Predicting the Time to Migrate Into Deadlock Using a Discrete Time Markov Chain

André B. Bondi
Software Performance and Scalability Consulting, LLC
Red Bank, New Jersery
andrebbondi@gmail.com

ABSTRACT

When processes join a common FCFS queue to acquire or release resources in an object pool of fixed size, deadlock occurs if the process at the head of the queue wishes to acquire a resource when the pool is empty, even if a process wishing to relinquish a resource is queued behind. We describe a state machine representation of this problem. We use the representation to develop a discrete time Markov chain analysis to identify the load conditions under which deadlock is most likely to occur and how soon it is likely to occur. We show that deadlock occurs almost surely regardless of the load, and that the time to the onset of deadlock depends on combinations of the request rate for resources in the pool, the average holding time of the resources, and the size of the pool. Calculations corroborate the intuition that deadlock will occur sooner at heavy loads or when the resource pool is small. A connection will be made between this problem and the problem of random walks with a single absorbing and a single reflecting barrier.

CCS CONCEPTS

• **Software and its engineering → Deadlocks**; **Software performance**; • **Theory of computation** → *Random walks and Markov chains*;

KEYWORDS

Deadlock prediction and prevention; Markov chain analysis; random walks

ACM Reference Format:
André B. Bondi. 2018. Predicting the Time to Migrate Into Deadlock Using a Discrete Time Markov Chain. In *ICPE '18: ACM/SPEC International Conference on Performance Engineering Companion , April 9–13, 2018, Berlin, Germany*. ACM, New York, NY, USA, Article 4, 6 pages. https://doi.org/10.1145/3185768.3186403

1 INTRODUCTION

Identifying the cause of deadlock in a computer system and the possible circumstances and events leading to its onset can be complicated because deadlock usually occurs in a non-deterministic manner. The conditions and events leading to deadlock may be hard to reproduce, even though the load conditions under which it is likely to occur might be readily identifiable in some instances. Our aim is to quantify the time from the onset of a particular load condition to the onset of deadlock, because of the insight this can give us into system reliability. This is difficult to do in general, because the numbers of way in which deadlock can occur is vast, and because the way in which deadlock occurs is sometimes obscure and complex.

In some cases, it is possible to describe sequences of events that lead the system into deadlock with the aid of a finite state machine. The states could correspond to the states of resources under contention, while the alphabet might correspond to events that trigger transitions. The probability that the next symbol in the alphabet takes one value or another is equivalent to the probability of a transition into state k given that the system is currently in state j. The transition probabilities define the mapping from a finite state machine to a discrete time Markov chain with the same states and topology.

Deadlock states are inherently trap states in the context of finite state machines and absorbing states in the context of discrete time Markov chains. Hence, our approach to predicting the likelihood and the time to the onset of deadlock consists of formulating a FSM description of how deadlock arises, mapping the FSM description into a discrete time Markov chain, and then computing the transition probabilities and first passage times of the Markov chain.

Here, we examine the onset of deadlock in a system in which the elements of a resource pool are acquired and released singly by processes that pass through a common FCFS queue to do one and then the other. We have seen this performance antipattern in a number of computer systems that we are not allowed to name. We call it the Museum Checkroom with FCFS queueing, because visitors to the Metropolitan Museum in New York join the same queues to leave and claim their coats in the checkrooms. In [6] and in [8], we showed that deadlock could be avoided by giving priority to those tasks that wish to return a resource to the pool. In the case of the Museum Checkroom, those wishing to claim their coats, thus freeing hangers, would be given priority over those wishing to leave coats, thus occupying hangers. There are separate queues for those claiming and leaving their coats at the Louvre in Paris.

Modeling the system by tracking the state of the checkroom queue involves keeping track of the history of arrivals of visitors leaving and claiming their coats. That makes the problem intractable. Instead, we develop a deterministic finite state machine model (FSM) and a discrete time Markov chain model with the same topology. They describe the changes in the occupancy of the hangers based on the order in which visitors are admitted. When

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ICPE '18, April 9–13, 2018, Berlin, Germany
© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.
ACM ISBN 978-1-4503-5629-9/18/04…$15.00
https://doi.org/10.1145/3185768.3186403

the admission sequence reflects FCFS queueing, deadlock can occur. When priority is given to those wishing to claim their coats, a sequence leading to a transition into the deadlock state cannot occur. The state of the system is entirely described by the number of occupied hangers. For the FSM, the alphabet of events is $\{L, C\}$, where L denotes leaving a coat and C denotes claiming it back.

Our simple model of an object pool whose state is solely the number of occupied resources has the following benefits: (1) it enables us to predict the time for deadlock to ensue under various load conditions by computing expected first passage times into the absorbing state; (2) it obviates the need to consider the possible states of the external queue or queues of jobs waiting to acquire and release the resources in the pool; (3) the predictions of the model can be used to specify load test scenarios in which deadlock could be provoked if its possibility is suspected.

Since our model tells us that deadlock occurs with probability one if one waits long enough, it cannot be used to predict the probability of deadlock occurring; only the expected amount of time until deadlock occurs. This value could be used in an analysis of the mean time to failure of the system in specified operating ranges.

The present work raises a question: is it possible to develop a generalized method for building state machine representations of the progression of an arbitrary system into deadlock, or at least a principle or technique for doing so? We suspect that this method could be applied to other deadlock problems, but that there is no generalized method. We conjecture that such a method could be reduced to the Halting Problem. We believe that it would be possible to derive specific FSM representations and hence discrete time Markov chain models for other systems in which one can describe sequences of inputs and actions leading to deadlock, as is the case here. We note in passing that other authors have approached the problem of predicting the time to the onset of deadlock using stochatic Petri nets [21],[18].

The remainder of this paper is organized as follows. We first describe our FSM representation of the problem, and map it to a discrete time Markov chain. Following an analysis of the algebra of the Markov chain representation, we present some numerical results. This is followed by a discussion of related work, conclusions, and directions for future work.

2 STATE MACHINE FORMULATION

Suppose there are N hangers in the checkroom. The set of checkroom states is the number of occupied hangers, together with the deadlocked state occurring when all hangers are occupied and the when the first customer in the queue wishes to leave a coat. Initially, at opening time, there are no coats in the checkroom, so state 0 is the starting state. We tag arrivals at the checkroom who wish to leave their coats with an L and those who are claiming their coats with a C. Since a coat cannot be claimed unless it has been left, we cannot have more Cs in any finite sequence of arrivals at the checkroom than there are Ls. Moreover, the number of potential coat claimants cannot exceed the number of hangers N.

Figure 1 shows that state transition diagram of the number of occupied hangers in the checkroom, and the deadlock state D. By inspection, the number of sequences that could lead to deadlock

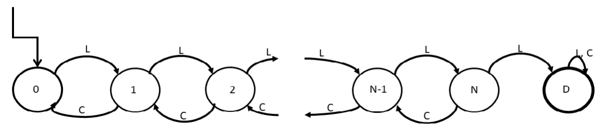

Figure 1: Finite state machine representation

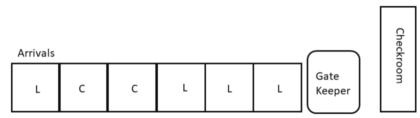

Figure 2: Queue of visitors with gate keeper

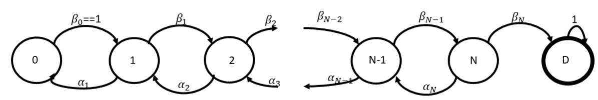

Figure 3: Discrete time Markov chain.

is infinite. This can be shown formally using Arden's Lemma [13]. To prevent deadlock, we must ensure that an arriving C is served ahead of any waiting Ls. We do this by installing a transducer or gatekeeper between the arrival stream and the checkroom that forces Ls to queue outside when all hangers are occupied, even if no Cs are present, and admits the Cs ahead of waiting Ls when they appear. In other words, we give the Cs priority over the Ls so that no admission sequence can occur that would be recognized by the FSM in Figure 1. The configuration is shown in Figure 2.

3 FORMULATION AS A DISCRETE TIME MARKOV CHAIN

Suppose there are N hangers in the checkroom. The set of checkroom states is the number of occupied hangers together with the deadlocked state $D = N + 1$. Initially, at opening time, there are no coats in the checkroom. Figure 3 shows the corresponding discrete time Markov chain in which deadlock occurs. When $N = 1$, this state transition graph has a topology similar but not quite identical to that of the spider trap in [12]. We ignore the order in which visitors wishing to leave or claim coats are queued in the system, and do not consider priorities at this stage. Visitors are assumed to arrive at the checkroom according to a Poisson process with rate λ. For ease of modeling, we assume that the time between leaving a coat and being admitted to the checkroom to claim it is exponentially distributed with rate μ or, equivalently, with mean $1/\mu$. The rate at which visitors return to the checkroom to claim their coats is $n\mu$ when n hangers are occupied, for $n = 0, 1, 2, \ldots, N$. The rate is slow when visitors who have left their coats spend more time in the museum, or more time queueing to claim their coats. Under FCFS queueing, $1/\mu$ would be increased for those claiming their coats, because they would have to queue behind some visitors leaving their coats. The hanger holding time would also be increased. This formulation reflects the notion that FCFS queueing causes self-expansion of the patrons' waiting times, thus undermining load scalability [6]. The number of occupied hangers can only decrease if a visitor arrives to claim a coat and is admitted to do so. It will only increase if a visitor arrives and is admitted to leave a coat. Since deadlock is possible, the Markov chain in Figure 3 is not irreducible and therefore cannot be ergodic. This is because it consists of two classes of states, $D' = \{0, 1, 2, \ldots, N\}$ and D. These two classes do not communicate, because sequences of transitions from D' to D are possible, but not from D to D'. The states in D' are transient, while D is absorbing. Transition into deadlock depends on the sequence of admissions of L and C visitors to the checkroom. We model the state of the hangers as a discrete time Markov chain in which a transition from state k to state $k+1$ occurs if the first admitted visitor after a transition into state k wishes to leave a coat. Similarly, a transition from state $k+1$ to k only occurs if the first visitor to arrive after a transition into state k wishes to claim a coat arrives and is admitted to the checkroom to do so. Let α_k denote the probability that a visitor returns and is admitted to claim a coat before another visitor arrives to leave a coat after the system has moved into state k. Let β_k be the probability of a visitor arriving and being admitted to leave a coat before one arrives to claim a coat. We have $\alpha_k + \beta_k = 1$. For simplicity, we assume that the time each coat spends on a hanger is exponentially distributed with rate μ. Now, if there are two Poisson processes of events of types A and B with rates a and b respectively, the probability that an event of type A occurs before an event of type B is $a/(a+b)$ [10]. Hence, since the rate at which visitors return to claim coats is $k\mu$ when k hangers are occupied, we have

$$\alpha_k = \frac{k\mu}{\lambda + k\mu} \quad , \quad k = 1, 2, ..., N \tag{1}$$

$$\beta_k = \frac{\lambda}{\lambda + k\mu} \quad , \quad k = 1, 2, ..., N \tag{2}$$

We have $\alpha_D = 0$ since deadlock has ensued. Recall that deadlock ensues if the first visitor to enter the checkroom after all hangers have been filled seeks to leave a coat. Notice also that $\beta_0 = 1$ because no returning visitor can arrive to claim a coat before the next new visitor if no coats are in the checkroom.

4 CHAPMAN KOLMOGOROV EQUATIONS

Let $\pi = [\pi_0, \pi_1, \ldots, \pi_N, \pi_D]$ denote the left eigenvector of the state transition probability matrix $\mathbf{P}$. Suppose that the admission policy is FCFS. It can be shown that $\pi = \pi\mathbf{P}$, where $\pi_D + \sum_{k=0}^{N} \pi_k = 1$ and $\mathbf{P}$ is the transition probability matrix given by

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & \ldots & \ldots & \ldots & 0 & 0 \\ \alpha_1 & 0 & \beta_1 & 0 & \ldots & \ldots & \ldots & 0 & 0 \\ 0 & \alpha_2 & 0 & \beta_2 & \ldots & \ldots & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \alpha_N - 1 & 0 & \beta_N - 1 & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & 0 & \alpha_N & \beta_N \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

Since $\mathbf{P}$ is stochastic, its dominant eigenvalue is 1. Since the chain is reducible, it is not ergodic and has either an infinity of solutions or none. By inspection, we see that the equation $\pi = \pi\mathbf{P}$ has a solution for which the entries sum to one,

$$\pi = [0, 0, \ldots, 0, 1] \tag{4}$$

This shows that the system eventually goes into the absorbing deadlock state. This value of π does not indicate the proportions of time that different numbers of hangers are occupied before deadlock occurs. We will see that it is consistent with the theorem that a random walk between a reflecting and an absorbing barrier crosses the barrier with probability one.

5 FIRST PASSAGE TIME INTO DEADLOCK

Recall [4] that if a Markov chain is reducible, its transition probability matrix $mathbf{P}$ can be partitioned into blocks of the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{5}$$

where $\mathbf{Q}$ is an $(N+1) \times (N+1)$ matrix of transition probabilities between the communicating transient states $0, 1, 2, \ldots, N$; $\mathbf{I}$ is the identity matrix, in this case of dimension 1, describing the transition probability from deadlock to itself every time a new visitor arrives to leave a coat, or every time a returning visitor arrives to claim one; block $\mathbf{0}$ denotes a row vector of zeros of dimension $N+1$; and $\mathbf{R}$ is an $(N+1)$-vector of transition probabilities from the transient states to the single absorbing deadlock state, $D = N + 1$. The transition probability into D from state N is β_N, and zero from states $0, 1, 2, \ldots, N-1$. Hence,

$$\mathbf{R} = [0, 0, \ldots, 0, \beta_N]^T \tag{6}$$

Let

$$\mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1} \tag{7}$$

The vector of mean first passage times from the each of transient states to the absorbing deadlock state is given by

$$\mathbf{t} = \mathbf{M}\mathbf{1} \tag{8}$$

where $\mathbf{1}$ is an $(N+1)$-vector of 1s. Finally, the absorption probability for state $D = N + 1$ given that the system started in state i is the (i, D)th entry in the matrix

$$\mathbf{B} = \mathbf{M}\mathbf{R} \quad . \tag{9}$$

We have already seen that $\mathbf{R}$ is a column vector with length $(N+1)$, while $\mathbf{M}$ is an $(N+1) \times (N+1)$ matrix, so $\mathbf{B}$ is also a column vector with length $(N+1)$. For our problem,

$$(\mathbf{I} - \mathbf{Q}) = \begin{bmatrix} 1 & -1 & 0 & 0 & \ldots & \ldots & \ldots & 0 \\ -\alpha_1 & 1 & -\beta_1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & -\alpha_2 & 1 & -\beta_2 & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & -\beta_{N-2} & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & -\alpha_{N-1} & 1 & -\beta_{N-1} \\ 0 & \ldots & \ldots & \ldots & 0 & 0 & -\alpha_N & 1 \end{bmatrix} \tag{10}$$

6 SYSTEM AND NUMERICAL PROPERTIES

6.1 I – Q is Ill Conditioned at Light Loads

$\mathbf{I} - \mathbf{Q}$ is ill conditioned when $\lambda << N\mu$. In the limit as the arrival rate λ tends to zero, $\alpha \rightarrow 1$ and $\beta \rightarrow 0$. Thus, in the limit, the first and second rows of $\mathbf{I} - \mathbf{Q}$ sum to zero. $\mathbf{I} - \mathbf{Q}$ is nearly singular and has a determinant with an absolute value close to zero in the limit. Therefore, it is ill conditioned when λ is small compared with $N\mu$. Now, a formula for the inverse of a matrix can be expressed as a matrix of cofactors divided by the determinant. It follows that the physical interpretation is that when the arrival rate is small compared with $N\mu$, the time that may be expected to elapse before the system goes into deadlock is very large. Our calculations illustrate that this is the case, provided that the corresponding system of linear equations is stable enough to be solved.

6.2 Deadlock Occurs Almost Surely

Our next result implies that deadlock is inevitable, whether or not it takes a long time to occur. This is because the equation for the vector of absorption probabilities is given by

$$\mathbf{B} = \mathbf{MR} \tag{11}$$

with $\mathbf{R} = [0, 0, \ldots, 0, \beta_N]^T$. It can be shown that

$$\mathbf{B} = [1, 1, \ldots, 1]^T. \tag{12}$$

This shows that deadlock is reached with probability one from any state in the system. Interestingly, this result only depends on the band structure of the transition probability matrix $\mathbf{P}$ and on that of $\mathbf{I} - \mathbf{Q}$. It is independent of the forms of α_n and β_n, so long as they both lie in $[0, 1]$ and so long as $\alpha_n + \beta_n = 1 \forall n$. When α_n and β_n are constant and independent of n, our problem is equivalent to a random walk with one absorbing barrier at state $N + 1 = D$ and a reflecting barrier at state 0. The theory of random walks tells us that the absorbing barrier is crossed eventually with probability one, but that the number of steps in the walk before the barrier is crossed depends on the drift [3] [22]. Our results show that this holds for more general forms of the probabilities of moving in one direction or the other, including those that depend on the current position of the process.

6.3 Other Properties

Space limitations prevent us from showing derivations or numerical examples of more properties of the system. In addition to the properties described above, we can also show and illustrate that the mean number of visits to a transient state before the onset of deadlock is independent of the prior occurrence of the occupancy of fewer hangers, and that the expected time for all hangers to reach full occupancy is independent of the initial state of the system [7].

7 NUMERICAL EXAMPLES

7.1 Mean Occupancy of One Hanger Sufficient for Load Specification

The parameters that drive the onset of deadlock are the number of coat hangers N, the arrival rate λ, and the mean service time $1/\mu$. Observe that

$$\alpha_k = \frac{k\mu}{\lambda + k\mu} = \frac{k}{\rho + k} \tag{13}$$

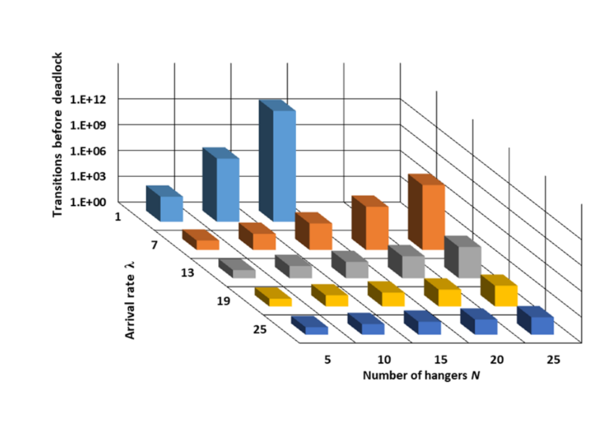

Figure 4: First passage time into deadlock when all hangers are empty.

where $\rho = \lambda/\mu$ is the mean occupancy of one hanger. Without loss of generality, we can set $\mu = 1$, since the values of α_k and β_k depend only on the ratio $\rho = \lambda/\mu$. We consider systems with the traffic intensity $1 \leq \rho \leq 25$ and the number of hangers N taking values between 5 and 25.

7.2 First Passage Times into Deadlock

We know from our results in Section 6.2 that deadlock occurs with probability one and that its onset is only a matter of time. The expected first passage time into deadlock indicates how long that time might be under various load conditions. It depends on the current number of occupied hangers. Figure 4 shows the first passage time into deadlock when all hangers are empty as a function of the number of hangers with $\mu = 1$ and $\lambda = \rho$varying from 1 to 25. We have not been able to compute the values the first passage times starting with all hangers empty for (ρ, N)=(1,20) and (1,25) because $(\mathbf{I} - \mathbf{Q})$ is ill conditioned there, as expected. The physical interpretation is that at this low traffic intensity, having 20 or 25 hangers might make the expected first passage times very long. Here, the expected first passage time from empty to deadlock would exceed 10^{12} transitions. Something similar is expected when all hangers are occupied initially, as shown in Figure 5. Figure 4 and Figure 5 show that the first passage time into deadlock is short for small numbers of hangers when the arrival rate is large compared with the time spent in the museum, and very large for larger numbers of hangers when the arrival rate is small, i.e. when the traffic intensity is small. Increasing the number of hangers in the checkroom at relatively high traffic volumes increases the expected time to deadlock, but does not eliminate the possibility of it occurring. Qualitatively, these results are consistent with intuition. Our numerical results show that the expected time to deadlock drops dramatically as the traffic intensity in the system increases. Interestingly, increasing the number of available hangers increases the expected time to deadlock much more when the traffic intensity is light than when it is heavy. The first passage time is never infinite, though it might be large. One practical implication of this is that deadlock might

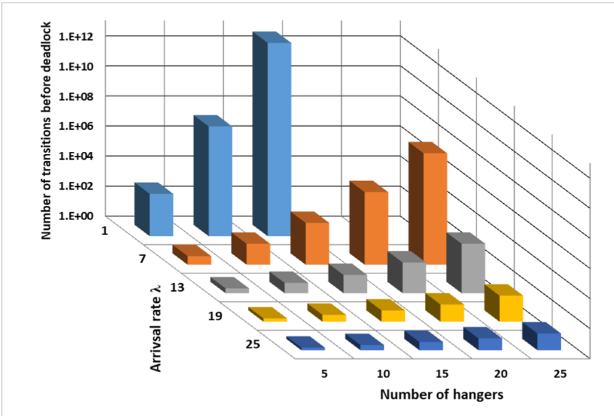

Figure 5: First passage time into deadlock when all hangers are occupied.

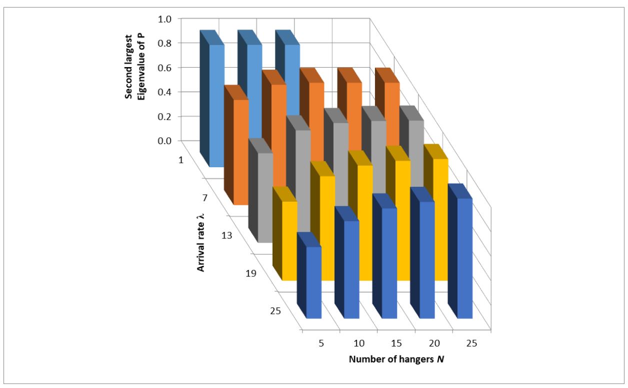

Figure 6: The second largest eigenvalue of P.

not occur in sunny-day light-load test scenarios, though it could be provoked to occur under heavy test loads.

7.3 Expected Time to Deadlock and Second Largest Eigenvalue of P

The smaller the second largest eigenvalue of **P**, the faster deadlock will occur. Figure 6 shows that **P** has a smaller second largest eigenvalue when the number of hangers is small and the arrival rate λ is large, and vice versa. Figure 7 shows the qualitative relationship between the second largest eigenvalue and the first passage time. The relationship is not strictly monotone because both quantities depend on λ and the number of hangers.

8 REVIEW OF PREVIOUS WORK

Our review of previous work falls into the following areas: queueing models of discrete object pools such as memory slots, deadlock, and Markov chain analysis and related topics.

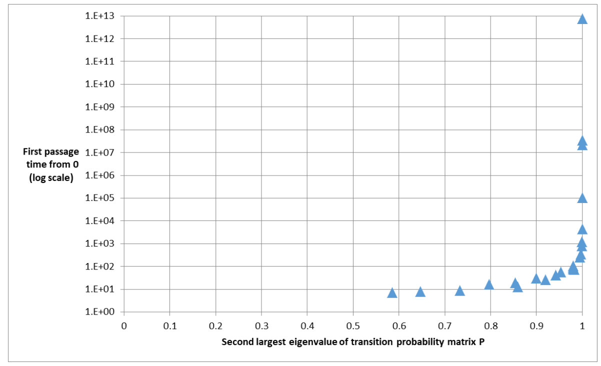

Figure 7: First passage time vs. second largest eigenvalue.

8.1 Constrained Memory Models of Discrete Object Pools

The museum checkroom is a metaphor for a system in which threads or processes acquire and release discrete objects that form a pool. Examples of this include memory partitions and Java database connections. The models we review here all describe systems in which processes queue for access to a memory partition, and release it without queueing before exiting the system. Avi-Itzhak and Heyman [1] and Latouche [14] modeled the queue for memory partitions. Neuts [16] used the caudal characteristic curve to study the parameter sets that make such systems more stable. Bondi [5] used Neuts' and Latouche's matrix geometric methods to study an admission policy that depends on the state of the CPU queue as well as on the number of jobs in the central subsystem [15]. None of these papers consider the case in which a job retains the discrete object when leaving the central subsystem and then queues to reenter the central subsystem to release it, as is the case with the present problem.

8.2 Deadlock

Deadlock and its prevention are discussed in many standard operating systems text books [9][11][19]. Bondi and Jin discuss the avoidance of deadlock in replicated database for tracking the location of cellular telephones [8]. Bondi discusses the present problem in the context of scalability in [6].

8.3 Markov Chain Analysis

The analysis in the present paper is mainly derived from presentations on stochastic processes in [4] and [10]. Avritzer and Weyuker use Markov chain models to plan performance and reliability tests in operating regions in which failure is more likely, thus shortening performance testing time [2]. Harchol-Balter describes a three-state Markov chain model of movements between linked web pages, culminating in a page with no links that corresponds to an absorbing state [12]. Popov uses a Markov chain with a small number of states to investigate the onset of deadlock when two processes request the same pair of resources in opposite orders [17]. Viswanadham *et al* discuss Markov chain and Petri net analyses applied to automated manufacturing systems [21]. Van Doorn and Pollett analyze discrete time reducible Markov chain models of progressive diseases in which there is a single absorbing or coffin state representing the patient's death [20]. The corresponding state transition probability matrix has zero entries everywhere except on the leading diagonal and on the diagonal just below it, unlike ours, which is tridiagonal. Reinecke *et al* use a stochastic Petri net to model the Dining Philosophers' problem [18].

9 CONCLUSIONS AND DIRECTIONS FOR FUTURE WORK

9.1 Conclusions

The stated goal of this research is to build a probabilistic model that predicts the expected amount of time taken for a component of system to migrate into deadlock under a given set of traffic and service conditions. The approach is to build a finite state machine that describes the sequences of events that can lead to deadlocked states. These states are trap states, or, in the terminology of Markov chains, absorbing states. The first passage time into deadlock can be computed from the probabilities of transitioning from a state to each of the adjacent ones. We have shown how to do this for a simple system with one deadlock state in which the transitions can be computed from customer arrival rates and resource holding times, provided that the arrival process is Poisson and provided that the resource holding times are exponentially distributed. Because this system corresponds to a one-dimensional random walk with a single absorbing barrier corresponding to the deadlocked state and a single reflecting barrier corresponding to the resource pool being fully unused, we know that the deadlock will occur with probability one. Our discrete time Markov chain model enables us to identify the circumstances under which the expected time for deadlock to occur is long or short. In our illustrative problem, the matrix to be inverted to compute the first passage times is ill conditioned in operating regions with the longest first passage times, i.e., the ones in which the expected time to enter deadlock is longest. Because our results show that deadlock eventually occurs with probability one, there is a clear case for designing the system component so that deadlock cannot occur.

9.2 Directions for Future Work

The circumstances under which deadlock occur in the example we have presented are fairly easy to describe. Future work could include methods of building generalized finite state machine descriptions of other circumstances in which deadlock can occur. If the causes of transitions and their probabilities can be identified, it will be straightforward to develop a discrete time Markov chain model that predicts the likelihood of deadlock and the most likely paths by which it occurs. It may be worthwhile to find circumstances under which the assumptions of Poisson arrivals and exponential resource holding times underlying our Markov chain model can be relaxed, e.g., by developing embedded Markov chain derivations of the state transition probabilities. The theory of Markov chains, with its classification of transient and absorbing states, supports the generalization of our method to systems with multiple deadlocked (absorbing) states and arbitrary initial states. When there is only a single absorbing state, we are only concerned with first passage times into that state. When there are multiple deadlocked states, each of which is reached by a distinct set of paths, we must evaluate and compare the first passage times for each of them. There is the added complication that there may be more than one set of equilibrium probabilities, because going into any absorbing state brings the evolution of the Markov chain to a halt. One could obviate this problem by merging all of the deadlocked states into a single one, but this would deprive us of the opportunity to determine which one has the shortest expected first passage time or to evaluate the consequences and risk associated with each. Moreover, the solution of the equation $\pi = \pi \mathbf{P}$ might not be unique, because there might be one for each absorbing state. We have not examined the effect of queueing for the transducer on the waiting times of coat claimants or new arrivals to the system. Delaying claimants increases the expected time to return a hanger to the pool, and shortens the expected first passage time into deadlock.

ACKNOWLEDGMENTS

This work was performed while the author was a visiting professor at the University of L'Aquila. The author has benefited from discussions with Vittorio Cortellessa and Pasquale Caianello. The author wishes to thank the referees for their useful comments.

REFERENCES

[1] B. Avi-Itzhak and D. P. Heyman. 1973. Approximate Queuing Models for Multiprogramming Computer Systems. *Opns. Res.* 21, 6 (1973), 1212–1230.
[2] A. Avritzer and E. Weyuker. 1995. The Automatic Generation of Load Test Suites and the Assessment of the Resulting Software. *IEEE Transactions on Software Engineering* 21, 9 (1995), 705–716.
[3] K. Baclawski. 2011. *Introduction to Probability with R.* Chapman and Hall.
[4] U. N. Bhat. 1972. *Elements of Applied Stochastic Processes.* Wiley, New York.
[5] A. B. Bondi. 1992. A Study of a State-Dependent Job Admissions Policy. *Performance Evaluation* 15 (1992), 133–153.
[6] A. B. Bondi. 2000. Characteristics of Scalability and Their Impact on Performance. In *Proc. WOSP2000 (Workshop on Software Performance).* 195–203.
[7] A. B. Bondi. 2017. *Predicting the Time to Migrate Into Deadlock Using a Discrete Time Markov Chain.* Technical Report TRCS 17/001. Dipartimento di Ingegneria e Scienze dell' Informazione e Matematica, Università degli Studi dell'Aquila, L'Aquila, Italy.
[8] A. B. Bondi and V. Y. Jin. 1996. A Performance Model of a Design for a Minimally Replicated Distributed Database for Database-Driven Telecommunications Services. *Distributed and Parallel Databases* 4 (1996), 295–318.
[9] E. G. Coffman and P.J. Denning. 1973. *Operating Systems Theory.* Prentice Hall.
[10] D. R. Cox. 1962. *Renewal Theory.* Methuen, London.
[11] A. N. Habermann. 1975. *Introduction to Operating System Design.* SRA, Chicago.
[12] Mor Harchol-Balter. 2013. *Performance Modeling and Design of Computer Systems: Queuing Theory in Action.* Cambridge University Press.
[13] P. J.Denning, J. B. Dennis, and J.E. Qualitz. 1978. *Machines, Languages, and Computation.* Prentice Hall.
[14] G. Latouche. 1981. Algorithmic Analysis of a Multiprogramming Multiprocessor Computer System. *JACM* 28, 4 (1981), 662–679.
[15] M. F. Neuts. 1981. *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach.* Johns Hopkins University Press.
[16] M. F. Neuts. 1986. The Caudal Characteristic Curve of Queues. *Advances in Applied Probability* 19, 1 (1986), 221–254.
[17] G. Popov. 2015. A Deadlock Probability Investigation Using Markov Chains. In *ESI Conference.*
[18] P. Reinecke, L. Bodrog, and A. Danilkina. 2012. Phase-Type Distributions. In *Resilience Assessment and Evaluation of Computing Systems*, K. Wolter, A. Avritzer, M. Vieira, and A. van Moorsel (Eds.). Springer, 85–114.
[19] A. C. Shaw. 1974. *The Logical Design of Operating Systems.* Prentice Hall, Englewood Cliffs, N.J.
[20] E. A. van Doorn and P. K. Pollett. 2009. Quasi-stationary distributions for reducible Markov chains in discrete time. *Markov. Processes and Relat. Fields* 15 (2009), 191–204.
[21] N. Viswanadham, T.L. Johnson, and Y. Narahari. 1990. Performance Analysis of Automated Manufacturing systems with Blocking and Deadlock.
[22] B. Weesakul. 1961. The Random Walk Between a Reflecting and an Absorbing Barrier. *Annals of Mathematical Statististics* 32, 3 (1961), 765–769.