Technological University Dublin

ARROW@TU Dublin

Conference papers   Conservatory of Music and Drama

2019

A Hybrid Process Mining Framework for Automated Simulation Modelling for Healthcare

Susan McKeever
Technological University Dublin, susan.mckeever@tudublin.ie

Waleed Abo Hamad
Technological University Dublin, waleed.abo-hamad@TUDublin.i

Mohammed Mesabbah
Technological University Dublin, mohammed.mesabbah@dit.ie

Follow this and additional works at: https://arrow.tudublin.ie/aaconmuscon

Part of the Music Commons

Recommended Citation

A HYBRID PROCESS MINING FRAMEWORK FOR AUTOMATED SIMULATION MODELLING FOR HEALTHCARE

Mohammed Mesabbah

Ernst and Young (EY)
Harcourt Centre
Harcourt Street, Dublin 2, IRELAND

Waleed Abo-Hamad

School of Management, College of Business
Technological University Dublin (TU Dublin)
Aungier Street, Dublin 2, IRELAND

Susan McKeever

School of Computer Science
Technological University Dublin (TU Dublin)
Kevin Street, Dublin 8, IRELAND

ABSTRACT

Advances in data and process mining algorithms combined with the availability of sophisticated information systems have created an encouraging environment for innovations in simulation modelling. Researchers have investigated the integration between such algorithms and business process modelling to facilitate the automation of building simulation models. These endeavors have resulted in a prototype termed Auto Simulation Model Builder (ASMB) for DES models. However, this prototype has limitations that undermine applying it on complex systems. This paper presents an extension of the ASMB framework previously developed by authors adopted for healthcare systems. The proposed framework offers a comprehensive solution for resources handling to support complex decision-making processes around hospital staff planning. The framework also introduces a machine learning real-time data-driven prediction approach for system performance using advanced activity blocks for the auto-generated model, based on live-streams of patient data. This prediction can be useful for both single and multiple healthcare units management.

1 INTRODUCTION

Overcrowding in emergency departments (EDs) has become a significant international crisis that negatively affects patient safety, quality of care, staff morale and patient satisfaction. In response, managerial interventions and clinical guidelines keep evolving to improve care processes and the quality of care delivery. However, the consequences of these such interventions are not tested prior to their introduction and ultimately guidelines are implemented differently across hospitals. In turn, more deviations from the normative process are created and patient outcomes are at risk in terms of lower patient throughput, increased length of stay, and higher treatment costs.

A key enabler of improving healthcare processes is process simulation (Liu et al. 2012). Complex processes, uncertain parameters, time critical activities and both budget and resource constraints typify an environment suitable for scenario testing. Rather than incurring the cost of "learning by doing", simulation modelling is one of the most powerful and cost-effective methods for system performance improvement (Hlupic and Robinson 1998). Simulated processes rely on solid information for derivation of an accurate conceptual model. In the typical top down approach, this information is gathered manually from interviews of knowledge holders, direct observation and procedural documentation. However, this

approach is time-consuming and liable to subjective or optimistic observations (Lynch et al. 2014). Process models can also be derived using a bottom-up approach through the distillation of information from actual system event logs (Mǎruşter and van Beest 2009). This exercise of process mining can be used to establish an unbiased picture of the actual executed processes.

Healthcare processes are increasingly underpinned with computer system capture of significant events – from the patient being checked in at reception, through each of the clinical checks. Recent approaches have used process mining of system event logs for conceptual modelling, as an input to the simulation modelling (Abohamad et al. 2017). This allows for the development of simulation models based on realistic representations of processes as logged in the underlying system logs. Mesabbah and McKeever (2018) presented a generic automated end-to-end simulation framework (Auto Simulation Model Builder or ASMB) that integrates process mining algorithms with data mining to automatically generate simulation models using system event logs. Various components were developed in ASMB that is capable of parsing system events log files and extracting agents' pathways, processing times, resources, and rate of agents' arrival to system. Such information is used as the input for generating a functional simulation model using a generic DES engine. ASMB also supports the execution of "what-if" scenarios on variations in resources, agents arrival patterns and/or process - in order to optimize the underlying processes through reduction of waiting times, queues lengths and higher system throughout. However, the framework was not restricted to a specific domain.

This paper presents an extension of the ASMB framework dovetailed to the health (hospital) sector. Three improvements are introduced to ASMB: Process Miner, Data Miner and Machine Learning activities. The modified framework (ASMB-H) introduces a machine learning data-driven prediction approach for system performance by combining real-data and advanced data analytics. An existing healthcare case study is used to demonstrate the fully implemented parts of the framework while the underlying architecture of the parts under development are explained.

This paper is structured as follows: Section 2 provides the background of the research topic through a literature review. Section 3 highlights the main components of the Auto-Simulation Model Builder for Healthcare (ASMB-H) framework, detailing each of the framework components. Finally, Section 4 provides a wrap up discussion and concluding remarks for the framework.

2   LITERATURE REVIEW

A variety of techniques are used in the domain of simulation modelling. In their review of simulation modelling in the healthcare domain, Katsaliaki and Mustafee (2011) identified Monte Carlo Simulation (MCS), Discrete-Event Simulation (DES), System Dynamics (SD) and Agent-Based Simulation (ABS) as the most widely used techniques. Each technique has a particular focus of simulation: MCS enables the testing of probability distributions for uncertain factors in the modelled system; DES models pre-defined systems with predictable interactions; ABS models the behavior of particular individual agents within a system where such agents have autonomous behavior; SD examines how the fundamental structure might influence the progressive behavior of a system including tangible elements and intangible elements (Taylor and Lane 1998). The success of each of these techniques is linked to the use of an accurate conceptual model upon which the simulation will be based.

It is suggested that hybridization across different methods and approaches can reduce the limitations of individual methods and increase their capabilities (Balaban et al. 2014). One way to achieve this is by combining more than one paradigm from one category of techniques. For instance, multiple simulation paradigms are integrated into many simulation studies to leverage the models' capabilities for addressing complex problems (Mesabbah et al. 2016). Hybridization can be also used by combining paradigms from different categories of techniques. For example, researchers have investigated the integration between simulation techniques and optimization approaches to facilitate complex evaluation functions for the systems that are optimized (Figueira and Almada-Lobo 2014). Recently, as organizational information systems advance to store and generate increasingly large volumes of data, researchers have been

encouraged to explore the integration between data and process mining and modelling and simulation techniques (Măruşter and van Beest 2009). Such hybridization can help to shorten the simulation model development life cycle and overcome the problems that are usually encountered during the data collection and analysis phases.

Process mining has emerged since the early 2000s to support the development of accurate unbiased conceptual models. The practice of process mining attempts to reconstruct complete process models from data logs that contain real process execution data (Tiwari et al. 2008). Through the mining of events logs, accurate portrayals of activities and events can be derived, thus providing a solid basis for overlaying simulation techniques. Măruşter and van Beest (2009) proposed a methodology to use the outputs of process mining as an input to simulation. Their approach provided detailed work of process flow derivations from mined event logs, and comparative performance of the "as-is" and "to-be" business processes. Liu et al. (2012) derived event graphs from process mining as a conceptual input to simulation models. They illustrated their method in a credit card scenario. Aguirre et al. (2012) produced a higher-level methodology for combining process mining with simulation modelling. Their evaluation focused on the domain of procurement processes in a university. Similarly, Abohamad et al. (2017) used process mining of system event logs in a health care domain to derive a conceptual model, which was in turn used as the basis for setting up a simulation model.

Recently, Machine Learning Models (MLMs) have been used in tandem with the simulation models to offer improved prediction capabilities of these models (Onggo et al. 2018). Historical data can be used to train prediction models that focus on the most critical aspects the modeled system. For example, MLMs have been used to model the crucial patient-related decisions made by the staff members such as determining the severity of patients to improve the predictability of staff resource requirements (Rashwan 2017). Similarly, Elbattah et al. (2018) have integrated simulation modelling and MLMs for designing pathways and evaluating the return on investments for the elderly healthcare hip-fracture care scheme in Ireland.

The novelty of the proposed framework lies in the data-driven approach simulation modelling and process load prediction. ASMB-H provides an approach to support rapid-build, accurate simulation modelling, system generated logs, flexible data upload modules, combined with predictive modelling of agents (patients) and their activities to support accurate outcome modelling and resource requirements.

3   DESIGN PRINCIPLES

The Auto-Simulation Model Builder for Healthcare (ASMB-H) framework is an enhanced and customized version of the generic Auto-Simulation Model Builder (ASMB) introduced by Mesabbah and McKeever (2018). The key approach of the ASMB is the reliance on system events log to extract the underlying process structure alongside the statistical analysis for patterns of occurrences for the different events and their processing times. This information is then fed into a Discrete Event Simulation (DES) engine that can be used to simulate the underlying system behavior under different types of scenarios including: 1) Changes in process design; 2) Change in resources settings; and 3) System agents behavior.

ASMB, however, operates under two particular simplifying assumptions. Firstly, it assumes that the resources associated for the system processes are fixed across the simulation duration. In reality resources and staffing may vary in numbers over different number of shifts as in the case of the Emergency Departments (ED) of the public hospitals (Arisha and Abo-Hamad 2013). Secondly, the systems agents (i.e. the patients in the healthcare case) that are presented during the simulation used to demonstrate the ASMB are assumed to be homogenous. This assumption was imposed because the framework uses only the information of events log data. relevant for 1) The process mining algorithm and 2) The data analyzer for arrival patterns and processing times. Advanced IT systems are currently able to capture more attributes for the system agents. For example, digital patient records in hospital systems can hold multiple attributes of patients (such as age, gender, historical diseases, etc.) alongside the tracking of their treatment within the hospital. Using advanced analytics to extract patterns based on the attributes of such

big data can be useful for reducing the abstraction of the simulation models and advancing their complexity to replicate real system behavior (Onggo et al. 2018). For example, heterogeneous agents can be presented during the simulation - and the characteristics and behavior of these agents will be derived from systems data records. ASMB-H is developed to relax these two ASMB simplifying assumptions, providing more advanced automated end-to-end simulation modeling for healthcare systems. Specifically, ASMB-H will provide support for flexible resources scheduling and will exploit advanced data analytics and machine learning algorithms to model the agents' behavior during the simulation.

Inherited from the generic ASMB, the proposed ASMB-H framework includes new and modified components alongside few unmodified components from the ASMB as illustrated in Figure 1. There are two key components that both drill into the hospital patients records to extract different types of information. The first component is the Process Miner (PM), which is responsible for retrieving the system events log data from the patients records. The PM is also responsible for extracting the underlying system process structure alongside the meta data of the patient pathways and the human resources. The second component is the Data Miner (DM) component, which is responsible for retrieving all relevant information about the patients. This information will be used to model patient behavior within the hospital processes (e.g. clinics) using simple Machine Learning Models (MLM) to predict the behavior at each activity unit within the hospital.

The outcomes of the Process Miner and Data Miner components are used as input for an embedded DES auto-modelling engine to construct the ultimate automated simulation model for the underlying healthcare system. This DES model is supported by a simple user interface for enabling two modes of using the simulation model. The first mode is simple investigating for the impact of proposed scenarios/ decisions (e.g., staff scheduling) on the hospital Key Performance Indicators (KPIs). The second mode is using the live-stream of patients' data– i.e. patients who are presenting at the hospital at any point of time – to predict how the hospital will perform for next few hours. For example, the model will predict the average time for patients discharge or the expected queue length for a certain caring service.

The following subsections present a more detailed description for each component of the ASMB-H framework, supported by an explanatory case study for a particular hospital dataset with illustration of outcomes for the fully developed components.

3.1   Case Study

Patients data for a hospital ED introduced by Abohamad et al. (2017) is used to demo the outcomes of the fully developed parts of the proposed framework. A real-time patient tracking information system was used to track the patients' journey within the ED. The hospital managers have provided nearly a year of historical data with anonymous patients' records. The dataset was provided in an event log structure with a total of 125,427 event logs representing 31,386 patients. Each record in the table represents an event (i.e. one process stage of the patient journey in the ED) with the following attributes (patient ID, Triage Category, Presenting Complaint, Age, Gender, Event ID, Tracking Step Name, Tracking Step Date Time, Location, Staff). Data pre-processing modules were used to convert the hospital dataset into a readable format defined for the ASMB-H framework.

3.2   Hospital Patients Logs Data

The Hospital Patients Logs Data (HPLD) is the input for the process miner component and is required in order for the ASMB-H to work properly. In addition to the basic events data designed for the generic ASMB (object ID, Activity Name, Activity Instance ID, Activity State, Resource ID, Time Stamp), the HPLD is designed to accept more data attributes related to the patients (e.g., gender, age, presenting complaint, etc.). These attributes will be used to build the machine learning models for the ML-Based Activities component in Figure 1. Currently, it is the user's decision to specify which patient attributes to include and which to exclude from the HPLD before importing them in the ASMB-H. For example, they

1094

user may decide to only consider patients gender and age for training the MLMs and exclude presenting complaints data.
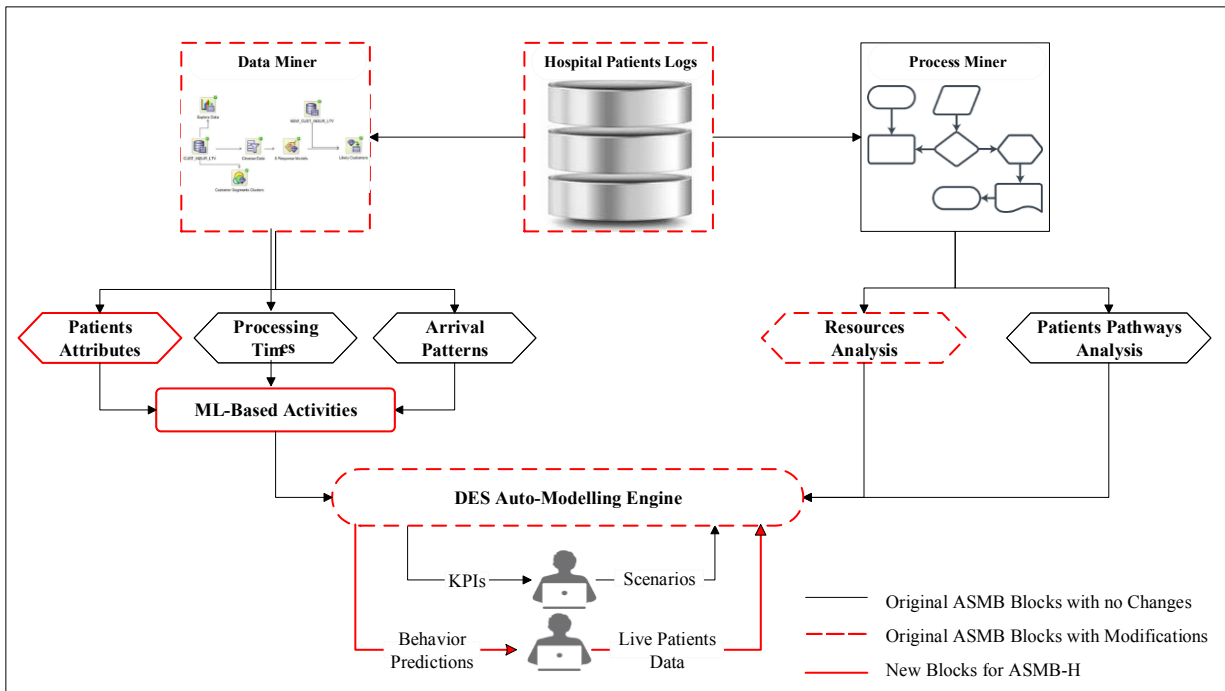

Figure 1: The ASMB-H Framework.

3.3   Process Miner Component

The purpose of this component is to extract the process flow structure of the underlying healthcare system, including the overall process maps and patient pathways alongside the resources and staffing distribution across the different treatment units within the system. The process miner parses the system patients events log data and applies a process mining algorithm in order to achieve that purpose. The PM also generates the system structure metadata such as patient frequencies through the different pathways and statistics for the processing and waiting times of patients at the different treatment units (e.g., triage room, x- ray room, etc.) within the system.

3.3.1 Patients Pathways Analysis

The discovered process map represents the work flow within the underlying healthcare system based on all the events that occurred and were recorded in the patient events log file. It is a graph that consists of a set of nodes and a set of arcs, where the nodes represent the treatment activity and the arcs represent the preceding orders for these activities. The PM can represent the discovered process map in either graph or matrix views as illustrated in Figures 2 and 3. Figure 2 shows the graphical process map for the ED of the case study while Figure 3 represents this graph in matrices view with frequencies of patients' movement from a treatment step to another.

The preceding matrix in Figure 3 along with the numerical information about the patients' movements within the healthcare process map will be used to construct: 1) activities adjacency matrix (A) and 2) the transition probability matrix (P). The two matrices and the way they are constructed are explained in detail by Mesabbah and McKeever (2018). Both matrices will be used later by the DES engine to simulate patient pathways within the underlying healthcare system.

1095

3.3.2 Resources Analysis

The process mining algorithm used for extracting the process map from the HPLD is the same algorithm used for the generic ASMB framework. This algorithm is able to capture the resources associated with the inputted event log data. However as previously mentioned, this original ASMB algorithm is only able to analyze the resources mapping for the overall period of the event log time horizon and it lacks the ability to analyse these resources over different time slots or different shifts. For example, the algorithm can identify how many nurses were assigned to the patients within the ED for the overall time period of the patients' event log data, but it cannot identify the nurses' assignment over different time shifts per day or week. This was important as mentioned earlier for hospital resources and staffing planning, where hospital mangers plan staffing levels on daily/ weekly shifts. We provide an enhanced algorithm in the ASMB-H framework to extract the assignment of different resources types across different time slots basis within the period of time covered in the HPLD. For example, this algorithm can dig into the HPLD and identify for each staff category (i.e., nurses, SHO, etc.) the number of a staff members assigned to patients for 24 hours of every day on weekly basis. This is important for two reasons. Firstly, for the verification of the simulation models that are generated by the ASMB-H. Secondly, to allow for the calculation of resources utilization for the proposed weekly staffing level proposed by the hospital managers. Figure 4 shows an example for identifying SHO (a type of junior doctor in Ireland called Senior Hospital Officer) assignments within the ED for two different weeks for the HPLD of the case study.

The process mining algorithm will be also used for the analysis of the behavior of resources/staff within the healthcare system at the level of resources-activities involvement. The ASMB-H uses this analysis to create the various resource pools and attach them to relevant activity blocks within the auto-developed simulation by the DES engine. For more details on resources-activities involvement analysis, see (Mesabbah and McKeever 2018).
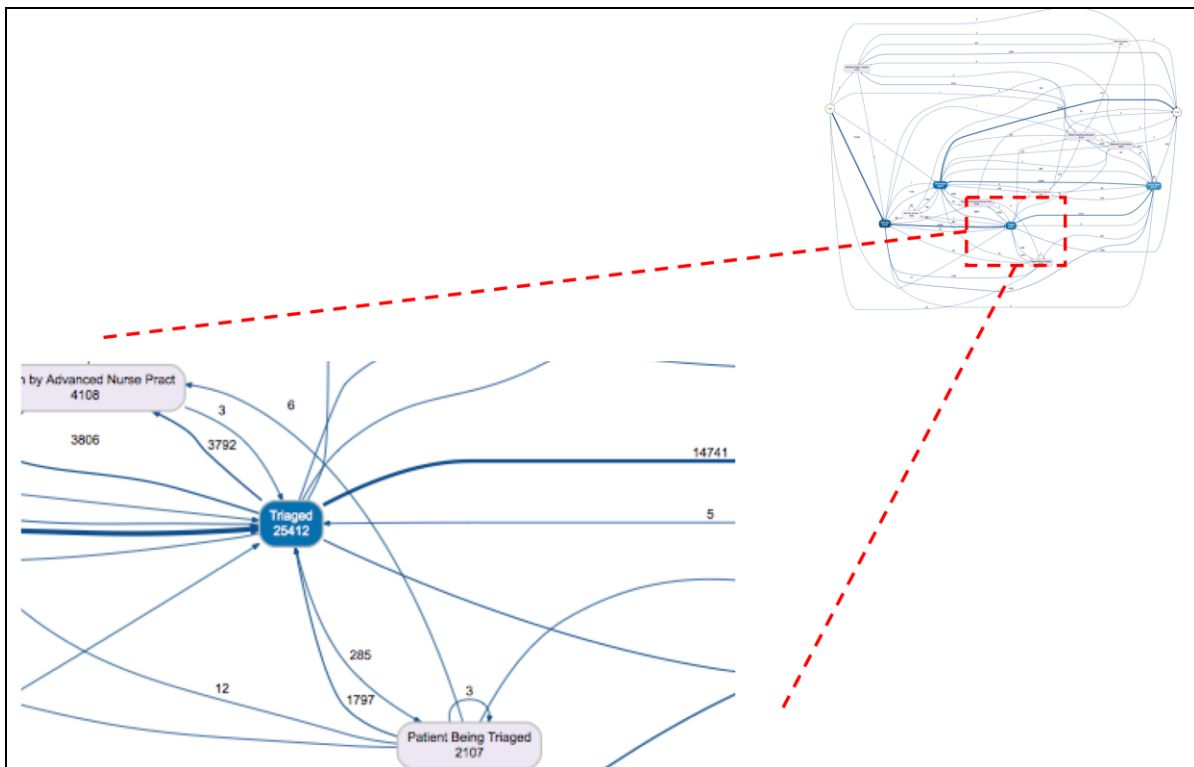

Figure 2: Graphical view of the ED process map.

3.4 Data Miner Component

This component is responsible for extracting all the information regarding patient attributes, pattern of arrival and processing times. In classical approaches for developing DES models, the arrival patterns (i.e., inter-arrival intervals between patients) and the processing times information are usually derived through direct observations via statistical sampling and measuring cycles techniques (Skoogh and Johansson 2008). But, in the ASMB-H the DM component can extract them from the HPLD more accurately.

For the inter-arrival times between patients, the time stamps for the earliest starting event for each patient is retrieved from HPLD. Then an array of the inter-arrival times between every two successive patients is calculated. The values of this array are then fitted to relevant statistical distributions to be used for simulating patient arrivals in the developed model for the hospital unit using the ASMB-H. Afterwards, each patient's attribute is analyzed to drive a statistical distribution to be used for simulating these attributes for the generated patients within the DES model.

	antecedent	consequent	n
1	Start	Attended	31386
2	Start	Discharged	1
3	Start	Doctor Seen	3
4	Start	Patient Being Triaged	24
5	Start	Triaged	44
6	Admitted Mater Hospital	Bed Allocated	2
7	Admitted Mater Hospital	End	3769
8	Admitted Mater Hospital	Patient Awaiting Admission	2
9	Attended	Did Not Answer	360
10	Attended	Discharged	1128
11	Attended	Doctor Seen	4306
12	Attended	Patient Being Triaged	1795
13	Attended	Referred for Admission	11
14	Attended	Seen by Advanced Nurse Pract	285
15	Attended	Triaged	23562

Figure 3: Matrix view of the ED process map (n here represents frequency of occurrences).

Week of 23rd of February 2009

resType	dayNum	date	weekDay	0	1	2	3	4	5		17	18	19	20	21	22	23
SHO	54	2009-02-23	Monday	3	1	3	1	1	1		2	2	2	2	1	1	2
SHO	55	2009-02-24	Tuesday	1	2	0	1	1	1		2	3	2	1	3	2	1
SHO	56	2009-02-25	Wednesday	3	2	0	2	0	1		1	1	2	3	0	3	1
SHO	57	2009-02-26	Thursday	3	1	2	1	2	3		1	0	3	2	3	5	3
SHO	58	2009-02-27	Friday	2	3	1	1	2	2		2	2	1	1	2	4	3
SHO	59	2009-02-28	Saturday	4	2	2	1	2	1		3	1	1	2	2	4	2
SHO	60	2009-03-01	Sunday	2	1	5	3	2	1		0	1	0	1	1	2	2

Figure 4: A sample of SHO staffing used within the ED case study extracted by ASMB-H.

Unlike the ASMB, the ASMB-H does not use the estimated statistical distributions of the processing times data to simulate the timing of the activity blocks within the generated model. Instead, different datasets are constructed for supervised machine learning models to predict: 1) the patients treatment time at any activity block in the extracted process map, and 2) the next transition within the process map after

1097

leaving the current one given the available set of attributes of the patients in the HPLD. More elaboration on how these datasets will be constructed and how the machine learning model will be built is given the following subsection.

3.4.1 Machine Learning Based Activities

The intrinsic idea behind using MLMs for supporting the auto-develop DES model activities is to pave the way for using the ASMB-H as a predictive tool of the healthcare system behavior given the live stream of patient data alongside using it as a tool for policy investigation (e.g., weekly staff planning). This will open the door for different types of applications for the framework over different levels of management for the healthcare system.

Figure 5 illustrates the overall workflow, which is not fully developed yet, for building the ML models for each activity within the hospital system process structure. All the information for starting and finishing times for each activity within the hospital unit are extracted in a separate dataset. The timestamps for the starting and finishing of the activity are used to calculate the processing time at each activity. The rows of this dataset represent all the patients who have been through this activity according to the HPLD, while the columns represent the patients attributes plus three more attributes that represent 1) the processing time for each patient at this activity; 2) the previous activity this patients coming from; 3) and the next activity the patients has visited after the current one. Both the processing time and the two activity columns provide stage labelling for each patient row, thus providing a training dataset for the creation of activity MLMs. For categorical attributes, all unique values are identified then replaced by numerical codes.

The function of the MLM for each activity is to predict the processing time and next activity (Dependent attributes) for each patient given the available attributes for the patients. An example for the X-Ray activity is illustrated in Figure 5. The MLM in this case predicts the time that the patient will take within the X-Ray given information about the patient's age, gender, triage category, presenting compliant and the previous step before the X-Ray (e.g., triage or doctor seen). It also predicts the next step inside the ED (e.g., Discharge or Admission).

3.5 DES Auto-Modelling Engine

This component is responsible for building the final simulation model that represents the underlying healthcare unit. The received inputs from PM and DM components and, then, automatically creates: 1) MLM-based Activity blocks according to the discovered process map; 2) Queues blocks to be placed before each activity; 3) Resource pools based on the identified activities-resources assignment. The engine also contains a patient agent creator block which injects simulated patients into the model based on the statistical distributions estimated for the inter-arrival times and patients' attributes generated by the DM component.

3.5.1 Scenarios Simulation Usage Mode

In this mode of using the ASMB-H, the user can investigate the impact of scenarios/ policies on the overall performance of the underlying healthcare unit. For this mode, the work flow of the DES engine is quite simple where it is triggered by the first patient arrival event. Then the simulated attributes for this patient are injected to the MLM of the starting activity. The MLM predicts the processing time and the next step after leaving this activity. The processing is commencing subject to the availability of relevant resource. If no resources are available, the patient will queue before this activity until a resource becomes available. These steps will be replicated at every following activity until the stopping criteria of the simulation model is achieved. While the simulation is running, a set of KPIs (e.g., average waiting times, resources utilization, etc.) are computed and reported at the end of the simulation. In this usage mode, the user can develop a range of scenarios to be tested against the Business As Usual (BAU) scenario. These scenarios may include decisions or parameters related to one or more of the following:
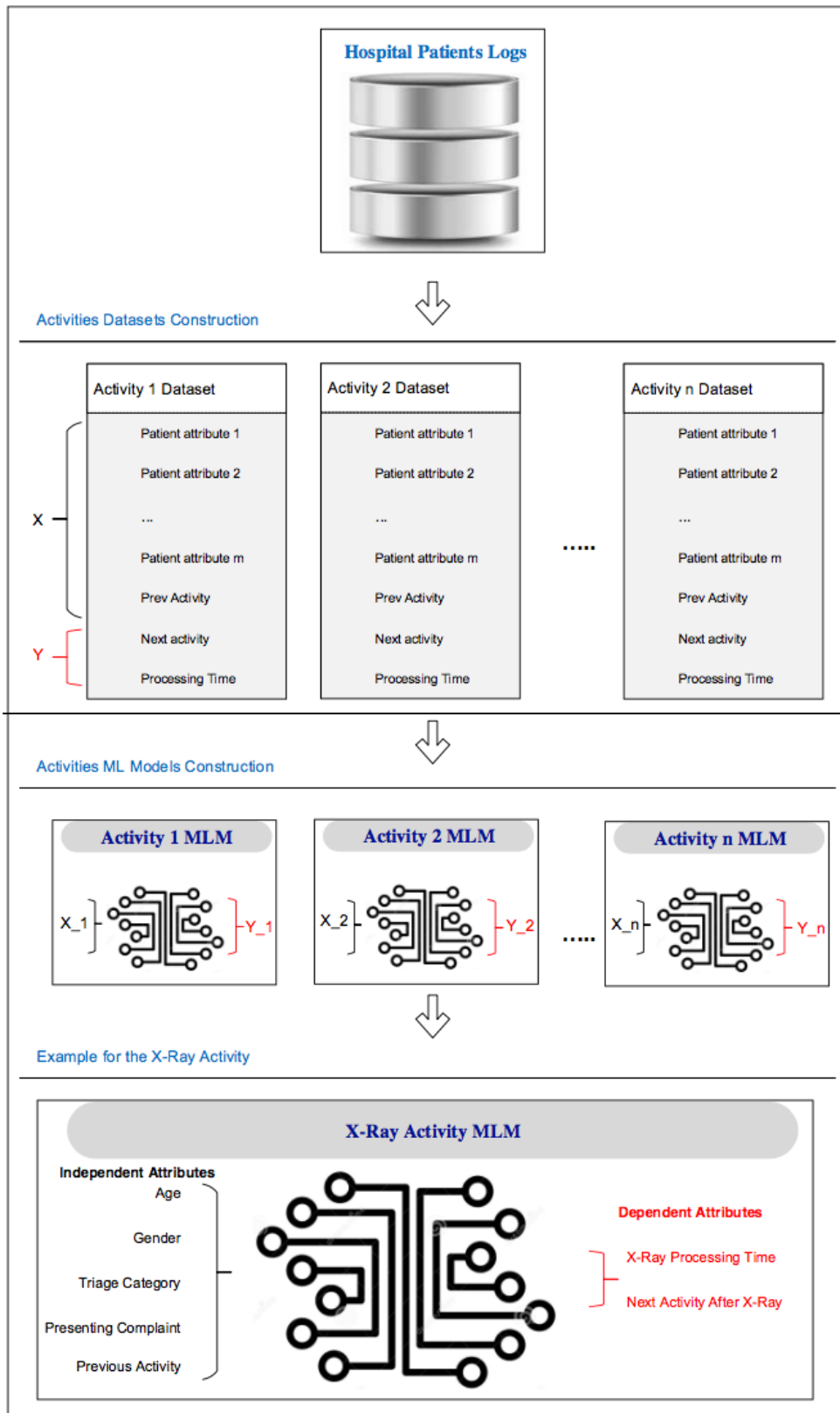
Figure 5: Workflow for building Activities ML Models.

1. **Staff Scheduling** – For this decision a set of matrices for resources availability of the daily shifts on weekly basis for each staff category. The initial values for these matrices are derived from the "resources analysis module" under the PM component.
2. **Patients Agent Arrival Rate** – This can be done by either: i) changing the parameters of the estimated distribution for inter-arrival times between patients or ii) using a different distribution to simulate patient arrivals to the system.
3. **Patients Agent Attributes** – This can be done by changing the parameters of the estimated distribution for each attribute. The rationale behind that is to investigate how changes in patients attributes (e.g., more elderly patients presenting in the system) will be reflected in the overall performance.

3.5.2 System Performance Prediction Usage Mode

In this mode, the auto-simulation model of the DES engine can be used to simulate the system performance given a live-stream of patient data being fed into the model. The patients creator component under this mode will not inject random patients. Rather it will inject real patients who arrived at the reception of the hospital system along with their available attributes. In this case the MLMs attached to the different activities within the hospital system will play the role of predicting the overall pathway for that patient until s/he leaves the system. Simultaneously, the system KPIs are updated and make known to the user (e.g., ED manager) based on the patient pathways and timestamp predictions.

Having said that the live-stream of patient data is used for system behavior prediction, it will be also stored into the HPDL. On pre-defined constant time intervals, the updates of the HPDL will be reflected on the PM and DM components and all the analysis modules within both of them, and accordingly the changes will take place in the auto DES model created.

4   DISCUSSION AND CONCLUDING REMARKS

The work in this paper proposes a framework for customizing the generalized ASMB presented by Mesabbah and McKeever (2018) for healthcare systems. The ASMB was develop to offer an accurate and auto-generated DES model for systems relying on their event log data. These models can be easily adjusted and updated given any changes that take place in the system structure by re-loading the most recent system event log data to the ASMB. The powerful value of this automation is saving time and effort needed every time by manually updating the DES model using the conventional DES modeling life cycle. However, the ASMB is criticized for assuming simple manner for managing resources during the simulation runtime. This might not be suitable for some complex system where the essence of using the simulation model is their ability to reflect the most complexity of the system structure – which is captured from the event log data – and the complexity of decisions that influence the behavior of that system. Resource planning is among the complex decisions that have a great impact on the system performance. Therefore, ASMB-H was proposed in the current paper to overcome the problem of resources planning. The resource analysis module within the ASMB-H is developed to allow modeling of each resource (i.e., hospital staff) with various availability for different shifts. After the simulation of resources, KPIs can be reflected on the same way to allow the decision maker to identify which days the resources are over/ under-utilized. This adds more value for the auto-generated DES model for staff planning of healthcare systems.

ASMB-H has introduced several key improvements for simulating the agents that arrive to the system during the simulation runtime. Instead of generating homogenous patients during the runtime, ASMB-H generates heterogeneous patients with different values of a various range of attributes associated with the patients in the actual HPDL. Then MLM-based activity blocks are used to reflect the impact of the patient generation heterogeneity on performance of the system. The MLMs are trained using tailored datasets

from the original HPLD for each activity block to predict patient behaviors at the current activity and the path for the next one.

Classically the auto-generated DES model using the ASMB-H can be utilized for investigating the healthcare system performance against different scenarios related to staff planning and/or changes in patient attributes or arrival patterns. Additionally, it can be used in a more advanced way to predict the system behavior given the actual staff plan and live-stream of data for the actual data of patients arriving to the healthcare system. Feeding these data into the trained MLM of each activity allows the prediction of the flow and processing times of newly arrived patients. In this case the reported KPIs by the model will be a live prediction for these KPIs of the healthcare system for the following hours (e.g., next 5 or 10 hours). Operational wise, this will give early warning for any expected bottlenecks inside the hospital and it will allow the manager to respond proactively by taking the decisions that prevent these bottlenecks from happening. Another business use of the prediction mode of the ASMB-H, is using it on higher planning level for multiple healthcare systems within one region. For example, if ASMB-H is used along with live patient data for EDs of different hospital in same geographical region, it will help the management to predict the overcrowded EDs for the following few hours and hence direct the emergency patients or ambulances to the EDs with a prediction of being less crowded.

In this paper we presented the different components of the ASMB-H and how they are supposed to work and be synchronized together to facilitate the development of the auto-generated DES model given the HPLD. In the following phase of this research – upon completing the development work for the unimplemented modules – the case study data will be used to fully test the functionally of the proposed framework. The outcomes of this application are planned to be presented in another research paper as a continuation for the work presented in the current one.

REFERENCES

Abohamad, W., A. Ramy, and A. Arisha. 2017. "A Hybrid Process-Mining Approach for Simulation Modeling". In *Proceedings of the 2017 Winter Simulations Conference*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Pages 1527-1538, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Aguirre, S., P. Carlos, J. Alvarado. 2012. "Combination of Process Mining and Simulation Techniques for Business Process Redesign: a Methodological Approach". In Proceedings of the 2012 International Symposium on Data-Driven Process Discovery and Analysis, (SIMPDA), Jun 2012, Campione d'Italia, Italy. pp.24-43.

Arisha, A. and W. Abo-Hamad. 2013. "Towards Operations Excellence: Optimising Staff Scheduling For New Emergency Department", Proceedings of the 20th International Annual EurOMA Conference "Operations Management at the Heart of the Recovery", 9-12, June 2013, Dublin, Ireland.

Balaban, M., P. Hester, and S. Diallo. 2014. "Towards a theory of multi-method M&S approach: part I." In 2014 Winter Simulations Conference, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, Pages 1652-1663, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Elbattah, M., O. Molloy, and B. P Zeigler. 2018. "Designing Care Pathways Using Simulation Modeling and Machine Learning". In Proceedings of the *2018 Winter Simulations Conference,* edited by M. Rabe, A.A. Juan, N. Mustafee, A, Skoogh, S. Jain, and B. Johansson, 1452-1463. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Figueira, G. and B. Almada-Lobo. 2014. "Hybrid Simulation-Optimization Methods: A Taxonomy and Discussion". *Simulation Modelling Practice and Theory* Volume 46: Pages 118-134.

Hlupic, V. and S. Robinson. 1998. "Business Process Modelling and Analysis Using Discrete-Event Simulation". In *Proceedings of the 1998 Winter Simulations Conference*, edited by D. J. Medeiros, E. F. Watson, J. S. Carson and M. S. Manivannan, Pages 1363-1369, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Katsaliaki, K., and N. Mustafee. 2011. "Applications of Simulation Within the Healthcare Context". *Journal of the Operational Research Society* 62 (8):1431-1451.

Liu, Y., H. Zhang, C. Li, and R. J. Jiao. 2012. "Workflow Simulation for Operational Decision Support Using Event Graph Through Process Mining". *Decision Support Systems 52*(3):685-697.

Lynch, C., J. Padilla, S. Diallo, J. Sokolowski, and C. Banks. 2014. "A Multi-paradigm Modeling Framework for Modeling and Simulating Problem Situations." In Proceedigs of the 2014 Winter Simulations Conference, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 1688-1699. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Măruşter, L. and N. van Beest. 2009. "Redesigning Business Processes: a Methodology Based on Simulation and Process Mining Techniques." *Knowledge and Information Systems* 21(3):267.

Mesabbah, M., A. Mahfouz, M. A. F. Ragab, and A. Arisha. 2016. "Hybrid Modeling for Vineyard Harvesting Operations." In *Proceedings of 2016 Winter Simulations Conference* ,edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, Pages 1642-1653. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Mesabbah, M, and McKeever, S. 2018. "Presenting a Hybrid Processing Mining Framework for Automated Simulation Model Generation." In *Proceedings of 2018 Winter Simulations Conference*, edited by M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, Pages 1370-1381., Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc

Onggo, B. S., Mustafee N., Smart A., Juan A.A., Molloy O. and Smart, A. 2018. "Symbiotic Simulation System: Hybrid Systems Model Meets Big Data Analytics." In *Proceedings of 2018 Winter Simulations Conference* , edited by N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y-J. Son, Pages 1358-1369, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rashwan, W. 2017. "An Integrated Framework For Staffing And Shift Scheduling In Hospitals", PhD Thesis, College of Business, Technological UniversityDublin, Aunier Street, Dublin, https://arrow.dit.ie/cgi/viewcontent.cgi?article=1032&context=busdoc, accessed April 2019

Skoogh, A., and B. Johansson. 2008. "A Methodology For Input Data Management in Discrete Event Simulation Projects." In *Proceedings of 2008 Winter Simulations Conference,* edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler, Pages 1727-1735, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Taylor, Kathryn, and David Lane. 1998. "Simulation Applied to Health Services: Opportunities for Applying the System Dynamics Approach", Journal of health services research & policy 3 (4):226-232.

AUTHOR BIOGRAPHIES

MOHAMMED MESABBAH, Ph.D is a Data Analytics Manager at Ernst and Young (EY) Dublin. Mohammed obtained his PhD in developing smart planning and decision support tools for complex systems from TU Dublin. He also holds a MSc in system dynamics from university of Bergen (UiB) and bachelor degree from Faculty of Computers and Information in Cairo University. His Research interests include Big Data Analytics, Simulation and Modelling, Machine Learning and Optimization Algorithms. His email address is mohamed.mesabbah@ie.ey.com

WALEED ABO-HAMAD, Ph.D is a Senior Lecturer in College of Business at the Technological University Dublin (TU Dublin) and a Team Leader in 3S Group. His research interests include Modeling and Simulation, Optimization, Computational Intelligence, Machine Learning and Analytics. His email address is waleed.abo-hamad@tudublin.ie.

SUSAN MCKEEVER, **Ph.D** is a Principal Investigator in the Centre for Applied Data Analytics technology group and a Senior Lecturer in the School of Computing at the Technological University Dublin (TU Dublin). Her research is in the domains of autonomous / sensor driven systems and machine learning for both structured and unstructured data. She holds a degree in Electronic Engineering from Trinity College Dublin, an MSc in IT for Strategic Management from TU Dublin and PhD from University College Dublin. Her email address is susan.mckeever@tudublin.ie.