:

Prof. Dr. Miklós Telek, BUTE

Budapest, Hungary
2010

1   Introduction

The polling model gained greater attention first in the 1970s, when it was used in the analysis of the time-sharing computer systems. In the 1980s the cyclic polling model has gained much attention in the performance analysis of the token passing protocols of local area networks. In the subsequent decades it has been applied to channel access protocols in metropolitan area networks, to land mobile and satellite radio communication networks and also to wireless communication networks. There was an explosive growth of research on polling systems, motivated by the increasing number of applications in the communications systems. Numerous variations and extensions of the basic polling model have been proposed and analyzed (see the survey of Takagi [27].

The classical polling model in its original form is a single-server continuous-time queueing system, in which the single server attends the stations in cyclic manner (see on Figure 1). Each station has a queue, which is served when the server visits that station. The time required for the server to travel from one station to the next one is called *switchover time*.
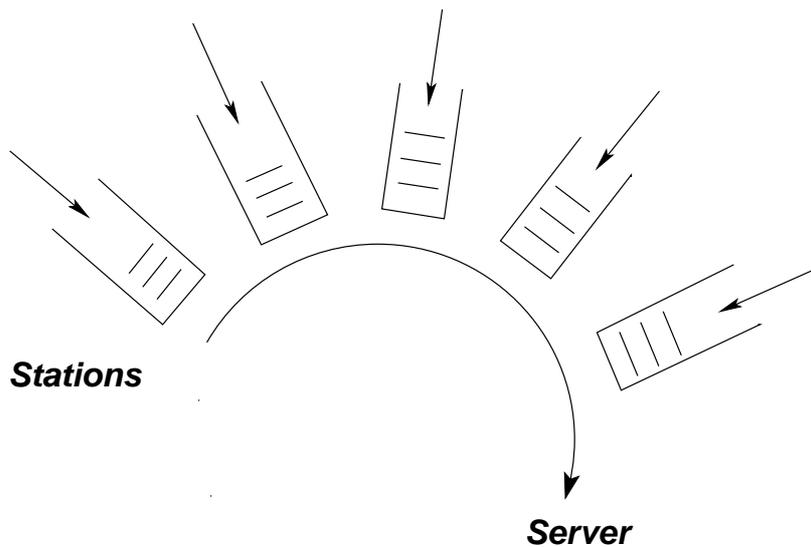

Figure 1: Classical polling model

The arrival of the server to a station and the departure of the server from a station are called *polling epoch* and *departure epoch*, respectively. The *cycle time* is the time between two successive visits of the same station. The *station time* is the duration of the server visit to a station. Another characterizing term of the model is the *intervisit time* of a station, which is the duration from the departure epoch of a station to the next polling epoch of the same station.

The polling models can be differentiated according to the *service discipline* used at the stations. It is also called as service policy. In general it gives the condition on the beginning and on the end of the service at a given station. The most commonly known disciplines are, e.g., the exhaustive, the gated, the non-exhaustive, the semi-exhaustive, the G-limited and the decrementing-K. In the exhaustive policy the server serves the station until its queue becomes empty. This is the discipline, which is used also in the ordinary queue. Under gated discipline during a visit to a station the server serves only the customers that are present at the polling epoch at that station. In case of G-limited discipline at most K customers are served among the customers that are present at the polling epoch at that station. Under decrementing-K discipline, the service at a station continues until the number of customers, that are present at the polling epoch, decreases by K or the station becomes empty. With $K = 1$ the non-exhaustive and the semi-exhaustive disciplines are special cases of the G-limited and the decrementing-K policies, respectively. For the analysis of the classical cyclic polling model with different disciplines we refer to the fundamental book of Takagi [26].

As the introduction of the numerous extensions and variations of polling models and their applications show, the driving force of the further development of the polling models is the rapid development of the modern telecommunication networks. There is a growing demand for more accurate modeling of the operation of such systems, while they become more complex. The advanced modeling makes their performance evaluation more precise, facilitates the optimization of their parameters alleviating the better tuning of these parameters to the requirements of the actual application scenario.

One way of further development of polling models is the use of more elaborated stochastic processes. Specifically the widespread introduction and application of data networks has played an important role in the development of advanced arrival processes. Among others the modeling of Internet traffic motivated and promoted the development, analysis and application of more powerful arrival process models. The objectives of this generalization are the analytical tractability, modeling the correlation among the interarrival times and capturing the internal structure of the real life traffic. These objectives resulted in the introduction of various types of new arrival processes, whose interarrival time is built up from different compositions of exponential distributions (e.g. the Markov-modulated Poisson process (MMPP) see in [19]). These arrival processes are special cases of the Markovian arrival process (MAP) [18]. A further generalization of MAP to the case of batch arrival is the batch Markovian arrival process (BMAP) introduced by Lucantoni ([16]). Due to the more realistic traffic modeling capability of BMAP, the extension of the arrival process of the classical polling system to batch Markovian arrival process results in a promising model.

In the BMAP/G/1 cyclic polling model at each of the stations batch of customers arrive according to BMAP process. The customer service times at each station are general independent and identically distributed. The switchover times of the consec-

utive cycles are general independent and identically distributed. The *vacation model* is a special case of the cyclic polling model, in which the number of stations is one. In this case the switchover time of the only station of the polling model becomes the *vacation period* of the vacation model.

The analysis of queueing models with BMAP attracted a great attention from the beginning of 1990s. The most of the analysis works on BMAP queueing models including vacation models are based on the standard matrix analytic-method pioneered by Neuts [20] and further extended by many others (see e.g. [17]). It provides the vector probability-generating function (vector GF) of the number of customers at departure or at an arbitrary epoch. The standard matrix analytic-method exploits the underlying M/G/1-type structure of the model, i.e., that the embedded Markov chain at the customer departure epochs is of M/G/1-type ([21]) in which the block size in the transition probability matrix equals to the number of phases of the BMAP. The easy identification of the exceptional boundary states makes the standard matrix analytic-method very suitable for analyzing models with exhaustive discipline ([5]). However this is not the case for a broad class of BMAP/G/1 queuing models ([C2]). As a consequence of it very few work is available in the literature on such BMAP/G/1 vacation models ([1, 10, 23, C2, C3]) and they apply different methods.

Relative few works can be found in the literature on MAP/BMAP queuing models with multiple queues. [22, 28] address the analysis of the nonpreemptive priority queue, which is slightly different from the zero-swithover times polling model with exhaustive service, since that model has common underlying arrival process forming a finite state continuous-time Markov chain (CTMC)

The analysis of polling models with MAP is an area hardly covered by literature so far ([2, 12, 15]). All these works also utilize the Markovian characteristic of the customer service time. I am unaware of any previous work on general service time polling models with MAP. Relative few results have been published on BMAP vacation models with disciplines others than the exhaustive ones and on BMAP priority queueing models. However to the best of my knowledge so far none of the polling models with BMAP has been analyzed.

Summarizing all these so far, although there were need for polling models with BMAP in more precise performance evaluation of modern telecommunication networks, such polling models have not yet been investigated. This motivated us to extend the classical cyclic polling model to BMAP and investigate it.

2   Research objectives

The principal goal of this thesis is to provide a unified analysis of cyclic polling models with BMAP. Only the continuous-time nonzero-switchover-times basic model with stations having infinite buffer queues is addressed. The unified character of the analysis ensures the applicability of the applied methodology and the obtained results for a broad class of service disciplines including the most common gated and

exhaustive policies and at least one more complex discipline, e.g. the G-limited one.

The first objective of the thesis is to establish stability results for this model for a group of service disciplines including the gated, the exhaustive and the G-limited disciplines. In the practical applications the most important stability result is the condition to ensure the stability of the system. Therefore the sufficient condition on the whole stability is targeted.

The major objective of the thesis is to present a service discipline independent analysis of the cyclic polling models with BMAP. Here the term "service discipline independent" means that the analysis results are valid for a broad class of service disciplines. The dependency on the individual disciplines are incorporated by discipline specific quantities. Such a unified analysis considerably simplifies the analysis of the system, since after having the service discipline independent results it is enough to determine only the discipline specific quantities in a second part of the analysis. Similar problem separation principles are the stochastic decomposition [13] or probability-generating function (PGF) factorization forms [6], which have already been successfully applied in the analysis of queueing models ([2, 5]).

Thus this kind of analysis results in a two-step methodology in which the service discipline independent analysis is embedded as first part. A suitable candidate for the methodology is the possible generalization of the one used by Borst and Boxma in [4] as it has already been successfully applied in the analysis of the cyclic polling systems with Poisson arrivals.

The main target of this objective is the set-up of a closed form expression for the vector GF of the stationary number of customers at an arbitrary instant in terms of discipline specific quantities. In the applications usually instead of PGF the mean of the stationary number of customers is the quantity of interest. Hence the second target of this objective is the determination of the vector mean of the stationary number of customers at an arbitrary instant.

The last objective of the thesis is the application of the service discipline independent results to the most common policies including the gated, the exhaustive and the G-limited disciplines. This can be achieved by the determination of the discipline specific quantities in the discipline independent results. Although this can be performed on discipline specific way, the investigation of a possibly unified method is targeted for realizing it.

Even though the analysis and the results can be extended or generalized to other variations and extensions to the basic cyclic polling models with BMAP, it is out of scope of this thesis to deal with them. Waiting time analysis of the considered model is not covered by this thesis. These topics are left for future research. Elaborating applications based on the results presented in the thesis also remains out of scope of the thesis. Instead it is a challenging future work.

3 Research methodology

3.1 Stability

The applied stability methodology is based on the identification of properly chosen embedded Markov chains. The basic idea of this methodology incorporates several elements from the work of Fricker and Jaïbi in [11], like the limited and unlimited type of disciplines and proper embedded Markov chains. However our proposed methodology requires only relaxed conditions on the service disciplines in comparison with the monotonicity property of [11]. The proposed methodology allows the generalization of the arrival process to BMAPs and a much simpler stability analysis than the existing ones based on monotonicity properties and dominance theorems like the [11].

Moreover the applied stability methodology can be applied for a fairly general set of service disciplines and hence it leads to a unified stability analysis.

3.2 The unified analysis approach

A behavior of a station during a server visit can be modeled by a vacation system in which the vacation period corresponds to the intervisit time of the considered station. Thus the BMAP/G/1 cyclic polling models can be seen as the generalization of both the corresponding vacation and priority queueing models. Hence they inherit the analytical difficulties of both models:

BMAP/G/1 vacation models: for these models with disciplines other than exhaustive the proper definition of the system states, which results in an $M/G/1$-type structure, would make the state space very complicated

BMAP priority queueing models: in contrast to the matrix analytic-method, which allows only one random variable having countable infinite space to describe the system dynamics, these models require mutually dependent random variables, each of which is defined in the countable infinite space, i.e. the number of customers in each priority class

To overcome these drawbacks we separate the analysis into two parts based on quantities at polling and departure epochs of the given station treating only one of the above mentioned difficulties in both of them. This results in simplification in the overall analysis. In the first part dealing only with the vacation model part of the problem enables to establish relation for the stationary number of customers at an arbitrary epoch in terms of the stationary number of customers at polling and departure epochs. The closed-form of these factorization results makes the analysis considerably easier. In the second part due to the problem separation the above quantities are to be determined at polling and departure epochs of the stations, which requires the description of the system dynamics only at these polling and departure epochs. This is achieved by the help of relating the joint PGFs of the stationary

number of customers and the phases of the BMAPs at that epochs by means of a unified method. This results in a simpler mathematical structure compared to the possible descriptions at other system epochs (like e.g. at customer departure times) or to the application of other methods (like e.g. the supplementary variable technique).

The results in the first part are valid for a group of disciplines since the dependency on the given discipline is incorporated by the quantities at polling and departure epochs. Hence this part is called as service discipline independent part of the analysis. In the second part the required quantities at server arrival and departure epochs and their determination are discipline specific. This part is called as service discipline dependent part of the analysis.

This two-steps methodology can be seen as a generalization of the one, which has also been used for analyzing cyclic polling models with Poisson arrival by Borst and Boxma in [4]. Besides of incorporating several elements of the matrix analytic-methods the whole methodology can be seen rather as the natural generalization of several methods used for the analysis of the classical cyclic polling model.

Since the service discipline independent part is valid for a broad class of disciplines and the discipline dependent part of the analysis is realized by a unified method, this methodology provides a unified approach for analysis of polling models with BMAP.

4   New results

The contributions of this thesis are grouped as follows. The first group of theses deals with the stability results. The second group of theses is about the service discipline independent analysis. The third group of theses discusses the application of the service discipline independent results to the gated, the exhaustive and the G-limited disciplines.

4.1   Stability of Cyclic Polling Models with BMAP

The customer who arrives to station i is called i-customer. Station i is any of the N stations, i.e. $i = 1, \ldots, N$. Similarly the station time of station i is called i-station time. Furthermore the polling epoch of station i is referred to as i-polling epoch.

Station i of the polling model is said to be stable, when the number of i-customers at i-polling epoch has proper limiting distribution and the limiting cycle time has a finite mean, while the number of cycles goes to infinity.

Note that the proper limiting distribution of the number of i-customers at i-polling epoch enables an infinite mean of the number of i-customers at i-polling epoch. This definition is different from the stability definition of Kuehn [14], since it excludes this case.

The polling model is said to be stable, when the number of customers at i-polling epoch have proper limiting distributions and the limiting cycle time has a finite mean, while the number of cycles goes to infinity. This stability definition of polling models is equivalent with the one of [11].

Unlimited type and *limited type* service disciplines and stations are differentiated. The service discipline at station i is unlimited type when $g_i^\infty = \infty$, where g_i^∞ is the mean number of customers served during the i-station time given that the number of i-customers at i-polling epoch goes to infinity. A station is of unlimited type, if that station has unlimited type service discipline. On the other hand the service discipline at station i is called limited type when $g_i^\infty < \infty$. A station is of limited type, if that station has limited type service discipline.

Thesis 1.1 (*Published in [EJOR]*) *I have shown that there are 3 possible stability states of the BMAP/G/1 cyclic nonzero-switchover-times polling model:*

- *Whole stability: all stations are stable.*

- *Partial stability: 1 or more limited type stations are instable, but the rest of the stations are stable.*

- *Instability: all stations are instable and the limiting mean cycle time is infinite.*

The statement of the Thesis 1.1 is a straightforward consequence of the following properties:

- All unlimited type stations share the same stability state.

- When the unlimited type stations are instable the limited type stations are instable as well, and the limiting mean cycle time is infinite.

- When the unlimited type stations are stable the limited type stations can be both stable and instable.

The proof of these properties are based on the *maximum limit* and *mean limit* properties of the service disciplines used in the model. The maximum limit property means that $g_i^\infty = g_i^{max}$, where g_i^{max} is the maximum of the mean number of customers, which can be served during an i-station time and the maximum is taken over the number of i-customers and the phases of the BMAP at station i at the i-polling epoch. The mean limit property states that if $g_i^\infty = \infty$, and the mean number of i-customers at the i-polling epoch goes to infinity, then the mean number of i-customers served during the i-station time also tends to infinity.

The importance of statement of the Thesis 1.1 lies in its overview character.

In the following the BMAP at station i is called as i-th BMAP. Let λ_i be the stationary arrival rate of the i-th BMAP. Let $\mathbf{D}_{i,\ell}$ denote the $\mathbf{D}_\ell$ matrix of the i-th BMAP, which governs the transitions with batch arrivals, in which the batch size is ℓ ($\ell \geq 0, i \in \{1, \ldots, N\}$). Let the traffic intensity be controlled by applying scaling parameter ξ, such that $\mathbf{D}_{i,\ell}(\xi) = \xi \mathbf{D}_{i,\ell}$, $l \geq 0$. This way $\lambda_i(\xi) = \xi \lambda_i$, and thus the relative ratios of station arrival rates remain fixed. It can be shown that scaling the traffic intensity from 0 to ∞ the stations gets instable in order $i_1, i_2, \ldots, i_N$, where

$$\frac{\lambda_{i_1}}{g_{i_1}^{max}} \geq \frac{\lambda_{i_2}}{g_{i_2}^{max}} \geq \ldots \geq \frac{\lambda_{i_N}}{g_{i_N}^{max}}. \tag{1}$$

Therefore the stations are indexed such that $\frac{\lambda_1}{g_1^{max}} \geq \frac{\lambda_2}{g_2^{max}} \geq \ldots \geq \frac{\lambda_N}{g_N^{max}}$. If there are N^l limited type stations, it follows from the statement of Thesis 1.1, that the first N^l indexes identify the limited type stations.

Thesis 1.2 (*Published in [EJOR]*) *I have proved that the necessary and sufficient condition of the stability of station i in the BMAP/G/1 cyclic nonzero-switchover-times polling model is*

$$g_i < g_i^{max}, \tag{2}$$

where g_i is the mean number of customers served in the i-station time.

Furthermore I have proved that station i of limited type ($i \leq N^l$) is stable if and only if

$$\sum_{k=i}^{N} \rho_k + \frac{\lambda_i}{g_i^{max}} \left(r + \sum_{k=1}^{i-1} g_k^{max} b_k \right) < 1, \tag{3}$$

and station i of unlimited type ($i > N^l$) is stable if and only if

$$\rho^u < 1, \tag{4}$$

where ρ_k and b_k is the utilization of station k and the mean customer service time at station k, respectively. r is the sum of the mean switchover times and ρ^u is the sum of the utilizations of the unlimited stations.

The proof of the statement (2) is based on the structure of the state spaces of properly chosen embedded Markov chains at i-polling epochs. Besides the mean limit and maximum limit service discipline properties the proof also utilizes a third discipline property, which is called *non-zero maximum* discipline property. This property states that the mean number of customers that can be served during the i-station time is greater than zero, $g_i^{max} > 0$.

The non-zero maximum and the maximum limit service discipline properties are similar to the assumptions of the model of Down [8].

An alternative form of the necessary and sufficient condition of the stability of station i is given as

$$a_i < g_i^{max}, \tag{5}$$

where a_i is the mean number of arriving i-customers between two consecutive i-polling epochs.

At the stability boundary $a_i = g_i = g_i^{max}$ holds and thus the station is still in statistical equilibrium. However the necessary and sufficient conditions of the stability of a particular station imply that at this boundary the station is already instable.

Stability	Unlimited type station i	Limited type station i
Stable $g_i < g_i^{max}$	$a_i = g_i$; $a_i < g_i^{max}$	
Instable $g_i = g_i^{max}$	**Stability boundary** $a_i = g_i$; $a_i = g_i^{max}$	
		Above stability boundary $a_i > g_i$; $a_i > g_i^{max}$

Table 1: Stability regions of a particular station

The overview of the stability regions of a particular station is given in Table 1.

The proof of the second and third statement of the Thesis 1.2 utilizes the expression of the mean cycle time in partial stability. If the first N^u limited type stations $(1 \le N^u \le N^l)$ are out of stability and the remaining $N - N^u$ stations are stable, then the mean cycle time (c) is given as

$$c = \frac{r + \sum_{k=1}^{N^u} g_k^{max} b_k}{1 - \sum_{k=N^u+1}^{N} \rho_k},\tag{6}$$

Starting from (6) the statements (3) and (4) can be shown.

Thesis 1.3 (*Published in [EJOR]*) *I have provided the necessary and sufficient conditions for the whole stability of the BMAP/G/1 cyclic nonzero-switchover-times polling system as*

$$\rho + \left(\frac{\lambda_1}{g_1^{max}}\right) r < 1,\tag{7}$$

where ρ is the overall utilization.

The statement of Thesis 1.3 comes from the second statement of the Thesis 1.2 by setting $i = 1$. This result is an extension of the one in the classical cyclic polling model with Poisson arrival process ([11]) to the cyclic polling model with BMAP.

Note that the statement of Thesis 1.1 and the third statement of the Thesis 1.2 implies that the necessary and sufficient conditions for instability of the BMAP/G/1 cyclic nonzero-switchover-times polling system is

$$\rho^u \ge 1.\tag{8}$$

The condition for the whole stability for a given discipline can be established by applying the service discipline specific g_1^{max} to the result on the whole stability (7). Table 2 summarizes the condition for the whole stability for the gated, the exhaustive and the G-limited disciplines. In case of the G-limited discipline K_1 stands for the discipline limit at station 1.

10

Service discipline	$g_1^{\max}$	Stability condition
gated	∞	$\rho < 1$
exhaustive	∞	$\rho < 1$
G-limited	K_1	$\rho + \left(\frac{\lambda_1}{K_1}\right) r < 1$

Table 2: Stability condition for several service disciplines

4.2 Discipline independent results for cyclic polling models with BMAP

The discipline independent results provided in this group of theses are valid for a broad class of service disciplines. This includes among others the most commonly known disciplines, e.g., the exhaustive, the gated, the non-exhaustive, the semi-exhaustive, the G-limited and the decrementing-K ones.

The departure epoch of station i is referred to as i-departure epoch.

Thesis 2.1 (*Published in [Queueing Systems]*) *I have provided the discipline independent expression for the vector GF of the stationary number of i-customers at an arbitrary instant $(\widehat{\mathbf{q}}_i(z))$ in the stable BMAP/G/1 cyclic nonzero-switchover-times polling model as*

$$\widehat{\mathbf{q}}_i(z)\widehat{\mathbf{D}}_i(z)\left(z\mathbf{I} - \widehat{\mathbf{A}}_i(z)\right) = \lambda_i(1 - \rho_i^S)(z-1)\frac{\widehat{\mathbf{f}}_i(z) - \widehat{\mathbf{m}}_i(z)}{f_i^{(1)} - m_i^{(1)}}\widehat{\mathbf{A}}_i(z), \quad |z| \le 1, \qquad (9)$$

where $\widehat{\mathbf{D}}_i(z)$ is the matrix generating function (matrix GF) of the i-th BMAP, $\mathbf{I}$ is the identity matrix, $\widehat{\mathbf{A}}_i(z)$ is the matrix GF of the number of i-th BMAP arrivals during an i-customer service time, $\rho_i^S = \lambda_i^S b_i$ and λ_i^S is the stationary arrival rate of the i-th BMAP during the i-station time.

Furthermore $\widehat{\mathbf{f}}_i(z)$ and $\widehat{\mathbf{m}}_i(z)$ stand for the vector GFs of the stationary number of i-customers at i-polling and i-departure epochs, respectively. $f_i^{(1)}$ and $m_i^{(1)}$ denote the mean of the stationary number of i-customers at i-polling and i-departure epochs, respectively.

The factorization form (9) expresses $(\widehat{\mathbf{q}}_i(z))$ in terms of $\widehat{\mathbf{f}}_i(z)$ and $\widehat{\mathbf{m}}_i(z)$ i.e., $\widehat{\mathbf{q}}_i(z) = \mathcal{F}\left(\widehat{\mathbf{f}}_i(z), \widehat{\mathbf{m}}_i(z)\right)$. The dependency on the applied service discipline in (9) is expressed by $\widehat{\mathbf{f}}_i(z) - \widehat{\mathbf{m}}_i(z)$.

It can be shown that the term $\frac{\left(1-\rho_i^S\right)}{\left(f_i^{(1)}-m_i^{(1)}\right)}$ in (9) can be computed on discipline independent way as $\frac{1-\rho}{\lambda_i r}$.

The proof of (9) relies on a general statement, called as fundamental relationship, which is derived by the help of the generalization of an argument from an early

11

work of Eisenberg [9]. Using it the expression of the vector GF of the stationary number of i-customers at i-customer departure epochs ($\widehat{\mathbf{q}}_i^d(z)$) is given in terms of $\widehat{\mathbf{f}}_i(z)$ and $\widehat{\mathbf{m}}_i(z)$. Applying it in the stationary relationship between the vector GF of the stationary number of i-customers at an arbitrary instant and at i-customer departure epochs provided by Takine and Takahashi in [29] ($\widehat{\mathbf{q}}_i(z) = \mathcal{H}\left(\widehat{\mathbf{q}}_i^d(z)\right)$) yields to the expression of the vector GF of the stationary number of i-customers at an arbitrary instant.

In the derivation of (9) we utilize only the evolution of the system from the i-polling epoch to the next i-departure epoch. During this period the service process is completely independent of the other stations. Therefore in this part we consider only the number of i-customers without representing the state of the other stations in the notation. Thus the statement of the Thesis 2.1 can be also interpreted as the solution of the vacation model corresponding to station i.

Besides several general model assumptions, the proof of the statement of Thesis 2.1 utilizes only the work-conservation and nonpreemptive properties of the service discipline. Moreover this result is also valid for the zero-switchover-times counterpart of the considered model. In this model the server off periods, while the system is empty, can be taken into account e.g. by applying the argument presented in [C1]. Thus the statement of Thesis 2.1 holds under more general settings.

Thesis 2.2 (*Published in [Queueing Systems]*) *I have provided the discipline independent recursive formula for computing the vector factorial moments of the stationary number of i-customers at an arbitrary instant in the stable BMAP/G/1 cyclic nonzero-switchover-times polling model.*

By applying the formula of Thesis 2.2 the vector factorial moments of the stationary number of i-customers ($\mathbf{q}_i^{(n)}$) can be computed recursively for $k = 1, \ldots, n$. In the recursive formula $\mathbf{q}_i^{(n)}$ depends on the vector factorial moments of the stationary number of i-customers and the phase probability vectors of the i-BMAP at i-polling epochs ($\mathbf{f}_i^{(k)}$) and i-departure epochs ($\mathbf{m}_i^{(k)}$) for $0 \leq k \leq n + 1$, which are all service discipline specific quantities.

The formula for computing $\mathbf{q}_i^{(n)}$, $n \geq 1$ can be derived from the expression of Thesis 2.1. For this derivation newly established properties of model specific key matrices ($\widehat{\mathbf{D}}_i(z)$ and $\left(z\mathbf{I} - \widehat{\mathbf{A}}_i(z)\right)$) are also utilized.

4.3 Application of discipline independent results to BMAP/G/1 cyclic polling models with specific disciplines

In this group of theses the determination of the vector factorial moments of the stationary number of i-customers and the phase probability vectors of the i-th BMAP at i-polling and i-departure epochs ($\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$ for $0 \leq k$) are discussed for polling models with specific disciplines. These are the quantities, which are needed for the

application of the discipline independent result (the recursive formula of Thesis 2.2) to these models.

Thesis 3.1 (*Published in [Queueing Systems]*) *I have established the governing equations for the stable BMAP/G/1 cyclic nonzero-switchover-times polling model with gated discipline. I also developed a numerical procedure to compute the required vector factorial moments and phase probability vectors at i-polling and i-departure epochs ($\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$ for $0 \leq k$). The governing equations are given as*

$$\widehat{\mathbf{m}}_i(z_1,\ldots,z_N) = \widehat{\mathbf{f}}_i\left(z_1,\ldots,z_{i-1},\widehat{\mathbf{A}}_i(z_1,\ldots,z_N),z_{i+1},\ldots,z_N\right),$$

$$\widehat{\mathbf{f}}_{i+1}(z_1,\ldots,z_N) = \widehat{\mathbf{m}}_i(z_1,\ldots,z_N)\,\widehat{\mathbf{U}}_i(z_1,\ldots,z_N),\ \ |z_1| \leq 1,\ldots,|z_N| \leq 1, \qquad (10)$$

where $\widehat{\mathbf{f}}_i(z_1,\ldots,z_N)$ and $\widehat{\mathbf{m}}_i(z_1,\ldots,z_N)$ are the joint PGFs of the stationary number of customers at every stations and the phases of every BMAPs at i-polling and i-departure epoch, respectively. $\widehat{\mathbf{A}}_i(z_1,\ldots,z_N)$ and $\widehat{\mathbf{U}}_i(z_1,\ldots,z_N)$ stand for the the hypermatrix GFs of the number of simultaneously arriving k-customers for $k = 1,\ldots,N$ during the service of one i-customer and during the switchover time after the service of station i, respectively. Furthermore $\widehat{\mathbf{f}}_i\left(z_1,\ldots,z_{i-1},\widehat{\mathbf{A}}_i(z_1,\ldots,z_N),z_{i+1},\ldots,z_N\right)$ denotes a substitution of hypermatrix $\widehat{\mathbf{A}}_i(z_1,\ldots,z_N)$ into the defining series of the hypervector GF $\widehat{\mathbf{f}}_i(z_1,\ldots,z_N)$.

The quantities $\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$ (for $0 \leq k$) are to be determined for the specific service discipline at i-polling and i-departure epochs. This requires the description of the system dynamics by mutually dependent discrete random variables only at i-polling and i-departure epochs. The governing equations of the system are setup at these epochs in terms of the joint PGFs of the stationary number of customers at every stations and the phases of every BMAPs at i-polling epoch ($\widehat{\mathbf{f}}_i(z_1,\ldots,z_N)$) and i-departure epochs ($\widehat{\mathbf{m}}_i(z_1,\ldots,z_N)$). This results in relations for transition $f_i \rightarrow m_i$ (the first equation of (10)) and $m_i \rightarrow f_{i+1}$ (the second equation of (10)).

The fundamental idea of the derivation of the governing equations is the generalization of the buffer occupancy method ([7]), which was also used in the analysis of the classical cyclic polling model.

Different to the classical polling model with the gated disciplines no closed-form system of linear equations can be derived for the factorial moments of the stationary number of i-customers at i-polling and i-departure epochs for this model. This is due to the complexity introduced by BMAP. Instead a system of linear equations are derived for the joint probabilities of the stationary number of customers and the phases of the BMAPs at i-polling and i-departure epochs by taking the appropriate derivatives of the governing equations of the system. This is the base for the numerical solution.

The main steps of the numerical procedure to compute $\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$, for $0 \leq k$, are given as

- Computation of matrices $\widehat{\mathbf{A}}_i(z_1, \ldots, z_N)$ for every $i = 1, \ldots, N$.

- Building up one large system of linear equations consisting of hypermatrices.

- Solving this system of linear equations.

- Computation of $\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$, for $0 \leq k$, from the joint probabilities.

The most computational intensive parts of the procedure are the computation of matrices $\widehat{\mathbf{A}}_i(z_1, \ldots, z_N)$ and building up the large system of linear equations consisting of hypermatrices.

The total number of operations required by the whole computational procedure is in the magnitude of $N^2 \, L^{3N}(X+1)^{3N}$, where L is the number of phases of the BMAPs and X is the highest number of i-customers taken into account at i-departure epoch.

Thesis 3.2 (*Published in [Queueing Systems]*) *I have established the governing equations for the stable BMAP/G/1 cyclic nonzero-switchover-times polling model with exhaustive discipline. I also developed a numerical procedure to compute the required vector factorial moments and phase probability vectors at i-polling and i-departure epochs ($\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$ for $0 \leq k$). The governing equations are given as*

$$
\begin{aligned}
&\widehat{\mathbf{m}}_i \left(z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_N \right) \\
&= \widehat{\mathbf{f}}_i \left(z_1, \ldots, z_{i-1}, \widehat{\mathbf{H}}_i(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_N), z_{i+1}, \ldots, z_N \right), \\
&\widehat{\mathbf{f}}_{i+1} \left(z_1, \ldots, z_N \right) \\
&= \widehat{\mathbf{m}}_i \left(z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_N \right) \widehat{\mathbf{U}}_i(z_1, \ldots, z_N), \quad |z_1| \leq 1, \ldots, |z_N| \leq 1, \quad (11)
\end{aligned}
$$

where $\widehat{\mathbf{H}}_i(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_N)$ denotes the hypermatrix GFs of the number of simultaneously arriving k-customers for $k = 1, \ldots, i-1, i+1, \ldots N$ during the decrement of the number of i-customers by one.

The transition $m_i \to f_{i+1}$ is discipline independent. However in case of exhaustive discipline the number of i-customers is 0 at i-departure epochs. Applying it to the second part of (10) results in the second relation of (11).

To describe the transition $f_i \to m_i$ we apply a similar line of arguments as we used for the model with gated discipline (Thesis 3.1). Due to the exhaustive discipline the evolution of the i-customers during the i-station time is exceptional compared to the evolution of the other k-customers ($k \neq i$). This requires the introduction of matrix $\mathbf{G}_i(t)$, $t \geq 0$, which is the phase transition probability matrix during the decrement of the number of i-customers by one when it occurs no later than time t. Then matrix $\mathbf{G}_i(t)$ is used for the definition of hypermatrix GF $\widehat{\mathbf{H}}_i(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_N)$.

The system of linear equations for the joint probabilities of the stationary number of customers and the phases of the BMAPs at i-polling and i-departure epochs, which

is the base of the numerical solution, is derived by taking the appropriate derivatives of the governing equations of the system.

The main steps of the numerical procedure to compute $\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$, for $0 \leq k$, are given as

- Computation of matrices $\widetilde{\mathbf{G}}_i(s)$ (for $Re(s) \geq 0$) for every $i = 1, \ldots, N$. They are calculated by means of the recursive algorithm. This is a generalization of an algorithm provided for computing matrices $\widetilde{\mathbf{G}}_i(0)$, which is based on the application of the concept of uniformization, see in [16].

- Computation of matrices $\widehat{\mathbf{H}}_i(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_N)$ for every $i = 1, \ldots, N$.

- Building up one large system of linear equations consisting of hypermatrices. This includes exhaustive discipline specific handling due to 0 elements in the matrices. This is because the number of i-customers is 0 at i-departure epochs.

- Solving this system of linear equations.

- Computation of $\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$, for $0 \leq k$, from the joint probabilities.

The most computational intensive parts of the procedure are the computation of matrices $\widehat{\mathbf{H}}_i(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_N)$ and building up the large system of linear equations consisting of hypermatrices.

The total number of operations required by the whole computational procedure is in the magnitude of $N^2 \, L^{3N}(X+1)^{3N-3}$.

Thesis 3.3 *I have established the governing equations for the stable BMAP/G/1 cyclic nonzero-switchover-times polling model with G-limited discipline. I also developed a numerical procedure to compute the required vector factorial moments and phase probability vectors at i-polling and i-departure epochs ($\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$ for $0 \leq k$).*

Due to the complexity of the G-limited discipline the governing equations are expressed in terms of joint probabilities of the stationary number of customers and the phases of the BMAPs at i-polling and i-departure epochs instead of the corresponding joint PGFs ($\widehat{\mathbf{f}}_i(z_1, \ldots, z_N)$ and $\widehat{\mathbf{m}}_i(z_1, \ldots, z_N)$).

The same line of argument is used for the derivation of the governing equations as in case of the model with gated discipline (Thesis 3.1).

The steps of the numerical procedure to compute $\mathbf{f}_i^{(k)}$ and $\mathbf{m}_i^{(k)}$, for $0 \leq k$, are also the same as the ones described for the model with gated discipline (Thesis 3.1). Therefore the total number of operations required by the whole numerical procedure is in the magnitude of $N^2 \, L^{3N}(X+1)^{3N}$.

5 Application of results

The major application area of the presented results on cyclic polling model with BMAP is the field of modern telecommunication networks, in which so far e.g. the classical polling models have been also applied (see [25, C4, C5, C6]). This is motivated by achieving more precise queueing model and more accurate results in performance evaluation. Such potential application examples are performance modeling to IEEE 802.11 and IEEE 802.16 systems or analysis of power saving mechanisms.

There are at least two clear ways of exploiting the advanced traffic modeling capabilities of BMAP in the practical applications of the results presented in this thesis. The first one is the fitting of BMAP to correlated traffic models in order to apply them in the performance analysis of the system under consideration. Such correlated traffic models have been elaborated for various data, voice and video traffic types e.g. for simulation based performance analysis of the IEEE 802.16 [24], [30]. In the second case the BMAP is constructed from given traffic parameters to allow traffic characteristics dependent performance evaluation of the studied system.

An important current limitation on the application of the results of the thesis is that both the above-mentioned fitting task and the problem of constructing an arrival process from given traffic characteristics are actually solved only for special classes of BMAPs, e.g. for two-phase MAPs (see in [3]).

An example for the investigation of the effect of traffic parameters in the performance evaluation of the considered system is the application of BMAP vacation model to IEEE 802.16e sleep mode mechanism ([C7, JIMO]). In the work [JIMO] optimization examples are also presented for determining the optimal sleep mode parameters under different constraints. This includes enforcing an upper bound on mean delay and a cost model taking into account more general Quality of Service (QoS) requirements. These optimizations facilitate the tuning of the sleep mode parameters to the requirements of the actual application scenario and thus they can be applied in network control.

6 Acknowledgement

References

[1] A. D. Banik, U. C. Gupta, and S. S. Pathak. BMAP/G/1/N queue with vacations and limited service discipline. *Applied Mathematics and Computation*, 180:707–721, 2006.

[2] D. Bertsimas and G. Mourtzinou. Decomposition results for general polling systems and their applications. *Queueing Systems*, 31:295–316, 1999.

[3] L. Bodrog, A. Heindl, G. Horváth, and M. Telek. A Markovian Canonical Form of Second-Order Matrix-Exponential Processes. *European Journal of Operational Research*, 190(2):459–477, 2008.

[4] S. C. Borst and O. J. Boxma. Polling models with and without switchover times. *Operations Research*, 45:536–543, 1997.

[5] S. H. Chang and T. Takine. Factorization and stochastic decomposition properties in bulk queues with generalized vacations. *Queueing Systems*, 50:165–183, 2005.

[6] S. H. Chang, T. Takine, K. C. Chae, and H. W. Lee. A unified queue length formula for BMAP/G/1 queue with generalized vacations. *Stochastic Models*, 18: 369–386, 2002.

[7] R. B. Cooper. Queues served in cyclic order: waiting times. *The Bell System Technical Journal*, 49:399–413, 1970.

[8] G. Down. On the stability of polling models with multiple servers, Tech. Report BS-R9605. Technical report, CWI, 1996.

[9] M. Eisenberg. Queues with Periodic Service and Changeover Time. *Operations Research*, 20:440–451, 1972.

[10] J. M. Ferrandiz. The BMAP/G/1 queue with server set-up times and server vacations. *Adv. Appl. Prob.*, 25:235–254, 1993.

[11] C. Fricker and M. R. Jaïbi. Monotonicity and stability of periodic polling models. *Queueing Systems*, 15:211–238, 1994.

[12] I. Frigui and A.-S. Alfa. Analysis of a time-limited polling system. *Computer Communications*, 21:558–571, 1998.

[13] S. W. Fuhrmann and R. B. Cooper. Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations. *Operations Research*, 33:1117–1129, 1985.

[14] P. J. Kuehn. Multiqueue Systems with Nonexhaustive Cyclic Service. *The Bell System Technical Journal*, 58:671–698, 1979.

[15] T. Y. S. Lee. Models for Design and Control of Single Server Polling Computer and Communication Systems. *Operations Research*, 46(4):515–531, jul-aug 1998.

[16] D. L. Lucantoni. New results on the single server queue with a batch markovian arrival process. *Stochastic Models*, 7:1–46, 1991.

[17] D. L. Lucantoni. The BMAP/G/1 queue: A tutorial. In L. Donatiello and R. Nelson, editors, *Models and Techniques for Performance Evaluation of Computer and Communications Systems*, pages 330–358. Springer Verlag, 1993.

[18] D. M. Lucantoni, K. S. Meier-Hellstern, and M. Neuts. A single server queue with server vacations and a class of non-renewal arrival processes. *Adv. Appl. Prob.*, sep 1989.

[19] K. Meier-Hellstern and W. Fischer. The Markov-modulated Poisson process (MMPP) cookbook. *Performance Evaluation*, 18:149–171, 1992.

[20] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The John Hopkins University Press, Baltimore, 1981.

[21] M. F. Neuts. *Structured stochastic matrices of M/G/1 type and their applications*. Marcel Dekker, New York, 1989.

[22] S. Nishimura. A spectral method for a nonpreemtive priority BMAP/G/1/N queue. *Stochastic Models*, 21:579–597, 2005.

[23] Y. W. Shin and C. E. M. Pearce. The BMAP/G/1 vacation queue with queue-length dependent vacation schedule. *J. Austral. Math. Soc.*, Ser. B 40:207–221, 1998.

[24] Standard IEEE 802.16.3c-01/30r1. Traffic Model for 802.16 TG3 MAC/PHY Simulations, mar 2001.

[25] H. Takagi. Analysis and Application of Polling Models. In *Lecture Notes In Computer Science: Performance Evaluation: Origins and Directions*, volume 1769, pages 423–442, 2000.

[26] H. Takagi. *Analysis of Polling Systems*. MIT Press, 1986.

[27] H. Takagi. Queueing analysis of polling models. *ACM Comput. Surveys*, 20:5–28, 1988.

[28] T. Takine. The nonpreemtive priority MAP/G/1 queue. *Operations Research*, 47:917–927, 1999.

[29] T. Takine and Y. Takahashi. On the relationship between queue lengths at a random instant and at a departure in the stationary queue with BMAP arrivals. *Stochastic Models*, 14:601–610, 1998.

[30] WiMAX Forum. WiMAX System Evaluation Methodology V2.0, dec 2007.

Publications

Journal papers

[EJOR] Zs. Saffer and M. Telek. Stability of periodic polling system with BMAP arrivals. *European Journal of Operational Research*, 197(1):188–195, 2009.

[Queueing Systems] Zs. Saffer and M. Telek. Unified analysis of BMAP/G/1 cyclic polling models. *Queueing Systems*, 64(1):69–102, 2010.

[JIMO] Zs. Saffer and M. Telek. Analysis of BMAP vacation queue and its application to IEEE 802.16e sleep mode. *Journal of Industrial and Management Optimization (JIMO)*, 6(3):661–690, aug 2010.

[Publ. Math. Debrecen] Zs. Saffer and M. Telek. Closed form results for BMAP/G/1 vacation model with binomial type disciplines. *Publ. Math. Debrecen*, 76(3):359–378, 2010.

[HÍRADÁSTECHNIKA] Saffer Zsolt. Digitális távközlő rendszerek hibaarány számítási módszereinek összehasonlítása. *HÍRADÁSTECHNIKA*, 4222–26, 1991.

Conference and workshop papers

[C1] Zs. Saffer. An introduction to classical cyclic polling model. In *Proc. of the 14th Int. Conf. on Analytical and Stochastic Modelling Techniques and Applications (ASMTA'07)*, pages 59–64, 2007.

[C2] Zs. Saffer and M. Telek. Analysis of BMAP/G/1 vacation model of non-M/G/1-type. In N. Thomas and C. Juiz, editors, *EPEW 2008. LNCS*, volume 5261, pages 212–226. Springer Verlag, Sept. 2008.

[C3] Zs. Saffer and M. Telek. Analysis of BMAP/GI/1 phase dependent vacation model with decrementing-K discipline. In *Queues: flows, systems, networks: Proceedings of the International Conference "Mathematical Methods for Analysis and Optimization of Information Telecommunication Networks" Minsk*, pages 193–198, jan 2009.

[C4] Zs. Saffer and S. Andreev. Delay analysis of IEEE 802.16 wireless metropolitan area network. In *15th International Conference on Telecommunications (ICT 2008 and MACOM 2008 Workshop Proceedings), Saint-Petersburg, Russia*, jun 2008.

[C5] S. Andreev, Zs. Saffer, and A. Anisimov. Overall Delay Analysis of IEEE 802.16 Network. In *IEEE International Conference on Communications (IEEE ICC-2009 and MACOM-2009 Workshop Proceedings, Dresden, Germany)*, June 2009.

[C6] S. Andreev, Zs. Saffer, A. Turlikov, and A. Vinel. Overall Delay in IEEE 802.16 with Contention-Based Random Access. In K. Al-Begain, D. Fiems, and G. Horváth, editors, *16th International Conference, ASMTA 2009. LNCS*, volume 5513, pages 89–102. Springer Verlag, June 2009.

[C7] Zs. Saffer and M. Telek. Waiting time analysis of BMAP vacation queue and its application to IEEE 802.16e sleep mode. In *4th International Conference on Queueing Theory and Network Applications (QTNA 2009). Singapore*, jul 2009.

[C8] S. Andreev, Zs. Saffer, A. Turlikov, and A. Vinel. Upper Bound on Overall Delay in Wireless Broadband Networks with Non Real-Time Traffic. In K. Al-Begain, D. Fiems, and W. Knottenbelt, editors, *to appear in 17th International Conference, ASMTA 2010. LNCS*. Springer Verlag, June 2010.

[C9] Zs. Saffer, S. Andreev, and Y. Koucheryavy. Modeling the influence of the Real-Time Traffic on the Delay of the Non Real-Time Traffic in IEEE 802.16 Network. *to appear in MACOM-2010 Workshop Proceedings, Barcelona, Spain.*, Sept. 2010.