Analyzing a security and reliability model using Krylov methods and matrix functions*

Giulio Masetti[1,2,✉] and Leonardo Robol[2,**]

[1] Department of Computer Science, Largo B. Pontecorvo 3, 56125, Pisa, Italy
[2] Institute of Science and Technology "A. Faedo", 56124, Pisa, Italy,
{giulio.masetti,leonardo.robol}@isti.cnr.it

Abstract. It has been recently shown how the computation of performability measures for Markov models can be recasted as the evaluation of a bilinear form induced by appropriate matrix functions. In view of these results, we show how to analyze a security model, inspired by a real world scenario. The model describes a mobile cyber-physical system of communicating nodes which are subject to security attacks. We take advantage of the properties of matrix functions of block matrices, and provide efficient evaluation methods.
Moreover, we show how this new formulation can be used to retrieve interesting theoretical results, which can also rephrased in probabilistic terms.

Keywords: Markov chains, Security model, Matrix functions, Krylov subspaces

1 Introduction

Consider an Homogeneous Continuous Time Markov Chain (CTMC) $X(t)$ with exponentially distributed transitions, and with infinitesimal generator Q. Assume that the process has $s > 0$ absorbing states, and that all these states are *final*, that is once the process arrives in any of those there will not be any more transitions. Such a CTMC is commonly encountered when studying reliability [17] and reliability/security models [12], and defines the context of our work.

Up to permuting the states, the Markov chain's generator Q can be partitioned as follows:

$$Q = \left[\begin{array}{c|c} \hat{Q} & V \\ \hline 0_{s\times(n-s)} & 0_{s\times s} \end{array}\right], \qquad V = \left[\begin{array}{c|c|c} v_{n-s+1} & \cdots & v_n \end{array}\right]. \tag{1}$$

* This research has been partially supported by the ISTI-CNR project "TAPAS: Tensor algorithm for performability analysis of large systems".

** The author is a member of the INdAM Research group GNCS. The research of this author has been supported by the Region of Tuscany (Project "MOSCARDO - ICT technologies for structural monitoring of age-old constructions based on wireless sensor networks and drones", 2016–2018, FAR FAS), and by the INdAM/GNCS project 2018 "Tecniche innovative per problemi di algebra lineare".

When the Markov chain is *acyclic*, which happens for instance when no "repairs" are possible in the system, the matrix $\hat{Q}$ can be permuted in upper triangular form as well. Nevertheless, for our analysis it is not necessary to consider acyclic Markov chains. We are concerned with the following questions:

1. What is the probability of being, at any positive time t, in a certain absorbing state k, for $n - s < k \leq n$ or, more generally, in a set $\mathcal{I}$ of absorbing states?
2. Given the same set $\mathcal{I}$ composed of absorbing states, what is the expected time that the process spends within this set in the time interval $[0, t]$?

Answering these questions allows us to compute, when the CTMC comes from a security and reliability model and a reward structure [17] is defined on the chain, measures such as: the probability that the system under analysis is in a failed state at a given time or the average time the system spent in a compromised state.

The answer to the first question yields a so-called *instantaneous* measure, whereas the second produces a *cumulative* measure. In view of the results in [13], we show that these measures can be rephrased in *matrix function* form [8], and that this yields an effective way of computing them exploiting the structure in the matrix Q, as well as some interesting theoretical characterizations.

In practice, we will show that the new formulation allows powerful algebraic manipulations of the measures, thus making easy to rephrase the problem in convenient ways. Matrix functions are becoming an increasingly popular tool to attack stiff ODEs [1, 10, 11], and allow for an easy exploitation of banded and more general structures which are often present in the infinitesimal generator of Markov chains [3, 14]. These results suggest that the reformulation of these measures in terms of matrix functions is natural and provides the base for further improvement to state-of-the art solvers.

Our main contributions in this paper are the following:

1. We reduce probabilities and cumulative measures of interest to the computation of bilinear forms whose matrices are indeed matrix functions of Q or $\hat{Q}$ (Theorem 1), where the latter is the restriction of Q to the non-absorbing states.
2. In view of this new formulation, we propose an efficient numerical method for their computation, based on the matrix function evaluation that extends the approach presented in [13].

Having at disposal the parallelism between security/reliability measures and special matrix functions evaluations allow us to import, we think for the first time, from the linear algebra community a toolbox of effective and efficient numerical methods into the dependability framework. In this framework, we impose no restrictions on $\hat{Q}$ except of being sparse. So, for instance, the CTMC can have cycles.

The paper is structured as follow. We start with a summary of definition and properties of matrix functions in Section 2. In particular we discuss how the block triangular structure of Q impact on the evaluation of $f(Q)$. Standard definition

of reward structures and measures on CTMCs are reported. In Section 3 we prove Theorem 1, that translate the measures of interest into the language of matrix functions, and show how to compute them. Section 4 is dedicated to the study of a concrete security/reliability model. Here we compare our methods to the well known uniformization [15] method in order to collect empirical data about performance. Finally, we drawn conclusions and discuss future work in Section 5.

2 Preliminaries

Here, we briefly summarize the definition of a matrix function. We refer the reader to [8] and the references therein for a more detailed discussion.

Definition 1 (Matrix function). *Let A be a matrix with spectrum $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$, and $f(z)$ a function that is analytic on the spectrum of A. Let $J = V^{-1}AV$ be the Jordan form of A, with $J = J_1(\lambda_{j_1}) \oplus \ldots \oplus J_k(\lambda_{j_k})$ being its decomposition in elementary Jordan blocks; we define the* matrix function $f(A)$ *as*

$$f(A) = V f(J) V^{-1}, \qquad f(J) = f(J_1) \oplus \ldots \oplus f(J_k)$$

where for an $m \times m$ Jordan block we have:

$$f(J(\lambda)) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \ldots & f^{(m)}(\lambda) \\ & \ddots & \ddots & \vdots \\ & & \ddots & f'(\lambda) \\ & & & f(\lambda) \end{bmatrix}$$

Even though the above definition needs to take the Jordan form into consideration for generality purposes, the intuitive definition of a matrix function is the action of of $z \mapsto f(z)$ on the *eigenvalues* of A. That is, when A is diagonalizable, $f(A)$ has the same eigenvectors of A, but the eigenvalues are $f(\lambda)$ in place of λ.

The most well-known examples of matrix functions are probably the inverse $f(z) = 1/z$, which corresponds to $f(A) = A^{-1}$, and the matrix exponential $f(A) = e^A$. As we will see, there are other functions $f(z)$ which are of interest.

Let us fix some notation. We denote by r a fixed weight vector of n elements, so that $r_{X(t)}$ is a process whose value at time t corresponds to entry of index $X(t)$ in r. The components of r are time independent. In most cases, r will be the *reward vector*, containing a "prize" assigned for being a certain state. We shall give two definitions of reward measures, to ease the discussion of their computation later on.

Definition 2. *Let r be a reward vector, and $X(t)$ a Markov chain with infinitesimal generator Q. Then, the number $\mathbb{E}[r_{X(t)}]$ is called* instantaneous reward measure at time t, *and is denoted by $M_{\text{inst}}(t)$. Similarly, the number*

$$M(t) = \mathbb{E}\left[\int_0^t r_{X(\tau)}\, d\tau\right] = \int_0^t \mathbb{E}\left[r_{X(\tau)}\right]\, d\tau = \int_0^t M_{\text{inst}}(\tau)\, d\tau$$

is called cumulative reward measure at time t.

Note that both definitions depend on the choice of the reward vector r. This dependency is not explicit in our notation to make it more readable — in most of the following examples the current choice of r will be clear from the context. Occasionally, we will make this dependency explicit by saying that a measure is associated with a reward vector r. Intuitively, the instantaneous reward measures the probability of being in a certain set of states (the non-zero entries of r), weighted according to the values of the components of r. The cumulative version is averaged over the time interval $[0, t]$.

The probability of being in a specific absorbing state k at a given time t and the expected time spend by $X(\tau)$ within a subset of absorbing states, weighted with r, for $\tau \in [0, t]$ can be derived with elementary probabilistic tools, and only depends on $\hat{Q}$ and v_k.

In fact, we can prove the following result.

Lemma 1. *Let $X(t)$ be a Markov chain with s absorbing states, and assume we have permuted the states as in (1). Let $\mathcal{I}$ be a set of absorbing states, that is $\mathcal{I} \subseteq \{k \mid n - s < k \leq n\}$, and r a vector such that $r_i \neq 0$ implies that $i \in \mathcal{I}$. Then,*

$$\mathbb{P}\{X(t) \in \mathcal{I}\} = \sum_{k \in \mathcal{I}} \sum_{j=1}^{n-s} \int_0^t \mathbb{P}\{X(\tau) = j\} \cdot (v_k)_j \; d\tau, \tag{2}$$

$$\mathbb{E}\left[r_{X(t)}\right] = \sum_{k \in \mathcal{I}} \sum_{j=1}^{n-s} r_k \int_0^t \mathbb{P}\{X(\tau) = j\} \cdot (v_k)_j \; d\tau, \tag{3}$$

$$\mathbb{E}\left[\int_0^t r_{X(\tau)} \; d\tau\right] = \sum_{k \in \mathcal{I}} \sum_{j=1}^{n-s} r_k \int_0^t \int_0^\theta \mathbb{P}\{X(\theta) = j\} \cdot (v_k)_j \; d\theta, \tag{4}$$

where $r_{X(t)}$ denotes the random variable that, at time t, has the value of the component of the vector r with index $X(t)$.

Proof. We note that, if we prove the result for (2) for the case $\mathcal{I} = \{k\}$ then the more general measures (2) and (3) follow immediately by linearity of $\mathbb{E}[\cdot]$. Therefore, without loss of generality, we first consider the case $\mathbb{P}\{X(t) = k\}$, where k is an absorbing state.

This probability can be written as the summation of the probability of the process jumping from state j to k at any time $\tau \leq t$. The probability of jumping from j to k at time τ has density $Q_{jk} \cdot \mathbb{P}\{X(\tau) = j\} \; d\tau$, since Q_{jk} is the rate of transitioning from j to k. Therefore, we can write

$$\mathbb{P}\{X(t) = k\} = \sum_{j=1}^{n-s} \int_0^t \mathbb{P}\{X(\tau) = j\} \cdot Q_{jk} \; d\tau,$$

and the thesis follows by $(v_k)_j = Q_{jk}$. Consider then

$$\mathbb{E}[r_{X(\tau)}] = \sum_{i=1}^{n} r_i \cdot \mathbb{P}\{X(\tau) = i\} = \sum_{k=n-s+1}^{n} r_k \cdot \mathbb{P}\{X(\tau) = k\},$$

and the thesis follows from the linearity of the integral, in fact

$$\int_0^t \mathbb{E}[r_{X(\tau)}]\ d\tau = \sum_{k\in\mathcal{I}} \sum_{j=1}^{n-s} r_k \int_0^t \int_0^\theta \mathbb{P}\{X(\theta)=j\}\cdot (v_k)_j\ d\theta,$$

where the last equality follows by the assumption that all the nonzero components of r lie in $\mathcal{I}$. □

The formula obtained by Lemma 1 do not yield a direct computation method especially in the cumulative measure case. In fact, the latter formula involve two integrals and two summations, so we would like to find a simpler evaluation procedure. As we will see in the next section, matrix functions will satisfy this request.

3 Rephrasing measures in matrix function form

The purpose of this section is to use the formalism of matrix function to rephrase the results of Lemma 1 in a way that is more suitable for computation. From now on, we will use `MATLAB` notation of matrix indexes, i.e., $A(3:7, 2:5)$ refers to the submatrix of A that consists of elements A_{ij} with $3 \le i \le 7$ and $2 \le j \le 5$. Moreover, we will often use the notation $\pi(t)$ to denote the probability of being in a certain state at time t, and π_0 the one encoding the probability of being in a certain state at time 0. With these hypothesis, $\pi(t)$ can be characterized by the ODE (known as Chapman–Kolmogorov equation)

$$\begin{cases} \dot{\pi}^T(t) = Q\pi(t)^T \\ \pi(0) = \pi_0 \end{cases}, \tag{5}$$

whose solution is $\pi(t)^T = \pi_0^T e^{tQ}$. Our task will be to prove the following result.

Theorem 1. *Let $X(t)$ be a Markov chain with n states and infinitesimal generator Q, and with s absorbing states, indexed as $n-s < k \le n$. Let π_0 be the vector containing the initial state of the chain at $t=0$, and assume that Q is partitioned as in (1). Let $\hat{\pi}_0$ be the vector containing the first $n-s$ entries of π_0, and assume that all the others are zero. Then, if r is a a reward vector with non-zero components on the absorbing states, we have*

$$\mathbb{E}[r_{X(t)}] = t\cdot \hat{\pi}_0^T \varphi_1(t\hat{Q})v, \quad \varphi_1(z) = \frac{e^z-1}{z}, \qquad v = Q(1:n-s,:)r,$$

$$\mathbb{E}\left[\int_0^t r_{X(\tau)}\ d\tau\right] = t^2\cdot \hat{\pi}_0^T \varphi_2(t\hat{Q})v, \quad \varphi_2(z) = \frac{\varphi_1(z)-1}{z}, \quad v = Q(1:n-s,:)r.$$

The definition of $f(\hat{Q})$ for $f(z)$ being either $\varphi_1(z)$ or $\varphi_2(z)$ is given in Definition 1. At this stage, however, it is not yet clear how to efficiently compute these matrix functions. In fact, the action of a matrix function on a vector can be efficiently approximated with $\mathcal{O}(n)$ flops, where n is the size of the matrix

$\hat{Q}$, by relying on Krylov methods, such as [7]. We refer the reader also to [13], where this strategy is applied precisely to Markov chains for the computation of performance and reliability measures.

We shall prove a property of matrix function of block upper triangular matrices that will be useful for our analysis. The result is a simple generalization of [8, Theorem 1.21], and the proof follows the same lines.

Lemma 2. *Let A be a 2×2 block upper triangular matrix partitioned as follows:*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ & A_{22} \end{bmatrix}, \qquad A_{11} \in \mathbb{C}^{m_1 \times m_1}, \qquad A_{22} \in \mathbb{C}^{m_2 \times m_2},$$

and assume that $A_{22} = \operatorname{diag}(d_1, \ldots, d_{m_2})$. Then, $f(A)$ has the same block structure, and can be expressed as

$$f(A) = \begin{bmatrix} f(A_{11}) & Y \\ & f(A_{22}) \end{bmatrix}, \qquad Y = \left[\begin{array}{c|c|c} g_1(A_{11})v_1 & \ldots & g_{m_1}(A_{11})v_{m_1} \end{array}\right],$$

where $v_j = A_{12}e_j$ and $g_j(z) = (f(z) - f(d_j))/(z - d_j)$.

Proof. We consider the case where the entry d_j are not eigenvalues of A_{11}. If that's not the case, one can still obtain the result by continuity, for instance approximating A with $A + \epsilon_j I$ with appropriately chosen $\epsilon_j \to 0$.

Since $f(A)$ commutes with A, we have $Af(A) = f(A)A$, and reading off this equation on the $(1,2)$ block yields $f(A_{11})A_{12} + YD = A_{11}Y + A_{12}f(D)$. Right multiplying by e_j yields $(A_{11} - d_jI)Ye_j = (f(A_{11}) - f(d_j)I)v_j$. Since $A_{11} - d_jI$ is nonsingular, this implies that

$$Ye_j = (A_{11} - d_jI)^{-1}(f(A_{11}) - f(d_j)I)v_j = g_j(A_{11})v_j,$$

and this concludes the proof. □

Since the vector $\pi(t)$, encoding the probability of being in each state at time t, solves the Chapman-Kolmogorov equation $\frac{d}{dt}\pi(t)^T = \pi(t)^T Q$, with initial condition $\pi(0) = \pi_0$, we have that $\pi(t)^T = \pi_0^T e^{tQ}$. In particular, we can express the quantities of interest for us as

$$\mathbb{E}[r_{X(t)}] = \pi_0^T e^{tQ} r \qquad \text{and} \qquad \mathbb{E}\left[\int_0^t r_{X(\tau)}\, d\tau\right] = \left(\int_0^t \pi_0^T e^{\tau Q}\, d\tau\right) \cdot r$$

The measure on the left is already in matrix function form. Concerning the one of the right, we rely on [13, Lemma 2.10], which can be stated as follows.

Lemma 3. *Let M be a cumulative reward measure, defined as*

$$M = \mathbb{E}\left[\int_0^t r_{X(\tau)}\, d\tau\right].$$

Then, $M(t) = \pi_0^T f(Q) r$, where π_0 is the initial probability distribution, r the reward vector, and

$$f(z) := t\varphi_1(tz) = \begin{cases} \frac{e^{tz}-1}{z} & z \neq 0 \\ t & z = 0 \end{cases}.$$

Now, we have all the ingredients to prove Theorem 1.

Proof (Theorem 1). Let us first consider the measure $\mathbb{P}\{X(t) = j\}$, with j being an absorbing state. We know that this is equal to $\pi_0^T e^{tQ} e_k$, in view of the previous remarks. Since Q is block upper triangular, we can apply Lemma 2 to obtain

$$e^{tQ} e_k = \exp\left(\begin{bmatrix} \hat{t}Q & tW \\ & 0 \end{bmatrix}\right) e_k = t \cdot \varphi_1(t\hat{Q})v, \qquad v := Q(1 : n-s, :)e_k,$$

where $\varphi_1(z) = (e^z - 1)/z$. This proves the first part of the claim. Concerning the second part, we can repeat the same steps replacing e^{tz} with $t\varphi_1 tz$, in view of Lemma 3. In particular, this yields

$$M = \mathbb{E}\left[\int_0^t r_{X(\tau)}\, d\tau\right] = \int_0^t \mathbb{E}[r_{X(\tau)}]\, d\tau = t \cdot \pi_0^T \varphi_1(tQ) r = t^2 \cdot \pi_0^T \varphi_2(t\hat{Q})v,$$

where $v = Q(1 : n-s, :) \cdot r$, thanks to a direct application of Lemma 2. □

3.1 A motivation for the use of the φ_j functions

For anyone that has experience with exponential integrators (see for instance [2, 10]) the appearance of the $\varphi_j(z)$ functions in the previous result will not come as a surprise. In fact, they appear in the explicit solution of the ODE

$$u'(t) = Au(t) + g(t, u(t)), \qquad t \geq 0, \qquad u(0) = u_0. \tag{6}$$

If we set $u_k = \frac{\partial^{k-1}}{\partial t^{k-1}} g(t, u(t))\Big|_{t=0}$, the solution can be written as

$$u(t) = e^{tA}u_0 + \sum_{k=1}^{\infty} \varphi_k(tA)t^k u_k,$$

where the $\varphi_j(z)$ functions are defined by recurrence as follows:

$$\varphi_{j+1}(z) = \frac{\varphi_j(z) - \frac{1}{j!}}{z}, \qquad \varphi_0(z) = e^z.$$

In particular, the $\varphi_1(z)$ and the $\varphi_2(z)$ are the ones we have considered in the previous section. Now, let us consider once more Lemma 1, and in particular the formulation

$$\mathbb{P}\{X(t) = k\} = \sum_{j=1}^{n-s} Q_{jk} \cdot \int_0^t \mathbb{P}\{X(\tau) = j\}\, d\tau.$$

Notice that if we define as $s(t) = \int_0^t \hat{\pi}(\tau)\ d\tau$ we have that, by substituting in the Chapman–Kolmogorov differential equation (5), we obtain

$$\dot{s}(t)^T = s(t)^T \hat{Q} + \hat{\pi}_0^T, \qquad s(0) = 0,$$

which is in the form of (6). Therefore, we can express the measure of interest as

$$s(t)^T = s(0)^T e^{t\hat{Q}} + t\hat{\pi}(0)^T \varphi_1(t\hat{Q}) = t \cdot \hat{\pi}(0)^T \varphi_1(t\hat{Q}).$$

We may then retrieve the same results given in Theorem 1 noting that the summation is in fact equivalent to the multiplication by the vector v. A similar statement can be given for the accumulated measure. In fact, this reasoning also provides an alternative proof to the one given in the previous section.

The construction of *exponential integrators*, which are a class of methods particularly effective for stiff ODEs, has been the main motivation behind the development of fast methods for the evaluations of $\varphi_j(A)b$ in recent years.

3.2 Computational remarks

In order to efficiently compute the matrix functions $\varphi_1(z)$ and $\varphi_2(z)$ on the matrix $\hat{Q}$, we rely on the tools developed in [13], which in turn heavily rely on the restarted Krylov methods for the evaluation of $f(A)b$ developed in [7]. The use of Krylov method for the evaluation of the action of matrix functions on a vector is in fact much older, and we refer the reader to [1, 2, 6, 9, 16] and the references therein for a list of similar approaches.

The choice of the method is motivated by the fact that it directly approximates the action of the matrix exponential, without integrating the Chapman–Kolmogorov differential equation. In particular, as shown in [13], this approach is much more resilient to stiffness in the problem with respect to the uniformization method used, for instance, in Möbius [5]. This will be apparent also in our numerical experiments, which we describe in Section 4.

The method presented in [7] allows to evaluate any matrix function that has an integral representation. Among these, we find entire functions (that is, holomorphic functions defined on $\mathbb{C}$), for which we have the Cauchy integral formula

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\hat{z})}{\hat{z} - z}\ d\hat{z}, \qquad z_0 \subseteq \mathrm{Int}(\Gamma),$$

where by $\mathrm{Int}(\Gamma)$ we denote the domain enclosed by a closed path Γ. All the functions we are interested in (namely, the φ_j functions) satisfy this constraint, and it is sufficient to choose the path Γ so that it encloses the spectrum of the matrix $\hat{Q}$. However, in [7] the authors present a particularly efficient choice of the parameters in the method that works particularly well for the approximation of $f(z) = e^z$. In order to exploit this work, we use [2, Theorem 2.1] that allows to rewrite $\varphi_j(A)b$ as $e^{\tilde{A}}\tilde{b}$ where $\tilde{A}$ is obtained by bordering A with j columns and rows. In the sake of conciseness, we shall not discuss the algorithmic details further, and we refer to [13] and to [7] for further details.

In the case of the instantaneous measure $\mathbb{P}\{X(t) = k\} = \pi_0^T e^{tQ} e_k$, it might seem odd to consider the formulation of the problem given in terms of a matrix exponential, transforming it into the evaluation of a $\varphi_1(z)$ matrix function in view of Theorem 1, and then actually computing it by going back to a matrix exponential. However, we stress that these are all different matrices. In fact, the original matrix Q has s absorbing states, whereas bordering $\hat{Q}$ only gives 1 absorbing state, and so the latter still has $s-1$ less columns and rows with respect to Q. We can give a precise characterization of the probabilistic interpretation of this procedure in the special case when the reward vector r is exactly equal to 1 on all the absorbing states.

Lemma 4. *Let $Q, \hat{Q}$ be the matrices as defined in (1), r a reward vector equal to 1 on the absorbing states $\mathcal{I}$. Consider the bordered matrix $\tilde{Q}$ given by*

$$\tilde{Q} := \begin{bmatrix} \hat{Q} & v \\ 0^T & 0 \end{bmatrix}, \qquad v = Qr.$$

Then, $\pi_0^T e^{t\tilde{Q}} e_{n-s+1} = t\varphi_1(tQ)r = \mathbb{P}\{X(t) \in \mathcal{I}\}$, and $\tilde{Q}$ is the infinitesimal generator of the Markov chain obtained by lumping together all the absorbing states.

Proof. We notice that the states in $\mathcal{I}$, being all absorbing and final, have all 0 outgoing transition rates, and in fact correspond to null rows in the matrix Q.

Moreover, the lumping process is obtained by considering a modified process $\tilde{X}(t)$ that coincides with $X(t)$ on the first $n - s$ states, and has an additional state $n - s + 1$ such that the rate of transitioning from a state $i \leq n - s$ to $n - s + 1$ is the same of the sum of all the rate of going from i to any state in $\mathcal{I}$ for $X(t)$. A direct verification reveals that these rates are exactly the entries of v, and therefore $\tilde{Q}$ is the infinitesimal generator of $\tilde{X}(t)$. □

If the reward vector r does not include all the absorbing states (or has more general weights), then the matrix $\tilde{Q}$ will not fall in the previous characterization. In fact, in general $\tilde{Q}$ will not even be stochastic, so it cannot be "transparently" associated with a Markov chain.

4 Reliability model for communication system attacks

We consider the mobile cyber-physical system model presented in [12], describing a collection of communicating nodes which are subject to attacks. The original study is based on a real-world architecture: there are N mobile nodes, each node using sensors for localization and measuring anomaly phenomena, and the system comprises an imperfect intrusion/detection functionality distributed to all nodes for dealing with both intrusion and fault tolerance. This mechanism is based on a voting system. Here a simplified version is discussed, we refer the reader to [12] for further details on the intrusion/detection functionality and the complete system description.

The model considers a node capture which involves taking control of a *good* node by deceiving the authentication and turning it into a *bad* node that will be able to generate attacks within the system. The attackers primary objective is to cause impairment failure by performing persistent, random, or insidious attacks. At each instant of time, the number of good and bad nodes are indicated as N_G and N_B, respectively, and N_E is the number of *evicted* nodes, i.e., nodes that have been detected as bad ones by the intrusion/detection mechanism. At the beginning, all nodes are considered good, i.e., $N_G = N$. Only bad nodes can perform internal attacks, and whenever one of this attacks have success the entire system fails, switching the value of N_F from 0, *ok*, to 1, *failed*.

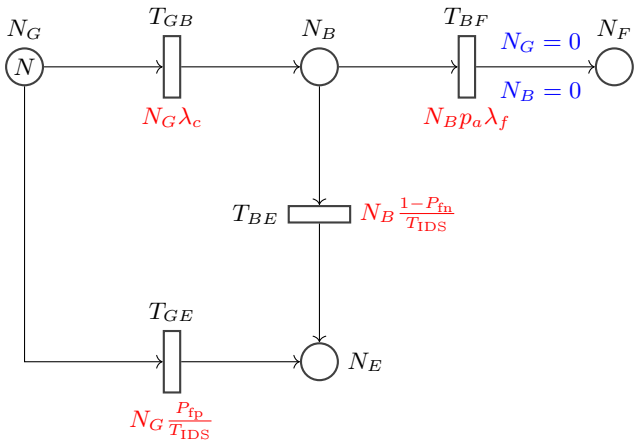


Fig. 1. Attack model for the cyber-physical communication system described in Section 4. This model is a simplified version of the model presented in [12]. Places are represented as circles, transitions are represented as rectangles. Place and transition names are in black, transition rates are in red and actions performed whenever transition T_{BF} completes are in blue.

The model is expressed through the definition of the Stochastic Reward Net [17] depicted in Figure 4, where places (circles) correspond to N_G, N_B, N_F, N_E, and determine the state of the system, and transitions (rectangles) define the behavior of the attack model:

T_{GB} is a transition of a node from good to bad, which represents the capture of a node by an attacker. The capture of a single node take place with rate λ_c, thus, being the capture of a node independent from the capture of other nodes, the rate of transition T_{GB} is $N_G\lambda_c$.

T_{BE} is a transition of a node from bad to evicted, that represents the correct detection of an attack. Calling P_{fn} the probability of intrusion/detection false negative, and T_{IDS} the period at which the intrusion/detection mechanism is exercised, the rate of T_{BE} is $N_B \frac{1-P_{\text{fn}}}{T_{\text{IDS}}}$.

T_{GE} is a transition of a node from good to evicted, that represent a false positive of the intrusion/detection mechanism. Calling P_{fp} the probability of intrusion/detection false positive, the rate of T_{GE} is $N_G \frac{P_{\text{fp}}}{T_{\text{IDS}}}$.

T_{BF} is a transition of a the entire system from ok to failed. When a node is captured it will perform attacks with a probability p_a and the success of attacks from N_B compromised nodes has rate λ_f, thus the rate of T_{BF} is $N_B p_a \lambda_f$. At completion of transition T_{BF} the entire system fails and then both N_G and N_B are set to 0 so that the Stochastic Reward Net reach a (failed) absorbing state.

The graph whose vertexes are all the feasible combinations of values within places and arcs correspond to transitions forms the Markov chain under analysis. For instance, with $N = 3$ The Stochastic Reward Net of Figure 4 produces the Markov chain depicted in Figure 4, where the notation (n_G, n_B, n_E, n_F) means $N_G = n_G, N_B = n_B, N_E = n_E$ and $N_F = n_F$. The parameters are chosen as

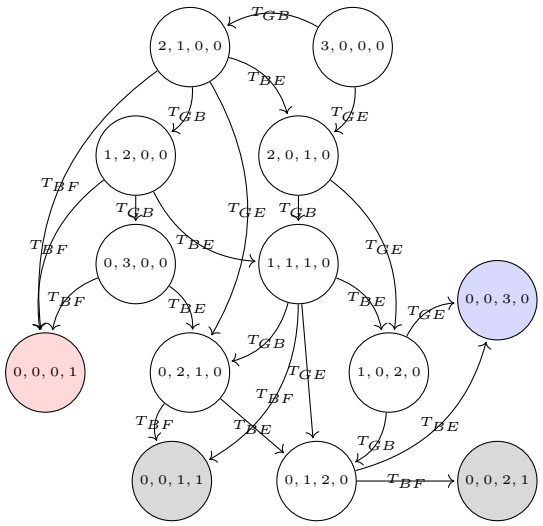


Fig. 2. Markov chain corresponding to Figure 4 with $N = 3$. Absorbing states are: failure without evicted nodes (red), failure with evicted (gray) and state with all evicted and without failure (blue).

follows

$$p_a = 0.7, \qquad P_{\text{fn}} = P_{\text{fp}} = 0.1, \qquad T_{\text{IDS}} = 15.0, \qquad \lambda_c = 0.1, \qquad \lambda_f = 0.2.$$

Fixed a time instant $t > 0$, the instantaneous measure of interest is the probability of system failure without evicted nodes at time t

$$P_{\text{failed-no-evicted}}(t) = \mathbb{P}\{X(t) = (0, 0, 0, 1)\}$$

An interesting cumulative measures is the averaged total time spent within states of system failure with some evicted nodes

$$F_{\text{some-evicted}}(t) = \frac{1}{t}\int_0^t \mathbb{E}\left[r_{X(\tau)}\right]\, d\tau, \qquad r = \mathbb{1}_{\left\{(0,0,k,1) \text{ for some } k \text{ s.t. } 1 \le k \le N-1\right\}}$$

Notice that $\lim_{t\to\infty} F_{\text{some-evicted}}$ is the probability that the process $X(t)$ ends up in an absorbing state where the system is failed and there are some evicted nodes [4], as can be seen in Figure 5.

We note that this can be proven also in the matrix function framework. In fact, in view of Lemma 3 the measure $F_{\text{some-evicted}}(t)$ can be written as $F_{\text{some-evicted}}(t) = \pi_0^T \varphi_1(tQ)r$. Since, for any z with negative real part we have $\lim_{t\to\infty} \varphi_1(tz) = \lim_{t\to\infty} e^{tz}$, and the spectrum of Q is contained in the left half plane, we immediately obtain

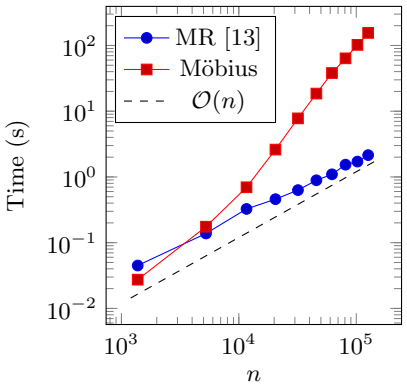
$$\lim_{t\to\infty} F_{\text{some-evicted}}(t) = \lim_{t\to\infty} \pi_0^T \varphi_1(tQ)r = \lim_{t\to\infty} \pi_0^T e^{tQ} r = \lim_{t\to\infty} \mathbb{P}\{X(t) \in \mathcal{I}\},$$

where $\mathcal{I}$ is the set where $r = 1$, corresponding to the states with some evicted nodes. Moreover, the fact that $e^{tz} - \varphi_1(tz)$ goes to zero linearly as $t \to \infty$ implies that the convergence above also behaves linearly. In fact, one can spot that the asymptotic convergence goes as $\frac{1}{t}$ in Figure 5.

We consider three examples to test the computational efficiency of the proposed approach. First, we use our algorithm to evaluate the instantaneous measure $P_{\text{failed-no-evicted}}(t)$ at the time $t = 30$. This is done using Theorem 1 that yields a formulation with the $\varphi_1(z)$ function. Then, we consider the average cumulative measure $F_{\text{some-evicted}}(t)$, and we repeat our test computing its value for $t = 30$. For this case, we also report the values of the measure for $t \in [0, 100]$ using a fixed value of N to show that it converges to a fixed value, namely the probability of being in a gray state.

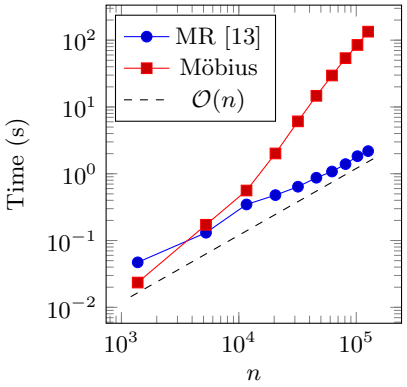
The results for the first experiments are reported in Figure 3, where the timings required to solve the problem with our approach (denoted by MR) and Möbius [5] are reported. The same is done for the second example in Figure 4. In both cases, we note that the complexity has a linear behavior as the number of states in the Markov chain increases. In contrast, the uniformization method suffers from the increasing stiffness of the problem, and therefore the timings required by Möbius are quadratic with respect to the number of states. This is a known limitation of the uniformization method, that is best suited to problems with balanced rates. The fact that the stiffness increases with n is due to the dependence of the rates by the marking, as visible in Figure 4.

Finally, we check that the averaged cumulative measure $F_{\text{some-evicted}}(t)$ tends to a limit values of t goes to infinity for the smaller case, which corresponds to $N = 50$ and has $n = 1376$ states, by sampling the measure for various values



n	t_{MR} (s)	$t_{\text{Möbius}}(s)$	$P_{\text{failed-no-evicted}}$
1,376	$4.49 \cdot 10^{-2}$	$2.74 \cdot 10^{-2}$	0.47
5,251	0.14	0.18	0.41
11,626	0.33	0.7	0.36
20,501	0.46	2.61	0.33
31,876	0.63	7.78	0.31
45,751	0.89	18.54	0.29
62,126	1.1	38.09	0.27
81,001	1.54	63.75	0.25
$1.02 \cdot 10^5$	1.71	102.2	0.24
$1.26 \cdot 10^5$	2.15	155.33	0.23

Fig. 3. On the left, the timings required to compute the measure $P_{\text{failed-no-evicted}}(t)$ at $t = 30$ using the approach in [13], denoted by MR, and the solver included in Möbius, which is based on the uniformization method. On the right, the timings and the computed measure for the same problem are reported in the table.



n	t_{MR} (s)	$t_{\text{Möbius}}(s)$	$F_{\text{some-evicted}}$
1,376	$4.71 \cdot 10^{-2}$	$2.35 \cdot 10^{-2}$	0.49
5,251	0.13	0.17	0.57
11,626	0.35	0.56	0.61
20,501	0.48	2.01	0.65
31,876	0.64	6.09	0.67
45,751	0.87	14.69	0.7
62,126	1.08	29.59	0.71
81,001	1.39	53.76	0.73
$1.02 \cdot 10^5$	1.83	84.77	0.74
$1.26 \cdot 10^5$	2.18	134.18	0.76

Fig. 4. On the left, the timings required to compute the averaged cumulative measure $F_{\text{some-evicted}}(t)$ at $t = 30$ using the approach in [13], denoted by MR, and the solver included in Möbius, which is based on the uniformization method. On the right, the timings and the computed measure for the same problem are reported in the table.

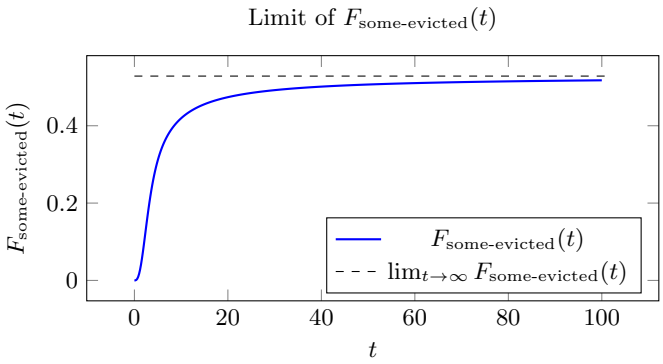


Fig. 5. The plot shows the value of $F_{\text{some-evicted}}(t)$ for different values of t ranging between 0 and 100. This case is the one with $N = 50$, and corresponds to a Markov chain of 1376 states.

of t between 0 and 100. The results are reported in Figure 5. By running our algorithm up to $t = 10^7$ we have extrapolated the limit for this case to be approximately 0.528483, which is also reported in the figure for comparison.

The results can be replicated using our public available *MATLAB* code [3].

5 Conclusions and future work

We have presented a way of analyzing a Markov model by heavily relying on matrix functions. We have confirmed the results reported in [13], which claim a computational advantage in treating the problem of computing these measures in matrix function form, in particular with respect to dealing with the stiffness in the model. We have tested our tools on a relevant example, but in fact we believe that they have far more general applicability. In particular, we suggest that the use matrix functions, which to the best of our knowledge is not part of the standard tools for the analysis of reliability and security models, enables the design of better algorithms, as well the achievement of new theoretical findings.

For the model under consideration, we have performed a quantitative analysis of attacks impact on reliability. The system logical architecture came from a real-world study but the parameters selected here are not taken from real measures.

The link with exponential integrators, which are sophisticated quadrature methods resilient to stiffness has been mentioned, and the tools needed to work with the measures using matrix functions directly have been presented. In particular, we have shown how measures originally formulated using the matrix exponential can be rephrased in terms of matrix functions with higher-order $\varphi_j(z)$ functions. The result given in Theorem 1 has been proven using matrix function theory, and a probabilistic interpretation as well. We have also provided

[3] https://github.com/numpi/markov-measures

the steps necessary to retrieve this result by direct integration of the underlying ODE defining the measure.

The numerical experiments confirm the effectiveness of Krylov methods for this task, presented in [13] for the computation of Markov measures, and in particular their resilience to stiffness in the ODE.

References

1. M. Afanasjew, M. Eiermann, O. G. Ernst, and S. Güttel. Implementation of a restarted Krylov subspace method for the evaluation of matrix functions. *Linear Algebra Appl.*, 429(10):2293–2314, 2008.
2. A. H. Al-Mohy and N. J. Higham. Computing the action of the matrix exponential, with an application to exponential integrators. *SIAM J. Sci. Comput.*, 33(2):488–511, 2011.
3. M. Benzi and P. Boito. Decay properties for functions of matrices over C^*-algebras. *Linear Algebra Appl.*, 456:174–198, 2014.
4. E. Çinlar. *Introduction to stochastic processes.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.
5. D. D. Deavours, G. Clark, T. Courtney, D. Daly, S. Derisavi, J. M. Doyle, W. H. Sanders, and P. G. Webster. The Möbius framework and its implementation. *IEEE Trans. on Softw. Eng.*, 28(10):956–969, 2002.
6. M. Eiermann and O. G. Ernst. A restarted Krylov subspace method for the evaluation of matrix functions. *SIAM J. Numer. Anal.*, 44(6):2481–2504, 2006.
7. A. Frommer, S. Güttel, and M. Schweitzer. Efficient and stable Arnoldi restarts for matrix functions based on quadrature. *SIAM J. Matrix Anal. Appl.*, 35(2):661–683, 2014.
8. N. J. Higham. *Functions of matrices. Theory and computation.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
9. M. Hochbruck and C. Lubich. On Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 34(5):1911–1925, 1997.
10. M. Hochbruck, C. Lubich, and H. Selhofer. Exponential integrators for large systems of differential equations. *SIAM J. Sci. Comput.*, 19(5):1552–1574, 1998.
11. M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numer.*, 19:209–286, 2010.
12. J. Martinez, K. Trivedi, and B. Cheng. Efficient computation of the mean time to security failure in cyber physical systems. In *Proceedings of the 10th EAI International Conference on Performance Evaluation Methodologies and Tools on 10th EAI International Conference on Performance Evaluation Methodologies and Tools*, pages 109–115, ICST, Brussels, Belgium, 2017. ICST.
13. G. Masetti and L. Robol. Computing performability measures in Markov chains by means of matrix functions. *arXiv:1803.06322*, 2018.
14. S. Massei and L. Robol. Decay bounds for the numerical quasiseparable preservation in matrix functions. *Linear Algebra Appl.*, 516:212–242, 2017.
15. A. Reibman and K. Trivedi. Transient analysis of cumulative measures of Markov model behavior. *Comm. Statist. Stochastic Models*, 5(4):683–710, 1989.
16. R. B. Sidje. Expokit: a software package for computing matrix exponentials. *ACM Trans. Math. Softw.*, 24(1):130–156, 1998.
17. K. S. Trivedi and A. Bobbio. *Reliability and Availability Engineering: Modeling, Analysis, and Applications.* Cambridge University Press, 2017.