Computers & Industrial Engineering 103 (2017) 300–309

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

A simulation and optimization based method for calibrating agent-based emergency department models under data scarcity

Zhengchun Liu [a,*], Dolores Rexachs [a], Francisco Epelde [b], Emilio Luque [a,*]

[a] Computer Architecture & Operating Systems, University Autonoma of Barcelona, Barcelona, Spain
[b] Medicine Department, Hospital Universitari Parc Tauli, Barcelona, Spain

ARTICLE INFO

Article history:
Received 4 March 2016
Received in revised form 9 September 2016
Accepted 28 November 2016
Available online 2 December 2016

Keywords:
Simulation-based optimization
Model calibration
Agent-based model
Emergency department

ABSTRACT

To tackle the problem of efficiently managing increasingly complex systems, simulation models have been widely used. This is because simulation is safer, less expensive, and faster than field implementation and experimenting. To achieve high fidelity and credibility in conducting prediction and exploration of the actual system with simulation models, a rigorous calibration and validation procedure should firstly be applied. However, one of the key issues in calibration is the acquisition of valid source information from the target system. The aim of this study is to develop a systematic method to automatically calibrate a general emergency department model with incomplete data. The simulation-based optimization was used to search for the best value of model parameters. Then we present a case study to particularly demonstrate the way to calibrate an agent-based model of an emergency department with real data scarcity. The case study indicates that the proposed method appears to be capable of properly calibrating and validating the simulation model with incomplete data.

1. Introduction

With the rapid growth of computational techniques, computational thinking brings researchers and practitioners into a new dimension of traditional modeling and simulation tasks. That is, the computational science transforms observed complex phenomena into conceptual models. Then the models are formulated into algorithms that can be executed to yield predictions and estimate hidden parameters. This generates an additional understanding of the phenomenon and leads to more specific models of the phenomenon (Sloot & Quax, 2012). From a theoretical computation perspective, the simulation of a system can be defined as an *"imitation (on a computer) of a system as it progresses through time"* (Robinson, 2004). Although a simulator is mostly designed for prediction, the simulator should firstly be able to imitate the real system. Generally, a simulator of a specific system is comprised of the following: input (X), the model or transformation function ($f(X)$), and output (Y). For an accurate simulator, when we put the same input as in a real system, the output of simulator should be close enough to the output of the actual system. Since $f(X)$ is based on abstractions, idealization, and many disputable assumptions, the model must be fine-tuned according to some historical input-output samples from the target system in order to get reliable simulation results.

The emergency department (ED) is a typical complex system, which serves essential needs in society, delivering emergency health care and simultaneously acting as a safety net provider (Hoot, Epstein, Allen, & Jones, 2009). In recent years, simulation has emerged as an increasingly effective tool to study ED related problems and support making decisions to efficiently manage the complex ED system. While these simulation models can be advantageous to engineers, the models must be calibrated and validated, i.e., the model should first be able to accurately imitate the real system. Advances in computational technology, along with the increased complexity of system design and management, have created an environment in which microscopic simulation models have become useful tools for managing complex system. Among which, the Agent-based Model (ABM) is one of the most important tools for exploring emergent behavior (a phenomenon that describes the behavior of a system, which cannot be explained alone by the sum of its parts Wagner, Cai, Lees, & Aydt, 2015), mostly because it can provide a way to see the forest through the trees and insight is often more important than sheer numbers (Gul & Guneri, 2015; Heard, Dent, Schifeling, & Banks, 2015; Rudin, 2014).

The agent-based simulation models encompass numerous independent parameters to describe the individual behavior of

* Corresponding authors at: Computer Architecture & Operating Systems, University Autonoma of Barcelona, Bellaterra 08193, Spain.
E-mail addresses: lzhengchun@caos.uab.es (Z. Liu), emilio.luque@uab.cat (E. Luque).

http://dx.doi.org/10.1016/j.cie.2016.11.036
0360-8352/

the system components. Reliable and complete real data from the target system is obviously the precondition for setting up an accurate simulator. Unfortunately, many of the parameters are either unavailable in historical data or difficult to measure in a real situation, yet they can have a substantial impact on the model's accuracy. Thus, when real data was incomplete to allow direct estimation of the model parameters, a calibration process (also known as tuning) has to be conducted to indirectly estimate good values for those unknown parameters. However, the calibration of model parameters for an ABM is a big challenge for standard calibration techniques, due to the large parameter search-space, long simulation runtime, uncertainties in the structural model design and different observation levels upon which the model needs to be calibrated (Fehler, Klügl, & Puppe, 2006). Given this, the model parameter calibration problem can be formulated as a stochastic programming problem whose objective function is an associated measurement of an experimental simulation. Nevertheless, the objective function is typically (a) subject to various levels of randomness, (b) not necessarily differentiable, and (c) computationally expensive to evaluate due to the complexity of the model.

Accordingly, conventional calibration, which is carried out manually by using the trial-and-error method, is time consuming and tedious. A systematic method to automatically search for the optimal value of model parameters is promising. The simulation-based optimization is an emerging field which integrates optimization techniques into simulation analysis. The primary goal of simulation-based optimization is to optimize the performance of a system through simulation. More specifically, it is a way to find the optimal set of parameters for a given criterion. Then the optimal parameter set will enable the model to achieve a specific function optimally or the results of the simulation are close enough to actual data. Therefore, if we set the model input the same as reality and we consider the unknown model parameters as variables, and the similarity between simulation output and actual system output as objective, the optimization is a model calibration process. When some of the model parameters are missing and impossible to get from the real system, this optimization process will be able to find the optimal values for setting up the model. Thus, the precondition for the calibration process is a set of reliable input-output pairs from the target system.

In this article, we will address a critical step in simulating a complex system - the systematic model calibration in the face of data scarcity. To the best of our knowledge, limited research has been conducted on this thorny and critical problem of estimation in the face of data scarcity. The simulation-based optimization was conducted by using an existing tool (Gray & Kolda, 2006; Griffin & Kolda, 2006; Kolda, 2006) developed by Sandia National Laboratory. According to the practical requirements of evaluating a simulation-based objective function, an initial distance-based lookup mechanism was proposed to further speed up the optimization. The rest of the paper is structured as follows: Section 2 gives a literature review on related work. The method to calibrate agent-based ED model is given in Section 3. With the presented method, Section 4 gives a case study which calibrates a general agent-based model of an ED to simulate the ED of the Hospital of Sabadell (a university tertiary level hospital in Barcelona, Spain) with incomplete data (missing duration of key services). This case study will thoroughly demonstrate the way to calibrate an agent-based ED model by using the presented method. Finally, Section 5 draws the conclusions.

2. Related work

Model calibration is the task of adjusting an already existing model to a reference system. Trucano et al. thoroughly discussed the relation of calibration and validation in reference Trucano, Swiler, Igusa, Oberkampf, and Pilch (2006). They identified some technical challenges that must be resolved for successful validation and calibration of a predictive modeling capability. Their findings proved the possibility of validation and highlighted great practical difficulties associated with model parameter calibration and validation. Hofmann (2005) introduced a formal approach to model calibration, within the frame of the presented formalism it is shown that the computational complexity of model calibration is NP-complete. The author addressed the issue that for huge model federations the complexity of parameter calibration could draw a serious line with respect to the validation of the federation and its cost-benefit ratio. This is mostly because in a huge model of a complex system, no single person has an overview of the whole simulation, and the interpretation of unexpected results is extremely difficult. Therefore, a manual trial-and-error method does not work for this kind of model (e.g., an agent-based model).

As described in Section 1, the model parameter calibration process can be easily formed as a simulation-based optimization process. Due to the complex behavior of the objective function, Evolutionary Algorithms (EAs) are often used to efficiently explore large parameter spaces. However, EA still takes a considerable amount of time because it requires a large number of simulation runs, and each run takes considerable length of time in simulation. To this end, Wagner et al. (2015) proposed the use of complexification to improve the performance of EAs as it emulates the natural way of evolution. This method has been used for parameter estimation of multi-agent based models. Zhong, Hu, Cai, Lees, and Luo (2015) proposed an evolutionary framework to automate the crowd model calibration process. In the proposed framework, a density-based matching scheme is introduced. By using the dynamic density of the crowd over time, and a weight landscape to emphasize important spatial regions, the proposed matching scheme provides a generally applicable way to evaluate the simulated crowd behaviors. Besides, the authors also proposed a hybrid search mechanism based on differential evolution to efficiently tune parameters of crowd models. And in reference Zhong and Cai (2015), Zhong et al. proposed another novel evolutionary algorithm named differential evolution with sensitivity analysis and Powell's method (DESAP) for model calibration. The proposed DESAP first applies an entropy-based sensitivity analysis operation to dynamically identify important parameters of the model. Then, Powell's method is performed periodically to fine-tune the important parameters of the best individual in the population. Finally, in each generation, the evolutionary operators are performed on a small number of better individuals in the population. Their new search mechanisms are integrated into the differential evolution framework to improve search efficiency. In summary, different from conventional mathematical models, the calibration of agent-based models has its own challenge and are attracting researchers' attention. However, all of these developed algorithms are mostly focused on solving agent-based crowd behavior model calibration problems with complete data, in which the system metrics and objective function are different from the requirement of tuning an agent-based ED model.

In contrast to traditional black box search methods, which only consider the input and output of simulation model, Fehler et al. (2006, 2005) proposed a promising white box calibration approach, which uses the knowledge of the agent-based model to improve the tuning process. In this, the idea is to reduce the parameter space by breaking down the model into smaller sub-models. Each of the sub-models is then calibrated before merging them back to form the model. However, in this method, the division and fusion operations are difficult steps and they require additional knowledge about the model, and this knowledge may not be available for simulation users (non-developer). Moreover, the fusion