SYMBIOTIC SIMULATION SYSTEM: HYBRID SYSTEMS MODEL MEETS BIG DATA ANALYTICS

Bhakti Stephan Onggo

Trinity Business School
Trinity College Dublin
College Green
Dublin 2, IRELAND

Navonil Mustafee
Andi Smart

The Business School
University of Exeter
Streatham Court, Rennes Drive
Exeter EX4 4ST, UK

Angel A. Juan

IN3 – Computer Science Department
Universitat Oberta de Catalunya
Rambla del Poblenou, 156
Barcelona, 08018, SPAIN

Owen Molloy

College of Engineering and Informatics
National University of Ireland Galway
University Road
Galway City, IRELAND

ABSTRACT

Symbiotic simulation is one of Industry 4.0 technologies that enables interaction between a physical system and the simulation model that represents it as its digital twin. Symbiotic simulation is designed to support decision making at the operational levels by making use of real- or near real- time data that is generated by the physical system, which is used as an input to the simulation model. From the modeling perspective, a symbiotic simulation system comprises a hybrid systems model that combines simulation, optimization and machine learning models as well as a data acquisition module and an actuator. The actuator is needed when the symbiotic simulation system is designed to directly control the physical system without human intervention. This paper reviews the components of a symbiotic simulation system from the perspective of hybrid systems modeling and highlights research questions needed to advance symbiotic simulation study.

1 INTRODUCTION

The use of a hybrid modeling approach to solving complex real-world problems has gained popularity among simulation researchers and practitioners (Mustafee et al. 2017; Eldabi et al. 2016). This stems from the realization that a single modeling approach is bound by its underlying methodological constraints, and with systems becoming more complex and multi-faceted, there is now a need to use a combination of analytical approaches to best represent the system of interest. In simulation literature, there are two types of hybrid modeling approaches. The first and most common one is *hybrid simulation*, which combines different simulation approaches (e.g. Discrete-Event Simulation, System Dynamics and Agent-Based Simulation) into one model (e.g. Viana et al. 2014; Djanatliev and German 2013). The second approach is a *hybrid systems modeling* and simulation approach (Powell and Mustafee 2016) which combines simulation with compatible modeling approaches including system thinking models (e.g. soft system methodology, cognitive mapping) and analytics models (e.g. mathematical model). Combining analytics models and simulation is not new (see Shanthikumar and Sargent 1983). Andreas Tolk in Mustafee et al. (2017) states that a useful hybrid model combines the characteristics of its components into something more useful (in comparison to using any of its components alone).

This paper is on symbiotic simulation and its representation as a hybrid systems model. In a *symbiotic simulation system* (SSS), there is a level of interaction between a simulation model and the physical system that it represents (Aydt et al. 2008). It is designed to support decision making at the operational levels by making use of real- or near real-time data (generated by the physical system), and which is streamed subsequent to the development of the simulation model. In addition to the simulation model, a symbiotic simulation system may comprise *data analytics and machine learning models*, *optimization models*, as well as *data acquisition module* and *actuators*. An actuator can be defined as a hardware component of the physical system which interfaces with both a control software and the physical machinery. The actuator thus allows the former to trigger actions on the latter. For example, taking the case of manufacturing, a signal from the symbiotic simulation model (software) may trigger actions such as slowing down the conveyer belt in a factory (physical system). However, hardware actuators are not a necessary element of the symbiotic simulation systems that are discussed in this paper. For example, in the services sector, any change to the physical system may need the intervention of the decision maker. In an Emergency Department (ED), for example, the output from the symbiotic system may suggest re-allocation of staff. However, for this to be reflected in the urgent care setting, the staff managers/ED matron will need to action this. Thus, for this study, the actuator can either be a piece of hardware (e.g., as in Industry 4.0) or a person (as in the case of a service context).

This paper makes a novel contribution to modeling & simulation (M&S) literature in the following ways: (1) It presents a hybrid systems modeling architecture for a symbiotic simulation system and defines its various components; (2) It discusses the components of the system with reference to existing research and software implementation, e.g., Airline Disruption Recovery (Rhodes-Leader et al. 2018), the *NHSquicker* platform (Mustafee et al. 2018); (3) It identifies future work to advance symbiotic simulation research and its representation through hybrid models. The remainder of this paper is organized as follows. Section 2 presents relevant work on symbiotic simulation. Section 3 discusses the building blocks of such a system and presents specific techniques (forecasting, clustering, optimization, machine learning) that can be combined with simulation to develop a hybrid systems model. Section 4 is on further research and draws this paper to a close .

2   SYMBIOTIC SIMULATION

The origin of the concept of symbiotic simulation can be traced back to Davis (1998), who then referred to such systems as online simulation. The term itself was introduced at the 2002 Dagstuhl seminar on the *Grand Challenges for Modeling and Simulation* (Fujimoto et al. 2002). The early definitions of online/symbiotic simulation emphasized the ability of simulation to control a physical system (this was heavily influenced by the concept of dynamic data-driven application systems). Aydt et al. (2008) provided a degree of flexibility to the original use of the term, by defining *symbiotic systems* as those having "a close association between a simulation system and a physical system, which is beneficial to at least one of them." This association is enabled through communication channels existing between the simulation and the physical system, and allowing for interaction in real- or near real-time. Aydt et al. (2008)'s use of this term is now generally accepted as the definition for symbiotic simulation.

To clarify the definitions used in this paper, a *symbiotic system* is composed of a *physical system* and the *symbiotic simulation system* (SSS) that represents the physical system (i.e. virtual system or digital twin). Such a system is implemented as a *hybrid systems model* by combining *SSS* with methods and techniques from the wider Operations Research and other disciplines (e.g., our architecture proposes the use of simulation with data analytics, machine learning and data streaming technology). The execution of the symbiotic simulation model is referred to as *symbiotic simulation*.

Figure 1 illustrates the high-level view of three symbiotic types (based on Aydt et al. 2008). The first and second types are formed by a closed loop between a physical system and the SSS that represents the physical system. In the first type (i.e. symbiotic simulation decision support system in Aydt et al. (2008)), at runtime, the SSS reads data from the physical system. Subsequently, the symbiotic simulation produces outputs that are used by a decision maker to control the physical system. Hence, the SSS indirectly controls

the physical system via a decision maker. As for the second type (i.e. symbiotic simulation control system in Aydt et al. (2008)), the SSS directly controls the physical system through an actuator without any human intervention. The third type is the open-loop symbiotic simulation system. In this case, both input data and output data (key performance indicators) from the physical system are read by the SSS. This type of SSS is used to predict the behavior of the physical system (i.e. similar to non-symbiotic simulation except that it has the ability to read real-time data from the physical system at runtime), to emulate the behavior of the physical system (hence, the most valid model can be used as a reference model) or to detect anomalies either in the physical system or in its reference model. Aydt et al. (2008) refer to these three applications as symbiotic simulation forecasting system, symbiotic simulation model validation system, and symbiotic simulation anomaly detection system, respectively. Hence, the SSS itself is not used to control the physical system. However, its output may be used by another system that is not integrated to the SSS to control the physical system.
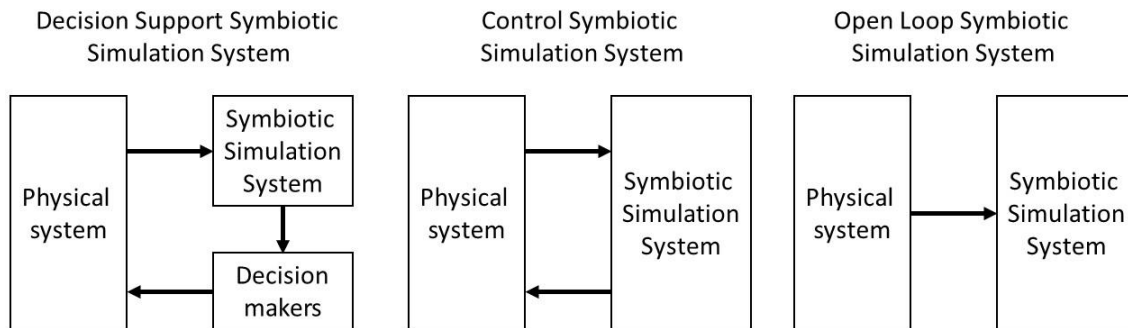

Figure 1: Three types of symbiotic system (adapted from Aydt et al. 2008).

We have adapted Aydt et al. (2008)'s original classification and have identified three types of symbiotic system based on the actual mechanism of communication feedback (output) from the SSS to the physical system.

- **Type 1** (*Decision Support Symbiotic Simulation System*) – Here the output from a scenario manager or an optimization model will be used to help a decision maker to control the physical system.
- **Type 2** (*Control Symbiotic Simulation System*) – Here the output will directly affect the physical system through an actuator.
- **Type 3** (*Open Loop Symbiotic Simulation System*) – No direct feedback mechanism exists.

Since 2005 researchers have implemented symbiotic simulation system in various physical systems. The earliest application is in manufacturing. Low et al. (2005) argue that symbiotic simulation is suitable for semiconductor manufacturing because it is constantly changing as part of the continuous improvement of the manufacturing process. Further, increasing customization of products has led to flexible manufacturing systems (FMS) that are subject to changes in demand, products and process flows. Symbiotic simulation is considered suitable for FMS (Kück et al. 2016). The early adoption of symbiotic simulation by manufacturing is enabled by the fact that many modern factories are equipped with sensors that collect real-time data to support monitoring and continuous improvement. Symbiotic simulation has also been implemented to control highly dynamic systems such as unmanned aerial vehicles (e.g. Mitchell and Yilmaz 2008; Holt et al. 2014), transportation (e.g. Sunderrajan et al. 2014; Rhodes-Leader et al. 2018) and data center operations (e.g. Liu et al. 2012; Abar et al. 2016). Recently, Tan et al. (2013) applied symbiotic simulation to support operational management decisions in an emergency department.

3 COMPONENTS OF A SYMBIOTIC SIMULATION SYSTEM

The components of a symbiotic system are shown in Figure 2. In all types of symbiotic system (Types 1, 2, 3, refer to Fig. 1), the data from the physical system are collected by a *data acquisition module* (arrow 1). The use of appropriate data analytics models (e.g. forecasting, data mining) can help the system make the best use of the combination of old and new data (arrow 2). For example, when there is a sudden change in the number of job arrivals in the system, an appropriate data analytics model may be able to decide whether the parameters in the scenario manager (arrow 3), optimization model (arrow 3) and/or symbiotic simulation model need updating (arrow 4). A scenario manager implements various what-if analyses using the symbiotic simulation model (arrows 5); the model returns results to the scenario manager (arrow 6). Alternatively, an optimization model can be used to produce an optimum solution. The output from the scenario manager and optimization model will be used by the machine learning model (arrow 9) to adapt the scenario manager or optimization model (arrow 10), symbiotic simulation model (arrow 11) and data analytics model (arrow 12). The results from the scenario manager/optimization model are communicated to an actuator or a decision maker (arrow 7), leading to changes being made to the physical system (arrow 8). The components of the SSS are described below.
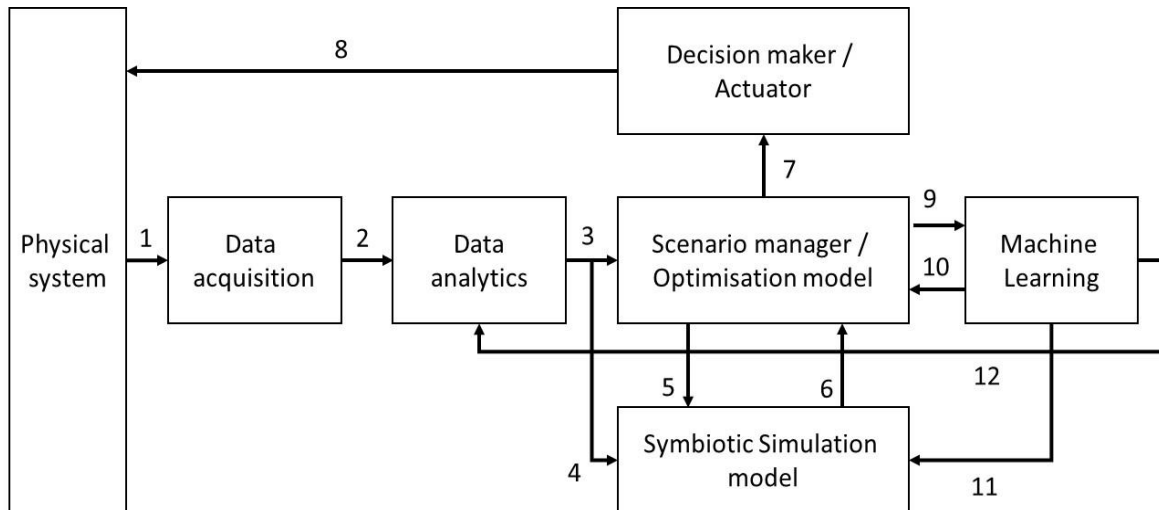

Figure 2: Physical system and symbiotic simulation system (SSS).

3.1 Data Acquisition Module

Data from a physical system is read by a data acquisition module (arrow 1). The data can be read online (direct communication from the sensors and the data acquisition module) or off-line (the data from sensors are stored in a file, and the file will be read by the data acquisition module). The data acquisition module can be implemented as a web service, web application or mobile application. For example, Rhodes-Leader et al. (2018) develop a SSS for airlines disruption recovery management and use web services to retrieve real-time data from FlightRadar24 to feed their SSS. Mustafee et al. (2008) discuss data specification and data acquisition for the *NHSquicker* platform; the platform displays real-time data on urgent care wait times from nearly 25 ED and Urgent Care Centers (UCC) that belong to several NHS Trusts in the South West of England. Future work includes the development of a wait-time prediction algorithm and ED models that are driven by real-time data. It is expected that this will take the form of symbiotic simulation decision support system (Type 1,Fig 1).

3.2 Data Analytics Model

One important characteristic of a SSS is that it can respond to real-time data to improve decisions. The data may come in different volumes, velocities, varieties and veracities. Hence, the analytics model(s) may need to determine which data should be used in what way given the possibility of receiving different volumes of data. Typically, the data analytics models used are time-series models (and forecasting models in general) and data mining models. For example, Mustafee et al. (2008) discuss the use of autoregressive integrated moving average (ARIMA) models for ED/UCC wait time prediction for the NHSquicker platform. The paper presents the *Right Hospital-Right Time (RH-RT)* data analytics framework and discusses how the data acquisition module (implemented as web-services – arrow 1) can be used with a data analytics model (implemented as predictive analytics using ARIMA forecasting models – arrow 2). Future work will involve the use of the ED/UCC wait time predictions (including expected number of patients in the queue) as input to simulation models (arrow 4) and scenario manager/optimization model (arrow 3). Use of forecasting with simulation is not restricted to symbiotic/real-time simulation only, for example, Harper et al. (2017) have implemented a hybrid approach using forecasting and discrete-event simulation for the planning of endoscopy services.

Another analytics technique that can be used here is clustering. Clustering is a very effective, intuitively simple technique, for analyzing data sets without prior assumptions regarding the data. It is very useful in identifying groups of related data points. The unsupervised nature of clustering techniques, and generally speaking, its ability to produce good results without prior assumptions, make it a valuable technique. Clustering techniques can generally be classified into two main families - Hierarchical and Partitioning (Rokach 2009). Hierarchical cluster methods take either a top-down (Divisive) or bottom-up (Agglomerative) approach to recursively divide the data into clusters. While the quality of the results obtained can be higher using hierarchical agglomerative clustering, it is computationally expensive for even moderate size data sets. In general, partitioning algorithms, such as k-means techniques offer a good compromise between quality and efficiency for larger datasets, whereas the accuracy of hierarchical clustering techniques is preferred for smaller data sets (Rokach 2009; Abbas 2008).

Within healthcare, clustering techniques have been largely used to identify patient groups. K-means clustering appears to be very widely used in identifying patient groups which have high degrees of similarity (Tomar and Agarwal 2013; Liao et al. 2016). To take one example, clustering aids in the identification of entity groups (e.g. types of patient, types of job), predicted behavior based on the entity groups (e.g. typical length of stay for major trauma patients), scenarios (i.e. groups of uncontrolled variables and parameters such as weather and seasons) to be used in a simulation, and clusters of process histories (which can identify previously unrecognized important care pathways through the healthcare process, see Delias et al. (2015)).

At the other end of the simulation process, the simulation outputs can be analyzed using clustering techniques to identify behaviors contributing to systems performance. For example, Johnstone et al. (2015), used clustering techniques to analyze simulation event logs to group entity flows based on paths taken through an airport baggage handling system, and thereby identifying key factors influencing handling time.

From the above examples, we can see that the promise of forecasting and clustering techniques within symbiotic simulation is their ability to direct the execution and design of the simulation model (arrow 4) and to identify the selection of scenarios for simulation experiments (arrow 3).

3.3 Scenario Manager

In a symbiotic simulation system, a scenario manager may be used instead of an optimization model, especially when we are interested in finding the best among a set of pre-defined policies. The role of a scenario manager in a symbiotic simulation system is the same as in a non-symbiotic simulation system. It controls how a simulation model is run to meet the objective of the experiment or analysis (arrows 5 and 6). Typical analysis such as sensitivity analysis and experiment design may be used here.

3.4 Optimization Model

Simulation methods allow decision makers to get insight into complex systems and compare different scenarios. However, simulation is not an optimization method, and thus it cannot provide by itself optimal or near-optimal solutions in complex scenarios with a large number of possible settings. Thus, there is an opportunity to develop hybrid approaches combining simulation methods with optimization methods (arrows 5 and 6). In symbiotic simulation, an optimization model is required when an optimal, near-optimal, or at least high-quality decision needs to be made regarding some of the parameters existing in the physical system. This is especially the case both in a symbiotic simulation decision support system (Type 1 – see Fig. 1, where human intervention is expected) as well as in a symbiotic simulation control system (Type 2, where a decision on system parameters are automatically set based on the outcome of the optimization process). The optimization model uses the symbiotic simulation model to estimate the objective function (the combined simulation and optimization model is also known as optimization-via-simulation or simulation optimization model). Fu (2015) classifies optimization-via-simulation based on whether the decision variables are continuous, discrete (ordered or unordered, finite or infinite) or a combination of the two. Abo-Hamad and Arisha (2011) further classify the techniques into several groups, i.e. gradient-based, meta-model-based, statistical and random search/meta-heuristics.

One important optimization problem is *combinatorial optimization problems* (COP), where a solution maximizing or minimizing a given objective function must be chosen among a discrete but vast amount of possible candidates. In addition, the solution must satisfy all the real-life constraints imposed in the problem formulation. Examples of such COPs are found in smart cities, logistics and transportation systems, manufacturing systems, healthcare systems, etc. In these scenarios, symbiotic simulation-optimization approaches might be decisive to support real-time decision making based on the outcomes of the optimization module. Thus, for example, during waste collection management in smart cities, an initial vehicle routing plan might require changes as new sensor-gathered data on the traffic and containers status arrives at the control centre in real-time. Similarity can be found in the case of a manufacturing process in a digital industry; an initial scheduling plan might require updating as real-time data on machine failures or bottlenecks is received. Similar examples can be found in many other areas.

Typically, small-scale and static versions of these COP can be optimally solved by employing mathematical programming methods. Mula et al. (2010) offer a review of mathematical programming models for transportation planning and supply chain production. They point out the need for integrating optimization, simulation, multi-agent systems and evolutionary algorithm techniques. Optimization models of real-life systems are frequently NP-hard in nature. In those cases, the use of heuristics or metaheuristics to obtain high-quality or near-optimal solutions in real-time is usually needed (Talbi 2009). With the increase of computing power and the possibilities that parallel and distributed computing offer, metaheuristics are rapidly becoming one of the most popular optimization methods. In particular, for symbiotic simulation-optimization approaches they might be advantageous due to the short computing times needed to make complex decisions. A particular case of such hybrid approach is called *'simheuristics'* (Juan et al. 2015), which integrates heuristics with simulation to solve stochastic optimization problems.

Another emerging approach that can be useful in the context of symbiotic simulation is *'learnheuristics'* (Calvet et al. 2017). It proposes the combination of heuristics with machine learning methods to cope with optimization problems with dynamic inputs, such as, variable traffic conditions, dynamic travelling times, or evolving customers' demands. Thus, for instance, in Calvet et al. (2016) the authors propose a learnheuristic approach to solve a transportation-related problem in which the customers' demands vary as the distribution plan is constructed. With the increased availability of real-time data and greater processing power, these hybrid simulation-optimization-learning approaches present opportunities for intelligent decision-making in real-time; these decisions would arguably inspire some confidence since they would have been first experimented using symbiotic simulation.

3.5   Simulation Model

A simulation model of the physical system is the core component in a symbiotic simulation system. The structure of a symbiotic simulation model is the same as a non-symbiotic simulation model with one main difference: a symbiotic simulation model must be able to read data at runtime and make an appropriate response as specified by the modeler. The response can be as simple as re-initializing the simulation using the latest state of the physical system. It can also use conditional distributions to inform the service times (i.e. when entities are already in the system at the point of simulation re-initialization).

One of the most challenging issues is the ability to adapt the structure of the simulation model in response to feedback from the physical system. One possible approach is to use an ensemble of simulation models that are controlled by a genetic algorithm to produce the best estimate (Mitchell and Yilmaz 2008). The genetic algorithm can adapt the configuration of the ensemble based on the most recent data. Another possible approach is to make the simulation model adaptive (Zhang and Zhao 2014). In this case, the model can change its structure at runtime based on the most recent data. This approach will require a mechanism for auto-verification and auto-validation of the adaptive symbiotic simulation model. The verification and validation suites in test-driven simulation modeling (Onggo and Karatas 2016) are one of the promising approaches for auto-verification and auto-validation of an adaptive symbiotic simulation model. Another possible alternative approach is to make the model adaptive by adding a machine learning model (which will be discussed in Section 3.6). This approach will also require the ability to auto-verify and auto-validate the adaptive symbiotic model.

3.6   Machine Learning Model

The benefit of a symbiotic simulation system can be enhanced by the use of a *machine learning* (ML) model. A ML model makes the simulation model and optimization model (when used) learn from past performance. For example, the ML model can compare the observed KPI from the physical system and the estimated KPI from the model and adjust the model to improve its accuracy.

As discussed in section 3.2, unsupervised techniques such as clustering can be used to both learn about the entities we are simulating as well as to determine important paths through the process being simulated. In SSS we can, therefore, use such techniques to assist in identifying important data and variables, building simulation models and monitoring of real process data to ensure our simulation model accurately reflects the current process. ML can be used to predict the behavior of important system variables which may not be easily expressed analytically.

A hybrid simulation-ML environment can also be used to enhance the development of mental (simulation) models with data-driven knowledge acquired through ML. In a thought-provoking study, Tolk (2015) envisions that the next generation of simulation models will be integrated with ML, and Deep Learning (LeCun et al. 2015) in particular. By automating the data capture and learning process using ML, it is possible to achieve fully adaptive simulation models which change behavior in response to changes in the system itself and its environment. Without this capability, there is likely to be an inevitable drift between the modelled and actual system. ML techniques are well suited to the task of predicting values for important simulation variables based on snapshots of the system state, which in turn will ensure the ongoing accuracy of the simulation model. For example, in Elbattah (2016), Random Forest techniques (Breiman 2001) were used to train regression models used to predict the in-patient 'Length Of Stay' and binary classifiers used for predicting 'Discharge Destination' for elderly hip fracture care patients.

ML can also be used to create a metamodel of a simulation model. For example, Rabelo et al. (2014) applied ML to data generated by discrete and systems dynamics simulations of Panama Canal operations. Neural network metamodels were built which can encapsulate the knowledge of the hybrid simulation model, thereby creating a fast, efficient, cost-effective alternative to the execution of simulation models.

Deist et al. (2018) propose another way of combining ML and simulation, applied to prediction of the effectiveness of cancer treatment. In their approach, simulation is used as a means of pre-processing sample data to improve the prediction performance of the ML, in the absence of sufficient patient data repositories.

Similarly, but in a very different field, another study on data-driven ML shows how simulation-generated training data can be used to train ML algorithms for mechanical fault detection in industrial machines, where limited historical data exists (Sobie et al. 2018). Wang and Marek-Sadowska (2015) reverse the process, using ML (clustering algorithms and support vector machines) to select important input samples for simulation in electronic circuit design analysis.

It is also noteworthy that a related discussion was featured in an industry workshop from PwC in the Winter Simulation Conference 2017 (Wallis and Paich 2017), which demonstrates that the potential for integration of Artificial Intelligence/ML and Simulation both as components of simulation models, but also as aids in building them. For example, simple neural networks are being integrated into agents as a dynamic replacement of manually coded rules governing agent behavior. This allows agents behavior to change dynamically in response to system conditions. In their comprehensive paper, Bohlmann et al. (2017) describe a complex symbiotic system to control robot movement in which data-driven ML (using evolutionary algorithms embedded in a multi-agent implementation) is used to generate candidate process models. These, in turn, drive a simulation system which starts without a pre-existing model.

ML provides a means of mitigating the limitations and inherent flaws of simulation models built on 'mental models' and sometimes limited causal knowledge. In this regard, ML can be utilized as an assistive artefact within the process of problem formulation. Further, ML provides the potential to create self-correcting as well as adaptive simulations, with in-built error checking between simulation and ML-predicted behavior. ML models can be embedded into simulation experiments to realize the learning factor or adaptive behavior.

4 RESEARCH CHALLENGES AND CONCLUSION

SSS, as discussed in this paper, differs from non-symbiotic (conventional) simulation in several ways. Some fundamental differences are as follows. (1) A symbiotic simulation is an implementation of hybrid systems modeling, whereas a non-symbiotic simulation does not necessarily have to be hybrid. (2) A symbiotic simulation will have, as one of its inputs, real-time or near-real-time data; there is no such data requirement for conventional models. (3) The time available for a symbiotic simulation to deliver results is usually shorter since it is mostly used for proactive/just-in-time operational decision making; the results from conventional simulations are usually used for medium to long-term decision making. (4) A symbiotic simulation (ideally) needs to adapt to changes in the physical system; there is no such requirement for simulations that are non-symbiotic. These characteristics pose several research challenges.

Need for integration frameworks: The hybrid nature of SSS requires a well-defined set of guidelines and frameworks to guide the integration of various analytics models (e.g. forecasting, simulation) as well as ML algorithms. The challenges are both conceptual (i.e., commensurability and philosophical compatibility) and technical (i.e., integration methodology, open access to software implementing different elements of the hybrid model). Balaban et al. (2014,2015) discussed the philosophical issues surrounding hybrid systems model.

Need for state synchronization during execution of symbiotic simulation: Advancement in data collection technologies makes it increasingly common for organizations to collect and publish data in real-time. While this will benefit symbiotic simulation, it also poses a research challenge. In a physical system where its states change quickly over time, the initial system state used in the scenario manager or optimization to find a solution may be significantly different from the state of the physical system when the solution is about to be applied. Hence, a technique to deal with this issue is needed. One approach proposed by Hanisch et al. (2005) involves maintaining a continuously synchronized parent model, from which any number of child models can be generated and run at any time. The use of historical data and real-time data raises a question on how a symbiotic simulation can make the best use of both data sources.

Need for multiple approaches based on time available for decision making: From an optimization perspective, in many real-life applications, a fundamental challenge may arise from the use of symbiotic simulation systems, The challenge could be stated as follows: how does the system provide high-quality solutions in 'real-time' for complex and large-scale optimization problems? In some business and life-

critical applications, the time available for computation (before a decision needs to be taken) is limited to a few seconds or less. This is the case, for instance, of symbiotic systems used to model the traffic in large cities, mobile telecommunication services, emergency situations in airports, or intermodal transportation systems in smart cities. In these scenarios, one promising research area is the use of biased-randomized heuristics, which are capable of generating 'good' solutions in just a few milliseconds by introducing non-uniform (oriented) randomization inside a heuristic logic and taking advantage of parallelization techniques (Grasas et al. 2017). Another promising approach is the use of multi-fidelity models (Rhodes-Leader et al. 2018). For decisions that could to be taken over longer period (in minutes rather than in seconds), better solutions can be obtained by utilizing metaheuristics, matheuristics – metaheuristics with mathematical programming (Boschetti et al. 2009), simheuristics, or leanheuristics (metaheuristics with machine learning). Once more, the use of distributed and parallel computing techniques might become a valuable resource to reduce the computing times necessary to generate 'high-quality' or even 'near-optimal' solutions (Juan et al. 2013).

Need for adapting the simulation model to reflect changes in the physical system: The current practice in simulation modeling assumes that the future system behavior is the same as the behavior during model development. Thus, a research challenge is to identify ML approaches which allow the structure of a simulation model to change at run-time and to select the most appropriate algorithms to trigger the changes in its model.

This paper has discussed the architecture of a symbiotic simulation system and its implementation using hybrid systems modeling. It has reviewed related work and has highlighted future research opportunities by identifying several challenges. The hybrid approach proposes to combine real-time data with forecasting, optimization, machine learning and computer simulation. Although such a symbiotic simulation system holds promise, many non-trivial theoretical and practical challenges exist. Addressing these challenges require, for example, new integration frameworks, algorithms capable of supporting informed decision making in short computing times, new approaches for validation and verification of models.

ACKNOWLEDGMENTS

REFERENCES

Abar, S., P. Lemarinier, G. Theodoropoulos, and G. O'Hare. 2016. "Automated Dynamic Resource Provisioning and Monitoring in Virtualized Large-Scale Datacenter". In *Proceedings of the 28th International Conference on Advanced Information Networking and Applications*, 961–970. Piscataway, NJ: IEEE.

Abo-Hamad, W., and A. Arisha. 2011. "Simulation–Optimization Methods in Supply Chain Applications: A Review". *Irish Journal of Management* 30(2):95–124.

Abu Abbas, O. 2008. "Comparisons Between Data Clustering Algorithms". *The International Arab Journal of Information Technology* 5(3):320–325.

Aydt, H., S. J. Turner, W. Cai, and M. Y. H. Low. 2008. "Symbiotic Simulation Systems: An Extended Definition Motivated by Symbiosis in Biology". In *Proceedings of the 22nd Workshop on Principles of Advanced and Distributed Simulation*, 109–116. Piscataway, NJ: IEEE.

Aydt, H., S. J. Turner, W. Cai, and M. Y. H. Low. 2009. "Research Issues on Symbiotic Simulation". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti et al., 1213–1222. Piscataway, New Jersey: IEEE.

Balaban, M., P. Hester, and S. Diallo. 2014. "Towards a Theory of Multimethod M&S Approach: Part I". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk et al., 1652–1663. Piscataway, New Jersey: IEEE.

Balaban, M., P. Hester, and S. Diallo. 2015. "Towards a Theory of Multimethod M&S Approach: Part III". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz et al., 1633–1644. Piscataway, New Jersey: IEEE.

Bohlmann, S., V. Klinger, and H. Szczerbicka. 2017. "Integration of a Physical System, Machine Learning, Simulation, Validation and Control Systems towards Symbiotic Model Engineering". In *Proceedings of the Symposium on Modeling and Simulation of Complexity in Intelligent, Adaptive and Autonomous Systems*. San Diego, CA: SCS.

Boschetti, M. A., V. Maniezzo, M. Roffilli, and A. B. Röhler. 2009. "Matheuristics: Optimization, Simulation and Control. In *Proceedings of the International Workshop on Hybrid Metaheuristics*, 171–177. Berlin, Heidelberg: Springer.

Breiman, L. 2001. "Random Forests". *Machine learning* 45(1):5–32.

Calvet, L., J. de Armas , D. Masip, and A. A. Juan. 2017. "Learnheuristics: Hybridizing Metaheuristics with Machine Learning for Optimization with Dynamic Inputs". *Open Mathematics* 15(1):261–280.

Calvet, L., A. Ferrer, M. L. Gomes, A. A. Juan, and D. Masip. 2016. "Combining Statistical Learning with Metaheuristics for the Multi-Depot Vehicle Routing Problem with Market Segmentation". *Computers & Industrial Engineering* 94:93–104.

Davis, W. 1998. "On-line Simulation: Need and Evolving Research Requirements. In *Handbook of Simulation*, edited by J. Banks, 46–516. Hoboken, NJ: John Wiley & Sons.

Delias, P., M. Doumpos, E. Grigoroudis, P. Manolitzas, and N. Matsatsinis. 2015. "Supporting Healthcare Management Decisions via Robust Clustering of Event Logs". *Knowledge-Based Systems* 84:203–213.

Deist, T., A. Patti, Z. Wang, D. Krane, T. Sorenson, and D. Craft. 2018. "Simulation Assisted Machine Learning". https://arxiv.org/abs/1802.05688, accessed 01.05.2018.

Djanatliev, A., and R. German. 2013. "Prospective Healthcare Decision-Making by Combined System Dynamics, Discrete-Event and Agent-Based Simulation". In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy et al., 270–281. Piscataway, New Jersey: IEEE.

Elbattah, M. and O. Molloy. 2016. "Coupling Simulation with Machine Learning: A Hybrid Approach for Elderly Discharge Planning". In *Proceedings of the 2016 Conference on Principles of Advanced Discrete Simulation*, 47–56. New York : ACM.

Eldabi, T., M. Balaban, S. Brailsford, N. Mustafee, R. E. Nance, B. S. Onggo, and R. G. Sargent. 2016. "Hybrid Simulation: Historical Lessons, Present Challenges and Futures". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T.M. Roeder et al., 1388–1403. Piscataway, New Jersey: IEEE.

Fu, M.C. 2015. *Handbook of Simulation Optimization*. 1st ed. New York, NY: Springer-Verlag.

Fujimoto, R., D. Lunceford, E. Page, and A. M. Uhrmacher. 2002. "Grand Challenges for Modeling and Simulation: Dagstuhl Report". Technical Report 350, Schloss Dagstuhl. Seminar No 02351.

Grasas A., A. A. Juan, J. Faulin, J. de Armas, and H. Ramalhinho. 2017. "Biased Randomization of Heuristics using Skewed Probability Distributions: A Survey and Some Applications". *Computers & Industrial Engineering* 110:216–228.

Hanisch, A., J. Tolujew, and T. Schulze. 2005. "Initialization of Online Simulation Models". In *Proceedings of the 37th Winter Simulation Conference*, edited by M. E. Kuhl et al. 1795–1803. Piscataway, New Jersey: IEEE.

Holt, J., S. Biaz, L. Yilmaz, and C. A. Aji. 2014. "A Symbiotic Simulation Architecture for Evaluating UAVs Collision Avoidance Techniques. *Journal of Simulation* 8:64–75.

Harper, A., N. Mustafee, and M. Feeney. 2017. "A Hybrid Approach Using Forecasting and Discrete-Event Simulation for Endoscopy Services." In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan et al., 1583–1594. Piscataway, New Jersey: IEEE.

Johnstone, M., V. Thanh Le, J. Zhang, B. Gunn, S. Nahavandi, and D. Creighton. 2015. "A Dynamic Time Warped Clustering Technique for Discrete Event Simulation-Based System Analysis". *Expert Systems with Applications* 42:8078–8085.

Juan A. A., J. Faulin, J. Jorba, J. Caceres, and J. M. Marquès. 2013. "Using Parallel & Distributed Computing for Real-Time Solving of Vehicle Routing Problems with Stochastic Demands". *Annals of Operations Research* 207(1):43–65.

Juan A. A., J. Faulin, S. E. Grasman, M. Rabe, and G. Figueira. 2015. "A Review of Simheuristics: Extending Metaheuristics to Deal with Stochastic Combinatorial Optimization Problems". *Operations Research Perspectives* 2:62–72.

Kück M., J. Ehm, T. Hildebrandt, M. Freitag, and E. M. Frazzon. 2016. "Potential of Data-Driven Simulation-Based Optimization for Adaptive Scheduling and Control of Dynamic Manufacturing Systems". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, 2820–2831. Piscataway, Piscataway, New Jersey: IEEE.

LeCun Y., Y. Bengio, and G. Hinton. 2015. "Deep Learning." *Nature* 521:436–444.

Liao, M., Y. Li, F. Kianifard, E. Obi, and S. Arcona. 2016. "Cluster Analysis and Its Application to Healthcare Claims Data: A Study of End-Stage Renal Disease Patients Who Initiated Hemodialysis". *BMC Nephrology* 17:25.

Low M. Y. H., K. W. Lye, P. Lendermann, S. J. Turner, R. T. W. Chim, and S. H. Leo. 2005. "An Agent-Based Approach for Managing Symbiotic Simulation of Semiconductor Assembly and Test Operation". In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems*, 85–92. New York, NY: ACM.

Liu Q., G. K. Theodoropoulos, D. da Silva, and E. S. Liu. 2012. "Towards an Agent-Based Symbiotic Architecture for Autonomic Management of Virtualized Data Centers". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque et al., 961-970. Piscataway, New Jersey: IEEE.

Mitchell B., and L. Yilmaz. 2008. "Symbiotic Adaptive Multisimulation: An Autonomic Simulation Framework for Real-Time Decision Support under Uncertainty". *ACM Transactions on Modeling and Computer Simulation* 19(1):2–31.

Mula J., D. Peidro, M. Díaz-Madroñero, and E. Vicens. 2010. "Mathematical Programming Models for Supply Chain Production and Transport Planning". *European Journal of Operational Research*, 204(3):377–390.

Mustafee, N., S. Brailsford, A. Djanatliev, T. Eldabi, M. Kunc, and A. Tolk. 2017. "Purpose and Benefits of Hybrid Simulation: Contributing to the Convergence of Its Definition". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan et al., 1631–1645. Piscataway, New Jersey: IEEE.

Mustafee, N., J. H. Powell, and A. Harper. 2018. "Right Hospital-Right Time: A Business Analytics Framework for Analysing Urgent Care/A&E Wait Time Data". In *Proceedings of the 2018 Simulation Workshop*, 105–115. Birmingham, UK: The Operational Research Society.

Onggo, B. S. and M. Karatas. 2016. "Test-Driven Simulation Modelling: A Case Study using Agent-Based Maritime Search-Operation Simulation". *European Journal of Operational Research* 254(2):517–531.

Powell, J. H. and N. Mustafee. 2016. "Widening Requirements Capture with Soft Methods: An Investigation of Hybrid M&S Studies in Healthcare". *Journal of the Operational Research Society* 68(10):1211–1222.

Rabelo, L., L. Cruz, S. Bhide, O. Joledo, J. Pastrana, and P. Xanthopoulos. 2014. "Analysis of the Expansion of the Panama Canal Using Simulation Modeling and Artificial Intelligence". In *Proceedings of the 2014 Winter Simulation Conference*, edited by Tolk et al., 910–921. Piscataway, New Jersey: IEEE.

Rhodes-Leader, L., B. S. Onggo, D. J. Worthington, and B. L. Nelson. 2018. "Airline Disruption Recovery Using Symbiotic Simulation and Multi-fidelity Modelling". In *Proceedings of the 2018 Simulation Workshop*, 146–155. Birmingham, UK: The Operational Research Society.

Rokach, L. 2009. "A Survey of Clustering Algorithms". In *Data Mining and Knowledge Discovery Handbook*, 269–298. Boston, MA: Springer.

Shanthikumar, J. G., and R. G. Sargent. 1983. "A Unifying View of Hybrid Simulation/Analytic Models and Modeling." *Operations Research* 31(6):1030–1052.

Sobie, C., C. Freitas, and M. Nicolai. 2018. "Simulation-Driven Machine Learning: Bearing Fault Classification." *Mechanical Systems and Signal Processing* 99:403–419.

Sunderrajan, A., H. Aydt, W. Cai and A. Knoll. 2014. "Map Stream: Initializing What-if Analyses for Real-Time Symbiotic Traffic Simulations. In *Proceedings of the 2014 Winter Simulation Conference*, edited by Tolk et al., 2896–2906. Piscataway, New Jersey: IEEE.

Talbi, E. G. 2009. *Metaheuristics: From Design to Implementation (Vol. 74)*. Hoboken, NJ: John Wiley & Sons.

Tan K. W., W. H. Tan, and H. C. Lau. 2013. "Improving Patient Length-of-Stay in Emergency Department through Dynamic Resource Allocation Policies". In *Proceedings of the International Conference on Automation Science and Engineering*, 984–989. Piscataway, New Jersey: IEEE.

Tolk, A. 2015. "The Next Generation of Modeling & Simulation: Integrating Big Data and Deep Learning". In *Proceedings of the Conference on Summer Computer Simulation*, 1–8. San Diego, CA: SCS.

Tomar, D., and A. Agarwal. 2013. "A Survey on Data Mining Approaches for Healthcare". *International Journal of Bio-Science and Bio-Technology* 5(5):241–266.

Viana, J., S. C. Brailsford, V. Harindra, and P. R. Harper. 2014. "Combining Discrete-Event Simulation and System Dynamics in a Healthcare Setting: A Composite Model for Chlamydia Infection". *European Journal of Operational Research* 237(1):196–206.

Vijay, J., and J. Subhashini. 2013. "An Efficient Brain Tumor Detection Methodology using K-means Clustering Algorithm". In *Proceedings of the International Conference on Communication and Signal Processing*, 653–657. Piscataway, New Jersey: IEEE.

Wallis, L., and M. Paich. 2017. "Integrating Artificial Intelligence with AnyLogic Simulation". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan et al., 4449–4449. Piscataway, New Jersey: IEEE.

Wang, Li-C., and M. Marek-Sadowska. 2015. "Machine Learning in Simulation-Based Analysis". In *Proceedings of the International Symposium on Physical Design*, 57–64. New York, NY: ACM.

Zhang, F., and Q. Zhao. 2014. "Runtime Reconstruction of Simulation Models for Dynamic Structure Systems". In *Proceedings of the AsiaSim Conference*, Part III, CCIS 325, 27–36. Berlin Heidelberg: Springer-Verlag.

AUTHOR BIOGRAPHIES

BHAKTI STEPHAN ONGGO is an Associate Professor of Data Analytics at Trinity Business School, Trinity College Dublin, Ireland. His email address is stephan.onggo@tcd.ie.

NAVONIL MUSTAFEE is Associate Professor of Operations Management and Analytics at University of Exeter Business School, UK. His email address is n.mustafee@exeter.ac.uk.

ANGEL A. JUAN is a Professor of Operations Research & Industrial Engineering in the Computer Science Dept. at the Open University of Catalonia. His email address is ajuanp@uoc.edu.

OWEN MOLLOY is a Lecturer in Information Technology at the National University of Ireland, Galway, Ireland. His e-mail address is owen.molloy@nuigalway.ie.

ANDI SMART is Professor of Operations Management at University of Exeter Business School, UK. His email address is p.a.smart@exeter.ac.uk.