Statistical Model Checking as an Effective Technology to Formally Analyze Industry-Relevant Cyber-Physical Systems

Angela Pappagallo

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

Abstract

Many autonomous Cyber-Physical Systems (*e.g.*, devices for Internet of Things, Unmanned Autonomous Vehicles, medical devices, etc) are mission-critical (*i.e.*, errors result in loss of money) or safety-critical (*i.e.*, errors result in damage or even death for humans). This motivates research on efficient formal verification methods for such Cyber-Physical Systems.

Unfortunately, this is not an easy task, as verifying a Cyber-Physical System entails evaluating a huge number of scenarios (*scenario explosion*). Furthermore, a unified mathematical model for the (discrete) cyber part and the (continuous) physical part is currently not available. Such obstructions may be mitigated by using Statistical Model Checking, which uses statistical methods to sample the set of scenarios while basing on possibly black-box models of the System Under Verification.

In this paper, we review 5 recent real-world and industry-relevant case studies from the literature that involved usage of Statistical Model Checking. Such case studies range on very different application areas, namely: i) intelligent services for peak shaving in smart grids, ii) In-Silico Clinical Trial for medical services, iii) applications for wireless sensor networks; iv) aircraft data networks; v) plug-in electric vehicles. This shows the maturity, feasibility and flexibility of Statistical Model Checking when applied to real-world case studies.

CEUR-WS.org/Vol-3065/paper5_151.pdf

1. Introduction

A Cyber-Physical System (CPS) is a system where a (continuous) physical system (*plant*) is controlled and/or monitored by a (discrete) software. The deployment of autonomous CPSs [3], such as, *e.g.*, devices for Internet of Things (IoT) [11, 88], Unmanned Autonomous Vehicles [35] and medical devices [20], has been speeding up for the last decades, with a projected 1.1 trillion USD global speding on IoT only [81]. For many of such CPSs, it is important to rule out errors [21, 22], especially bugs in the software part, since such bugs may lead to:

- loss of money in *mission-critical systems* [8]. This is the case, *e.g.*, in aerospace: as an example, in 1996 the Ariane 5 [4] rocket was destroyed after launch due to a type conversion error in the software, resulting in a 500 M$ loss;

IPS-RCRA 2021: 9th Italian Workshop on Planning and Scheduling and 28th International Workshop on Experimental Evaluation of Algorithms for Solving Problems with Combinatorial Explosion

© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

- death of serious injury for people in *safety-critical systems* [66]. This is the case, *e.g.*, for medical devices.

As standard testing could not provide the required degree of correctness assurance, this motivates research on efficient formal verification methods [16]. There are multiple challenges to overcome when formally verifying a CPS [17], *e.g.*, the huge number of scenarios to be evaluated (*scenario explosion*, *e.g.*, [49, 53, 52, 54]), which is hard to tackle also using High-Performance Computing (HPC) [55, 56, 51, 60]. Furthermore, much research must still be done in order to find a unified mathematical model for the discrete cyber part and the continuous physical part [38, 45]. Such issues make it hard to apply analytical approaches based on logics (*e.g.*, [14, 26, 47, 13, 48]) or automata (*e.g.*, [19, 46, 62]).

Statistical Model Checking (SMC) [41] holds the promise to overcome this obstacle by using statistical methods to sample the set of scenarios up to desired accuracy and precision [27, 28, 18], while possibly relying on black-box models of the System Under Verification (SUV) (*i.e.*, the full system encompassing both the software and the plant) [2, 5].

In this paper, we review 5 recent real-world and industry-relevant case studies from the literature that involved usage of SMC. Such case studies range on very different application areas, namely:

- verification of an intelligent service for peak shaving in smart grids;
- generation of Virtual Patients (VPs) to enable In-Silico Clinical Trial (ISCT) for medical services (Virtual Physiological Human [31, 36]);
- parameter estimation for an application to stream audio in wireless sensor networks;
- computation of network latency under different system parameters for an aircraft network;
- computation of confidence intervals for the probability of failures in the recharging process of a Plug-in Electric Vehicle (PEV).

This shows the feasibility and flexibility of SMC when applied to real-world case studies. A preliminary version of this paper has been presented in [68]. Here we discuss more case studies, by also providing more details about methodologies and results. For a complete survey of SMC methodologies themselves, see, *e.g.*, [69, 1, 74, 7].

2. Real-World Case Studies for Statistical Model Checking

This section discusses some recent real-world and industry-relevant problems that have been solved by using SMC or SMC-based methodologies. Namely, Section 2.1 shows an application in the field of intelligent services for smart grids, Section 2.2 presents an SMC-based methodology used for enabling ISCT in Virtual Physiological Human (VPH),

Section 2.3 illustrates how SMC may be used in the field of wireless sensors networks, Section 2.4 discusses results on verifying network latency of an aircraft data network and finally Section 2.5 computes the probability of failures in PEV recharging.

2.1. Peak Shaving in Smart Grids

An Electric Distribution Network (EDN) [71] is composed of several substations, where each substation serves a set of residential houses. By using the measurements taken from the home electricity mains (Advanced Metering Infrastructure, AMI), 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 some or all substations of the EDN at times of peak demand (*peak shaving* [73]). In fact, this reduces costs of buying energy from the market at times of peak electricity price (which involves usage of peak power plants [65]), and reduces overloading of network components during times of peak demand (thus reducing substations aging), or during periods when the system is weakened due to line/transformer maintenance or other outages [83].

Many work in the literature address the problem above, see, *e.g.*, [34, 23, 86, 32, 77]. In this paper, we focus on the methodology in [29, 58, 59, 64], for which a verification based on SMC techniques is available. Namely, in that line of research the problem of achieving peak shaving is counteracted by proposing the two following intelligent services (for an high-level schema, see Figure 1).

1. The first service (EDN Virtual Tomography, EVT) computes time-varying upper bounds for the aggregated electricity demand resulting from the residential houses U connected to a given EDN substation s. As a result, if the aggregated demand of s is kept below such upper bounds, the DSO will save in the maintenance costs for s, as well as in energy production costs.

2. The second service (Demand-Aware Price Policy, DAPP) computes individualised time-varying upper bounds for each residential house in U. If a residential user keeps its demand below the bounds computed by DAPP, then a low energy tariff is applied, otherwise an high tariff is applied. Note that, in order to do this, residential users must perform *load shifting*, by consuming more electricity when the bound is high and less electricity when the bound is low. As a result, if all residential users succeeds in keeping their demand below the given bounds, the aggregated demand on s will be below the bound computed by EVT.

However, there is no guarantee that residential users will be able to perform load shifting so as to stay below the bounds computed by DAPP. In [57], a domain-specific statistical model checker named Aggregated Power Demand-Analyzer (APD-A) is designed, in order to compute the probability of violations of the bounds on the aggregated demand on s, given probabilistic deviations from the expected power demand (again, computed by DAPP) of each single house. More in detail, APD-A takes in input:

1. the time T on which to perform the evaluation (usually, one month divide in time-slots of one hour);

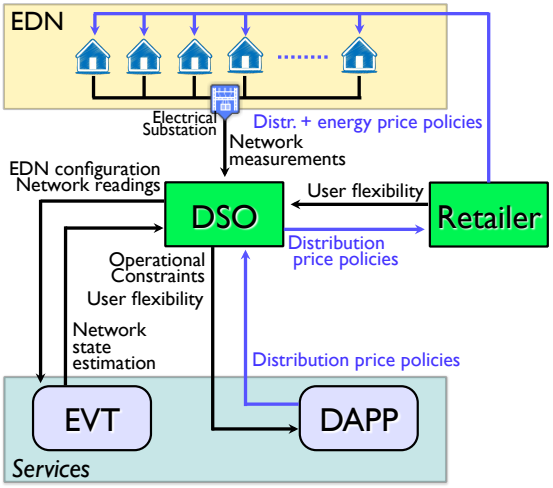

Figure 1: Intelligent systems for smart grids [29]

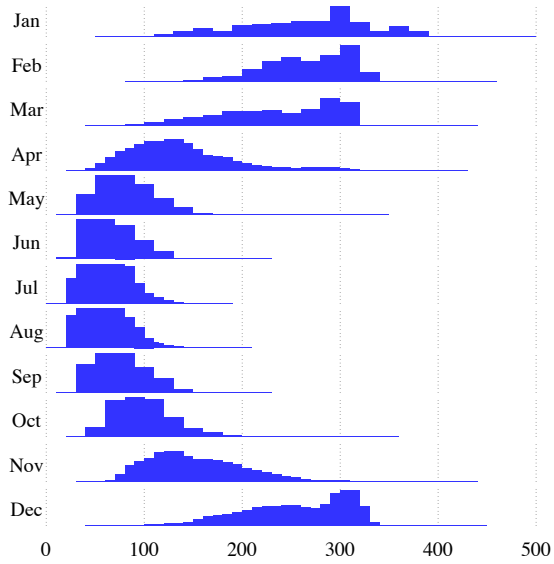

Figure 2: Results for Aggregated Power Demand-Analyzer from [5

2. for each user $u \in U$, the *Expected Power Profile (EPP)* $p_u : T \to \mathbb{R}$, i.e., a function taking as input a time-slot in $t \in T$ and returning the power demand $p_u(t)$ (in kW) of user u in t; such demand is a further output of DAPP and is always below the power bound for u in t (i.e., $p_u(t) \leq P_u(t)$, being $P_u(t)$ the upper bound output by DAPP, for all $u \in U, t \in T$;

3. a probabilistic model dev_u for users deviations from deviations from EPPs, i.e., $\int_a^b dev_u(x)dx =$ is the probability that actual power demand of u in any time-slot $t \in T$ is in $[(1+a)p_u(t), (1+b)p_u(t)]$ (e.g., $\int_{-0.02}^{0.02} dev_u(x)dx =$ probability that actual power demand of u in any time-slot $t \in T$ deviates at most by 2% from EPP of u);

4. the substation safety requirements, i.e., $p_s : T \to \mathbb{R}$ s.t., for each $t \in T$, the DSO wants the aggregated demand on s to be below $p_s(t)$;

5. parameters for the output probability distribution $0 < \delta, \varepsilon < 1$ and $\gamma \in \mathbb{R}$, i.e., the output values must be correct up to tolerance ε with statistical confidence $1 - \delta$, and the output probability distribution is discretized with step γ.

As an output, APD-A returns the probability distribution for the aggregated demand on s resulting from EPPs disturbed with the given probabilistic disturbance model dev_u. To this aim, APD-A relies on a parallel version (for cluster of computers with distributed memory) of the Optimal Approximation Algorithm (OAA) from [28]. Figure 2 shows the resulting output of APD-A for a group of 186 real-world houses in Denmark.

2.2. Virtual Patients for In-Silico Clinical Trials

One of the most complex problems in Medicine is assessing safety and efficacy of pharmacological drugs, medical devices and, more in general, treatment strategies [76]. In the

last years, a wide research area called ISCT has been developed [70, 6], with the aim to approach such a problem via Computer Science techniques. By prioritizing the successive *in vivo* experimentations, this would decrease time and cost of the overall process, reduce animal and human testing, and enable precision medicine [37, 24, 84, 85].

A key enabler to carry out an ISCT is the availability of a population of VPs, *i.e.*, a set computational models of the physiology of interest and of the Pharmacokinetics/Pharmacodynamics (PK/PD) of the relevant pharmacological compounds on which to perform computer simulations.

However, to guarantee compelling evidence of safety and efficacy of the therapy under assessment, such a population of VP must be *representative* of the entire spectrum of human phenotypes. This includes the possible individual differences in physiology and the different possible reactions to the external stimuli (*e.g.*, drug administrations).

Such computational, *quantitative, personalized* models of the human physiology and drugs PK/PD are typically derived in two steps. First, quantitative *inter-individual* VPH models are derived from *qualitative* knowledge from, *e.g.*, available repositories [33, 25], and are often formalized in terms of systems of *parametric* differential equations (for continuous-time models) or different equations (for discrete-time models). Different assignments to such (*real valued*) parameters yield different time courses (aka trajectories) of the modeled biological quantities, and different reactions to the same stimuli. Thus, a quantitative VPH model combined with a parameter assignment is regarded as a *Virtual Patient (VP)*, representing a *human phenotype*. Such VPs can then be *simulated* (typically as black-box systems via numerical simulators, given the complexity of the differential equations) to assess the values of proper metrics of the therapy of interest, *e.g.*, expected safety and efficacy (*In-Silico Clinical Trials*, ISCT).

Unfortunately, computing VPs is all but easy. Indeed, most of the legal assignments to a VPH model do yield model trajectories which clearly *violate* human physiology. This is because such models are often over-parameterised, and *unknown* inter-dependency constraints among the various parameters do exist. Also, parameters are often introduced to model not-well-understood biological mechanisms (see, *e.g.*, [82, 61]), or to abstract away details that are not needed to be modeled accurately to perform the planned verification activity. Also in this case, a random assignment to such parameters would yield, with very high probability, an overall model behavior which is clearly non-admissible from a biological standpoint.

The major obstacle is thus to automatically recognize whether a model parameter assignment is a (physiologically admissible) VP, and to *search* for such VPs in the (typically huge real-valued) space of model parameter assignments.

However this is not enough. Indeed, since, in order to carry out an ISCT we need a population of VPs *representative* of the entire spectrum of the phenotypes entailed by the VPH model, we need to search for *all* VPs satisfying the physiological admissibility criterion. Furthermore, since complex VPH models are often non-identifiable, it is often the case that several parameter assignments yield VPs which have *indistinguishable* (with respect to some given tolerance) trajectories under *all* time series of external stimuli (*e.g.*, drug administrations). The presence, in the computed population, of such indistinguishable VPs would be a major source of redundancy, hence inefficiency of the

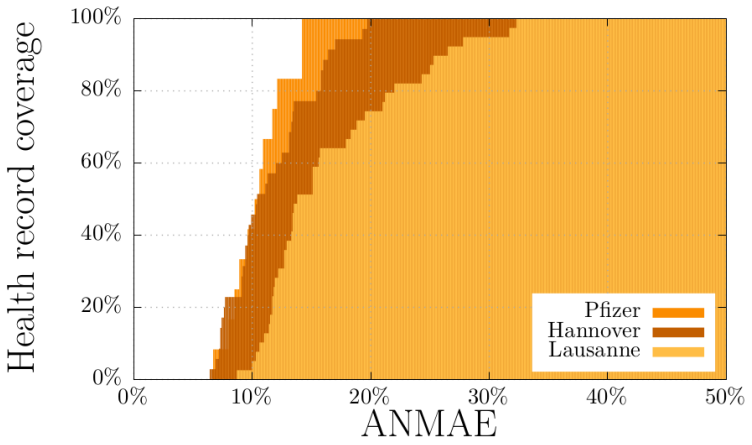

Figure 3: Results for Virtual Patients coverage from [79]

verification process, and should be avoided.

In [79], SMC-based techniques are used to drive *global search* (intelligently guided by an *heuristic*) in the VPs parameters space. Namely, starting from a (non-identifiable) VPH model and suitable biological and medical knowledge elicited from experts to formally define what a *physiologically admissible* trajectory is, such techniques compute a population of VPs which is representative of the entire spectrum of phenotypes entailed by the model and does not contain indistinguishable VPs, up to the user-requested *statistical guarantees*. Namely, given user-defined constants $\varepsilon, \delta \in (0, 1)$, when the algorithm terminates, the probability that further sampling will yield a VP showing an *unknown* phenotype (*i.e.*, a phenotype not already included in the population computed so far) is $\leq \varepsilon$ with statistical confidence $\geq 1 - \delta$.

The effectiveness of such approach has been proven on GynCycle [75], a non-identifiable model of the female Hypothalamic Pituitary Gonadal (HPG) axis, consisting of 33 highly non-linear stiff ordinary differential equations. Namely, a population of 4,830,264 VPs (each one being an assignment to 75 real-valued parameters) was generated and stratified into 7 levels (at different granularity of behaviours). The representativeness of such VPs was assessed against 86 retrospective health records from Pfizer, Hannover Medical School and University Hospital of Lausanne. Figure 3 shows that the datasets are respectively covered by such VPs within Average Normalised Mean Absolute Error (ANMAE) of 15%, 20%, and 35%.

The computed population of VPs was then used in [50, 80] 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, 30, 39]. 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 [75] was driven [78] via intelligent backtracking on such digital twins.

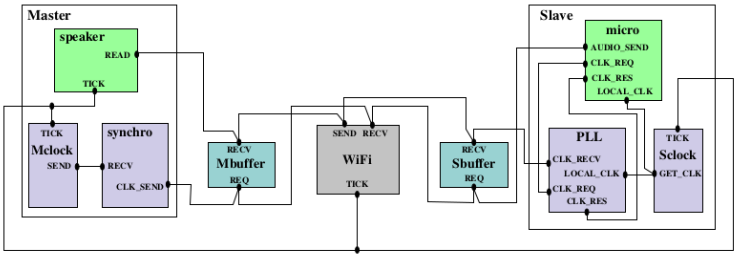

Figure 4: SBIP model of the wireless sensor network from [67]

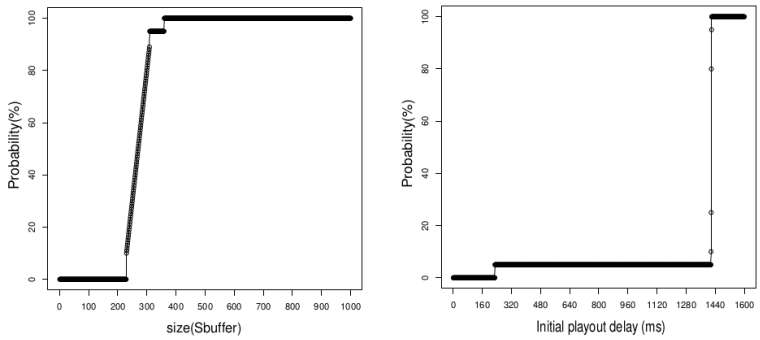

Figure 5: Results for properties ϕ_1 and ϕ_2 from [67]

2.3. Wireless Sensors Network

In this section we discuss 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 [43, 44]. 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 [42, 67] a Phase Locked Loop (PLL) synchronization master-slave protocol [15] 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, as well as to perform a *parameter tuning* of the main protocl parameters, the SBIP statistical model checker [63] is used, which is based on the Behaviour, Interaction, Priority (BIP) framework [10]. A schema of the SBIP model for the PLL is shown in Figure 4. Namely, the following three properties were verified:

ϕ_1 the size of the slave buffer must be below its maximum capacity (no overflow);

ϕ_2 the master buffer must be not empty (no underflow);

ϕ_3 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

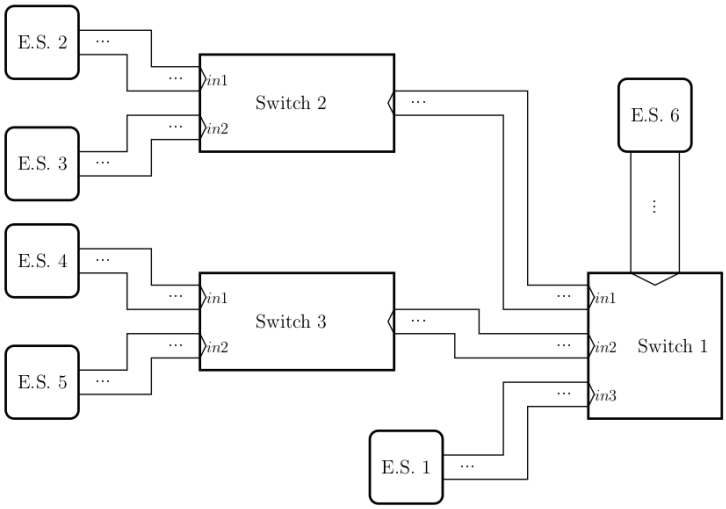

Figure 6: An AFDX architecture from [9]. E.S. stands for "end-system", which is a source of information

and accuracy.

For ϕ_1, many verifications were run, by varying the size of the slave buffer. Analogously, for ϕ_2 the initial playout (*i.e.*, the delay after which the master starts consuming from its buffer), was varied. The results are shown in Figure 5, as a function of the size of the buffer and of the initial delay, respectively. Of course, in both cases we are interested in the probability of ϕ_1 and ϕ_2 to be as close to 100% as possible. From the verification results it is possible to conclude that the optimal value for the slave buffer is 400 slots, while the optimal value for the inital playout is 1430 ms.

Finally, as for ϕ_3, again many verifications were run, by varying Δ as the specific time bound between the master clock and the software clock. The obtained result was that, for the considered setting, the smallest bound that ensures the synchronisation is $\Delta = 76\mu s$.

2.4. Avionics Full-Duplex Switched Ethernet

Avionics Full-Duplex Switched Ethernet (AFDX) is a network architecture used in many aircrafts, such as Airbus, Boeing, AgustaWeistland, Comac and many others. It has been patented by Airbus in order to provide data network connection inside aircrafts while maintaining deterministic quality of service.

In [9], SMC is used in order to estimate AFDX performances under various assumptions (scenarios), by focusing on a given AFDX architecture (see Figure 6, though the method may be easily generalized). The main idea is to model the network by replacing the network switches with a probability distribution for the delay experienced by a packet which traverses a given switch. Then, the network latency is estimated, when varying the following network parameters:

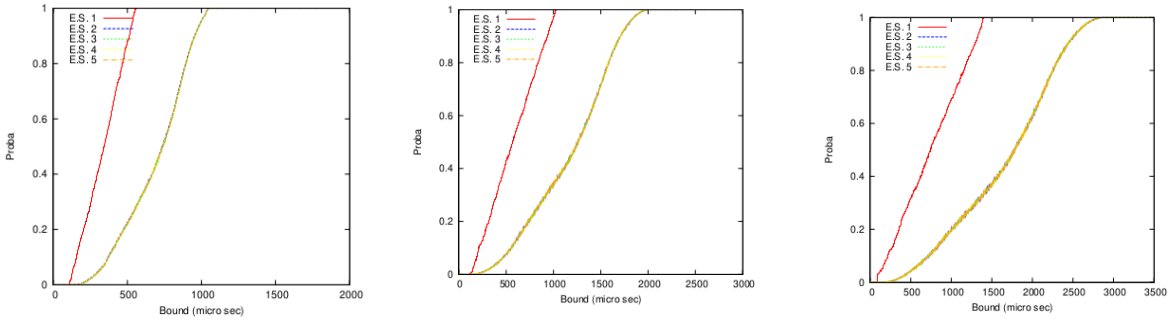

Figure 7: Results for AFDX verification (scenario 2 from [9]): left is for 10 VLs, center is for 20 VLs and right is for 30 VLs

- number of Virtual Links (VLs), *i.e.*, of logical unidirectional connections from one transmitter end-system to one or many receiver end-systems;
- number of frames (*i.e.*, packets);
- size of the Bandwidth Allocation Gap (BAG), *i.e.*, the time interval allocated for the transmission of one packet.

We review one of the main results of SMC usage on such case study from [9]. Three experiments are performed with increasing number of VLs (let X be such number), namely 10, 20, and 30 links. For all VLs we have that BAG is 4 ms while the frame size varies from 100 (for end-system number 1) to 500 bytes (for end-system number 5). Note that end-system number 6 is not considered in this scenario. For each $X \in \{10, 20, 30\}$, the task is to compute the probability that the total delivery time for packets (network latency) is smaller than a given bound, until we reach probability one. The results are given in Figure 7. We note that, for end-system 1 (which only traverses one switch), the results are better, *i.e.*, the network latency is always below 500 μs ($X = 10$), 1000 μs ($X = 20$) and 1500 μs ($X = 30$). All other end-systems have the same (worse) results, as they traverse two switches. Namely, the network latency is always below 1100 μs ($X = 10$), 2000 μs ($X = 20$) and 3000 μs ($X = 30$).

2.5. Recharging of a Plug-in Electric Vehicle

PEVs are being increasingly used in the last decade all over the world [87]. Being able to efficiently recharge PEVs is of great importance, and has impacts on the smart grids field again.

In this case study, we consider a Tesla Model S with a battery capacity of 90kWh which must be charged at a charging station. The model consider we consider here (see [12, 72]) provides several features, such as:

1. charging may be probabilistically delayed, modeling that the grid is currently congested;

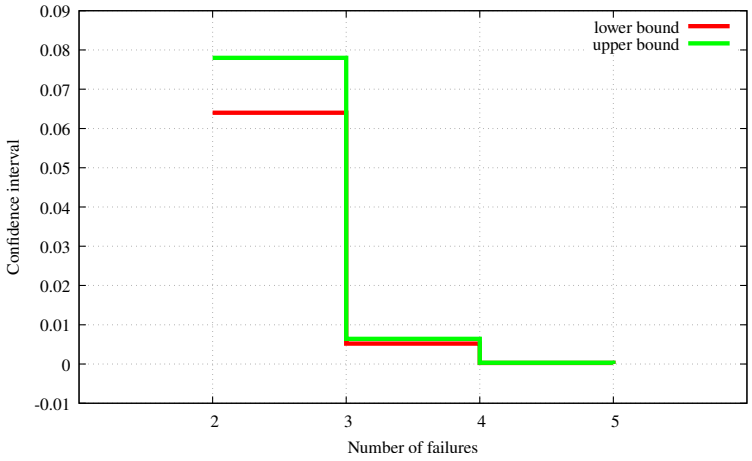

Figure 8: Results for the confidence intervals (lower and upper bounds) of the PEV case study from [12]

2. the amount of time after which the PEV is disconnected is unknown, so it is modeled as a normal probability distribution;
3. the charging processes starts from an empty battery, goes through a "good" charging state after a given time interval and end up in a "full" charging state after another time interval;
4. it is possible to charge the PEV multiple times;
5. one entire week of operation is considered, where recharging is also started in the night and the PEV has to be found fully recharged on the next morning (at an unknown time, as discussed above);
6. we want to compute the probability that the recharging process fails at least $n_{fail} \in \{2, 3, 4, 5\}$ times.

Results for the confidence intervals of the resulting probability is shown in Figure 8, as a function of the number of failures n_{fail}. Note that probabilities decrease very fast when increasing n_{fail}.

3. Conclusions

In this work, we have reviewed some recent real-world problems that were solved using SMC-based techniques. Such problems were taken from very different application areas:

- smart grid intelligent services, in order to compute the probability of EDN substations to be overloaded, when residential users may deviate from their expected power profiles;
- Virtual Physiological Human, to generate a population of VPs for ISCT of drugs, medical devices and treatment strategies, s.t. such population is complete and not over-representative;

- wireless sensor networks, in order to find the smallest bound for clock synchronization accuracy of an audio streaming application;
- aircraft data network (AFDX, Avionics Full-Duplex Switched Ethernet), in order to estimate network latency under different system parameters such as frame size, BAG and number of VLs;
- Plug-in Electric Vehicles, in order to estimate the probability of failures during the recharging process.

This results show that SMC is a mature methodology which can be successfully applied to real-world meaningful problems.

References

[1] G. Agha and K. Palmskog. A survey of statistical model checking. *ACM Trans. Model. Comput. Simul.*, 28(1), 2018.

[2] B. K. Aichernig and M. Tappler. Probabilistic black-box reachability checking (extended version). *Formal Methods Syst. Des.*, 54(3):416–448, 2019.

[3] R. Alur. *Principles of Cyber-Physical Systems.* MIT, 2015.

[4] ARIANE 5 Flight 501 Failure, hiips://www-users.cse.umn.edu/ arnold/disasters/ariane5rep.html, 1996.

[5] P. Ashok, P. Daca, J. Křetínský, and M. Weininger. Statistical model checking: Black or white? In T. Margaria and B. Steffen, editors, *Leveraging Applications of Formal Methods, Verification and Validation: Verification Principles*, pages 331–349, Cham, 2020. Springer International Publishing.

[6] Avicenna Project. *In silico* clinical trials: How computer simulation will transform the biomedical industry. hiip://avicenna-isct.org/wp-content/uploads/2016/01/AvicennaRoadmapPDF-27-01-16.pdf, 2016.

[7] M. Bakir, M. Gheorghe, S. Konur, and M. Stannett. Comparative analysis of statistical model checking tools. In *CMC 2016*, 2016.

[8] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee, and S. K. S. Gupta. Ensuring safety, security, and sustainability of mission-critical cyber–physical systems. *Proceedings of the IEEE*, 100(1):283–299, 2012.

[9] A. Basu, S. Bensalem, M. Bozga, B. Delahaye, A. Legay, and E. Sifakis. Verification of an afdx infrastructure using simulations and probabilities. volume 6418, pages 330–344, 11 2010.

[10] A. Basu, M. Bozga, and J. Sifakis. Modeling heterogeneous real-time components in bip. In *Fourth IEEE International Conference on Software Engineering and Formal Methods (SEFM'06)*, pages 3–12, 2006.

[11] B. Bordel, R. Alcarria, T. Robles, and D. Martín. Cyber–physical systems: Extending pervasive sensing from control theory to the internet of things. *Pervasive and Mobile Computing*, 40:156–184, 2017.

[12] C. E. Budde, P. R. D'Argenio, A. Hartmanns, and S. Sedwards. A statistical model checker for nondeterminism and rare events. In D. Beyer and M. Huisman, editors, *Tools and Algorithms for the Construction and Analysis of Systems*, pages 340–358, Cham, 2018. Springer International Publishing.

[13] M. Cadoli and T. Mancini. Combining relational algebra, SQL, constraint modelling, and local search. *TPLP*, 7(1-2), 2007.

[14] M. Cadoli, T. Mancini, and F. Patrizi. SAT as an effective solving technology for constraint problems. In *ISMIS 2006*, volume 4203 of *LNCS*. Springer, 2006.

[15] K. Choi and H. Liu. *Phase-Locked Loop and Synchronization*, pages 135–150. Wiley-IEEE Press, 2016.

[16] E. Clarke and J. M. Wing. Formal methods: state of the art and future directions. *Comp. Surv.*, 28(4), 1996.

[17] E. Clarke and P. Zuliani. Statistical model checking for cyber-physical systems. In *ATVA 2011*, volume 11. Springer, 2011.

[18] P. Dagum, R. Karp, M. Luby, and S. M. Ross. An optimal algorithm for Monte Carlo estimation. *SICOMP*, 29(5), 2000.

[19] G. Della Penna, B. Intrigila, I. Melatti, E. Tronci, and M. Venturini Zilli. Bounded probabilistic model checking with the Murφ verifier. In *FMCAD 2004*. IEEE, 2004.

[20] N. Dey, A. S. Ashour, F. Shi, S. J. Fong, and J. M. R. S. Tavares. Medical cyber-physical systems: A survey. *Journal of Medical Systems*, 42(74), 2018.

[21] K. Ding, S. Ding, A. Morozov, T. Fabarisov, and K. Janschek. On-line error detection and mitigation for time-series data of cyber-physical systems using deep learning based methods. In *2019 15th European Dependable Computing Conference (EDCC)*, pages 7–14, 2019.

[22] B. Dowdeswell, R. Sinha, and S. G. MacDonell. Finding faults: A scoping study of fault diagnostics for industrial cyber–physical systems. *Journal of Systems and Software*, 168:110638, 2020.

[23] O. Erdinc, A. Tascikaraoglu, N. G. Paterakis, and J. P. S. Catalao. Novel incentive mechanism for end-users enrolled in dlc-based demand response programs within stochastic planning context. *IEEE Transactions on Industrial Electronics*, 66(2):1476–1487, 2019.

[24] European Medicines Agency. Reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation, 2019. EMA/CHMP/458101/2016.

[25] A. Fabregat, S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati, R. Haw, B. Jassal, F. Korninger, B. May, M. Milacic, C. Roca, K. Rothfels, C. Sevilla, V. Shamovsky, S. Shorser, T. Varusai, G. Viteri, J. Weiser, G. Wu, L. Stein, H. Hermjakob, and P. D'Eustachio. The Reactome pathway knowledgebase. *Nucl. Ac. Res.*, 46(D1), 2018.

[26] G. Gottlob, G. Greco, and T. Mancini. Conditional constraint satisfaction: Logical foundations and complexity. In *IJCAI 2007*, 2007.

[27] R. Grosu and S. Smolka. Quantitative model checking. In *ISoLA 2004*, 2004.

[28] R. Grosu and S. Smolka. Monte Carlo model checking. In *TACAS 2005*, volume 3440 of *LNCS*. Springer, 2005.

[29] B. Hayes, I. Melatti, T. Mancini, M. Prodanovic, and E. Tronci. Residential demand

management using individualised demand aware price policies. *IEEE Trans. Smart Grid*, 8(3), 2017.

[30] 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, and 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. *J. Psycho. Res.*, 99, 2017.

[31] P. Hunter, P. Coveney, B. de Bono, V. Díaz-Zuccarini, J. Fenner, A. Frangi, P. Harris, R. Hose, P. Kohl, P. Lawford, K. Mccormack, M. Mendes, S. Omholt, A. Quarteroni, J. Skår, J. Tegner, S. R. Thomas, I. Tollis, I. Tsamardinos, and M. Viceconti. A vision and strategy for the virtual physiological human in 2010 and beyond. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 368:2595–614, 06 2010.

[32] A. Jindal, B. S. Bhambhu, M. Singh, N. Kumar, and K. Naik. A heuristic-based appliance scheduling scheme for smart homes. *IEEE Transactions on Industrial Informatics*, 16(5):3242–3255, 2020.

[33] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima. Kegg: New perspectives on genomes, pathways, diseases and drugs. *Nucl. Ac. Res.*, 45(D1), 2017.

[34] C. E. Kement, H. Gultekin, and B. Tavli. A holistic analysis of privacy-aware smart grid demand response. *IEEE Transactions on Industrial Electronics*, 68(8):7631–7641, 2021.

[35] W. Koch, R. Mancuso, R. West, and A. Bestavros. Reinforcement learning for uav attitude control. *ACM Trans. Cyber-Phys. Syst.*, 3(2), Feb. 2019.

[36] P. Kohl and D. Noble. Systems biology and the virtual physiological human. *Molecular Systems Biology*, 5(1):292, 2009.

[37] M. R. Kosorok and E. B. Laber. Precision medicine. *Annual Review of Statistics and Its Application*, 6(1):263–286, 2019.

[38] E. A. Lee. Fundamental limits of cyber-physical systems modeling. *ACM Trans. Cyber-Phys. Syst.*, 1(1), Nov. 2016.

[39] B. Leeners, T. Krüger, K. Geraedts, E. Tronci, T. Mancini, M. Egli, S. Röblitz, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, and F. Ille. Associations between natural physiological and supraphysiological estradiol levels and stress perception. *Front. Psycol.*, 10, 2019.

[40] B. Leeners, T. Kruger, K. Geraedts, E. Tronci, T. Mancini, F. Ille, M. Egli, S. Roeblitz, L. Saleh, K. Spanaus, C. Schippert, Y. Zhang, and M. Hengartner. Lack of associations between female hormone levels and visuospatial working memory, divided attention and cognitive bias across two consecutive menstrual cycles. *Front. Behav. Neuro.*, 11, 2017.

[41] A. Legay, B. Delahaye, and S. Bensalem. Statistical model checking: An overview. In *RV 2010*, volume 6418 of *LNCS*. Springer, 2010.

[42] A. Lekidis, P. Bourgos, S. Djoko-Djoko, M. Bozga, and S. Bensalem. Building distributed sensor network applications using bip. In *Proceedings of 2015 IEEE Sensors Applications Symposium, Zadar, Croatia*, 2015.

[43] A. Lekidis, M. Bozga, D. Mauuary, and S. Bensalem. A model-based design flow for can-based systems. In *13th International CAN Conference, iCC13*, 11 2013.

[44] A. Lekidis, E. Stachtiari, P. Katsaros, M. Bozga, and C. K. Georgiadis. Using BIP to reinforce correctness of resource-constrained iot applications. In *10th IEEE International Symposium on Industrial Embedded Systems, SIES 2015, Siegen, Germany, June 8-10, 2015*, pages 245–253. IEEE, 2015.

[45] F. Maggioli, T. Mancini, and E. Tronci. SBML2Modelica: Integrating biochemical models within open-standard simulation ecosystems. *Bioinformatics*, 36(7), 2020.

[46] T. Mancini and M. Cadoli. Detecting and breaking symmetries by reasoning on problem specifications. In *SARA 2005*, volume 3607 of *LNCS*. Springer, 2005.

[47] T. Mancini, M. Cadoli, D. Micaletto, and F. Patrizi. Evaluating ASP and commercial solvers on the CSPLib. *Constraints*, 13(4), 2008.

[48] T. Mancini, P. Flener, and J. Pearson. Combinatorial problem solving over relational databases: View synthesis through constraint-based local search. In *SAC 2012*. ACM, 2012.

[49] T. Mancini, F. Mari, A. Massini, I. Melatti, F. Merli, and E. Tronci. System level formal verification via model checking driven simulation. In *CAV 2013*, volume 8044 of *LNCS*. Springer, 2013.

[50] T. Mancini, F. Mari, A. Massini, I. Melatti, I. Salvo, S. Sinisi, E. Tronci, R. Ehrig, S. Röblitz, and B. Leeners. Computing personalised treatments through in silico clinical trials. A case study on downregulation in assisted reproduction. In *RCRA 2018*, volume 2271 of *CEUR W.P.* CEUR, 2018.

[51] T. Mancini, F. Mari, A. Massini, I. Melatti, I. Salvo, and E. Tronci. On minimising the maximum expected verification time. *Inf. Proc. Lett.*, 122, 2017.

[52] T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. Anytime system level verification via random exhaustive hardware in the loop simulation. In *DSD 2014*. IEEE, 2014.

[53] T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. System level formal verification via distributed multi-core hardware in the loop simulation. In *PDP 2014*. IEEE, 2014.

[54] T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. SyLVaaS: System level formal verification as a service. In *PDP 2015*. IEEE, 2015.

[55] T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. Anytime system level verification via parallel random exhaustive hardware in the loop simulation. *Microprocessors and Microsystems*, 41, 2016.

[56] T. Mancini, F. Mari, A. Massini, I. Melatti, and E. Tronci. SyLVaaS: System level formal verification as a service. *Fundam. Inform.*, 149(1–2), 2016.

[57] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, and L. Elmegaard. Parallel statistical model checking for safety verification in smart grids. In *SmartGridComm 2018*. IEEE, 2018.

[58] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, M. Prodanovic, and L. Elmegaard. Demand-aware price policy synthesis and verification services for smart grids. In *SmartGridComm 2014*. IEEE, 2014.

[59] T. Mancini, F. Mari, I. Melatti, I. Salvo, E. Tronci, J. Gruber, B. Hayes, M. Pro-

danovic, and L. Elmegaard. User flexibility aware price policy synthesis for smart grids. In *DSD 2015*. IEEE, 2015.

[60] T. Mancini, I. Melatti, and E. Tronci. Any-horizon uniform random sampling and enumeration of constrained scenarios for simulation-based formal verification. *IEEE TSE*, 2021.

[61] T. Mancini, E. Tronci, I. Salvo, F. Mari, A. Massini, and I. Melatti. Computing biological model parameters by parallel statistical model checking. In *IWBBIO 2015*, volume 9044 of *LNCS*. Springer, 2015.

[62] F. Mari, I. Melatti, I. Salvo, and E. Tronci. Model based synthesis of control software from system level formal specifications. *ACM TOSEM*, 23(1), 2014.

[63] B. L. Mediouni, A. Nouri, M. Bozga, M. Dellabani, A. Legay, and S. Bensalem. SBIP 2.0: Statistical Model Checking Stochastic Real-time Systems. In *ATVA 2018 - 16th International Symposium Automated Technology for Verification and Analysis*, pages 536–542, Los Angeles, CA, United States, Oct. 2018. Springer.

[64] I. Melatti, F. Mari, T. Mancini, M. Prodanovic, and E. Tronci. A two-layer near-optimal strategy for substation constraint management via home batteries. *IEEE Trans. Ind. Elect.*, 2021.

[65] J. Milewski, A. Szczęśniak, and J. Lewandowski. Dynamic characteristics of auxiliary equipment of sofc/soec hydrogen peak power plant. *IERI Procedia*, 9:82–87, 2014. International Conference on Environment Systems Science and Engineering (ESSE 2014).

[66] R. Mitchell and I.-R. Chen. Behavior rule specification-based intrusion detection for safety critical medical cyber physical systems. *IEEE Transactions on Dependable and Secure Computing*, 12(1):16–30, 2015.

[67] A. Nouri, B. L. Mediouni, M. Bozga, J. Combaz, S. Bensalem, and A. Legay. Performance Evaluation of Stochastic Real-Time Systems with the SBIP Framework. *International Journal of Critical Computer-Based Systems*, pages 1–33, 2018.

[68] A. Pappagallo. Statistical model checking for the analysis of mission- and safety-critical cyber-physical systems. In *3rd Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis (OVERLAY 2021)*, volume to appear of *CEUR Workshop Proceedings*. CEUR-WS.org, 2021.

[69] A. Pappagallo, A. Massini, and E. Tronci. Monte carlo based statistical model checking of cyber-physical systems: A review. *Information*, 11(12), 2020.

[70] F. Pappalardo, G. Russo, F. Tshinanu, and M. Viceconti. *In silico* clinical trials: concepts and early adoptions. *Brief. Bioinform.*, 20(5), 2019.

[71] D. R. Patrick and S. W. Fardo. *Electrical Distribution Systems, Second Edition*. Pearson Professional Education, 2009.

[72] C. Pilch and A. Remke. Statistical model checking for hybrid petri nets with multiple general transitions. In *2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 475–486, 2017.

[73] A. J. Pimm, T. T. Cockerill, and P. G. Taylor. The potential for peak shaving on low voltage distribution networks using electricity storage. *Journal of Energy Storage*, 16:231–242, 2018.

[74] D. Reijsbergen, W. de Boer, P.T.and Scheinhardt, and B. Haverkort. On hypothesis

testing for statistical model checking. *Int. Soft. Tech. Trans.*, 17(4), 2015.

[75] S. Röblitz, C. Stötzel, P. Deuflhard, H. Jones, D.-O. Azulay, P. van der Graaf, and S. Martin. A mathematical model of the human menstrual cycle for the administration of GnRH analogues. *J. Theor. Biol.*, 321, 2013.

[76] W. Rogers and K. Hutchison. *Evidence-Based Medicine in Theory and Practice: Epistemological and Normative Issues*, pages 851–872. Springer Netherlands, Dordrecht, 2017.

[77] A. Saad, T. Youssef, A. T. Elsayed, A. Amin, O. H. Abdalla, and O. Mohammed. Data-centric hierarchical distributed model predictive control for smart grid energy management. *IEEE Transactions on Industrial Informatics*, 15(7):4086–4098, 2019.

[78] S. Sinisi, V. Alimguzhin, T. Mancini, and E. Tronci. Reconciling interoperability with efficient verification and validation within open source simulation environments. *Simul. Model. Pract. Theory*, 109, 2021.

[79] S. Sinisi, V. Alimguzhin, T. Mancini, E. Tronci, and B. Leeners. Complete populations of virtual patients for in silico clinical trials. *Bioinformatics*, 36(22–23), 2020.

[80] S. Sinisi, V. Alimguzhin, T. Mancini, E. Tronci, F. Mari, and B. Leeners. Optimal personalised treatment computation through in silico clinical trials on patient digital twins. *Fundam. Inform.*, 174(3–4), 2020.

[81] Statistics on IoT Spending, hiips://www.statista.com/topics/2637/internet-of-things/, 2021.

[82] 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, and F. Ille. Patient-specific models from inter-patient biological models and clinical records. In *FMCAD 2014*. IEEE, 2014.

[83] M. Uddin, M. F. Romlie, M. F. Abdullah, S. A. Halim], A. H. A. Bakar], and T. C. Kwang]. A review on peak load shaving strategies. *Renewable and Sustainable Energy Reviews*, 82:3323 – 3332, 2018.

[84] U.S.A. Food and Drug Administration. Reporting of computational modeling studies in medical device submissions, 2016. FDA-2013-D-1530.

[85] U.S.A. Food and Drug Administration. Physiologically based pharmacokinetic analyses – format and content guidance for industry, 2018. FDA-2016-D-3969.

[86] N. Zhang, B. D. Leibowicz, and G. A. Hanasusanto. Optimal residential battery storage operations using robust data-driven dynamic programming. *IEEE Transactions on Smart Grid*, 11(2):1771–1780, 2020.

[87] Y. Zhou, M. Wang, H. Hao, L. Johnson, and H. Wang. Plug-in electric vehicle market penetration and incentives: a global review. *Mitigation and Adaptation Strategies for Global Change*, 20:777–795, 06 2015.

[88] M. Zimmerling, L. Mottola, P. Kumar, F. Ferrari, and L. Thiele. Adaptive real-time communication for wireless cyber-physical systems. *ACM Trans. Cyber-Phys. Syst.*, 1(2), Feb. 2017.