A Generalized Agent-Based Model to Simulate Emergency Departments

Zhengchun Liu, Eduardo Cabrera, Dolores Rexachs and Emilio Luque

Computer Architecture and Operating Systems Department
Universitat Autònoma de Barcelona
Bellaterra, Barcelona, Spain
Email: {zhengchun, ecabrera}@caos.uab.es, {dolores.rexachs, Emilio.Luque}@uab.es

Abstract—Computer simulation based methods have enjoyed widespread use in healthcare system investigation and improvement in recent years. Healthcare systems are based on human interactions and Emergency Departments (ED) are one of the key components of the healthcare system. The efficiency and quality of service in ED have a great influence on the whole healthcare system. The first step to intensively study the emergency department, to find its underlying problem or to provide the best service with limited budget, should be to create a realistic computational model of the ED. Agent-Based Modeling and Simulation (ABMS) is an excellent tool to deal with complex system like ED. This research introduces a generalized ABMS-based computational model of ED. The model has been implemented and verified in a Netlogo modeling environment and can be used to simulate different EDs through a tuning process.

Keywords—Emergency Department; Healthcare; Agent-Based Modeling and Simulating; Complex system.

I. INTRODUCTION

The Emergency Department (ED), a medical treatment facility specializing in acute care of patients who arrive without prior appointment, needs to operate 24 hours per day, 365 days per year. Hundreds of people attend ED per day looking for healthcare services. It is an important entry point to access the healthcare service system. Patients frequently arrive with unstable conditions. Some of them arrive unconscious, and their information such as their medical history, allergies, and blood type may be unavailable. Thus, they should be treated quickly.

In order to set which patient should be visited first, it is mandatory to classify them. Triage is the process of determining the priority of patients’ treatments based on the severity of their condition. This can efficiently improve patients’ treatment process when resources are insufficient for everyone to be treated immediately [1]. The Spanish scale of triage is very similar to the worldwide Canadian one; it consists of 5 levels, with 1 being the most critical (resuscitation), and 5 being the least critical (non-urgent). The triage process also determines the order and priority with which the patient must be attended and the treatment area where they will be treated. This research has been performed with the participation of the ED Staff of the Hospital of Sabadell (a University tertiary level hospital in Barcelona, Spain that provides care service to a catchment area of 500,000 people, and attends 160,000 patients per year in the ED). The model and the simulator will be verified and validated with the data taken from this Hospital.

In general, there are two separate treatment areas (labeled as A and B in this study) in some big EDs to provide diagnosis and treatment service after the triage process. Area A is for those patients with acuity levels 1, 2 and 3 whereas area B is for patients with acuity levels 4 and 5. Area A is occupied by the most urgent patients and is made up of careboxes. A carebox is a small room which contains essential medical equipment and supplies that could be used for patients’ treatment in ED. Patients attended in area A will stay in their own carebox during all the diagnosis and treatment process. Area B is for patients with an acuity level of 4 and 5, which for the Hospital of Sabadell represents 60% of the patients attended in the ED. In area B, there are 3 or 4 attention boxes in which doctors and nurses interact with patients, and a large waiting room in which all patients will remain while not having interaction with the ED Staff. Area A occupies more space than area B.

As for the category of the research object, although the term complexity may have different definitions, according to the definition by Tan et al. [2], a complex system consists of interconnected components that work together, interchange resources and information with the environment in order to meet an objective. This kind of system exhibits several major characteristics: a large number of interactive parts; interactive complexity and self-organization. Thus, there is no doubt that the Emergency Department is a specific case of a complex system. There are no standard models to describe such systems, and analytical models cannot easily represent the complex system caused by random events. With the development of high performance computing techniques, computer based simulation could be one of the best solution to study this kind of system.

The purpose of this work is to develop a general model and simulator that could be used to simulate any emergency department by using an Agent-Based Modeling and Simulation (ABMS) approach. The final objective is to develop a simulator that, used as a decision support system, aids the managers of the EDs, to analyze risks, facilitate coordination implementation, allocate the resources and identify weaknesses in service of resource. In addition, it could be used for studying other related problems in the healthcare service system and as a sensor of ED to generate data concerning different simulation scenarios for finding some unusual knowledge of
the healthcare service by using big-data and data mining techniques. These are three ongoing research lines of the High Performance Computing for Efficient Application and Simulation (HPC4EAS) research group at Universitat Autònoma de Barcelona research group based on the ED simulator. Our previous studies have created the simulator of area B for patients with acuity level 4 and 5 [3][4]. This research is continuing with the previous work to create the model of area A in the ED.

The rest of this article is organized as follows. Section II gives the literature review, a brief introduction to our previous work and the main improvements of this study. Section III is the main part of this article, which has four parts: Section III.A describes the modeling approach for this kind of complex system; Section III.B gives the structure of the model, the definition of the agents and agent behaviors; Section III.C explains the way to model the interactions between agents, and section III.D presents the mathematical-computational model of the diagnosis and treatment phase in detail. Finally, Section IV closes the article with conclusion and future work.

II. RELATED WORK

Rising et al. [5] are among one of the earliest publications on using computer modeling and simulation for improving healthcare service. The authors use the Monte Carlo simulation model for analyzing the effects of alternative decision rules for scheduling appointment periods during the day to increase patient throughput and physician utilization. Hancock et al. [6] developed a computer-based simulator of the hospital systems, which is used for predicting the size of nursing staff configurations under different scenarios.

Concerning the development of the computational model of ED, Paulussen et al. [7] describe a multi-agent based approach for patient scheduling in hospitals. In such a system, patients and hospital resources are implemented as autonomous agents in which the resource agents view the patients as entities to be treated, and the patient agents view the medical actions as tasks that need to be performed. The coordination of patients is achieved through a market mechanism. Patient agents negotiate with each other over scarce hospital resources, using state health dependent cost functions to compute bid and ask prices for time slots. Within this concept, stochastic processing times and variable pathways are considered. Unfortunately, the system does not take into consideration patient variety or the different kinds of healthcare staff. But in fact, the variety of patients and staff has great influence on the performance of ED.

As the use of simulation approach for studying EDs, Badri and Hollingsworth [8] developed an Emergency Room (ER) simulation model incorporating the major activities. The model allows the evaluation of “what if?” questions through changing the values of the variables and simulating the results. The ER simulation model determines the effects of changes in the scheduling practices, allocation of scarce resources, patient demand patterns, and priority rules for serving patients. In the study of Gove and Hewett [9], they examined the problem of capacity in hospitals and proved that: due to the complexity of the hospital and its departments, simulation was an ideal choice to study. Moreover, Diefenbach et al. [10] found that varying the number of beds, physical layouts, access to radiology and pathology services etc. in the ED has an exponential effect on expression of the system. The simulation results under the change of the configurations can provide valuable reference for management decision making. Kuljis et al. [11] compared the healthcare system with business and manufacturing, and provided the feasibility of using modeling and simulation methods to improve the Quality of Service (QoS) in healthcare system.

For the use of ABMS approach for simulating EDs, Macal et al. [12] gave a tutorial to create an agent-based model for the complex system, and they suggested that ABMS promises to have far-reaching effects in the future on how to use computers to support decision making. As for the reason for choosing ABMS approach for simulating ED, Escudero-Marín et al. [13] gave the reason that ABMS is better for modeling the EDs than others. The authors also provided a general description of the possible potential use of ABMS in healthcare application.

The previous studies in our HPC4EAS research group mainly included creating the simulator of area B in ED [3], balancing between the budget and QoS, finding the optimal and sub-optimal resource configurations of ED to achieve better QoS with limited budget by using K-means methods and pipeline scheme [4][14].

Unlike area B, area A is the area dedicated to the critical patient. It is more complex and quite different with area B mainly because the patients in this area usually cannot move by themselves; consequently, the doctor and other auxiliary staff need attend these patients in their carebox. These cases lead to a greater amount of restrictions and interactions between the agents in ED. Compared with our previous model, the main improvements and contribution of this article include: considering some more agents, modifying the behavior of some exist agents, introducing a new way to define and simulate the interactions between agents and state transition of the agents, and providing an easy-tuning model to simulate the diagnosis and treatment phase.

The model created in this study is a generalized model. A tuning process is necessary before simulation. In this study, the tuning process is a process to adapt the generalized model to the real ED to be simulated. It is done by using the historical data of the given ED. The difference between the simulation results and the real data will correct the value of the model’s internal parameters. After a series of feedback and iterations, when the difference reaches to an acceptable value, the model is adapted to the real ED.

III. SIMULATION MODEL

Conducting a valid simulation is both an art and a science. One of the main challenges when developing a general simulation model is to keep a model as simple as possible whilst including all the key system information to achieve the objectives of the simulation. One feasible way to do this is through the following three steps: (1) survey multitude real models; (2) analyze the concept structures of these real models; and (3) abstract and generalize from these real models to develop a reusable generic pattern model. This section detailed the general model of EDs.

A. Modeling approach

When faced this kind of complex system, it is almost impossible to model all its functionality directly because there
are large numbers of factors that can affect the result and need to consider. A good way to model is by using a bottom-up-modeling approach. Starting from the bottom subsystem (agents, agents’ behavior and interactions between agents), the execution of the simulator will cause a large amount of interactions between these agents, and then these interactions will emerge the functionality of the emergency department indirectly.

As shown in Figure 1, it works by modeling the agents, their behaviors and interactions between them. Then, when executing the simulator, the state of agents will be changed by their interactions, and the queue for the interactions and their Length of Waiting Time (LoWT) will emerge, by such analogy, the functionality of the ED will emerge indirectly through the execution of the simulator. Furthermore, the QoS can be evaluated through the results of different simulation scenarios.

Discrete Event Simulation (DES), System Dynamics (SD) and ABMS are the three main approaches used when simulating this kind of complex system. There is a large body of literature describing the use of DES models in ED studies, whilst there is considerably less literature on the use of ABMS simulating this kind of complex system. There is a large body of literature describing the use of DES models in ED studies, whilst there is considerably less literature on the use of ABMS for this purpose. As healthcare systems are based on human actions and interactions, combined with our experience and requirement, it can be more properly to model with ABMS [13]. ABMS models can offer ways to provide a deep insight view and to generate hypotheses about system behavior by representing this as a result of the interaction between the agents.

B. Agents and Agent Behaviors

In agent-based terms, a system is modeled as a set of heterogeneous agents that will create the overall behavior of the system through their interactions in the execution process of model. When developing ABMS, it is crucial to represent the two main parts of an agent-based model: agents and agent behaviors.

The definition of agents should include their capabilities, the actions they can perform and the characteristics of the environment that surrounds them. In this article, each kind of agents is defined as:

$$A = \{ V \cup B \}$$  \hspace{1cm} (1)

where $V$ is a set of state variable to represent the agents’ characteristics:

$$V = \{ V_1, V_2, \ldots, V_m \}$$  \hspace{1cm} (2)

The states of the agents are indicated by the values of the state variables. And set $B$ contains all the behaviors of the agents in this category:

$$B = \{ B_1, B_2, \ldots, B_n \}$$  \hspace{1cm} (3)

Each kind of agent has its own definition of state variable $V$ and each state variable has a set of possible values in its range:

$$V_i = \{ Y_1, Y_2, Y_3, \ldots, Y_{K_i} \} (0 \leq i \leq m)$$  \hspace{1cm} (4)

In Table I, each kind of agent has its own behavior. The behaviors are generalized that do not represent one specific action, instead, the combination of the value of their state variables and the generalized behavior will represent the real action. For example, if the value of state variables indicates that the patient stays in the waiting room, waiting for service after admission, which means that they are waiting for triage service instead of other services.

As to the diagnosis service, there are different kinds of tests, for example blood test, x-ray, B ultrasonic, and so on. In reality, most of the time spent on diagnosis was on waiting,
waiting for the service and waiting for the result. The LoWT depends on the length of the waiting queue which emerged as a result of the agents’ interactions. The length of the waiting queue depends on the number of patients, their acuity level distribution and the amount of available resources in ED. That is the reason why the LoWT emerged by executing the simulation instead of being modeled directly. From this point of view, it is better to take the test service as agents. In this manner, the patients need to interact with these agents to know their body condition, because of the limited number of these agents, the patient usually needs to wait for the interaction, as shown in Figure 1, the queues will emerge in this way. Therefore, in Table I, all these tests were classified into three types according to their interactive mode:

Laboratory test: It is a kind of test performed in a carebox and laboratory, when the doctor ordered this kind of test, the nurse will go to the carebox to take some samples and then send them to the laboratory (or analysis by nurse directly), after a period of time, the doctor will get the result. Due to the test being an agent, the LoWT could emerge from simulation. The main characteristic of this kind of test is that the patients do not need to move, hence there is no interaction with the auxiliary staff, but instead, the nurse will have some interactions to perform with the patient and laboratory.

Internal test: The internal test means the diagnostic equipment is owned and used only by ED, thus the length of the waiting queue for the service is under-control and can emerge through simulation. Unlike a laboratory test, in order to do this kind of test, the patients need to move, so they need to interact with the auxiliary staff, which may cause longer LoWT if the configuration of auxiliary staff is inadequate.

External test: In ED, there are some types of diagnostic equipment shared with the hospital wards or even shared with other EDs, therefore the length of the waiting queue cannot emerge because part of the agents outside ED who need to interact with these test agents but do not appear in the simulator. One way to simulate is by using a period of time delay (based on statistical data and following probability distribution obtained from tuning process) to model this kind of test. As with an Internal test, in order to do this kind of test, the patients need to interact with auxiliary staff to move them to the corresponding test room.

In addition to this, the ambulance and hospital wards were also considered because the behavior of these two agents also have obvious effects on the functionality of ED:

Ambulance: Some patients come by ambulance, especially the patients in area A. Part of them will do admission and triage in the ambulance, when the patient is critical enough, and on arrival they will go to a carebox directly or stay in the ambulance until a free carebox is available. At the same time, some patients may need to go home by ambulance. But usually the ambulance has arrival delay. Under this circumstance, they will keep using the carebox. Hence, the quality of ambulance service is one of the factors that may cause overcrowding in ED.

Hospital ward: A hospital ward is a main exit way for the patients in area A. It is common that the hospital ward does not have enough free beds, thus the patient will keep using the carebox even though this is not necessary, so the throughput of hospital wards also has direct influence on the performance of ED. For simulating the hospital ward, it is similar to the external test agent. The number of free beds and available time will be simulated through the probability distribution and the parameters of the distributions are obtained from analyzing the real data in the tuning process.

C. Model of the Interactions

The functional behavior of any system can be specified by a state machine (also called an object) [15]. In this research, to model the interactions between the agents and their states, the Finite State Machine (FSM) was used.

According to the definition of agents through (1)-(4), the state of agents are presented by the value of their state variables and each state variable has a set of possible values. Based on the actual situation, the transition of agents’ state is caused by interaction with other agents or in some cases with time elapse. Thus, the value of the state variables are changed by one of their behaviors or time elapsing, as in (5):

$$Y_{K_i} = f(B_j, T)(0 \leq i \leq m, 0 \leq j \leq n)$$

where $B_j$ represents the corresponding behavior with other agents, it is an element of the behavior set $B$. $T$ represents the elapsing of the time because sometimes the state of the agents, e.g., patients’ body condition after medicating, can change with time goes on without any interactions.

As shown in Figure 2, the state machine accepts commands and produces outputs, which means that when the agents interact with other agents and/or with the time elapsing (accept input), the value of one or several variables will be changed (because of the outputs produced). Any one of the variables’ value changing will represent the state transition.

$$S = \{S_0, S_1, S_2, \ldots, S_t\} = \{V_1 \times V_2 \times V_3 \times \ldots \times V_n\}$$

$$(0 \leq t \leq \prod_{i=1}^{n} K_i) (6)$$

Therefore, the set of one kind of agents’ states is the cartesian produce of each state variable’s possible value set(see (6)). The state set of a specific agent in this type is a subset of $S$, which is determined by the specific configuration of the agent.

Figure 2. A typical patient’s conceptual state transfer model.
TABLE II. A PART OF A PATIENT’S STATE TRANSITION.

<table>
<thead>
<tr>
<th>State index</th>
<th>Source State</th>
<th>Destination state</th>
<th>Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Sf</td>
<td>Waiting for service (free carebox)</td>
<td>Waiting for service (Doctor’s diagnosis)</td>
<td>Notice from IS with a free care box.</td>
</tr>
<tr>
<td>Sf+1</td>
<td>Waiting for service (doctor’s diagnosis)</td>
<td>Accepting Service(meet with doctor)</td>
<td>Doctor arrive at patient's carebox.</td>
</tr>
<tr>
<td>Sf+2</td>
<td>Accepting Service(meet with doctor)</td>
<td>Waiting for service (X-Ray test service)</td>
<td>Doctor order X-Ray test for patient.</td>
</tr>
<tr>
<td>Sf+3</td>
<td>Waiting for service (X-Ray test service)</td>
<td>Accepting Service(X-Ray test service)</td>
<td>X-Ray service available.</td>
</tr>
<tr>
<td>Sf+4</td>
<td>Accepting Service(X-Ray test service)</td>
<td>Waiting for service (Doctor’s review of the test result)</td>
<td>X-Ray service finished.</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

By combining with Table I about the generalized behavior of agents, (1) - (6) for the definition of agents and Figure 2 about the conceptual state transfer, it is feasible to list all the agents’ evolution during the stay in ED, Table II gives a part of one patient’s state transition. In Table II, although some of the state is the same as Waiting for service, the value of its state variables will determine the specific service the patient waiting for.

Above all, by means of defining the agents through state variables, it will be feasible to deal with the huge amount of states of the agents. At the same time, it will be easy to add/remove states simply by adding/removing elements in the set of possible values of the state variables and their corresponding behavior. For the study of other ED related problems, for example, the virus propagation in ED, some new state variables and their possible values will become easy to be added to indicate some more states. With the same approach, by the execution of the model, some new functionality of the research object will emerge from these new states.

D. Diagnosis and Treatment Phase

Before the diagnosis and treatment phase, for the patients, they need to do admission and triage, actually these two phases take very little time in reality. Figure 3 indicates the common process in the emergency department. For the patients, especially the patients with acuity levels 1, 2 and 3, most of their Length of Stay (LoS) in the emergency department is spent in the carebox for having various kinds of tests, receiving treatment and waiting for the treatment to take effect. This is the most important part of the model because most of the state transitions take place during this phase.

For the patients in area A (with acuity levels 1, 2 and 3), they always stay in their own carebox during the diagnosis and treatment phase. The general process in the carebox is that the patient takes some tests (x-ray, ultrasound, blood test and so on), then the doctor reviews the test result and provides one treatment plan or asks to do further tests. After that, the nurse will carry out the treatment plan or take some test samples if a laboratory test is ordered. After a period of time, some state variables of the patient will be changed or the patient unfortunately die. With the change of the patients’ body condition, the doctor will decide what the patient need to do next: to go the hospital ward, go home or continue with diagnosis and treatment. In order to generalize the process of all the patients, the next state will be decided by probability distribution during simulating. The distribution model of the probability was based on the statistical data from the real EDs.

Figure 3 indicates the general process-transfer strategy during the patients’ stay in EDs.

In Figure 3, $P_1(\%)$, $P_2(\%)$, $P_3(\%)$ and $P_4(\%)$ represent the probability of the next state transition separately. $P_1(\%)$ and $P_2(\%)$ represent the decision of the doctor after reviewing the test results and body condition of the patients in probability. All of the probabilities follow some probability distributions. The probability density function of the distribution is decided by several key parameters based on the statistical analysis of doctors’ decision and patients’ behavior, the value of these parameters are estimated by a tuning process from real historical data of the specified ED. The uniform forms of the density functions are:

$$P_i = f(LoS, age, level)$$  \hspace{1cm} (7)

$$\sum_{i=1}^{4} P_i = 100\%$$  \hspace{1cm} (8)

$$P’ = f’(ToT, age, level)$$  \hspace{1cm} (9)

$$\sum_{i=1}^{2} P’_i = 100\%$$  \hspace{1cm} (10)

where $LoS$ is the patient’s length of stay in the carebox. $age$ is the age of the patient, which also has big influence to the probability of state transition. $level$ is the acuity level of the patient. And $ToT$ is the type of test service or diagnosis by doctor.

The function $f$ and $f’$ are the probability density function. These functions will be implemented by analyzing real historical data in tuning process. This work can only be done in the tuning process because different EDs have different characteristics, it is a part of the simulator instead of the general model. Therefore, combined with (1) - (10), every patient will show different behavior during the execution of the model because of the probability distribution and their own differences in body condition. But the statistical property of agents will reflect their common behavior.

IV. CONCLUSION AND FUTURE WORK

Simulation methods have long been used to model elements of healthcare systems with a view to analyzing new system designs, retrofitting to existing systems and proposing changes to operating rules. The Emergency Department (ED) is a typical complex system. To perform intensively study, a realistic computational model is compulsory. An approach to modeling this kind of system is by using agent-based modeling and simulation, which is a kind of bottom-up modeling approach. This paper presents a generalized agent-based model of the emergency departments. It was designed based on the survey of different EDs and with the participation of sanitary staff in ED. This model has been implemented and verified in a Netlogo modeling environment. It is not dedicated with one
specific ED, which can be used to simulate different EDs through tuning process.

This research is a progress of our previous work. To model the critical part of the ED (here, we call it area A), we defined some new agents also added some new state variables to extend the behavior of the previous agents. A new way to simulate the interaction between the agents and the state transition of the agent was provided, an easy tuning model was created for diagnosis and treatment phase in area A. In reality, most of the emergency department works like area A. For those big EDs, they have both area A and B, hence with both the model of A and B, we have the model of the whole EDs.

Creating the computational model of the object is the first step of simulation. Model verification is the task to ensure that the model behaves as intended, some basic experiments has been done to verify the functionality of the model. But, in order to validate the simulator, tuning for some real EDs is mandatory. Therefore, the first step of future work should be validation. Some real historical data of EDs will be asked to perform the tuning process. Moreover, during the tuning process, due to the large number of parameters for the model, and the large number of agents and interactions between them. To increase the number of studied scenarios and reduce execution time as well, the use of high performance computing will be mandatory.

In addition, the ED is the main entrance to the healthcare service; some problems of the healthcare service system are caused by the performance of ED. However, the ED is not independent, all the departments of healthcare system influence each another. Thus, our future work also include creating the simulator of other departments, for example the hospital wards to close the simulation loop of the whole healthcare service system. After that, the simulators of these departments will work as the sensor of the healthcare service system. The data generated from these sensors will be analyzed through data-mining and big-data techniques to find some unusual knowledge of the system to provide smarter service to patients.

ACKNOWLEDGEMENTS

This research has been supported by the MINECO (MICINN) Spain under contract TIN2011-24384 and partially supported by China Scholarship Council (CSC) under reference number: 201306290023.

REFERENCES