Statistical Model Checking for the Analysis of Mission- and Safety-Critical Cyber-Physical Systems

Angela Pappagallo

Computer Science Dept., Sapienza University of Rome, Italy

Abstract

Many autonomous Cyber-Physical Systems (CPSs) (*e.g.*, autonomous vehicles, IoT or medical devices, etc.) are mission- or safety-critical (*i.e.*, errors may result in, resp., loss of money or human deaths). This motivates research for their efficient *formal* verification. Unfortunately, verifying a CPS entails evaluating a *prohibitively huge* number of scenarios. In this short paper, we show the maturity, feasibility and flexibility of Statistical Model Checking by reviewing 3 recent case studies of its successful application to real-world mission- and safety-critical CPSs in areas as diverse as smart grids, *in silico* medicine, wireless sensor networks.

1. Introduction

In a Cyber-Physical System (CPS), a (continuous) physical system (*plant*) is controlled and/or monitored by a (discrete) software. The deployment of autonomous CPSs [1], such as, *e.g.*, devices for Internet of Things (IoT) [2], Unmanned Autonomous Vehicles [3] and medical devices [4], has been speeding up for the last decades, with a projected $1.1 trillion global speding on IoT only [5]. For many of such CPSs, it is important to rule out errors [6], especially in the software part, since they may lead to *a*) loss of money in *mission-critical systems* [7] (*e.g.*, the 1996 Ariane 5 rocket incident, due to a software type conversion error, resulted in a $500 million loss); *b*) death or serious injury for people in *safety-critical systems* [8] (*e.g.*, clinical treatments or medical devices).

Unfortunately, standard testing could not provide the required degree of correctness assurance, and this motivates research on efficient formal verification methods. There are multiple challenges to overcome when formally verifying a CPS, *e.g.*, the complexity of the system dynamics, the huge number of scenarios to be evaluated (*scenario explosion*, *e.g.*, [9, 10, 11, 12, 13, 14, 15]) and the lack of a unified mathematical model for both the discrete cyber and the continuous physical parts (*e.g.*, [16, 17]). Such issues make it hard to apply analytical approaches based on logics or automata, *e.g.*, [18, 19, 20, 21, 22].

Statistical Model Checking (SMC) [23, 24] aims at overcoming such obstructions by using statistical methods to *sample* the set of scenarios up to desired accuracy and precision, while possibly relying on *black-box models* of the System Under Verification (*i.e.*, the full system encompassing both the software and the plant), for example available only via a *simulator*.

OVERLAY 2021: 3rd Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, September 22, 2021, Padova, Italy

In this short paper, we review 3 recent real-world case studies, from very diverse application areas (smart grids, *in silico* medicine, wireless sensor networks), which were successfully addressed via SMC. This shows the feasibility and flexibility of SMC when applied to real-world mission- and safety-critical systems. A complete survey of SMC methodologies can be found in, *e.g.*, [23, 24].

2. Peak Shaving in Smart Grids

An Electric Distribution Network (EDN) [25] is composed of several substations, each serving a set of residential households. By using the measurements taken from the home electricity mains, we know each house power demand, with periodicity at least one hour. Our objective is to reduce costs for the Distribution System Operator (DSO), by limiting the demand drawn at each substation at times of peak demand (*peak shaving*). This reduces the amount of electricity purchased on the market at peak prices, and reduces overloading of network components (hence, substation ageing).

Many works address this problem. Here, we focus on the methodology in [27, 28, 29, 26, 30], for which SMC-based verification is proposed. Namely, the problem of achieving peak shaving is solved by proposing two intelligent services. Whilist the *EDN Virtual Tomography (EVT)* service computes time-varying upper bounds for the Aggregated Power Demand (APD) of the households U connected to a substation s yielding low operational costs for the DSO, the *Demand-Aware Price Policy (DAPP)* service computes *individualised time-varying* upper bounds for the demand of *each* household in U. If a household keeps its demand below such bounds, a low energy tariff is applied, otherwise an high tariff is applied. If all households succeed in keeping their demand below their bounds (by performing *load shifting*), the APD on s will be below the bound computed by EVT. However, there is no guarantee that such an *indirect steering* of the demand of each household will be successful. In [26], a domain-specific highly parallel statistical model checker (Aggregated Power Demand Analyzer, APD-A) is designed, which estimates the *probability distribution* of the APD given probabilistic deviations from the expected demands of each household. Figure 1 shows the APD-A results for a set of 186 houses in Denmark.

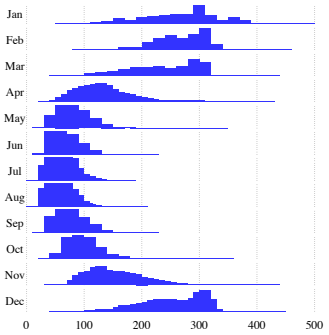

Figure 1: APD prob. [26]

3. Virtual Patients for In Silico Clinical Trials

A major problem in medicine is assessing safety and efficacy of pharmacological drugs, medical devices and treatment strategies. In the last years, a research area called In Silico Clinical Trials (ISCT) has emerged [31], with the aim of using Computer Science techniques to decrease time and cost for the experimentations, reducing animal and human testing, prioritising *in vivo* clinical trials, and enabling precision medicine. A

cornerstone of ISCT is the *simulation* of the therapy/device under assessment on a *population* of Virtual Patients (VPs). VPs are typically computed by parameterising quantitative mechanistic Virtual Physiological Human (VPH) models, in turn defined by encoding qualitative knowledge of the human physiology of interest [32, 33]. For an ISCT to provide *compelling evidence* of the safety/efficacy of a therapy, such populations of VPs must be *complete*, i.e., representative of the *entire spectrum* of behaviours deemed of interest (in order not to skip significant behaviors).

In [35, 36, 34], SMC is used to drive intelligent global search in the VPs parameter space to compute a complete population of VPs starting from a (non-identifiable) VPH model and suitable biological knowledge. The effectiveness of the approach was proven on a (differential equation–based) model of the female hypothalamic-pituitary-gonadal axis [37], for which a population of as many as 4,830,264 VPs stratified into layers at different level of granularity of behaviours was generated (Figure 2 shows the possible time courses of the Estradiol hormone in the different VP layers), and whose completeness was evaluated against retrospective health records. Such VPs were then used in [38, 39] to compute, again *in silico*, optimal robust personalised treatments for assisted reproduction, an area currently showing many factors that can be hardly kept under full control [40, 41, 42]. Namely, *digital twins* of human patients were computed by selecting those VPs best matching clinical measurements on them, and a black-box simulator of the VPH model in [37] was driven [43] via intelligent backtracking on such digital twins.

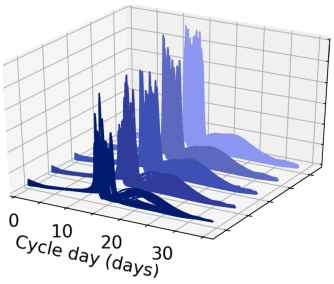

Figure 2: Estradiol [34]

4. Wireless Sensors Network

The last case study consists of a low-level engineering application, namely an audio streaming application over a Wi-Fi network. Such an application is representative of a wide area of applications on networked systems [44, 45]. In such a network, several nodes are equipped with microphones which produce different audio streams and are transmitted to a base station equipped with a speaker to play the received audio. The goal is to ensure the synchronization between the different nodes of the network, in order to guarantee a consistent audio output. To this extent, in [46, 47] a Phase Locked Loop (PLL) synchronization protocol [48] is designed so that all nodes in the network agree on a synchronized clock, within a $1\mu s$ tolerance. In order to show that the PLL synchronization protocol fulfills the main design requirement, the SBIP statistical model checker [49], which is based on the Behaviour, Interaction, Priority (BIP) framework [50], is used. Namely, the following property were verified: it must hold that the difference between the Master clock θ_m and the software clock, computed in every Slave θ_s, must be within a given bound Δ with high probability and accuracy. The obtained result was that, for the considered setting, the smallest bound that ensures the synchronisation is $\Delta = 76\mu s$.

5. Conclusions

In this short paper, we showed the maturity, feasibility and flexibility of Statistical Model Checking in the verification of real-world mission- and safety-critical CPSs, by reviewing some of its recent applications to real-world case studies, in areas as diverse as smart grids, *in silico* medicine, wireless sensor networks.

References

[1] R. Alur, Principles of Cyber-Physical Systems, MIT, 2015.
[2] M. Zimmerling, *et al.*, Adaptive real-time communication for wireless cyber-physical systems, ACM TCPS 1 (2017).
[3] W. Koch, *et al.*, Reinforcement learning for UAV attitude control, ACM TCPS 3 (2019).
[4] N. Dey, *et al.*, Medical cyber-physical systems: A survey, J Med Sys 42 (2018).
[5] Statistics on IoT Spending, https://www.statista.com/topics/2637/internet-of-things, 2021.
[6] B. Dowdeswell, *et al.*, Finding faults: A scoping study of fault diagnostics for industrial cyber–physical systems, J Sys Softw 168 (2020)
[7] A. Banerjee, *et al.*, Ensuring safety, security, and sustainability of mission-critical cyber–physical systems, Proc IEEE 100 (2012).
[8] R. Mitchell, *et al.*, Behavior rule specification-based intrusion detection for safety critical medical cyber physical systems, IEEE Trans Dep Sec Comp 12 (2015).
[9] T. Mancini, *et al.*, System level formal verification via distributed multi-core hardware in the loop simulation, in: PDP 2014, IEEE, 2014.
[10] T. Mancini, *et al.*, SyLVaaS: System level formal verification as a service, in: PDP 2015.
[11] T. Mancini, *et al.*, Anytime system level verification via random exhaustive hardware in the loop simulation, in: DSD 2014, IEEE, 2014.
[12] T. Mancini, *et al.*, Anytime system level verification via parallel random exhaustive hardware in the loop simulation, Micropr Microsys 41 (2016).
[13] T. Mancini, *et al.*, SyLVaaS: System level formal verification as a service, Fund Inf 149(1–2) (2016).
[14] T. Mancini, *et al.*, On minimising the maximum expected verification time, IPL 122 (2017).
[15] T. Mancini, *et al.*, Any-horizon uniform random sampling and enumeration of constrained scenarios for simulation-based formal verification, IEEE TSE (2021). To appear.
[16] E. A. Lee, Fundamental limits of cyber-physical systems modeling, ACM TCPS 1 (2016).
[17] F. Maggioli, *et al.*, SBML2Modelica: Integrating biochemical models within open-standard simulation ecosystems, Bioinf. 36 (2020).
[18] G. Della Penna, *et al.*, Bounded probabilistic model checking with the Murφ verifier, in: FMCAD 2004, IEEE, 2004.
[19] M. Cadoli, *et al.*, SAT as an effective solving technology for constraint problems, in: ISMIS 2006, *LNCS* 4203, Springer, 2006.
[20] M. Cadoli, *et al.*, Combining relational algebra, SQL, constraint modelling, and local search, Theor Pract Logic Programm 7 (2007).
[21] T. Mancini, *et al.*, Combinatorial problem solving over relational databases: View synthesis through constraint-based local search, in: SAC 2012, ACM, 2012.
[22] F. Mari, *et al.*, Model based synthesis of control software from system level formal specifications, ACM TOSEM 23 (2014).
[23] G. Agha, *et al.*, A survey of statistical model checking, ACM Tran Mod Com Sim 28 (2018).

[24] A. Pappagallo, *et al.*, Monte carlo based statistical model checking of cyber-physical systems: A review, Information 11 (2020).
[25] D. R. Patrick, *et al.*, Electrical Distribution Systems, 2nd Ed., Pearson, 2009.
[26] T. Mancini, *et al.*, Parallel statistical model checking for safety verification in smart grids, in: SmartGridComm 2018, IEEE, 2018.
[27] T. Mancini, *et al.*, Demand-aware price policy synthesis and verification services for smart grids, in: SmartGridComm 2014, IEEE, 2014.
[28] T. Mancini, *et al.*, User flexibility aware price policy synthesis for smart grids, in: DSD 2015.
[29] B. Hayes, *et al.*, Residential demand management using individualised demand aware price policies, IEEE Trans Smart Grid 8 (2017).
[30] I. Melatti, *et al.*, A two-layer near-optimal strategy for substation constraint management via home batteries, IEEE Trans Ind Elect (2021). To appear.
[31] F. Pappalardo, *et al.*, *In silico* clinical trials: concepts and early adoptions, Brief Bioinf 20 (2019).
[32] M. Kanehisa, *et al.*, Kegg: New perspectives on genomes, pathways, diseases and drugs, Nucl Ac Res 45 (2017).
[33] A. Fabregat, *et al.*, The Reactome pathway knowledgebase, Nucl Ac Res 46 (2018).
[34] S. Sinisi, *et al.*, Complete populations of virtual patients for in silico clinical trials, Bioinf 36 (2021).
[35] E. Tronci, *et al.*, Patient-specific models from inter-patient biological models and clinical records, in: FMCAD 2014, IEEE, 2014.
[36] T. Mancini, Computing biological model parameters by parallel statistical model checking, in: IWBBIO 2015, *LNCS* 9044, Springer, 2015.
[37] S. Röblitz, *et al.*, A mathematical model of the human menstrual cycle for the administration of GnRH analogues, J Theor Biol 321 (2013).
[38] T. Mancini, *et al.*, Computing personalised treatments through in silico clinical trials. A case study on downregulation in assisted reproduction, in: RCRA 2018, *CEUR* 2271, 2018.
[39] S. Sinisi, *et al.*, Optimal personalised treatment computation through in silico clinical trials on patient digital twins, Fund Inf 174 (2020).
[40] B. Leeners, *et al.*, Lack of associations between female hormone levels and visuospatial working memory, divided attention and cognitive bias across two consecutive menstrual cycles, Front Behav Neur 11 (2017).
[41] M. Hengartner, *et al.*, Negative affect is unrelated to fluctuations in hormone levels across the menstrual cycle: Evidence from a multisite observational study across two successive cycles, J Psych Res 99 (2017).
[42] B. Leeners, *et al.*, Associations between natural physiological and supraphysiological estradiol levels and stress perception, Front Psycol 10 (2019).
[43] S. Sinisi, *et al.*, Reconciling interoperability with efficient verification and validation within open source simulation environments, Sim Mod Pract Theory 109 (2021).
[44] A. Lekidis, *et al.*, A model-based design flow for can-based systems, in: iCC 2013.
[45] A. Lekidis, *et al.*, Using BIP to reinforce correctness of resource-constrained iot applications, in: SIES 2015, IEEE, 2015.
[46] A. Lekidis, *et al.*, Building distributed sensor network applications using BIP, in: SAS 2015.
[47] A. Nouri, *et al.*, Performance Evaluation of Stochastic Real-Time Systems with the SBIP Framework, Int J Critical Comp-Based Sys (2018).
[48] K. Choi, *et al.*, Phase-Locked Loop and Synchronization, Wiley-IEEE, 2016.
[49] B. L. Mediouni, *et al.*, SBIP 2.0: Statistical Model Checking Stochastic Real-time Systems, in: ATVA 2018, Springer, 2018.
[50] A. Basu, *et al.*, Modeling heterogeneous real-time components in BIP, in: SEFM 2006.