

A logic for model-checking of mean-field models

Anna Kolesnichenko

Anne Remke

Pieter-Tjerk de Boer

Centre for Telematics & Information Technology,
University of Twente, Enschede, The Netherlands

a.v.kolesnichenko@utwente.nl

a.k.i.remke@utwente.nl

p.t.deboer@utwente.nl

Boudewijn R. Haverkort

Embedded Systems Institute, Eindhoven, The Netherlands;
and Centre for Telematics & Information Technology,
University of Twente, Enschede, The Netherlands

b.r.h.m.haverkort@utwente.nl

1 Introduction

Recently, many systems consisting of a large number of interacting objects were analysed using the mean-field method, see [2, 4]. Key advantages of the mean-field method are: avoidance of the state space explosion problem; the speed of calculation; and accuracy. The accurate analysis of large systems can be done in seconds using the mean-field method, whereas simulation of such system behaviour will take hours or even days [4].

So far, the mean-field method has only been used for performance evaluation. In this short paper, we apply it to model checking. The properties we want to check are overall properties of large systems composed of many identical interacting objects. These overall properties are defined in terms of properties of the individual objects which in turn are expressed using CSL [1]. The glue between the overall and the individual properties is provided by new *expectation operators*.

In this paper we define logic, which allows to describe the overall properties. We first provide a brief overview of the mean-field method in section 2, and then introduce MF-CSL and its new operators in section 3, followed by an example in section 4. Section 5 concludes the paper.

2 Mean-Field analysis

The mean-field method allows to compute the exact limiting behaviour of an infinite population of identically behaving objects, and suggests an approximation when the number of components is sufficiently large. Mean-field models are obtained in the following way:

- The individual CTMC model \mathcal{M} , describing the behaviour of one object, is constructed as a tuple (S^l, \mathbf{Q}, L) that consists of a countable and finite set of states $S^l = \{s_1, s_2, \dots, s_K\}$; the transition rate matrix \mathbf{Q} which may depend on the overall system state S^o , described below: $S^l \times S^l \times S^o \rightarrow \mathbb{R}$; and the labelling function $L : S^l \rightarrow 2^{AP}$ that assigns atomic propositions from a fixed finite set AP to each state.
- The overall mean-field model \mathcal{M}^o , describes the behaviour of N (possibly even infinitely many) similar objects, each modelled by \mathcal{M} , and is defined as a tuple $(S^o, \mathbf{Q}(\bar{m}(t)))$, that consists of an infinite set of states $S^o = \left\{ \bar{m} = (m_1, m_2, \dots, m_K) : \forall j \in \{1, \dots, K\}, m_j \in [0, 1] \mid \sum_{j=1}^K m_j = 1 \right\}$, where \bar{m}

is called occupancy vector; m_j denotes the fraction of the individual objects in state s_j of the local model \mathcal{M} at time t ; and the transition rate matrix $\mathbf{Q}(\bar{m}(t))$ of the local model \mathcal{M} , where $\mathbf{Q}_{s,s'}(\bar{m}(t))$ now describes the transition rates of objects (fractions of objects) from state s to state s' .

In the following we briefly recall the mean-field result, for more information see [3, 4, 5]. The mean-field approximation is based on the assumptions of Theorem 1 from [3], which states that the occupancy vector $\bar{m}(t)$ of the overall behaviour tends to be deterministic in distribution and satisfies the following differential equation:

$$\frac{d\bar{m}(t)}{dt} = \bar{m}(t)\mathbf{Q}(\bar{m}(t)), \text{ with initial distribution } \bar{m}(0). \quad (1)$$

The transient analysis of the overall system behaviour is obtained using the system of differential equations (1). The stationary behaviour of the overall system can be approximated using stationary points of the deterministic *fluid limit*, see [6]. The stationary distribution of the overall model \bar{m}_{st} is approximated as follows:

$$\bar{m}_{st}\mathbf{Q}(\bar{m}_{st}) = 0. \quad (2)$$

In the following section, we assume that the stationary distribution of the overall system exists and can be calculated using (2).

Equations (1) and (2) allow us to compute “standard” transient and steady-state fractions of the mean-field model. In the following section we explore how to express more intricate properties using mean-field models in combination with model checking techniques.

3 A continuous stochastic logic for mean-field models

We now introduce an extra layer “on top of CSL” to come to the logic *CSL for mean-field models* (MF-CSL) which allows us to reason at the level of the overall model in terms of expected probabilities. We first recall the CSL logic definition from [1].

Definition 1. Syntax of CSL. Let $p \in [0, 1]$ be a real number, $\bowtie \in \{\leq, <, >, \geq\}$ a comparison operator, $I \subseteq \mathbb{R}_{\geq 0}$ a non-empty time interval and AP a set of atomic propositions with $ap \in AP$. **CSL state formulas** Φ are defined by:

$$\Phi ::= tt \mid ap \mid \neg\Phi \mid \Phi_1 \wedge \Phi_2 \mid \mathcal{S}_{\bowtie p}(\Phi) \mid \mathcal{P}_{\bowtie p}(\phi),$$

where ϕ is a **path formula** defined as:

$$\phi ::= \chi^I \Phi \mid \Phi_1 U^I \Phi_2.$$

□

Definition 2. Syntax of MF-CSL. Let $p \in [0, 1]$ be a real number, and $\bowtie \in \{\leq, <, >, \geq\}$ a comparison operator. **MF-CSL formulas** Ψ are defined as follows:

$$\Psi ::= tt \mid \neg\Psi \mid \Psi_1 \wedge \Psi_2 \mid \mathbb{E}_{\bowtie p}(\Phi) \mid \mathbb{ES}_{\bowtie p}(\Phi) \mid \mathbb{EP}_{\bowtie p}(\phi),$$

where Φ is a **CSL state formula** and ϕ is a **CSL path formula**.

□

As mentioned in Section 2, in some cases the local CTMC model \mathcal{M} might be non-homogeneous, i.e. transition rates might depend on the state of the overall model \mathcal{M}^σ , which makes model-checking on the local level non-trivial.

Here we have introduced three expectation operators: $\mathbb{E}_{\bowtie p}(\Phi)$, $\mathbb{ES}_{\bowtie p}(\Phi)$ and $\mathbb{EP}_{\bowtie p}(\phi)$, with the following interpretation:

- $\mathbb{E}_{\bowtie p}(\Phi)$ is the *expected* fraction of objects that are in a (local) state that satisfying a general CSL state formula Φ ;
- $\mathbb{ES}_{\bowtie p}(\Phi)$ is the *expected* fraction of objects that satisfy Φ in steady state, starting from the given distribution of objects;
- $\mathbb{EP}_{\bowtie p}(\phi)$ is the *expected* probability of taking a ϕ -path from a given distribution of objects over local states.

The formal definition of the MF-CSL semantics is as follows:

Definition 3. Semantics of MF-CSL. The satisfaction relation \models for MF-CSL formulas and states $\bar{m} \in S^o$ is defined by:

$$\bar{m} \models tt \quad \forall \bar{m} \in S^o, \quad \left| \begin{array}{l} \bar{m} \models \mathbb{E}_{\bowtie p}(\Phi) \quad \text{iff} \quad \left(\sum_{j=1}^K m_j \cdot \text{Ind}_{(s_j \models \Phi \mid \bar{m})} \right) \bowtie p, \\ \bar{m} \models \neg \Psi \quad \text{iff} \quad \bar{m} \not\models \Psi, \\ \bar{m} \models \Psi_1 \wedge \Psi_2 \quad \text{iff} \quad \bar{m} \models \Psi_1 \text{ and } \bar{m} \models \Psi_2, \end{array} \right. \quad \left| \begin{array}{l} \bar{m} \models \mathbb{ES}_{\bowtie p}(\Phi) \quad \text{iff} \quad \left(\sum_{j=1}^K m_j \cdot \pi^{\mathcal{M}}(s_j, \text{Sat}(\Phi, \bar{m})) \right) \bowtie p, \\ \bar{m} \models \mathbb{EP}_{\bowtie p}(\phi) \quad \text{iff} \quad \left(\sum_{j=1}^K m_j \cdot \text{Prob}^{\mathcal{M}}(s_j, \bar{m}, \phi) \right) \bowtie p, \end{array} \right.$$

where $\text{Sat}(\Phi, \bar{m})$ is a satisfaction set of the CSL formula Φ for a given state of the overall model; $\pi^{\mathcal{M}}(s, \text{Sat}(\Phi, \bar{m})) = \sum_{s_j \in \text{Sat}(\Phi, \bar{m})} \pi^{\mathcal{M}}(s, s_j)$, describes the local steady-state probability of being in a state from $\text{Sat}(\Phi, \bar{m})$, given initial state s and overall state \bar{m} ; $\text{Prob}^{\mathcal{M}}(s, \bar{m}, \phi)$ is the probability measure of all paths $\sigma \in \text{Path}(s)$ that satisfy ϕ when starting in state s for a given overall state of the system, that is, $\text{Prob}^{\mathcal{M}}(s, \bar{m}, \phi) = \Pr\{\sigma \in \text{Path}^{\mathcal{M}}(s) \mid \sigma \models \phi, \text{ and } \bar{m}\}$; and $\text{Ind}_{(s_j \models \Phi \mid \bar{m})}$ is an indicator function, which shows whether a local state $s_j \in S^l$ satisfies formula Φ for a given overall state \bar{m} :

$$\text{Ind}_{(s_j \models \Phi \mid \bar{m})} = \begin{cases} 1, & \text{if } s_j \models \Phi, \text{ given } \bar{m}, \\ 0, & \text{if } s_j \not\models \Phi, \text{ given } \bar{m}. \end{cases}$$

□

Algorithms of the MF-CSL model checking are currently being developed, and will be presented at the workshop.

4 Example

Figure 1 shows a simplified version of the model used in [4], which describes the spread of a computer virus. States represent the states of an individual computer, which can be *not-infected* $\{s_1\}$, *infected and active* $\{s_2\}$ or *infected and inactive* $\{s_3\}$. This results in the finite state space $S^l = \{s_1, s_2, s_3\}$ with $|S^l| = K = 3$ states. The set of atomic properties $AP = \{\text{infected}, \text{active}, \text{inactive}\}$, as indicated in the figure. The transitions represent *infection*, *recovery* and the change between being *active* and *inactive*. The transition rates are not constant in this example, since the infection rate $\mathbf{Q}_{s_1, s_2}(\bar{m}(t))$ depends on the number of computers in state s_3 , which are actively spreading the infection.

Given a system of N such computers we can model the overall average behaviour of the system using a mean-field model, which has the same structure as the individuals model (see Figure 1), however, with state space $S^o = \bar{m} = \{m_1, m_2, m_3\}$, where m_1 denotes the fraction of not-infected computers, and m_2 and m_3 denote the fraction of active and inactive infected computers, respectively. A system without infected computers then is in state $\bar{m} = (1, 0, 0)$.

To illustrate the expressivity of MF-CSL for mean-field models, consider the following MF-CSL formulas for the virus spreading model:

1. To define the atomic propositions on a level of mean-field model the operator $\mathbb{E}_{\times p}(ap)$ can be used. If the system is considered *infected* if more than 80% of the computers are infected, this can be expressed as $\mathbb{E}_{\geq 0.8}(\textit{infected})$.

2. The property “at least 80% of the computers will not have more than 10% probability of being infected in the long run”, is expressed as $\mathbb{E}_{\geq 0.8}(\mathcal{S}_{\leq 0.1}(\textit{infected}))$.

3. The property “the probability of a random computer to be infected in steady-state is higher than 10%” is expressed as follows: $\mathbb{ES}_{\geq 0.1}(\textit{infected})$. This property might be rephrased as “the fraction of computers, which are infected in steady state is at least 10%”.

4. The property “the probability of a random computer to recover (that is change state from infected to not-infected) within five time units is less than 40%” is expressed as $\mathbb{EP}_{< 0.4}(\textit{infected } U^{[0,5]} \textit{not-infected})$.

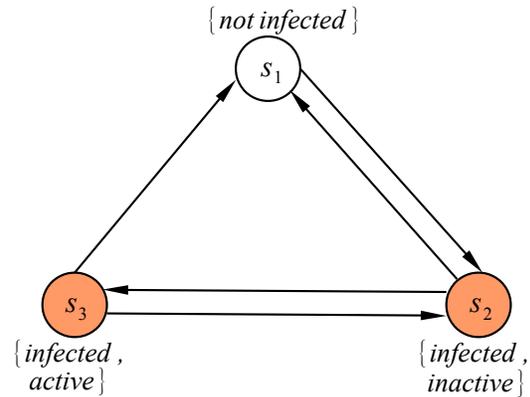


Figure 1: Example of the CTMC describing computer virus spread.

5 Conclusions and Future work

In this short paper, we have introduced the new logic MF-CSL, for use with models consisting of a large number of identical objects. MF-CSL adds a layer on top of regular CSL, with new operators describing the mean number of objects that satisfy CSL formulas. Algorithms for checking MF-CSL properties using the mean-field method are currently under development.

References

- [1] C. Baier, B. Haverkort, H. Hermanns & J. P. Katoen (2003): *Model-checking algorithms for continuous-time Markov chains*. *IEEE Trans. Softw. Eng.* 29(7), pp. 524–541.
- [2] R. Bakhshi, L. Cloth, W. Fokkink & B. R. Haverkort (2011): *Mean-field framework for performance evaluation of push-pull gossip protocols*. *Perform. Eval.* 68, pp. 157–179.
- [3] A. Bobbio, M. Gribaudo & M. Telek (2008): *Analysis of Large Scale Interacting Systems by Mean Field Method*. In: *Proceedings of the 2008 Fifth International Conference on Quantitative Evaluation of Systems*, IEEE Computer Society, Washington, DC, USA, pp. 215–224.
- [4] A. Kolesnichenko, A. Remke, P. T. de Boer & B.R. Haverkort (2011): *Comparison of the mean-field approach and simulation in a peer-to-peer botnet case study*. In: *8th European Performance Engineering Workshop (EPEW’11)*, LNCS 6977, Springer, pp. 133–147.
- [5] J.-Y. Le Boudec, D. McDonald & J. Munding (2007): *A generic mean field convergence result for systems of interacting objects*. In: *4th Int. Conference on Quantitative Evaluation of SysTems (QEST’07)*, IEEE CS Press, pp. 3–18.
- [6] J.Y. Le Boudec (2010): *The Stationary Behaviour of Fluid Limits of Reversible Processes is Concentrated on Stationary Points*. In: *arXiv:1009.5021v2 [math.DS]*, submitted.