Formal Certification of Surrogate Models for Cyber-Physical Systems Verification

Marco Esposito[1], Leonardo Picchiami[1]

[1] Computer Science Dept., Sapienza University of Rome, via Salaria 113, 00198, Italy

Abstract

In this short paper, we propose a method based on Statistical Model Checking to formally verify the prediction accuracy of surrogate models of Cyber-Physical Systems learned from simulation data. We show how surrogate models, trained with any desired Machine Learning algorithm and certified via our approach, can aid simulation-based formal verification techniques by greatly reducing the overall total number of model simulations needed. Our preliminary experimental evaluation over a Modelica model of a water pumping system shows that the proposed approach is viable in real-world scenarios.

Keywords

AI, Formal Methods, Statistical Model Checking, Surrogate Models, Verification, Cyber-Physical Systems

1. Introduction

The task of formally verifying properties of Cyber-Physical Systems (CPSs) is a crucial one in systems engineering. Unfortunately, real-world CPSs can often be analysed only by simulation, as they are either too complex to make symbolic analyses feasible, or are protected by intellectual property, thus only treated as black-boxes. Simulation-based verification methods, however, suffer from the fact that the number of simulations to perform, *i.e.,* the number of operational scenarios to consider, is typically so large that its full exploration is impossible or infeasible [1, 2]. In [3], we proposed an approach to verify properties of CPSs using Statistical Model Checking (SMC) in a fully simulation-based fashion. Such an approach exploits $\mathcal{AA}$ [4], a Monte Carlo sequential estimation algorithm, to produce a statistically accurate estimation of the expected value of simulation outputs. Although promising, such a method needs a large number of simulations (hence, long computation times) to produce accurate estimates. In this short paper, we propose a method to formally certify a surrogate model M, *i.e.,* an approximation of the system under verification learned from simulation data via SMC so that the property can be statistically verified over M using the approach from [3] with formal guarantees over the result. Preliminary experimental results show that the whole process of learning, certifying and verifying the surrogate model requires drastically fewer simulations than our previous fully simulation-based verification approach while guaranteeing, under fair assumptions over the surrogate model prediction accuracy, the same statistical guarantees.

OVERLAY 2022: 4th Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, December 28, 2022, Udine, Italy
esposito@di.uniroma1.it (M. Esposito); picchiami@di.uniroma1.it (L. Picchiami)
0000-0003-4543-8818 (M. Esposito); 0000-0001-5477-6419 (L. Picchiami)

2. Related Works

In this paper, we extend [3] by investigating new computational methods based on surrogate models and Statistical Model Checking (SMC) to verify safety-/mission-critical Cyber-Physical Systems [5] such as, *e.g.*, smart grids [6, 7, 8, 9], automotive systems [10, 11, 12] and biological systems [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24]. The limitations deriving from well-known formal approaches such as numerical techniques, logics or automata [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35] require the usage of SMC as enabling strategy to make the verification feasible for industrial practice. SMC is Monte Carlo-based technique that relies on *Hypothesis Testing* [36, 37, 38], *Estimation* [7, 39, 40, 4], and *Bayesian analysis* [41, 42] to *sample* new operational scenarios until statistical assurances on desired safety properties are provided. In such a way, we can counteract typical limitations such as the huge number of operational scenarios, namely *scenario explosion* [1, 43, 44, 45, 46, 47] or the system's complexity to evaluate *quantitative* and *quantitative* properties of interest. The literature (see, *e.g.*, [48, 49]) presents several simulation-based tools [50, 51, 52, 53, 54] that need specific system modelling through some kind of structure (*e.g.*, *Discrete Time Markov Chain*, *Continuous Time Markov Chain* [55, 56], *Probabilistic Timed Automata* [57]) to generate on demand all needed trajectories for the verification. To the best of our knowledge, no existing approaches and tools integrate a surrogate model as a system approximation to carry out verification activities. The formal certification of surrogate models falls within the field of Probably Accurately Correct (PAC) function learning, which has been studied extensively in the last decades (see, *e.g.*, [58, 59]). Existing methods (see, *e.g.*, [60, 61, 62, 63]), however, do not aim at minimising the total number of function samples (*i.e.*, simulations) and take a pre-defined number of samples derived from theoretical statistical bounds such as the Chernoff-Hoeffding Bound [64]. Finally, [65] proposes a method to perform Statistical Model Checking of CPSs using surrogate models and conformal inference. While such a method shares a similar goal with ours, it suffers from a combinatorial explosion in high dimensions as it tries to learn accurate surrogate models in subregions of the input space. Our approach, on the other side, leaves the burden of sampling the input space to the $\mathcal{AA}$ algorithm, which, in turn, guarantees the accuracy of the estimation while minimising the number of samples, independently of the input dimension.

3. Estimation-based Verification of Cyber-Physical Systems via Statistical Model Checking

In this section, we summarise the work done in [3] on the verification of Cyber-Physical Systems via Statistical Model Checking and numerical simulation. The proposed approach aims at establishing if the expected value of a given KPI exceeds or not a user-defined threshold. We exploited an *optimal* (ε, δ)-approximation algorithm that provides an estimate $\hat{\mu}$ of the desired expected value μ in which the accuracy has a (multiplicative) relative error of at most ε with probability at most $1 - \delta$. Such an algorithm guarantees the usage of the minimum number of required samples up to a constant factor. It avoids computing the expected value over the (possibly infinite) complete set of all relevant operational scenarios. We developed a message-passing based parallel implementation of the optimal Approximation Algorithm

($\mathcal{AA}$) [4] described by 1 orchestrator and n workers. Each worker produces new samples via numerical simulation of a given simulable model of a CPS, whereas the orchestrator updates the $\mathcal{AA}$ algorithms as soon as a new sample is available and handles all new inputs needed by each worker. We evaluated the viability of the approach on the Pumping System (PS), a Modelica system deployed via the Modelica Standard Library. PS is a pumping control system for drinking water described by an ingoing source pumped by a pump into a tank and outgoing sink water that models the users. The control component outputs the pump engine's rotational speed to regulate the tank's water level so that the system can keep the water level around 2.2 meters. In our experimental evaluation, we used the Mean Relative Absolute Error (MRAE) of the water level w.r.t. a reference value as the KPI and compared the computational performance of our method with several values for ε and δ.

4. Formal Certification of Surrogate Models

This section describes our surrogate-based approach to reduce the number of simulations needed to perform the verification task described in Section 3. Let W be a set of scenarios in which the system under verification operates and $p(w)$ be the probability density of scenario $w \in W$. Let S be the function that computes the KPI value (a real number) for a given scenario by simulating the model of the system. Given ε and δ in $(0, 1]$, our goal is to compute an (ε, δ)-approximation $\hat{\mu}$ of the expected value of the KPI, in order to statistically verify whether μ is lower than or equal to a given threshold P. We assume the availability of a surrogate model M of S, *i.e.*, a real function that approximates S over its whole domain. Many techniques exist to learn such a surrogate model in a simulation-efficient way; in this context, we are only interested in the model itself and the number of pairs $\langle w, S(w)\rangle$, say N_{train}, used to train it. Our goal is to formally certify M and its prediction performance in such a way that it can be used instead of S to prove that $\mu \leq P$ by computing its expected value μ_M over W. We define the Relative Absolute Error of M w.r.t. S for a given $w \in W$ as $r(w) = \frac{|M(w)-S(w)|}{S(w)+\zeta}$, where ζ is a small constant used to avoid division by zero. As the expected value of $r(w)$ over W is $\rho = \int_w r(w)p(w)dw$, it is easy to show that $\mu(1-\rho) \leq \mu_M \leq \mu(1-\rho)$. However, neither μ_M nor ρ can be computed exactly in finite time (unless with very strict assumptions over M), as the number of operational scenarios is infinite. Hence, we use $\mathcal{AA}$ twice to compute two approximations of such values. First, we compute an $(\varepsilon_{cert}, \delta_{cert})$-approximation $\hat{\rho}$ of ρ, for ε_{cert} and δ_{cert} in $(0, 1]$ provided by the user. Intuitively, such parameters will determine the statistical accuracy in the estimation of the expected relative absolute error of the surrogate, so they will influence the final error bounds over the estimation of μ. Once $\hat{\rho}$ is obtained, we compute an (α, β)-approximation $\hat{\mu}_M$ of μ_M, choosing α and β in $(0, 1]$ such that

$$\varepsilon \geq \varepsilon' = \frac{1}{2}\left(2\alpha + \hat{\rho}\left(\frac{1-\alpha}{1+\varepsilon_{cert}} + \frac{1+\alpha}{1-\varepsilon_{cert}}\right)\right) \tag{1}$$

and $1-\delta \leq (1-\delta_{cert})(1-\beta)$. It is easy to prove (we omit the proof for brevity) that $(1-\varepsilon)\,\mu \leq (1-\varepsilon')\,\mu \leq \hat{\mu_M} \leq (1+\varepsilon')\,\mu \leq (1+\varepsilon)\,\mu$, so $\hat{\mu}_M$ is an (ε, δ)-approximation of μ. This proves that the surrogate model can be safely used to solve the verification problem. Finally, we note, from eq. (1), that ε' tend to $\hat{\rho}$ as ε_{cert} and α tends to 0. This indicates that, no matter the statistical errors employed in the formal certification of the surrogate model and for the estimation of its expected value, the final error bound ε over μ cannot be stricter than the prediction accuracy $\hat{\rho}$ of the surrogate model. So, our method fails when $\hat{\rho} > \varepsilon$, reporting to the user that the surrogate model is not accurate enough for the verification task.

5. Experimental Evaluation

We evaluated the proposed approach through a comparison with the strategy presented in [3]. Along the same lines, we compared average values of $n = 10$ experiments for $\varepsilon = \delta = 0.02$ using the two approaches. The fully simulation-based method required, on average, around 6 hours and 60610.3 simulations to produce an (ε, δ)-approximation of the expected value of the KPI (see Section 3) $\hat{\mu}$ equal to 0.147. We trained a Support Vector Regressor (SVR) surrogate model using a dataset of $N_{train} = 1000$ simulation samples (sampled uniformly at random). On average the training phase required almost 6 minutes for simulations and 10 seconds for fitting the SVR model. For the formal certification phase, we chose $\varepsilon_{cert} = 0.05$ and $\delta_{cert} = 0.0049$; the surrogate model certification with $\mathcal{AA}$ required on average 4471 samples and 26 minutes to produce an estimation $\hat{\rho}$ of the model relative absolute error ρ equal to 0.0114. We chose $\alpha = 0.006$ and $\beta = 0.01$ for the estimation of the expected value of the surrogate prediction to get an $\varepsilon' \approx 0.018 < \varepsilon$ (according to eq. (1)) and $\delta' = 0.0149 < \delta$. The $\mathcal{AA}$ run on the surrogate model took, on average, 82162.8 prediction samples and 4 seconds, yielding an estimate $\hat{\mu}_M$ equal to 0.148. Hence, the total time required by the proposed surrogate-based approach was, on average, almost 32 minutes, *i.e.*, a reduction of approximately 91% w.r.t. the fully simulation-based approach.

6. Conclusions

In this short paper, we introduced an approach based on Statistical Model Checking to the *formal certification* of surrogate models of Cyber-Physical Systems. Our approach exploits the $\mathcal{AA}$ SMC algorithm to verify the accuracy of the surrogate model while minimising the total number of model simulations needed. We showed how such certification enables the adoption of surrogate models to formally verify properties of CPSs and demonstrates the performance improvement over our previous fully simulation-based method on a real-world case study. In future work, we plan to extend the proposed method to deal with more complex verification problems and evaluate it on higher-dimensional problems.

Acknowledgments

This work was partially supported by: Italian Ministry of University and Research under grant "Dipartimenti di eccellenza 2018–2022" of the Department of Computer Science of Sapienza University of Rome; INdAM "GNCS Project 2020"; Sapienza University projects RG12117A8B393BDC, RG11816436BD4F21, RG11916B892E54DB, RP11916B8665242F, AR1221816C974186, AR1221816C545113; Lazio POR FESR projects E84G20000150006, F83G17000830007.

References

[1] T. Mancini, F. Mari, A. Massini, I. Melatti, F. Merli, E. Tronci, System level formal verification via model checking driven simulation, in: Proceedings of 25th International Conference on Computer Aided Verification (CAV 2013), volume 8044 of *Lecture Notes in Computer Science*, Springer, 2013, pp. 296–312. doi:10.1007/978-3-642-39799-8_21.

[2] T. Mancini, I. Melatti, E. Tronci, Any-horizon uniform random sampling and enumeration of constrained scenarios for simulation-based formal verification, IEEE Transactions on Software Engineering (2021). doi:10.1109/TSE.2021.3109842.

[3] M. Esposito, L. Picchiami, Estimation based verification of cyber-physical systems via statistical model checking, in: Joint Proceedings of the 1st International Workshop on HYbrid Models for Coupling Deductive and Inductive ReAsoning (HYDRA 2022) and the 29th RCRA Workshop on Experimental Evaluation of Algorithms for Solving Problems with Combinatorial Explosion (RCRA 2022), volume 3281 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022.

[4] P. Dagum, R. Karp, M. Luby, S. M. Ross, An optimal algorithm for Monte Carlo estimation, SIAM Journal on Computing 29 (2000) 1484–1496. doi:10.1137/S0097539797315306.

[5] R. Alur, Principles of Cyber-Physical Systems, MIT Press, 2015.

[6] B. Hayes, I. Melatti, T. Mancini, M. Prodanovic, E. Tronci, Residential demand management using individualised demand aware price policies, IEEE Transactions on Smart Grid 8 (2017). doi:10.1109/TSG.2016.2596790.

[7] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, M. Prodanovic, L. Elmegaard, Demand-aware price policy synthesis and verification services for smart grids, in: Proceedings of 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm 2014), IEEE, 2014, pp. 794–799. doi:10.1109/SmartGridComm.2014.7007745.

[8] I. Melatti, F. Mari, T. Mancini, M. Prodanovic, E. Tronci, A two-layer near-optimal strategy for substation constraint management via home batteries, IEEE Transactions on Industrial Electronics 69 (2022) 8566–8578. doi:10.1109/TIE.2021.3102431.

[9] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, M. Prodanovic, L. Elmegaard, User flexibility aware price policy synthesis for smart grids, in: Proceedings of 18th Euromicro Conference on Digital System Design (DSD 2015), IEEE, 2015, pp. 478–485. doi:10.1109/DSD.2015.35.

[10] D. Goswami, R. Schneider, A. Masrur, M. Lukasiewycz, S. Chakraborty, H. Voit, A. Annaswamy, Challenges in automotive cyber-physical systems design, in: Proceedings of International Conference on Embedded Computer Systems (SAMOS 2012), IEEE, 2012, pp. 346–354. doi:10.1109/SAMOS.2012.6404199.

[11] S. Chakraborty, M. Al Faruque, W. Chang, D. Goswami, M. Wolf, Q. Zhu, Automotive cyber–physical aystems: A tutorial introduction, IEEE Design & Test 33 (2016) 92–108. doi:10.1109/MDAT.2016.2573598.

[12] L. Zhang, Modeling automotive cyber physical systems, in: Proceedings of 12th International Symposium on Distributed Computing and Applications to Business, Engineering & Science (DCABES 2013), IEEE, 2013, pp. 71–75. doi:10.1109/DCABES.2013.20.

[13] M. Hengartner, T. Kruger, K. Geraedts, E. Tronci, T. Mancini, F. Ille, M. Egli, S. Roeblitz,

R. Ehrig, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, B. Leeners, Negative affect is unrelated to fluctuations in hormone levels across the menstrual cycle: Evidence from a multisite observational study across two successive cycles, Journal of Psychosomatic Research 99 (2017) 21–27. doi:10.1016/j.jpsychores.2017.05.018.
[14] B. Leeners, T. Krüger, K. Geraedts, E. Tronci, T. Mancini, M. Egli, S. Röblitz, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, F. Ille, Associations between natural physiological and supraphysiological estradiol levels and stress perception, Frontiers in Psychology 10 (2019) 1296. doi:10.3389/fpsyg.2019.01296.
[15] F. Maggioli, T. Mancini, E. Tronci, SBML2Modelica: Integrating biochemical models within open-standard simulation ecosystems, Bioinformatics 36 (2020) 2165–2172. doi:10.1093/bioinformatics/btz860.
[16] T. Mancini, F. Mari, A. Massini, I. Melatti, I. Salvo, S. Sinisi, E. Tronci, R. Ehrig, S. Röblitz, B. Leeners, Computing personalised treatments through in silico clinical trials. A case study on downregulation in assisted reproduction, in: Proceedings of 25th RCRA International Workshop on Experimental Evaluation of Algorithms for Solving Problems with Combinatorial Explosion (RCRA 2018), volume 2271 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018.
[17] S. Sinisi, V. Alimguzhin, T. Mancini, E. Tronci, B. Leeners, Complete populations of virtual patients for in silico clinical trials, Bioinformatics 36 (2020) 5465–5472. doi:10.1093/bioinformatics/btaa1026.
[18] S. Sinisi, V. Alimguzhin, T. Mancini, E. Tronci, F. Mari, B. Leeners, Optimal personalised treatment computation through in silico clinical trials on patient digital twins, Fundamenta Informaticae 174 (2020) 283–310. doi:10.3233/FI-2020-1943.
[19] M. Esposito, L. Picchiami, Simulation-based synthesis of personalised therapies for colorectal cancer, in: Proceedings of 3rd Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis (OVERLAY 2021), volume 2987 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2021, pp. 109–113.
[20] M. Esposito, L. Picchiami, Intelligent search for personalized cancer therapy synthesis: an experimental comparison, in: Proceedings of 9th Italian workshop on Planning and Scheduling (IPS 2021) and the 28th RCRA International Workshop on Experimental Evaluation of Algorithms for Solving Problems with Combinatorial Explosion (RCRA 2021), volume 3065 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2021, pp. 69–84.
[21] M. Esposito, L. Picchiami, A comparative study of AI search methods for personalised cancer therapy synthesis in copasi, in: Proceedings of 21st International Conference of the Italian Association for Artificial Intelligence, (AI*IA 2022), volume 13196 of *Lecture Notes in Computer Science*, Springer, 2022, pp. 638–654.
[22] D. Teutonico, F. Musuamba, H. Maas, A. Facius, S. Yang, M. Danhof, O. Della Pasqua, Generating virtual patients by multivariate and discrete re-sampling techniques, Pharmaceutical research 32 (2015) 3228–3237.
[23] R. Allen, T. Rieger, C. Musante, Efficient generation and selection of virtual populations in quantitative systems pharmacology models, CPT: Pharmacometrics & Systems Pharmacology 5 (2016) 140–146. doi:10.1002/psp4.12063.
[24] P. Balazki, S. Schaller, T. Eissing, T. Lehr, A quantitative systems pharmacology kidney model of diabetes associated renal hyperfiltration and the effects of sglt inhibitors, CPT:

Pharmacometrics & Systems Pharmacology 7 (2018) 788–797. doi:10.1002/psp4.12359.

[25] G. Della Penna, B. Intrigila, I. Melatti, M. Minichino, E. Ciancamerla, A. Parisse, E. Tronci, M. Venturini Zilli, Automatic verification of a turbogas control system with the murphi verifier, in: Proceedings of 6th International Workshop on Hybrid Systems: Computation and Control (HSCC 2003), volume 2623 of *Lecture Notes in Computer Science*, Springer, 2003, pp. 141–155.

[26] G. Della Penna, B. Intrigila, I. Melatti, E. Tronci, M. Venturini Zilli, Finite horizon analysis of Markov chains with the Murphi verifier, International Journal on Software Tools for Technology Transfer 8 (2006) 397–409. doi:10.1007/s10009-005-0216-7.

[27] M. Cadoli, T. Mancini, F. Patrizi, SAT as an effective solving technology for constraint problems, in: Proceedings of 16th International Symposium on Foundations of Intelligent Systems (ISMIS 2006), volume 4203 of *Lecture Notes in Computer Science*, Springer, 2006, pp. 540–549.

[28] M. Cadoli, T. Mancini, Combining relational algebra, SQL, constraint modelling, and local search, Theory and Practice of Logic Programming 7 (2007) 37–65. doi:10.1017/S1471068406002857.

[29] T. Mancini, P. Flener, J. Pearson, Combinatorial problem solving over relational databases: View synthesis through constraint-based local search, in: Proceedings of ACM Symposium on Applied Computing (SAC 2012), ACM, 2012, pp. 80–87. doi:10.1145/2245276.2245295.

[30] G. Gottlob, G. Greco, T. Mancini, Conditional constraint satisfaction: Logical foundations and complexity, in: Proceedings of 20th International Joint Conference on Artificial Intelligence (IJCAI 2007), 2007, pp. 88–93.

[31] T. Mancini, M. Cadoli, D. Micaletto, F. Patrizi, Evaluating ASP and commercial solvers on the CSPLib, Constraints 13 (2008) 407–436.

[32] L. Bordeaux, M. Cadoli, T. Mancini, CSP properties for quantified constraints: Definitions and complexity, in: Proceedings of 20th National Conference on Artificial Intelligence (AAAI 2005), AAAI, 2005, pp. 360–365.

[33] T. Mancini, E. Tronci, A. Scialanca, F. Lanciotti, A. Finzi, R. Guarneri, S. Di Pompeo, Optimal fault-tolerant placement of relay nodes in a mission critical wireless network, in: Proceedings of 25th RCRA International Workshop on Experimental Evaluation of Algorithms for Solving Problems with Combinatorial Explosion (RCRA 2018), volume 2271 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018.

[34] Q. Chen, A. Finzi, T. Mancini, I. Melatti, E. Tronci, MILP, pseudo-boolean, and OMT solvers for optimal fault-tolerant placements of relay nodes in mission critical wireless networks, Fundamenta Informaticae 174 (2020) 229–258. doi:10.3233/FI-2020-1941.

[35] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, An efficient algorithm for network vulnerability analysis under malicious attacks, in: Proceedings of The 24th International Symposium on Methodologies for Intelligent Systems (ISMIS 2018), Springer, 2018.

[36] R. Grosu, S. Smolka, Monte Carlo model checking, in: Proceedings of 11th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2005), volume 3440 of *Lecture Notes in Computer Science*, Springer, 2005, pp. 271–286.

[37] T. Mancini, E. Tronci, I. Salvo, F. Mari, A. Massini, I. Melatti, Computing biological model

parameters by parallel statistical model checking, in: Proceedings of 3rd International Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2015), volume 9044 of *Lecture Notes in Computer Science*, Springer, 2015, pp. 542–554.

[38] E. Tronci, T. Mancini, I. Salvo, S. Sinisi, F. Mari, I. Melatti, A. Massini, F. Davi', T. Dierkes, R. Ehrig, S. Röblitz, B. Leeners, T. Krüger, M. Egli, F. Ille, Patient-specific models from inter-patient biological models and clinical records, in: Proceedings of 14th International Conference on Formal Methods in Computer-Aided Design (FMCAD 2014), IEEE, 2014, pp. 207–214.

[39] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, L. Elmegaard, Parallel statistical model checking for safety verification in smart grids, in: Proceedings of 2018 IEEE International Conference on Smart Grid Communications (SmartGridComm 2018), IEEE, 2018. doi:10.1109/SmartGridComm.2018.8587416.

[40] V. Mnih, C. Szepesvári, J. Audibert, Empirical bernstein stopping, in: Proceedings of 25th International Conference on Machine Learning (ICML 2008), Ass. Comp. Mach., 2008, pp. 672—679. doi:10.1145/1390156.1390241.

[41] P. Zuliani, A. Platzer, E. Clarke, Bayesian statistical model checking with application to Stateflow/Simulink verification, Formal Methods in System Design 43 (2013) 338–367. doi:10.1007/s10703-013-0195-3.

[42] L. Bortolussi, D. Milios, G. Sanguinetti, Smoothed model checking for uncertain continuous-time markov chains, Information and Computation 247 (2016) 235–253. doi:https://doi.org/10.1016/j.ic.2016.01.004.

[43] T. Mancini, F. Mari, A. Massini, I. Melatti, E. Tronci, SyLVaaS: System level formal verification as a service, in: Proceedings of 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP 2015), IEEE, 2015, pp. 476–483. doi:10.1109/PDP.2015.119.

[44] T. Mancini, F. Mari, A. Massini, I. Melatti, E. Tronci, System level formal verification via distributed multi-core hardware in the loop simulation, in: Proceedings of 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP 2014), IEEE, 2014, pp. 734–742. doi:10.1109/PDP.2014.32.

[45] T. Mancini, F. Mari, A. Massini, I. Melatti, E. Tronci, SyLVaaS: System level formal verification as a service, Fundamenta Informaticae 149 (2016) 101–132. doi:10.3233/FI-2016-1444.

[46] T. Mancini, F. Mari, A. Massini, I. Melatti, I. Salvo, E. Tronci, On minimising the maximum expected verification time, Information Processing Letters 122 (2017) 8–16. doi:10.1016/j.ipl.2017.02.001.

[47] M. Esposito, AI-guided optimal deployments of drone- intercepting systems in large critical areas, in: Proceedings of 3rd Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis (OVERLAY 2021), volume 2987 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2021, pp. 97–101.

[48] G. Agha, K. Palmskog, A survey of statistical model checking, ACM Transactions on Modeling and Computer Simulation 28 (2018) 6:1–6:39. doi:10.1145/3158668.

[49] A. Pappagallo, A. Massini, E. Tronci, Monte Carlo based Statistical Model Checking of Cyber-Physical Systems: a Review, Information 11 (2020) 588. doi:10.3390/info11120588.

[50] S. Sebastio, A. Vandin, MultiVeStA: Statistical model checking for discrete event simulators, in: Proceedings of 7th International Conference on Performance Evaluation Methodologies and Tools (ValueTools 2013), ICST/ACM, 2013, pp. 310–315.
[51] M. Kwiatkowska, G. Norman, D. Parker, Prism 4.0: Verification of probabilistic real-time systems, in: Proceedings of 23rd International Conference on Computer Aided Verification (CAV 2011), volume 6806 of *Lecture Notes in Computer Science*, Springer, 2011, pp. 585–591.
[52] A. David, K. Larsen, A. Legay, M. Mikučionis, D. Poulsen, Uppaal smc tutorial, International Journal on Software Tools for Technology Transfer 17 (2015) 397–415. doi:10.1007/s10009-014-0361-y.
[53] P. Ballarini, B. Barbot, M. Duflot, S. Haddad, N. Pekergin, HASL: a new approach for performance evaluation and model checking from concepts to experimentation, Performance Evaluation 90 (2015) 53–77. doi:10.1016/j.peva.2015.04.003.
[54] H. Younes, Ymer: A statistical model checker, in: Proceedings of 17th International Conference on Computer Aided Verification (CAV 2005), volume 3576 of *Lecture Notes in Computer Science*, Springer, 2005, pp. 429–433. doi:10.1007/11513988_43.
[55] H. Younes, R. Simmons, Probabilistic verification of discrete event systems using acceptance sampling, in: Proceedings of 14th International Conference on Computer Aided Verification (CAV 2002), volume 2404 of *Lecture Notes in Computer Science*, Springer, 2002, pp. 223–235. doi:10.1007/3-540-45657-0_17.
[56] C. Baier, B. Haverkort, H. Hermanns, J.-P. Katoen, Model-checking algorithms for continous-time markov chains, IEEE Transactions on Software Engineering 29 (2003) 524–541. doi:10.1109/TSE.2003.1205180.
[57] G. Norman, D. Parker, J. Sproston, Model checking for probabilistic timed automata, Formal Methods in System Design 43 (2013) 164–190. doi:10.1007/s10703-012-0177-x.
[58] L. Valiant, A theory of the learnable, Communications of the ACM 27(11) (1984) 1134–1142.
[59] S. A. Goldman, R. H. Sloan, Can pac learning algorithms tolerate random attribute noise?, Algorithmica 14 (1995) 70–84.
[60] P. Jiang, Q. Zhou, X. Shao, Verification methods for surrogate models, in: Surrogate Model-Based Engineering Design and Optimization, Springer, 2020, 89–113.
[61] M. Pedergnana, S. G. García, et al., Smart sampling and incremental function learning for very large high dimensional data, Neural Networks 78 (2016) 75–87.
[62] B. Xue, M. Fränzle, H. Zhao, N. Zhan, A. Easwaran, Probably approximate safety verification of hybrid dynamical systems, in: Proceedings of Formal Methods and Software Engineering - 21st International Conference on Formal Engineering Methods (ICFEM 2019), volume 11852 of *Lecture Notes in Computer Science*, Springer, 2019. doi:https://doi.org/10.1007/978-3-030-32409-4_15.
[63] S. Hanneke, The optimal sample complexity of pac learning, The Journal of Machine Learning Research 17 (2016) 1319–1333.
[64] W. Hoeffding, Probability inequalities for sums of bounded random variables, in: The collected works of Wassily Hoeffding, Springer, 1994, pp. 409–426.
[65] X. Qin, Y. Xian, A. Zutshi, C. Fan, J. V. Deshmukh, Statistical verification of cyber-physical systems using surrogate models and conformal inference, in: 2022 ACM/IEEE 13th International Conference on Cyber-Physical Systems (ICCPS), IEEE, 2022, pp. 116–126.