SCIENCE CHINA
Information Sciences

• LETTER •

May 2020, Vol. 63 159203:1–159203:3
https://doi.org/10.1007/s11432-018-9559-9

Pattern-based validation metric for simulation models

Yuanjun LAILI[1], Lin ZHANG[1*] & Yongliang LUO[2]

[1]*Beijing Advanced Innovation Center for Big Data-based Precision Medicine, School of Automation Science and Electrical Engineering, Beihang University, Beijing* 100191*, China;*
[2]*China State Shipbuilding Corporation System Engineering Research Institute, Beijing* 100191*, China*

Received 12 April 2018/Accepted 29 June 2018/Published online 27 September 2019

Citation Laili Y J, Zhang L, Luo Y L. Pattern-based validation metric for simulation models. Sci China Inf Sci, 2020, 63(5): 159203, https://doi.org/10.1007/s11432-018-9559-9

Dear editor,
Simulation is one of the most important technologies that aid in the study of real-world phenomena. It helps engineers in making better decisions related to complex systems when mathematical models are partially or even completely unknown. A simulation model is an analog of real-world objects. Only appropriate and qualified models can ensure applicable and valuable simulation processes and results [1].

Model credibility defines the quality that elicits trust in the simulation results associated with a simulation objective [2]. Verification and validation, i.e., V&V [3], measure whether a model is correctly implemented according to its specifications, and whether it performs close enough to the actual requirement, respectively. These measures are two crucial means of assessing the credibility of a model. Verification is usually accomplished by using various logical checking methods with an interactive debugger introduced through software engineering [4]. Therefore, most research attention in modeling and simulation, M&S, has been focused on model validation [5]. Researchers search for suitable experiment design and sampling strategies that reduce the duration of simulation, design appropriate domain-related hierarchical indicators and questionnaires, and introduce various statistical methods and expert scoring methods to evaluate model credibility.

With the growth in simulation objects and the increase in collaborative operations in a complex system, simulation models with multiple underlying components can no longer function in a pre-defined procedure. Multi-disciplinary components are gradually integrated together to act autonomously and randomly. Uncertainties in real-world circumstances can therefore be fully demonstrated through individual activities and local undeterministic logics. Such decentralized individual actions governs several additional dynamic characteristics and evolving behaviors of the model [6]. The one-to-one error analysis of simulation results compared with a particular series of reference data can no longer be convincing because the model and the simulation target change over time [7, 8]. Consequently, traditional guidelines and evaluation processes for various types of models become irrelevant.

This study presents a pattern-based validation metric to complementarily evaluate the credibility of a simulation model based on an overall similarity perspective. Our basic idea is to automatically organize the implicit regularity of the simulation target and compare the patterns produced by the simulation model to this regularity. The reference data is firstly clustered into several groups. The simulation results of the model under specific input configuration is then compared to the members of the closest cluster. Following this, an evaluation criterion is developed to assess the degree to which the simulation model is adaptable in terms of its

* Corresponding author (email: zhanglin@buaa.edu.cn)

input and output patterns.

Methodology. In traditional model validation methods, one can simply assume the simulation output at a specific time as $x = \{x_1, x_2, \ldots, x_n\}$ and the reference data from the real world as $y = \{y_1, y_2, \ldots, y_n\}$, in which n represents the number of time slots being observed. Apparently, the input configuration of the model should be exactly the same as the target object. The simulation length should be scaled by following the length of the reference data as well. The main objective is to compare the consistency of the two sequences.

However, validating the simulation output under only one input configuration does not ensure that the model is credible. Because of evolving dynamics in real-world objects, more and more simulation tests should be carried out repeatedly because specific models perform differently in each test. It is cumbersome to validate the model on every input configuration with multiple tests. More importantly, lack of sufficient reference data for every possible configuration makes it challenging to perform the one-to-one comparisons between the simulation model and the target object.

In most cases, we can only obtain a limited number (J) of reference data $\boldsymbol{Y} = \{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_J\}$, which contains some unknown noises and dynamics. Each data $\boldsymbol{Y}_j$, $j \in [1, J]$ can either be a single-dimensional time series or a multi-dimensional time series shown as follows:

$$\boldsymbol{Y}_j = \begin{bmatrix} y_{11}^{(j)} & y_{12}^{(j)} & \cdots & y_{1t}^{(j)} \\ y_{21}^{(j)} & y_{22}^{(j)} & \cdots & y_{2t}^{(j)} \\ \vdots & \vdots & & \vdots \\ y_{D1}^{(j)} & y_{D2}^{(j)} & \cdots & y_{Dt}^{(j)} \end{bmatrix}, \tag{1}$$

where D and t represent the dimension and the number of time steps, respectively. In addition, we can produce a group of simulation outputs $\boldsymbol{X} = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_I\}$, where I represents the number of simulation outputs that correspond to some special input configurations. Each output $\boldsymbol{X}_i$, $i \in [1, I]$ is also represented as a single-dimensional time series or a multi-dimensional vector shown as follows:

$$\boldsymbol{X}_i = \begin{bmatrix} x_{11}^{(i)} & x_{12}^{(i)} & \cdots & x_{1t'}^{(i)} \\ x_{21}^{(i)} & x_{22}^{(i)} & \cdots & x_{2t'}^{(i)} \\ \vdots & \vdots & & \vdots \\ x_{D1}^{(i)} & x_{D2}^{(i)} & \cdots & x_{Dt'}^{(i)} \end{bmatrix}. \tag{2}$$

On account of the input configuration, we define the pattern of a simulation model as a combinational regularity between an input configuration and the corresponding output features. Let $\boldsymbol{S} = \{s_1, s_2, \ldots, s_m\}$ be the parameter setting for starting a simulation, where m represents the number of conditional parameters of the model. The patterns extracted from the simulation model and the reference data are denoted as $\boldsymbol{H}_i^{(\boldsymbol{X})} = \{\boldsymbol{S}_i^{(\boldsymbol{X})}, \boldsymbol{X}_i\}$ and $\boldsymbol{H}_j^{(\boldsymbol{Y})} = \{\boldsymbol{S}_j^{(\boldsymbol{Y})}, \boldsymbol{Y}_j\}$, respectively. It should be noted that we do not need to strictly match the simulation patterns in $\boldsymbol{H}_i^{(\boldsymbol{X})}$ with the reference data in $\boldsymbol{H}_j^{(\boldsymbol{Y})}$ as the original pattern matching techniques [9].

The procedure for calculating the pattern-based validation metric is shown in Figure 1. The procedure includes two flexible steps. The first step is to organize the reference data into clusters according to their similarity. Because a pattern is represented by a real vector, the similarity between two patterns can be calculated according to their Euclidean distance. By using incremental clustering techniques, the reference clusters can be updated through an online mechanism, thereby allowing additional reference data to be imported. The second step is the validation of patterns in accordance with the reference clusters. We define the pattern match degree F as the degree of compatibility of a simulation model with a group of reference patterns. Let u be the number of compatible patterns with different input configurations in $\boldsymbol{H}^{(\boldsymbol{X})}$, which can be obtained using

$$F = u/\|\boldsymbol{H}^{(\boldsymbol{X})}\|. \tag{3}$$

In the beginning u is set to 0. For each simulation result in $\boldsymbol{H}^{(\boldsymbol{X})}$ that corresponds to a particular input configuration, we must first determine the closest reference pattern g from $\boldsymbol{H}^{(\boldsymbol{Y})}$ according to their Euclidean distances. To assess whether the given result is close enough to g, we designate a similarity threshold T_g according to the cluster $\boldsymbol{C}_g$ that g belongs to, as shown in Eq. (4).

$$T_g = \begin{cases} \max\limits_{g' \in \boldsymbol{H}^{(\boldsymbol{Y})}/\boldsymbol{C}_g} d(g, g'), & |\boldsymbol{C}_g| \leqslant 1, \\ \max\limits_{g' \in \boldsymbol{C}_g} d(g, g'), & \text{otherwise.} \end{cases} \tag{4}$$

In other words, when the size of the corresponding cluster $\boldsymbol{C}_g$ is larger than 1, the largest distance between g and other members in $\boldsymbol{C}_g$ is considered the threshold. Otherwise, the smallest distance between g and other members in $\boldsymbol{H}^{(\boldsymbol{Y})}$ is calculated as the threshold.

If the distance between the imported simulation result and g is within the threshold, we consider the result to be compatible with the reference data, and therefore $u = u+1$ is performed. When all the patterns in $\boldsymbol{H}^{(\boldsymbol{X})}$ are processed, the total number of compatible patterns can be obtained, such that

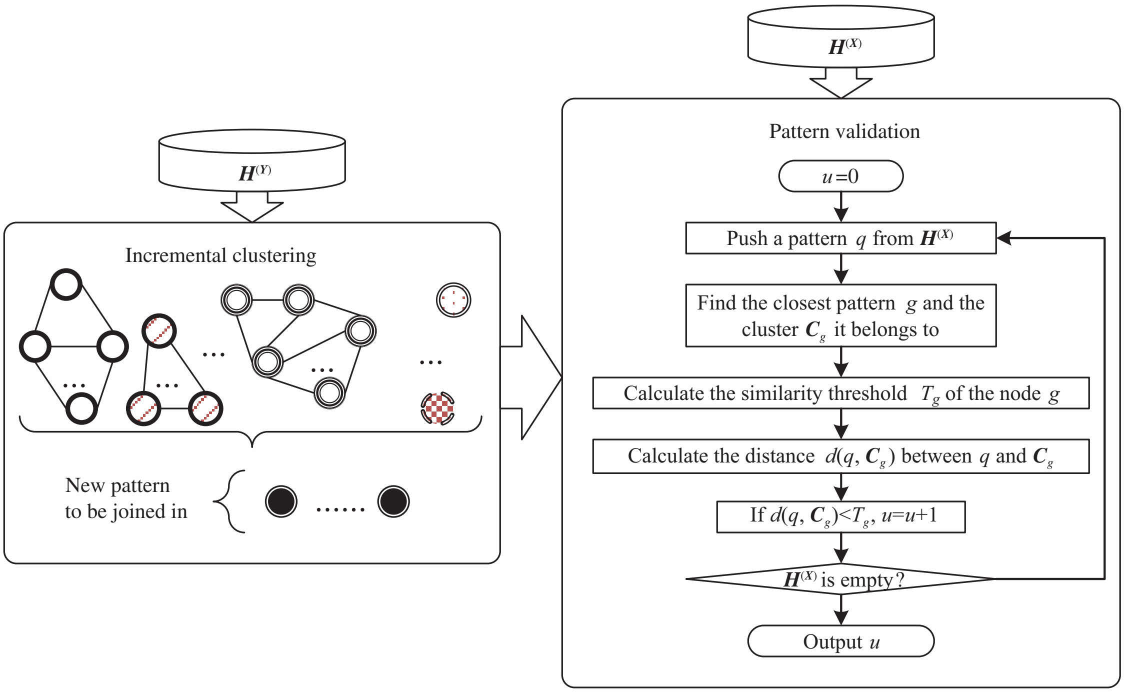

Figure 1 (Color online) The framework of the pattern-based validation scheme.

the pattern matches the degree, which forms the pattern-based validation metric.

Conclusion. We performed our research on the validation of simulation models from a new perspective. A pattern-based metric was proposed to validate the simulation model with uncertain activities. A similarity threshold was designed based on the clustering of the reference data obtained from the real world. The number of compatible patterns produced by a simulation model then represents its degree of matching to a target dataset. In the proposed validation scheme, the size of the reference data and that of the simulation results are not necessarily the same. Even if the reference data for a specific input configuration is insufficient, the degree of compatibility of a model can still be calculated according to other similar data organized in the clusters.

The patterns that exist in a series of simulation results are critical to the evaluation of model credibility. With increasing dynamic behaviors, the length of the simulation pattern and reference data will grow exponentially. Our future work will primarily focus on the compression of simulation results and the application of the proposed metric for different types of models.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61703015).

References

1 Mehta U B, Eklund D R, Romero V J, et al. Simulation Credibility: Advances in Verification, Validation, and Uncertainty Quantification. NASA Technical Report, NASA/TP-2016-219422, JANNAF/GL-2016-0001, ARC-E-DAA-TN35719, 2016
2 National Aeronautics and Space Administration. Standard for models and simulations. NASA-STD-7009, 2016
3 Babuska I, Oden J T. Verification and validation in computational engineering and science: basic concepts. Comput Method Appl Mech Eng, 2004, 193: 4057–4066
4 Berard B, Bidoit M, Finkel A, et al. Systems and Software Verification: Model-Checking Techniques and Tools. Berlin: Springer, 2013
5 Tantithamthavorn C, McIntosh S, Hassan A E, et al. An empirical comparison of model validation techniques for defect prediction models. IEEE Trans Softw Eng, 2017, 43: 1–18
6 Liu H, Zhang J. High dynamic adaptive mobility network model and performance analysis. Sci China Ser F-Inf Sci, 2008, 51: 1154–1166
7 Li B H, Song X, Zhang L, et al. CoSMSOL: complex system modeling, simulation and optimization language. Int J Model Simul Sci Comput, 2017, 08: 1741002
8 Ao D, Hu Z, Mahadevan S. Dynamics model validation using time-domain metrics. J Verif Valid Uncert, 2017, 2: 011004
9 Wu Y X, Shen C, Jiang H, et al. Strict pattern matching under non-overlapping condition. Sci China Inf Sci, 2017, 60: 012101