Improving graph-based methods for computing qualitative properties of Markov decision processes

Mohammadsadegh Mohagheghi¹, Khayyam Salehi²
¹Department of Computer Science, Vali-e-Asr University of Rafsanjan, Iran
²Department of Computer Science, University of Tabriz, Iran

Article Info

ABSTRACT

Probabilistic model checking is a formal verification method, which is used to guarantee the correctness of the computer systems with stochastic behaviors. Reachability probabilities are the main class of properties that are proposed in probabilistic model checking. Some graph-based pre-computation can determine those states for which the reachability probability is exactly zero or one. Iterative numerical methods are used to compute the reachability probabilities for the remaining states. In this paper, we focus on the graph-based pre-computations and propose a heuristic to improve the performance of these pre-computations. The proposed heuristic approximates the set of states that are computed in the standard pre-computation methods. The experiments show that the proposed heuristic can compute a main part of the expected states, while reduces the running time by several orders of magnitude.

Keywords:
- Graph-based pre-computation
- Markov decision processes
- Probabilistic model checking
- Qualitative reachability probabilities

Corresponding Author:
Mohammadsadegh Mohagheghi,
Department of Computer science,
Vali-e-Asr University of Rafsanjan, Iran.
Email: mohagheghi@vu.ac.ir

1. INTRODUCTION

Formal methods are mathematical-based approaches that are used in software engineering and hardware design. The goal of these methods is to guarantee the correctness of the qualitative or quantitative properties of the desired systems [1, 2]. Model checking is an automated formal method that uses graph-based structures for modelling the underlying systems and logic-based propositions to specify the system properties [1, 3]. A model checker is a software tool that decides the satisfiability of the specified properties against the proposed model [1, 4].

Because of the stochastic behaviors of many computer systems, probabilistic structures are more useful for modeling such systems [4-6]. Markov chains and Markov decision processes (MDPs) are well-known structures for modelling stochastic systems and are widely used in artificial intelligence, economy, operations research and software engineering [7-9]. Several examples of the stochastic systems and their modelling are available in [3, 4, 7]. For this class of systems, probabilistic model checking is a good technique to verify the quantitative or qualitative properties of the systems. PRSIM [10], STORM [11] and IscasMC [12] are state of the art probabilistic model checkers.

The main classes of properties that are used in probabilistic model checking contain reachability probabilities, i.e. the maximal or minimal probability of reaching a set of goal states in the MDP model. A standard approach for computing reachability probabilities is to use iterative graph-based and numerical methods [13]. Graph-based methods (also called pre-computation) compute the set of states for which the probability of reaching a goal state is exactly zero or one [4]. An iterative numerical method starts from an initial vector of values and iteratively updates the values until satisfying the stopping criterion [1].
The main challenge of model checking in all variants is the state explosion problem, i.e. the number of states of the model grows exponentially in the number of its components. Several techniques are proposed to cope with this problem [1, 14-17]. In the case of probabilistic model checking the running time of iterative methods is the main problem that limits the scalability of the method [3, 18, 19]. Several techniques have been proposed to improve the performance of the standard probabilistic model checking methods. These techniques are used to reduce the running time of the graph-based [20, 21] or numerical computations [22-24] or both methods [16, 25]. Although the proposed methods show promising results in the running time of probabilistic model checking, experimental results show more improvement is needed for iterative graph-based of numerical methods [26]. In this paper, we consider the running time of graph-based methods as an important problem in probabilistic model checking and propose a new approach to reduce this running time.

The forward and backward approaches are two alternatives for graph-based computations in the probabilistic model checking of MDPs [1]. The forward approach is normally faster than the backward approach, but its main drawback is the memory overhead, which limits its scalability to relatively small models [1, 20]. As a result, most prominent model checkers (such as PRISM and IscasMC) use the forward approach for the pre-computations to avoid memory overhead, which is essential to overcome the state explosion problem [15, 27]. The motivation of this paper is to improve the performance of the graph-based pre-computation methods. To avoid memory overhead, we focus on the forward approach for pre-computation. As the main contribution of the paper, we propose a heuristic to reduce the running time of pre-computation in the forward approach. This heuristic reduces the number of iterations by approximating the set of expected states in an improved order. Although this heuristic does not guarantee to detect all states that the standard pre-computation algorithms do, the experiments show that it can reduce the overall run-time in most cases and improves the performance of probabilistic model checking of MDPs.

2. FORMAL PROBLEM DEFINATION AND BACKGROUND

In this section, we propose a brief review of the main concepts of probabilistic model checking that are used in this paper. More details about probabilistic model checking and their methods are available in [1, 3, 4].

2.1. Probabilistic Models and Reachability Probabilities

Markov decision processes are used in probabilistic model checking to model both nondeterministic and probabilistic aspects of a system [7].

Definition 1. (Markov Decision Process) A Markov Decision Process (MDP) is defined as a tuple $M = (S, \bar{s}, Act, \rightarrow, G)$ where $S$ is a finite set of states, $\bar{s} \in S$ is the initial state, $Act$ is a finite set of actions. For every state $s \in S$ the set of enabled actions of $s$ is denoted by $Act(s)$. We use $|Act(s)|$ as the size of this set and $\rightarrow$ is a probabilistic transition function and is defined as a subset of $S \times Act \times Dist(S)$. For each state $s \in S$ and enabled action $a \in Act(s)$, exactly one transition $s \rightarrow P$ exists. The notation $s \rightarrow P$ means that $(s, a, P)$ is an element of $\rightarrow$. We use $P(s, a, s')$ to denote the probability of reaching from $s$ to $s'$. By the action $a, G \subseteq S$ is the set of goal states. MDPs are widely used to model decision making problems in stochastic environments. Transitions of an MDP $M$ show the behavior of the related system [1]. For a state $s \in S$, a transition is performed in two steps: First, an action $a \in Act(s)$ is selected non-deterministically. Next, the destination state $s'$ is selected randomly with probability $P(s, a, s')$. We use $Post(s)$, $Pre(s)$, $Post(s, a)$ for the set of successor and predecessor and $\alpha$-successor states of $s$:

\[
Post(s) = \{s' \in S | \exists a \in Act(s), P(s, a, s') > 0\} \tag{1}
\]

\[
Pre(s) = \{s' \in S | s \in post(s')\} \tag{2}
\]

\[
Post(s, a) = \{s' \in S | P(s, a, s') > 0\} \tag{3}
\]

We use $|M|$ for the size of $M$ and define it as the number of states and transitions of $M$. A finite path of $M$ is a possible sequence of actions and transitions of $M$ and is defined as $\pi = s_0 \rightarrow s_1 \rightarrow \ldots \rightarrow s_n$ where $s_i \in S, a_i \in Act(s_i)$ and $s_{i+1} \in Post(s_i, a_i)$ for all $0 \leq i < n$. We use last($\pi$) for the last state of $\pi$.

Reachability probabilities are one of the main properties of probabilistic models that are computed in probabilistic model checking [3, 4]. A reachability probability is defined as the probability of reaching a goal state from the initial state of the model. In the case of MDPs, reachability probabilities are defined as the optimal (minimal or maximal) probability of reaching one of the goal states from the initial state $\bar{s}$. We use
Improving graph-based methods for computing qualitative properties of M. Iterative numerical approaches (such as value iteration [18] and policy iteration [4]) can be used to approximate the reachability probabilities.

2.2. Qualitative Reachability

Some graph-based computations can determine the set of states for which the extremal reachability probabilities are exactly 0 or 1. In the case of maximum reachability probability, these sets are denoted by $S^0_{\text{max}}$ and $S^1_{\text{max}}$ and are defined as:

$$S^0_{\text{max}} = \{ s \in S | P^\text{max}_s(\text{reach}(G)) = 0 \}$$

$$S^1_{\text{max}} = \{ s \in S | P^\text{max}_s(\text{reach}(G)) = 1 \}$$

We define $S^1_{\text{max}} = S \setminus (S^0_{\text{max}} \cup S^1_{\text{max}})$ for the remaining states. For the case of minimum reachability probabilities, we use $S^0_{\text{min}}$, $S^1_{\text{min}}$ and $S^\perp_{\text{min}}$ and define them in an analogous way [4]. The computations of these sets are called qualitative reachability analysis and are used as pre-computation in probabilistic model checking. The main aims of pre-computations are to simplify the iterative computations by focusing on $S^1_{\text{max}}$ and to increase the precision of computations [4, 18, 20, 25]. Algorithm 1 and Algorithm 2 show the standard pre-computation methods for the $S^0_{\text{max}}$ and $S^1_{\text{max}}$ sets [4].

**Algorithm 1** Pre-Computation for $S^0_{\text{max}}$

1. $R := G$
2. do
3. $R' := R$
4. $R := R' \cup \{ s \in S | \text{Post}(s) \subseteq R' \}$
5. while $R \neq R'$
6. return $S \setminus R$

**Algorithm 2** Pre-Computation for $S^1_{\text{max}}$

1. $R := G$
2. do
3. $R' := R$
4. $R := R' \cup \{ s \in S | \exists s' \in \text{Post}(s), (\forall s' \in S, (P(s, s', G) > 0 \implies s' \in R')) \wedge (\text{Post}(s, s) \subseteq R') \}$
5. while $R \neq R'$
6. return $R$

Algorithm 1 iteratively computes the set $R$ of states that can reach to one of the goal states $G$. In each iteration, the algorithm adds a state $s \in S$ to $R$ if at least one state $s' \in \text{Post}(s)$ has been added to $R$ in the previous iteration. For any state $s \in R$, we have $P^\text{min}_s(\text{reach}(G)) > 0$. The remaining states cannot reach to any of the goal states and are returned as the $S^0_{\text{max}}$ set.

Algorithm 2 uses a nested while loop to compute the $S^1_{\text{max}}$ set. The outer loop starts from $S$ and successively removes those states $s \in S$ for which we are sure $P^\text{min}_s(\text{reach}(G)) < 1$. It induces a sequence of $R_i$ sets, where $S = R_0 \supseteq R_1 \supseteq \ldots \supseteq R_i = S^1_{\text{max}}$. To compute the $R_i$ sets, the inner loop starts from $G$ (line 6 of Algorithm 2) and iteratively adds each state $s' \in S$ to $R_i$ if $s'$ can reach to one of the goal states with probability one via the states of $R_{i-1}$. For the remaining states (the states in $S \setminus R_i$), we are sure that they do not belong to $S^1_{\text{max}}$. The proof of correctness of these algorithms is available in [18].

3. RESEARCH METHOD

The time complexity of Algorithm 1 is in $O(|M|)$ and the time complexity of Algorithm 2 is in $O(|M|^2)$. In these cases, we suppose that for each state $s \in R_i$, the algorithms can determine any states of...
**Pre** \( s \) in \( O(1) \) [1]. For this purpose, the method should restore the information of the model in the backward approach, i.e., for each state \( s \in S \), the list of states in \( \text{Pre}(s) \) should be restored. The main drawback of restoring the information of a model is its memory overhead which is a main challenge in the state explosion problem. On the other hand, the forward implementation of these algorithms (as is used in PRISM [10]) need not any additional memory, but may increase the running time of the computations. To improve the performance of the pre-computations in the forward manner, we propose a heuristic to reduce their running time. As the first part of the heuristic, we modify Algorithm 1 for the computation of the \( S_{\text{max}}^1 \) set. As the second part (the main contribution of the work), we propose a new approach to approximate the \( S_{\text{max}}^1 \) set.

### 3.1. Improving Method for Computing \( S_{\text{max}}^0 \)

The idea of Algorithm 1 (which is used in PRISM and IscasMC) is to add a state \( s \in S \) to \( R \) if at least one states in \( \text{Post}(s) \) has been added in the previous iteration. To reduce the number of iterations of this algorithm in the forward approach, it can use only one set \( (R) \) to store the set of states that can reach to \( G \). In each iteration \( k \), a state \( s \in S \) is added to \( R \) if at least one state in \( \text{Post}(s) \) has been added to \( R \) in the iteration \( k \) or \( k - 1 \). Note that in Algorithm 1, \( s \) is added to \( R \) only if one state in \( \text{Post}(s) \) has been added to \( R \) in the previous iteration and not in the current one.

### 3.2. Improving Method for Approximating \( S_{\text{max}}^1 \)

In this section we propose our heuristic to compute a set \( T \subseteq S_{\text{max}}^1 \) as an approximation of the \( S_{\text{max}}^1 \) set. This heuristic starts from \( T_1 = G \) and iteratively compute the sets \( T_1 \subseteq T_2 \subseteq \cdots \subseteq T_{n-1} \subseteq T_n \subseteq S_{\text{max}}^1 \) where \( T_n \) is the fixed point set of the computations. For each iteration \( i \) of the heuristic, a state \( s \in S \) should be added to \( T_i \) if there exists an action \( a \in \text{Act}(s) \) for which \( \text{Post}(s,a) \subseteq T_{i-1} \). In this case, all \( a \)-successor states of \( s \) are in \( T_{i+1} \), which means for all of these states, we are sure that their maximum reachability probability is one. This heuristic is explained in Algorithm 3.

#### Algorithm 3 Approximating \( S_{\text{max}}^1 \)

**Input:** An MDP \( M = (S, \varepsilon, \text{Act}, \rightarrow, G) \)

**Output:** The set \( S_{\text{max}}^1 = \{ s \in S | \mu_{\max}(s) = 0 \} \)

1. \( T := G; \)
2. do
3. \( R := T; \)
4. \( T := T \cup \{ s \in S | \exists a \in \text{Act}(s), (\forall s' \in S, (P(s,a,s') > 0 \rightarrow s' \in R)) \}; \)
5. while \( R \neq T \);
6. return \( T; \)

The running time of Algorithm 3 is \( n \times |M| \) where \( n \) the number of iterations of the algorithm is and \( |M| \) is the size of the model. In worst case, \( n = |S| - 1 \) and the worst case time complexity of the algorithm is in \( O(|S| \times |M|) \). However, in most cases, the algorithm terminates after a few numbers of iterations. Note that the proposed heuristics in this section (Subsections 3.1 and 3.2) do not need any additional memory.

### 4. RESULTS AND ANALYSIS

To compare the running time of our heuristic for pre-computation with the standard method, we have implemented them in PRISM. We use 5 classes of standard case studies which are used in several previous works [4, 15, 19, 20, 26, 28]. For each class, we consider 4 models. More details about these case studies are available in [10]. In Table 1 we propose the name, parameter values, the number of states of each model and the experimental results. We use the sparse engine of PRISM for running the standard and improved pre-computation methods. We propose the number of states in the \( S_{\text{max}}^1 \) and \( S_{\text{max}}^0 \) sets and the approximation of \( S_{\text{max}}^1 \) (columns \( |S_{\text{max}}^0| \) and \( |S_{\text{max}}^1| \) in Table 1. We also propose the running times of standard pre-computations and our proposed methods. All times are in seconds.

For most cases (except the coin ones) our heuristic finds all states of the \( S_{\text{max}}^1 \) set. On the other hand, our heuristic reduces the running time of pre-computations for all cases. The best result is for CSMA (\( n=3,k=6 \)) which the running time of the standard pre-computation method is more than 8 hours (24603 seconds), while the running time of our improved method is less than 30 seconds. For other cases of CSMA and most cases of zeroconf, wlan and wlan collide our methods reduce the running times of pre-computation by two orders of magnitude. These results show that our heuristic presents a significant improvement in the performance of graph-based iterative methods in probabilistic model checking.
Improving graph-based methods for computing qualitative properties of...

(Mohammadsadegh Mohagheghi)
5. CONCLUSION

The paper proposed a heuristic to reduce the running time of pre-computation for probabilistic model checking of MDPs. The idea of the proposed methods in subsection 3.1 and 3.2 is to reduce the number of iterations of the graph-based pre-computations. Experimental results show that the proposed heuristic reduces the running times by several orders of magnitude and outperforms the standard and previous improved methods for probabilistic model checking. For future works, other probabilistic structures (such as probabilistic timed automata or stochastic hybrid automata) can be considered and one can study the impact of the proposed heuristic on the performance of model checking methods for these structures.

REFERENCES

[1] Christel Baier and Joost-Pieter Katoen. "Principles of model checking." MIT press, 2008.
[2] Mungkasi, Sudi. "Formal expansion method for solving an electrical circuit model." TELKOMNIKA (Telecommunication, Computing, Electronics and Control), vol. 17(3), pp. 1338-1343, 2019.
[3] Christel Baier, et al., "Probabilistic Model Checking". Dependable Software Systems Engineering. 2016, vol. 45, pp. 1-23.
[4] V. Forejt et al., "Automated Verification Techniques for Probabilistic Systems". In SFM 2011, Vol. 11, pp. 53-113.
[5] Alostad, Jasem M., "Improved probabilistic distance based locality preserving projections method to reduce dimensionality in large datasets." International Journal of Electrical and Computer Engineering (IJEECS), vol. 9(1), pp. 593-599, 2019.
[6] EzatulAkma Abdullah, SitiMeriamZahari, S.SarifahRadiahShariff, Muhammad Asmua Abdul Rahim. "Modelling volatility of Kuala Lumpur composite index (KLCI) using SV and garch models". Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), vol. 13(3), 2019.
[7] ML. Puterman"Markov decision processes: discrete stochastic dynamic programming", John Wiley & Sons; 2014 Aug 28.
[8] Lalaoiu, Mohamed, Abdellatif El Afia, and RaddouaneChihieb. "A self-tuned simulated annealing algorithm using hidden markov model." International Journal of Electrical and Computer Engineering (IJEECS), vol. 8(1), pp. 291-298, 2018.
[9] Marbun, Musa Partahi, NgapululirmeeSinisuka, and NanangHarriyanto. "The use of Markov Chain method to determine spare transformer number and location." International Journal of Electrical and Computer Engineering (IJEECS), vol. 9(1), pp. 1-8, 2019.
[10] Kwiatkowska M, Norman G, and Parker D. “PRISM 4.0: Verification of probabilistic real-time systems”. InInternational conference on computer aided verification 2011 Jul 14 (pp. 585-591). Springer, Berlin, Heidelberg.
[11] Dehnert C, et al., "A storm is coming: A modern probabilistic model checker". InInternational Conference on Computer Aided Verification 2017 Jul 24 (pp. 592-600). Springer, Cham.
[12] Hahn, Ernst Moritz, Yi Li, Sven Schewe, Andrea Turrini, and Lijun Zhang. "ISCASMC: A Webbased Probabilistic Model Checker." In International Symposium on Formal Methods, pp. 312-317. Springer, Cham, 2014.
[13] G. Norman and D. Parker. Quantitative verification:"Formal guarantees for timeliness, reliability and performance". Technical report, The London Mathematical Society and the Smith Institute, 2014.
[14] Hartmanns. "On the analysis of stochastic timed systems". PhD thesis, Saaarland University, 2015.
[15] Joachim Klein, et al., "Advances in probabilistic model checking with PRISM: variable reordering, quantiles and weak deterministic Bchi automata". International Journal on Software Tools for Technology Transfer, vol. 20(2), pages 179-194. 2018.
[16] Mateusz Ujma. "On Verification and Controller Synthesis for Probabilistic Systems at Runtime" PhD thesis, University of Oxford, 2015.
[17] Guldstrand Larsen. "Statistical Model Checking the 2018 Edition!" In: Margaria T., Steffen B. (Eds) Leveraging Applications of Formal Methods, Verification and Validation. Verification. Isola 2018. Lecture Notes in Computer Science, vol 11245. Springer, Cham.
Improving graph-based methods for computing qualitative properties of... (Mohammadsadegh Mohagheghi)

Mohammadsadegh Mohagheghi received his Ph.D in computer science from University of Tabriz in 2019, Ms in computer science from Sharif university of technology in 2008 and Bs in software engineering from Shahidbeheshty university in 2006. He is currently a faculty member of computer science in Vali-e-asr university of Rafsanjan, Iran. His main research interests include formal verification of stochastic and real-time systems, probabilistic model checking and machine learning.

Khayyam Salehi received his M.Sc. degree in Computer Science from Sharif University of Technology in 2010 and his Ph.D. in Computer Science at University of Tabriz in 2019. His research interests include quantification of information leakage, formal methods, and AI in computer security.

BIOGRAPHIES OF AUTHORS