On Computing Minimal Critical Subsystems for DTMCs

Ralf Wimmer, Nils Jansen, Erika Ábrahám, and Bernd Becker

Theory of Hybrid Systems
RWTH Aachen University, Germany

Computer Architecture
University of Freiburg, Germany

August 3, 2011

AlgoSyn Seminar
Contents

1 Preliminaries

2 Motivation

3 Minimal Critical Subsystems

4 Experimental Results

5 The Complexity Issue...
**Markov Chain**

**Definition**

*Discrete-time Markov Chain (DTMC)*

Andrey Andreyevich Markov (1856-1922)

Image source: wikipedia.org
Markov Chain

Definition
A **Discrete-time Markov Chain (DTMC)** is a tuple $M = (S, s_{init}, P, L)$ with finite state space $S$, initial state $s_{init}$, state labeling $L : S \rightarrow 2^{AP}$ and transition probability matrix $P : S \times S \rightarrow [0, 1]$. 

![Diagram of a DTMC with states $S_1$ to $S_5$ and transition probabilities labeled on the edges.](attachment:image.png)
Property: $\mathbb{P}_{\leq 0.02}(\Diamond s_5)$

Model Checking result: $\Pr_M(s_1, \Diamond s_5) = 0.0263$
Counterexamples

\[ \Pr_M(s_1, \diamond s_5) = 0.0263 \]

\[ \mathbb{P}_{\leq 0.02}(\diamond s_5) \]

Property:

\[ \pi_1 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \] with \[ \Pr_{\text{fin}}(\pi_1) = 0.02 \]

Model Checking result:

\[ \Pr_M(s_1, \diamond s_5) = 0.0263 \]
Property:

\[ \mathbb{P}_{\leq 0.02} (\Diamond s_5) \]

- \( \pi_1 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \) with \( Pr_{fin}(\pi_1) = 0.02 \)
- \( \pi_2 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \) with \( Pr_{fin}(\pi_2) = 0.004 \)

Model Checking result: \( Pr_M(s_1, \Diamond s_5) = 0.0263 \)
Counterexamples

Property: \( P \leq 0.02(\Diamond s_5) \)

- \( \pi_1 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \) with \( Pr_{fin}(\pi_1) = 0.02 \)
- \( \pi_2 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \) with \( Pr_{fin}(\pi_2) = 0.004 \)
- \( Pr_{fin}(\pi_1) + Pr_{fin}(\pi_2) = 0.024 > 0.02 \Rightarrow Property \ is \ false! \)

Model Checking result: \( Pr_M(s_1, \Diamond s_5) = 0.0263 \)
Counterexamples

\[ \mathbb{P}_{\leq 0.02}(\Diamond s_5) \]

Property:

- \( \pi_1 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \) with \( Pr_{\text{fin}}(\pi_1) = 0.02 \)
- \( \pi_2 : s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4 \rightarrow s_2 \rightarrow s_3 \rightarrow s_5 \) with \( Pr_{\text{fin}}(\pi_2) = 0.004 \)
- \( Pr_{\text{fin}}(\pi_1) + Pr_{\text{fin}}(\pi_2) = 0.024 > 0.02 \leadsto \text{Property is false!} \)

\( \leadsto \) Set \( C = (\pi_1, \pi_2) \) is a counterexample for the property.

Model Checking result: \( Pr_M(s_1, \Diamond s_5) = 0.0263 \)
1 Preliminaries

2 Motivation

3 Minimal Critical Subsystems

4 Experimental Results

5 The Complexity Issue...
Counterexamples

Model Checking

- Shows *correctness* of a system
- Reveals *defectiveness* of a system
Counterexamples

Model Checking

- Shows correctness of a system
- Reveals defectiveness of a system

Counterexamples for LTL properties

- Are delivered by Model Checking for defective systems
- Consist of single traces through a system
Counterexamples

Model Checking
- Shows correctness of a system
- Reveals defectiveness of a system

Counterexamples for LTL properties
- Are delivered by Model Checking for defective systems
- Consist of single traces through a system

Counterexamples in the probabilistic setting
- Are not computed during Model Checking
- Consist of (large or infinite) sets of paths
Probabilistic Counterexamples

Some state-of-the-art methods

- Search for paths in order of their probability *(Damman, Han, and Katoen 2008)*
- Find minimal counterexamples
- Use the abstraction of SCCs *(Andrés, D’Argenio and van Rossum, 2008)*
Some state-of-the-art methods

- Search for paths in order of their probability \cite{Damman2008}
- Find minimal counterexamples
- Use the abstraction of SCCs \cite{Andres2008}

Counterexamples are represented

- By enumeration of the paths
- By regular expressions
Hierarchical Counterexample Generation

Method

- SCC-based Model Checking
- If property was falsified:
  - Search for counterexample on abstract system
  - Hierarchical concretization

[QEST’10] [ATVA’11]
Hierarchical Counterexample Generation

Method

- SCC-based Model Checking
- If property was falsified:
  - Search for counterexample on abstract system
  - Hierarchical concretization

Advantages

- Compact representation
- Usability
- Abstract counterexamples
- Treatment of large systems
- Hierarchical approach
- Omission of system parts

[QEST’10] [ATVA’11]

Representation

- Critical subsystem
Hierarchical Counterexample Generation

Method

- SCC-based Model Checking
- If property was falsified:
  - Search for counterexample on abstract system
  - Hierarchical concretization

[QEST’10]

[ATVA’11]

Representation

- Critical subsystem

Advantages

- Compact representation $\rightsquigarrow$ Usability
- Abstract counterexamples $\rightsquigarrow$ Treatment of large systems
- Hierarchical approach $\rightsquigarrow$ Omission of system parts
Question: How good are the critical subsystems generated by our approaches?
Quality of Critical Subsystems

**Question:** How good are the critical subsystems generated by our approaches?

Compute the **minimal** critical subsystem in terms of
- states
- transitions

for certain case studies.
Quality of Critical Subsystems

**Question:** How good are the critical subsystems generated by our approaches?

Compute the **minimal** critical subsystem in terms of

- states
- transitions

for certain case studies.

**How?**

- NP-hard problem (?)
- Large problem instances
Contents

1 Preliminaries
2 Motivation
3 Minimal Critical Subsystems
4 Experimental Results
5 The Complexity Issue...
Given

- **DTMC** $M = (S, s_{\text{init}}, P, L)$,
- Set $T \subseteq S$ of **target states**, and
- Real-valued **probability bound** $\lambda \in [0, 1]$

such that

- $\Pr_{M}(s_{\text{init}}, \Diamond T) > \lambda$
A subsystem of $M$ is a DTMC $M' = (S', s_i', P', L')$ such that
- $(S' \setminus \{s_\perp\}) \subseteq S$
- $P'(s, s') = \begin{cases} 
  P(s, s') & \text{if } s \neq s_\perp \text{ and } s' \neq s_\perp, \\
  1 - \sum_{s' \in S' \setminus \{s_\perp\}} P(s, s') & \text{if } s \neq s_\perp \text{ and } s' = s_\perp, \\
  1 & \text{if } s = s' = s_\perp, \\
  0 & \text{otherwise.} 
\end{cases}$
Subsystems

A **subsystem** of $M$ is a DTMC $M' = (S', s'_i, P', L')$ such that

- $(S' \setminus \{s_\bot\}) \subseteq S$
- \[ P'(s, s') = \begin{cases} 
P(s, s') & \text{if } s \neq s_\bot \text{ and } s' \neq s_\bot, \\
1 - \sum_{s' \in S' \setminus \{s_\bot\}} P(s, s') & \text{if } s \neq s_\bot \text{ and } s' = s_\bot, \\
1 & \text{if } s = s' = s_\bot, \\
0 & \text{otherwise.} 
\end{cases} \]

![Diagram of subsystems](image)

$\Pr_M(s_1, \diamond s_5) = 0.0263$
A subsystem of $M$ is a DTMC $M' = (S', s'_i, P', L')$ such that

1. $(S' \setminus \{s_\perp\}) \subseteq S$

2. $P'(s, s') = \begin{cases} 
  P(s, s') & \text{if } s \neq s_\perp \text{ and } s' \neq s_\perp, \\
  1 - \sum_{s' \in S' \setminus \{s_\perp\}} P(s, s') & \text{if } s \neq s_\perp \text{ and } s' = s_\perp, \\
  1 & \text{if } s = s' = s_\perp, \\
  0 & \text{otherwise.}
\end{cases}$

Pr$_M(s_1, \diamond s_5) = 0.0263$
A subsystem of $M$ is a DTMC $M' = (S', s_1', P', L')$ such that

1. $(S' \setminus \{s_\perp\}) \subseteq S$
2. $P'(s, s') = \begin{cases} P(s, s') & \text{if } s \neq s_\perp \text{ and } s' \neq s_\perp, \\ 1 - \sum_{s' \in S' \setminus \{s_\perp\}} P(s, s') & \text{if } s \neq s_\perp \text{ and } s' = s_\perp, \\ 1 & \text{if } s = s' = s_\perp, \\ 0 & \text{otherwise.} \end{cases}$

Pr$_{M'}(s_1, \diamond s_5) = 0.025$
A subsystem $M' = (S', s'_i, P')$ of $M$ is critical if

- $S' \cap T \neq \emptyset$, $s_{init} = s'_i$, and
- $Pr_{M'}(s_i, \Diamond (S' \cap T)) > \lambda$. 
A subsystem $M' = (S', s'_I, P')$ of $M$ is critical if

- $S' \cap T \neq \emptyset$, $s_{\text{init}} = s'_I$, and
- $\Pr_{M'}(s_I, \Diamond (S' \cap T)) > \lambda$.

$T = \{s_5\}$, $\lambda = 0.02$
A subsystem $M' = (S', s'_i, P')$ of $M$ is critical if

- $S' \cap T \neq \emptyset$, $s_{\text{init}} = s'_\text{init}$, and
- $\Pr_{M'}(s_I, \Diamond (S' \cap T)) > \lambda$.
Task: Compute the minimal critical subsystem in terms of

- States: \( \min | S' \cap S | \)
- Transitions

Idea: Use modern solver technologies

- \text{Sat Modulo Theories}
- \text{Mixed Integer Linear Program}
The probability to reach $T$ from $s$ is the unique solution of the linear equation system

$$p_s = \begin{cases} 1 & \text{if } s \in T, \\ \sum_{s' \in S} P(s, s') \cdot p_{s'} & \text{otherwise}. \end{cases}$$
Introduce variables $x_s \in \{0, 1\}$. For all $s \in S$:
- $x_s = 0$: $s$ does not belong to the subsystem
- $x_s = 1$: $s$ belongs to the subsystem
Introduce variables $x_s \in \{0, 1\}$. For all $s \in S$:  
- $x_s = 0$: $s$ does not belong to the subsystem  
- $x_s = 1$: $s$ belongs to the subsystem

Introduce variables $p_s \in [0, 1]$. For all $s \in S$:  
- Probability to reach $T$ from $s$ within the subsystem
Introduce variables $x_s \in \{0, 1\}$. For all $s \in S$:
- $x_s = 0$: $s$ does not belong to the subsystem
- $x_s = 1$: $s$ belongs to the subsystem

Introduce variables $p_s \in [0, 1]$. For all $s \in S$:
- Probability to reach $T$ from $s$ within the subsystem

$$\land_{s \in T} (x_s = 0 \land p_s = 0)$$

$$\land \land_{s \in S \setminus T} (x_s = 0 \land p_s = 0)$$
Introduce variables $x_s \in \{0, 1\}$. For all $s \in S$:
- $x_s = 0$: $s$ does not belong to the subsystem
- $x_s = 1$: $s$ belongs to the subsystem

Introduce variables $p_s \in [0, 1]$. For all $s \in S$:
- Probability to reach $T$ from $s$ within the subsystem

$$\forall s \in T \quad ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = 1))$$

$$\land \forall s \in S \setminus T \quad ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = \sum_{s' \in S} P(s, s') \cdot p_{s'}))$$
Introduce variables $x_s \in \{0, 1\}$. For all $s \in S$:

- $x_s = 0$: $s$ does not belong to the subsystem
- $x_s = 1$: $s$ belongs to the subsystem

Introduce variables $p_s \in [0, 1]$. For all $s \in S$:

- Probability to reach $T$ from $s$ within the subsystem

$$\bigwedge_{s \in T} ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = 1))$$

$$\bigwedge_{s \in S \setminus T} ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = \sum_{s' \in S} P(s, s') \cdot p_{s'}))$$

$$p_{s_{\text{init}}} > \lambda$$
Introduce variables $x_s \in \{0, 1\}$. For all $s \in S$:
- $x_s = 0$: $s$ does not belong to the subsystem
- $x_s = 1$: $s$ belongs to the subsystem

Introduce variables $p_s \in [0, 1]$. For all $s \in S$:
- Probability to reach $T$ from $s$ within the subsystem

\[
\bigwedge_{s \in T} ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = 1))
\]

\[
\bigwedge_{s \in S \setminus T} ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = \sum_{s' \in S} P(s, s') \cdot p_{s'}))
\]

\[
p_{s_{\text{init}}} > \lambda
\]

\[
\sum_{s \in S} x_s = c.
\]
Minimal Critical Subsystem - SMT Instance

Computing the *minimal* system

- (Non-optimizing) SMT Solver
Computing the **minimal** system

- (Non-optimizing) SMT Solver
- Binary search for optimal $c$ over $\{1, \ldots, |S|\}$

\[
\bigwedge_{s \in T} \ ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = 1))
\]

\[
\land \bigwedge_{s \in S \setminus T} ((x_s = 0 \land p_s = 0) \lor (x_s = 1 \land p_s = \sum_{s' \in S} P(s, s') \cdot p_{s'}))
\]

\[
\land \quad p_{s_{init}} > \lambda
\]

\[
\land \quad \sum_{s \in S} x_s = c.
\]
Forward Cuts

- Each selected non-target state $s$ has a selected successor state.

$$x_s = 1 \implies \bigvee_{s' \in \text{succ}_M(s)} x_{s'} = 1$$
Forward Cuts

- Each selected non-target state \( s \) has a selected successor state.

\[
\begin{align*}
x_s &= 1 \quad \Rightarrow \quad \bigvee_{s' \in \text{succ}_M(s)} x_{s'} = 1
\end{align*}
\]

Backward Cuts

- Each selected non-initial state \( s \) has a selected predecessor state.

\[
\begin{align*}
x_s &= 1 \quad \Rightarrow \quad \bigvee_{s' \in \text{pred}_M(s)} x_{s'} = 1
\end{align*}
\]
Initial and target states

- The initial state is selected.
- At least one target state is selected.

\[ x_{s_{\text{init}}} = 1 \]

\[ \sum_{s \in T} x_s \geq 1 \]
Intermediate Summary and Further Work

Approaches

- SMT problem ✓
- MILP problem ✓
Intermediate Summary and Further Work

Approaches

- SMT problem ✓
- MILP problem ✓

Minimality in terms of

- States ✓
- Transitions ✓
Intermediate Summary and Further Work

Approaches
- SMT problem √
- MILP problem √

Minimality in terms of
- States √
- Transitions √

Models
- DTMCs √
- Markov Decision Processes (solving the nondeterminism) √
- Markov Reward Models (computing costs)
1 Preliminaries

2 Motivation

3 Minimal Critical Subsystems

4 Experimental Results

5 The Complexity Issue...
Solvers

SMT:
- Yices (version 1.0.29)
- Z3 (version 2.19)

MILP:
- glpk (version 4.45)
- lp_solve (version 5.5.2.0)
- Cbc (version 2.6.4)
- Scip (version 2.0.1)
Case Study - The Crowds Protocol

- Protocol for anonymous communication in networks
- \( n \) users divided into good and bad members
- Random forwarding of messages to destination or other member
- A run of \( r \) message deliveries is modelled
- Our models are parameterized by \( n \) and \( r \)
- Fixed good-to-bad ratio, fixed forwarding probabilities
- Model Checking: Probability that a member is identified (i.e. not anonymous)?
### Crowds Instances

| Model     | $|S|$ | $|E_M|$ | $|T|$ | $\lambda$ | $S_{\text{min}}$ | $P_{\text{min}}$ |
|-----------|-----|-------|------|-----------|-----------------|------------------|
| crowds2-2 | 77  | 101   | 3    | 0.09      | 22              | 27               |
| crowds2-3 | 183 | 243   | 26   | 0.09      | 22              | 27               |
| crowds2-4 | 356 | 476   | 85   | 0.09      | 22              | 27               |
| crowds2-5 | 612 | 822   | 196  | 0.09      | 22              | 27               |
| crowds3-3 | 396 | 576   | 37   | 0.09      | 37              | 51               |
| crowds3-4 | 901 | 1321  | 153  | 0.09      | 37              | 51               |
| crowds3-5 | 1772| 2612  | 425  | 0.09      | 37              | 51               |
| crowds5-4 | 3515| 6035  | 346  | 0.09      | 72              | 123              |
| crowds5-6 | 18817| 32677 | 3710 | 0.09      | 72              | 123              |
Runtimes

<table>
<thead>
<tr>
<th>Model</th>
<th>with forward cuts</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cbc</td>
<td>glpk</td>
<td>lp_solve</td>
<td>SCIP</td>
<td>Yices</td>
</tr>
<tr>
<td>crowds2-2</td>
<td>0.10</td>
<td>0.02</td>
<td>0.01</td>
<td><strong>0.01</strong></td>
<td>0.04</td>
</tr>
<tr>
<td>crowds2-3</td>
<td>0.82</td>
<td>0.14</td>
<td>0.19</td>
<td><strong>0.17</strong></td>
<td>3.98</td>
</tr>
<tr>
<td>crowds2-4</td>
<td>2.98</td>
<td>0.29</td>
<td>3.75</td>
<td><strong>0.46</strong></td>
<td>906.67</td>
</tr>
<tr>
<td>crowds2-5</td>
<td>4.57</td>
<td>0.43</td>
<td>19.54</td>
<td><strong>0.63</strong></td>
<td>– TL –</td>
</tr>
<tr>
<td>crowds3-3</td>
<td>3.12</td>
<td>10.69</td>
<td>278.12</td>
<td><strong>0.63</strong></td>
<td>714.06</td>
</tr>
<tr>
<td>crowds3-4</td>
<td>12.76</td>
<td>49.59</td>
<td>3483.02</td>
<td><strong>2.65</strong></td>
<td>– TL –</td>
</tr>
<tr>
<td>crowds3-5</td>
<td>28.56</td>
<td>103.16</td>
<td>– TL –</td>
<td><strong>5.45</strong></td>
<td>– TL –</td>
</tr>
</tbody>
</table>

TL = Time limit (2 hours) exceeded
## Results: SCIP

<table>
<thead>
<tr>
<th>Model</th>
<th>no cuts</th>
<th>forward cuts</th>
<th>backward cuts</th>
<th>both cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>crowds2-3</td>
<td>0.16</td>
<td>0.17</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td>crowds2-4</td>
<td>0.58</td>
<td>0.46</td>
<td>0.33</td>
<td>0.25</td>
</tr>
<tr>
<td>crowds2-5</td>
<td>0.77</td>
<td>0.63</td>
<td>0.62</td>
<td>0.48</td>
</tr>
<tr>
<td>crowds3-3</td>
<td>0.53</td>
<td>0.63</td>
<td>0.28</td>
<td>0.21</td>
</tr>
<tr>
<td>crowds3-4</td>
<td>2.59</td>
<td>2.65</td>
<td>1.45</td>
<td>1.01</td>
</tr>
<tr>
<td>crowds3-5</td>
<td>8.86</td>
<td>5.45</td>
<td>4.26</td>
<td>3.06</td>
</tr>
<tr>
<td>crowds5-4</td>
<td>92.97</td>
<td>50.71</td>
<td>21.76</td>
<td>20.18</td>
</tr>
<tr>
<td>crowds5-6</td>
<td>– TL –</td>
<td>3164.26</td>
<td>826.66</td>
<td>840.81</td>
</tr>
</tbody>
</table>
Contents

1 Preliminaries
2 Motivation
3 Minimal Critical Subsystems
4 Experimental Results
5 The Complexity Issue...
Our problem  Given a DTMC $M = (S, s_{\text{init}}, P, L)$, a set of target states $T \subseteq S$ and a probability bound $\lambda$ satisfying

$$\Pr_M(s_{\text{init}}, \Diamond T) > \lambda.$$ 

Is there a subsystem $M' = (S', s_{\text{init}}, P', L')$ with $c \in \mathbb{N}$ states still satisfying

$$\Pr_{M'}(s_{\text{init}}, \Diamond (S' \cap T)) > \lambda?$$

To justify the application of solvers, the problem should be NP-hard!

@AlgoSyn: Is it?
One (failed) approach to prove the NP-hardness:

- Reduction from the KNAPSACK problem:

\[
\begin{align*}
\text{max} \sum_{i=1}^{n} v_i \cdot x_i \\
\text{such that} \sum_{i=1}^{n} w_i \cdot x_i \leq W
\end{align*}
\]

with integer values \( v_i \) and real-valued weights \( w_i \).
One (failed) approach to prove the NP-hardness:

- Reduction from the KNAPSACK problem:

\[
\begin{align*}
\text{max} & \sum_{i} v_i \cdot x_i \\
\text{such that} & \sum_{i} w_i \cdot x_i \leq W
\end{align*}
\]

with integer values \( v_i \) and real-valued weights \( w_i \).