The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/2398-7294.htm

Received 25 January 2019
Accepted 4 March 2019

A survey on VV&A of large-scale simulations

Yanan Wang and Jianqiang Li
Beijing University of Technology, Beijing, China

Sun Hongbo
Yantai University, Yantai, China

Yuan Li
Beijing University of Technology, Beijing, China

Faheem Akhtar
Faculty of Information Technology, Beijing University of Technology, Beijing, China and Department of Computer Science, Sukkur IBA University, Sukkur, Pakistan, and

Azhar Imran
School of Software Engineering, Beijing University of Technology, Beijing, China

Abstract

Purpose – Simulation is a well-known technique for using computers to imitate or simulate the operations of various kinds of real-world facilities or processes. The facility or process of interest is usually called a system, and to study it scientifically, we often have to make a set of assumptions about how it works. These assumptions, which usually take the form of mathematical or logical relationships, constitute a model that is used to gain some understanding of how the corresponding system behaves, and the quality of these understandings essentially depends on the credibility of given assumptions or models, known as VV&A (verification, validation and accreditation). The main purpose of this paper is to present an in-depth theoretical review and analysis for the application of VV&A in large-scale simulations.

Design/methodology/approach – After summarizing the VV&A of related research studies, the standards, frameworks, techniques, methods and tools have been discussed according to the characteristics of large-scale simulations (such as crowd network simulations).

Findings – The contributions of this paper will be useful for both academics and practitioners for formulating VV&A in large-scale simulations (such as crowd network simulations).

Originality/value – This paper will help researchers to provide support of a recommendation for formulating VV&A in large-scale simulations (such as crowd network simulations).

Keywords Large-scale simulation, Crowd network simulations, VV&A, Credibility

Paper type Literature review

This work is supported by the National Key R&D Program of China (Grant No. 2017YFB1400105).

International Journal of Crowd Science
Vol. 3 No. 1, 2019
pp. 63-86
Emerald Publishing Limited
2398-7294
DOI 10.1108/IJCS-01-2019-0004

1. Introduction

Due to rapid advancements in computer hardware and software development techniques, simulation technology is widely used in numerous application domains that include national defense, geology (Miller, 1981), human anatomy and biology (Snyder *et al.*, 1968; Andrews and Graef, 1970), electronics (Shichman and Hodges, 2003), military (Kheir and Holmes, 1978) and railway systems (Goodman *et al.*, 1987). It has also succeeded to achieve worldwide attention over the past several years due to its multiple features that include cost-efficiency, elimination of prototyping, better risk assessment, increased speed and optimized accuracy. Early concepts and theories of simulations and their applications are first introduced in previous works (Evans *et al.*, 1967; A, 1967; Martin, 1968; McLeod, 1968; Rothenberg, 1989). According to Rothenberg (1989), simulation can broadly be defined as a behavioral or phenomenological approach to modeling: This means, it is an active and behavioral analog of its referent. Since the end of the twentieth century, to deal with simulation objects' complexities and improving user's requirements in the simulation application domains, it has entered into a new stage of large-scale simulations. This also brings new challenges to the credibility of evaluation due to the underlying characteristics of large-scale simulations that include multiple levels, multiple structures, multiple relationships, multiple models and large scale (Zhang *et al.*, 2012).

For practical problems, simulations are of value only if the results of these simulations are reliable and accurate. Simulations that lack sufficient credibility are meaningless and will bring about catastrophic consequences (Zha and Kedi, 1997; Howard, 2011; Lent *et al.*, 2003). The credibility analysis of simulations not only validates simulation and its results but also reduces the risks caused by the application of simulation. Moreover, it also helps developers to find out the shortcomings of the simulation during the design process (Li, 2012). Therefore, credibility analysis is an essential parameter for the acceptance and validation of simulation system results (Muessig, 2001). Since 1962, Biggs and Wigan conducted research on the credibility analysis of simulation to fully evaluate the simulation (Abrahamson, 1980; Fishman and Kiviat, 1967). In the mid-1970s, the American Society of Computer Simulation established a Technical Committee on Model Credibility to build verification terms. The credibility is defined as the degree of trust of a simulation system in the simulation and the output of the simulation under a specific application purpose (Li, 2012), which mainly includes four aspects of nature, i.e. purpose relevance, objectivity, comprehensiveness and hierarchy (Zhang and Wang, 2001).

Initially, the concept of verification, validation and accreditation (VV&A) of simulation models is defined to improve the credibility of simulations and to make simulations efficient considering user requirements. Therefore, the research on VV&A and credibility of the simulation is essential for every simulation application (Zhang *et al.*, 2012; Figure 1). The term VV&A first appeared in the US Department of Defense (DoD) Modeling and

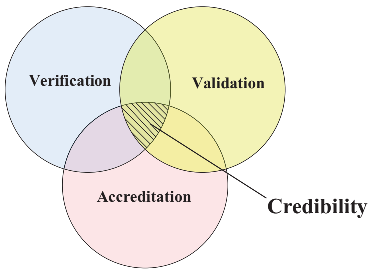

Figure 1.
The connections and effects of VV&A

Simulation Office (DMSO) in 1996 for setting up a technical support team, named Military Simulation VV&A Technical Support Team (TST; DoD, 1996). The role of that team was to enhance the scientific and usable military simulations. Moreover, VV&A activities can also assist in reducing development and integration risk with considerable improvement in the simulation credibility (Muessig, 2001; Tang *et al.*, 2006). Therefore, it cannot be accepted without the implementations of credibility analysis, or it cannot be used without a qualification certificate of VV&A (Zhang *et al.*, 2012; Figure 1).

VV&A is a significant mean to improve the credibility of simulations, especially during large-scale simulations. It is successfully affirmed in various studies of large-scale simulations to accredit and validate simulations results. Recently, numerous research surveys are published on the application of large-scale simulations. However, these research studies are either extremely broader to specifically discuss VV&A or extremely limited on the use of VV&A in a specific and relatively narrower application domain. Moreover, all these research studies provide an excellent review and details of current verification and validation (V&V) techniques but lack in providing origin and history of the development of these techniques. This research reviews the origin and development of VV&A theory, standards, frameworks, methods, techniques and tools from a specific application to large-scale simulation applications (like multi-agent simulations). This research will also focus on which features and characteristics of VV&A can be utilized to a large-scale simulation problem.

The rest of this paper is organized as follows: Section 2 introduces the details of VV&A. Section 3 reviews the VV&A standards. It explores VV&A frameworks and also discusses which VV&A framework is appropriate for the large-scale simulation platforms. Method and techniques of VV&A are presented in Section 4. This section also discusses about how to apply technology and methods to VV&A processes. Sections 5 and 6 explain tools of a VV&A framework and also suggests appropriate tools for large-scale simulations. Moreover, Section 7 summarizes this research (Figure 2).

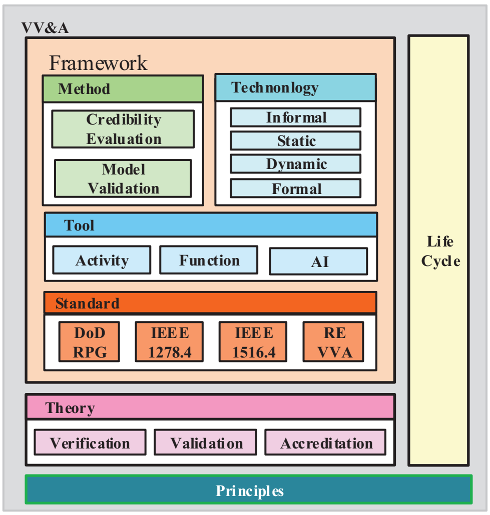

Figure 2.
The content structure of VV&A

2. The research on VV&A theory of simulations

Since the inception of modeling and simulations (M&S), the concept of VV&A in simulation models has been defined using life cycle simulations and model development processes (Borko, 1962).

Sargent (1997) presented a model for the model development process and described the relationship between models and reality (Figure 3). For simulations, Balci stated that V&V should be a constant ongoing process throughout the life cycle simulations. Later, in 1994, he introduced testing function into the verification process of simulation and concluded that VV&T is not a phase or step in the life cycle of a simulation study but a continuous activity throughout the entire life cycle (Figure 4). Furthermore, he provided a more exhaustive taxonomy for the VV&A process and talked about the necessity of verification in the formulation of simulation models (Balci, 1997; Balci and Nance, 1985).

The spectral analysis method was applied to the validation work of the missile simulation. Balci and Sargent (1984) believed that the confidence interval method can also perform the relevant validation work of the simulation. In 1996, the DMSO set up the Military Simulation VV&A TST to draft the Recommended Practice Guide (RPG) 5000.61. In 1996, the DoD issued the Ministry of Defense VV&A recommendations (Department of Defense Instruction [DoDI], 2003; Sanders, 1996). This recommendation guideline divides the VV&A work in the life cycle of simulations into seven main stages: determining VV&A requirements, VV&A plan design, concept model verification, system design verification, system application verification and system acceptance.

IEEE 1278.4 was initiated by the IEEE DIS (Distributed Interactive Simulation) Committee and completed in 1997, which provides VV&A guidelines for DIS emulation users and developers (Interactive and Committee, 1998). In 2004, the International Test Operations Procedure (ITOP) released the first version of the "General Procedure for M&S V&V Information Exchange" (ITOP 01-01-002). In 2007, IEEE 1516.4 practice defines the processes and procedures that should be followed to implement VV&A for federations being developed using the high-level architecture (HLA) using the Federation Development and Execution Process (FEDEP) (Interoperability *et al.*, 2007). The development history of VV&A is presented in Figure 5.

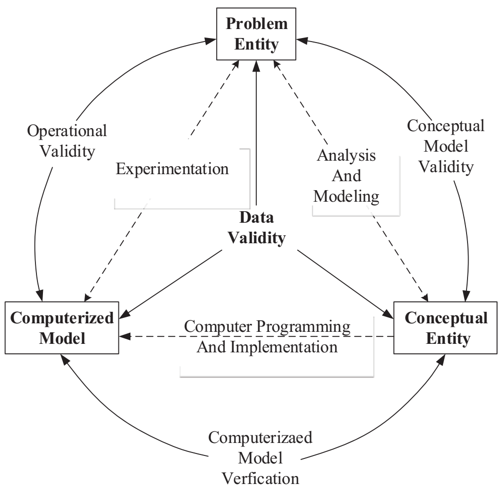

Figure 3.
The model development process

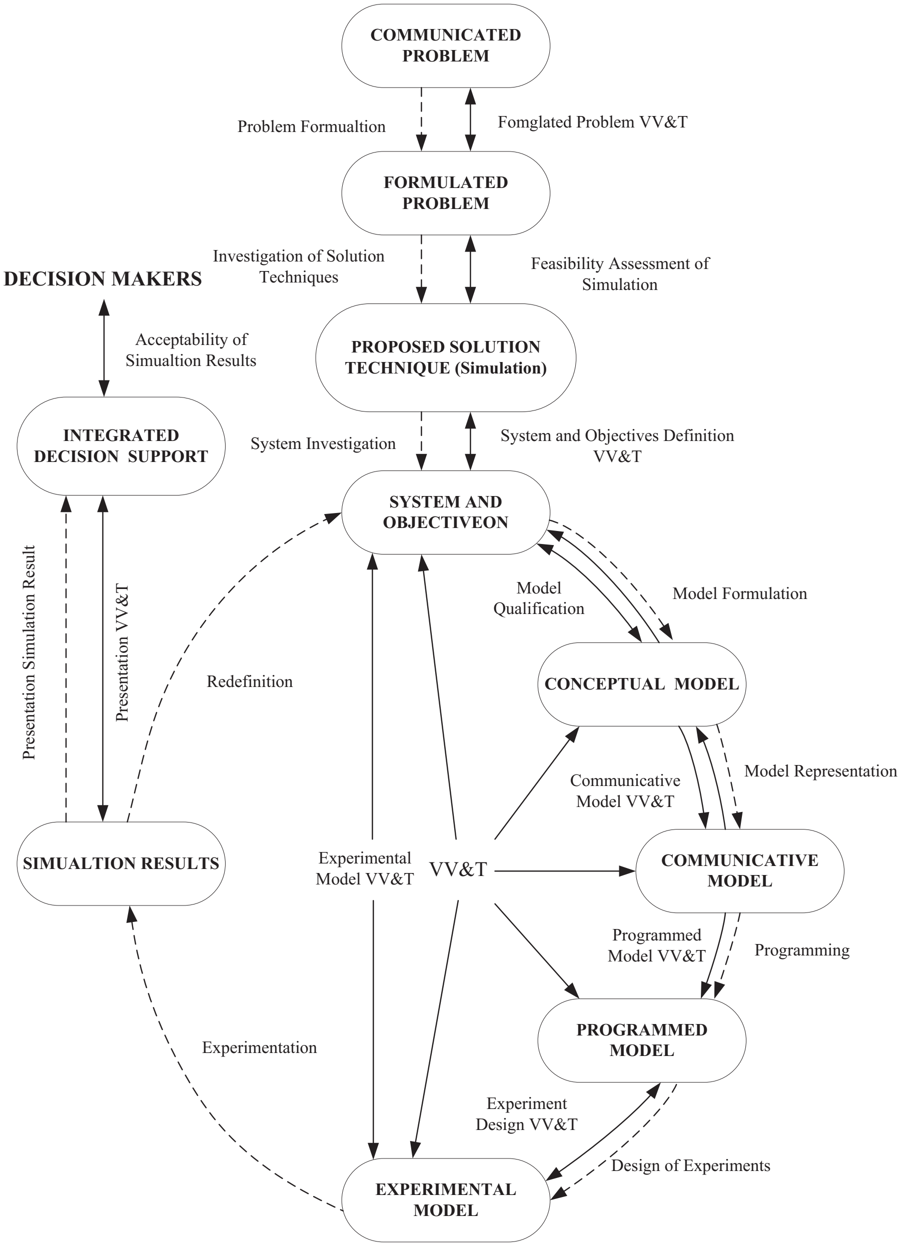

Figure 4.
The life cycle of VV&A

Different scholars and organizations also give relevant definitions and requirements to VV&A. The economist Machlup (1955) provides the first formalized definition of the concept of verification as "Verification in research and analysis may refer to several things that includes correctness of mathematical and logical arguments, the applicability of formulas and equations". Sornette *et al.* (2007) defined validation as "the process of determining the degree to which a model is an accurate representation of the real world from the perspective of its intended uses". This definition is identical to Sargent's: a model must be validated to ensure some minimal degree of accuracy for a given system or problem entity. The only difference is that the system/problem entity is a physical phenomenon in the real world instead of a generalized process. In previous works (Oberkampf and Barone, 2006; Oberkampf and Barone, 2007), V&V is again defined as "Verification is the assessment of the accuracy of the solution to a computational model. Validation is the assessment of the accuracy of a computational simulation with experimental data". This concept has widely been accepted by most of the scholars. The DoD defines accreditation in DoDI 5000.61 as "the official certification that a model or simulation is acceptable for use for a specific application" (Sanders, 1996).

Verification ensures that the model works as expected. Verification is a software-level process that does not necessarily require information about the model's outputs. It is more of a check that the model's equations are correctly coded and implemented. The ultimate goal of validation is to establish credibility in a model such that the model can also be used to predict problem entity behaviors on unseen and untested cases. Validation is the testing of model outputs against experimental data to see if the model yields accurate outputs. Model accreditation determines if a model satisfies a specified model according to a specified process.

Finally, this section introduces the concepts of VV&A, the evolution process of VV&A and the VV&A process in the simulation life cycle, which can help readers to have a preliminary understanding of VV&A.

3. The research and establishment on VV&A standards of large-scale simulations

Does the establishment of VV&A standards of a large-scale simulation in the life cycle guide the work effectively and orderly? To answer this question, we can observe the following VV&A standards and specifications.

3.1 DoD VV&A RPG

DoD VV&A RPG is a basic framework for the overall guidance of the subordinate Army, Navy and Air Force according to the DoDI 5000.61 recommended guidelines, allowing subordinate organizations to make corresponding adjustments and changes according to their specific circumstances (Glasow *et al.*, 1996; Glasow *et al.*, 2000; Glasow *et al.*, 2005).

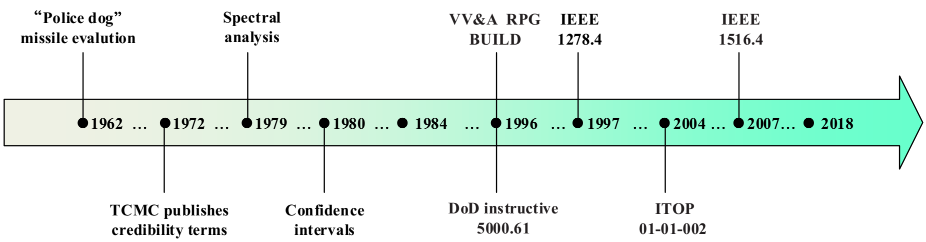

Figure 5.
The development history of VV&A

3.2 IEEE 1278.4

IEEE1278.4 was initiated by the IEEE DIS Committee and completed in 1997. IEEE 1278.4 provides VV&A guidelines for DIS emulation users and developers. It meets user requirements by flexibly defining and applying the DIS emulation VV&A basic principles, so different simulation purposes will result in completely different VV&A processes (Interactive and Committee, 1998).

3.3 IEEE 1516.4

IEEE 1516.4 practice defines the processes and procedures that should be followed to implement VV&A for federations being developed using the HLA with FEDEP, which provides implementation-level guidance to VV&A practitioners (Interoperability *et al.*, 2007).

A summary of VV&A standards is presented in Table I. The most essential content of DoD RPG is to put VV&A behavior into the entire development life cycle of M&S. The advantages include easy operating documentation and so on. Whereas, the disadvantage is that the specification is special and cannot be applied independently to a special M&S process. They are more like a tutorial on the management and education of VV&A than scientific achievement. The IEEE 1278.4 protocol is primarily intended for the DIS protocol. IEEE 1516.4 is the VV&A standard of distributed interactive simulation systems based on HLA. IEEE 1278.4 and IEEE 1516.4 are aimed at standards performing the VV&A work for a typical system with special structure, so neither of them is universal.

The research establishment on VV&A standards is the important content in the simulation technology. It is a work of great significance, and it can improve the level of creditability of large-scale simulations and accelerate the standardization, intelligence, integration and automation of the reliability assessment. It can speed up the process of normalization and standardization of modeling and simulation and can satisfy the demands of VV&A standards of large-scale simulations. The next section will focus on the validation framework of VV&A.

4. The research on VV&A framework of large-scale simulations

Heritage techniques for VV&A (Sargent, 2005) cannot easily be transferred for simulations. It needs time and efforts (Terano, 2007; Klügl, 2008; Niazi *et al.*, 2009; Pengfei *et al.*, 2011; Railsback and Grimm, 2011), but these studies do not directly deal with the model verifying process. Validation techniques and their guidelines are addressed in most of the modeling textbooks and have even been instantiated in the form of a validation process for general

Standard	DoD RPG	IEEE 1278.4	IEEE 1516.4
Application scenarios	Military project	Distributed interactive simulation	The high-level architecture simulation
The role of VV&A	Accreditation agent	VV&A agent VV&A team	VV&A agent VV&A team Accreditation agent
Time	1996	1997	2007
Purpose	Establish VV&A methodology	Guide VV&A operations	Provides VV&A implementation-level guidance
Method	Define the VV&A framework	Define the VV&A framework	Define the VV&A framework
Kernel ideas	Roles phase	Roles process	Roles responsibilities

Table I. Compare the standards of VV&A

agent-based models (Law, 2007; Klügl, 2009). However, such techniques are still too general to provide a concrete, practical methodology for the key validation step. At the same time, agent-based modeling and simulation (ABMS) is an important branch of large-scale simulations (Li and Sun, 2007). It involves multiple agent co-simulations and has diverted researchers' attention during the past decades. Furthermore, the weak validation and verification of agent-based simulation models make ABMS hard to trust. Each of these agent-based models needs to be validated separately, but how to validate the overall simulation of these combined models remains an unanswered question. It is generally well accepted that it is difficult to validate large-scale simulations sufficiently then to trust their results (Sargent *et al.*, 2000). In many cases, the cost of trying to achieve complete validation is neither practical nor worthwhile (Shannon, 1975). In fact, it has been shown in previous research studies (Weisel *et al.*, 2003; Weisel, 2004) that separately validated models can produce invalid outputs when combined. As per above discussion, it can be concluded that a VV&A simulation verification framework is needed to solve how to verify the simulations.

The first formal framework for V&V was proposed by Oberkampf (1994). This framework is proposed for building confidence in CFD (computational fluid dynamics) code predictions that overcomes some of the difficulties of past procedures and delineates, i.e. the causes of uncertainty in CFD predictions. Easterling (2001) provided a conceptual framework for quantifying the uncertainty in model predictions, which is shown in Figure 6.

Bayarri *et al.* (2007) presented the conceptual framework for V&V of physics-based simulations. This framework quantifies multiple sources of error and uncertainty in computer models, combines multiple sources of information and updates validation assessments and acquires new information. Li-Ping and Xiao-Ping (2007) combined fuzzy similarity theory, fuzzy analytic hierarchy process (FAHP) and fuzzy comprehensive evaluation method to propose a comprehensive fuzzy credibility evaluation framework, which is applicable to the effectiveness of crowd simulation in loops. Mehrabadi *et al.* (2014) proposed the verification, validation, and uncertainty quantification (VV&UQ) framework, i.e. applicable to power electronic systems. This framework is used to gather all of the uncertainties during the simulation and modeling process. It gathers model form uncertainty, model inputs uncertainty and uncertainty due to the numerical approximations

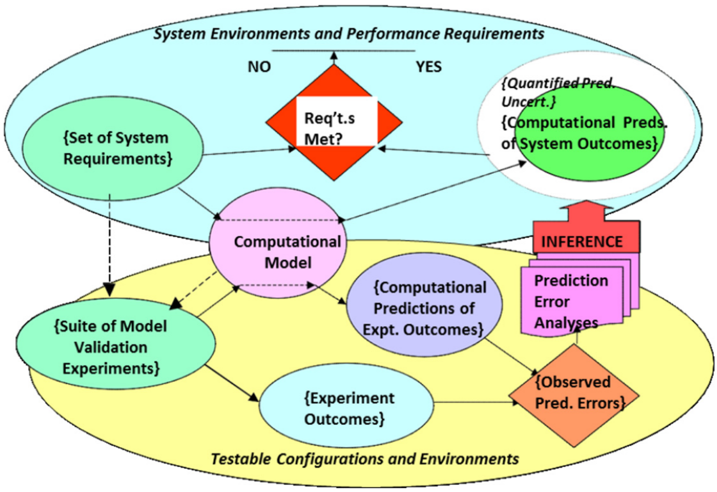

Figure 6.
The first formal framework for V&V

for quantitatively assessing the reliability of the model. Drchal *et al.* (2016) propose a six-step validation framework entitled VALFRAM (validation framework for activity-based models) that allows exploiting historical real-world data to assess the validity of activity-based models. Wang *et al.* (2016) present a novel rescheduling framework of the crude oil operations based on a continuous-time representation. Abnormal events and uncertainties in the crude oil tank farm area are considered and analyzed in this framework with an objective to improve the robustness of the final crude oil operations plan. In the same year, Barnes *et al.* (2016) presented a new simulation framework for wireless sensor networks based on QEMU and SystemC that aims at validating the binary code of wireless protocols by checking that the protocol's implementation complies with its property specifications during the simulation process. Based on our experimental study, we tried to show the correct functionality of our hardware platform model by comparing with real frame exchange traces and the verification of one of the protocol's properties during the simulation process.

VV&A verification framework not only improves the reliability of general simulation but also plays a significant role in large-scale simulations. The multi-agent simulations are from one of the main forms of large-scale simulations. There are many works about verification and validation of multi-agent simulations (Terano, 2006; Klügl, 2008; Niazi *et al.*, 2009; Pengfei *et al.*, 2011). However, these studies do not directly deal with the model testing process and never ever proposed a model testing framework to conduct validation and verification using the model testing process. In response to these problems, some scholars also proposed some validation frameworks for multi-agent simulation systems. Zoumpoulaki *et al.* (2010) proposed a framework for designing evacuation simulations that are based on a multi-agent BDI architecture enhanced with the OCEAN model of personality and the OCC model of the emotions. Furthermore, this paper (Gurcan *et al.*, 2011) presents our testing framework in detail and demonstrates its effectiveness by showing its applicability on a realistic agent-based simulation case study. Moreover, they propose a generic framework for the automated execution of these requirements defined at each level. Furthermore, this body of work presents the design of a novel generic framework for the automated model testing of agent-based simulation models.

Table II summarizes few of the validation framework schemes. As per Table II discussions, although these frameworks meet the verification requirements of their simulation systems in combination with different methods, none of them is universal. Therefore, in the face of large-scale simulation, scholars need to develop a suitable simulation framework based on actual application scenarios.

Therefore, establishing a universal verification framework in the life cycle of VV&A will be helpful to apply VV&A methods and techniques to the simulation models. It can obviously improve the credibility of simulations, especially large-scale simulations like multi-intelligent simulations. The following section provides details on VV&A methods and technology. The next section will introduce the main technologies and methods in the VV&A process.

5. The research on VV&A technology and method of large-scale simulations

In every step of VV&A activities, proper VV&A methods and techniques should be selected and used as the specific operation methods. Therefore, it is extremely important and necessary to expand the research on VV&A methods and techniques of large-scale simulations.

Scheme	Framework	Objective	Application
Zheng and Liu (2007)	The comprehensive fuzzy credibility evaluation framework	Used fuzzy analytic hierarchy process (FAHP) and fuzzy comprehensive evaluation method to support the credibility of crowd simulation in the loops	Crowd simulation in loops
Mehrabadi *et al.* (2014)	The verification, validation and uncertainty quantification (VV&UQ) framework	Used uncertainty quantification method to assess the confidence in modeling and simulation quantitatively	Power electronic systems
Drchal *et al.* (2016)	The six-step validation framework for activity-based models (VALFRAM) (Wigan, 1972)	Exploited historical real-world data (three real-world activity-based transport models) to assess the validity of the activity-based models	Daily activity schedules
Wang *et al.* (2016)	The framework of the crude oil operations based on a continuous-time representation	Used multi-agent based simulator to support dynamic optimization of crude oil operations	Crude oil operations
Barnes *et al.* (2016)	The framework for wireless sensor networks based on QEMU and System C	Validated binary code of wireless protocols by checking whether the protocol's implementation complies with its property specifications during the simulation process	Wireless sensor networks
Zoumpoulaki *et al.* (2010)	The framework for designing evacuation of the simulations	Used BDI (Belief-Desire-Intention) model to enhance the OCEAN model of personality and the OCC model of the emotions	Human behaviors in stressful situations
Gurcan *et al.* (2011)	The generic testing framework for agent-based simulation models to conduct validation and verification of the models	Showed testing framework applicability on a realistic agent-based simulation case study	Agent-based modeling and simulation (ABMS)

Table II.
Summary of VV&A framework

5.1 Methodology

Since the late 1960s, foreign countries began to entertain VV&A during the simulation process. Fishman and Kiviat (1967) used the spectral analysis method to evaluate the equivalence between the simulation model and the actual system. In fact, it was the first time when the concept of verification and validation of the simulation model was proposed. Moreover, Sargent (2001) summarizes the verification methods of related simulation models, including Turing test, sensitivity analysis, extreme condition test, statistical test and subjective validity test. Combined with the research of related scholars, this research reviews VV&A methods considering two main aspects, i.e. credibility evaluation method and model validation method.

5.1.1 Credibility evaluation method. Credibility calculation is considered the most important task of VV&A. It can further be divided into two main steps, which are credibility testing method and credibility calculation method (Yu and Xiao, 2018).

5.1.1.1 The first research focuses on the credibility testing method. Based on the whole life cycle trusted process guarantee model, Lv (2016) proposed the Web application software credibility verification model. He proposed a software credibility testing method under the support of the verification model. Furthermore, considering the content structure of the trusted behavior statement and the structural features and trusted features of REST application software, Liu (2017) proposed a RESTful Web application credibility testing method based on the behavioral declaration. Moreover, Yu and Xiao (2018) also proposed a new method of generating credibility test cases based on the immune algorithm.

5.1.1.2 The next research method is the credibility calculation method. Wright (1972) proposed a graph comparison method for credibility analysis. Balci and Sargent (1984) believe that the confidence interval method can be used to perform the relevant verification work of the simulations. Kheir *et al.* (1986) proposed using Theil's inequality to analyze the relevant data of the missile's simulation system and the actual system in flight experiments. Montgomery and Conard (1980) studied the spectrum analysis method and applied it to the verification of missile simulations. Han (2013) proposed a hybrid metric design method for application software credibility using static hash metrics and dynamic behavior values as evaluation criteria. Xiong *et al.* (2016) proposed a multi-attribute decision-making modeling method to design a strategy to build a credible indicator tree. This method is based on on-demand driving and using dynamic methods. Zhao *et al.* (2014) used the factor analysis method to construct a credibility evaluation index system for Web software. The structural entropy method is applied to the weight calculation of the credibility index. Qi *et al.* (2018) proposed a credible evaluation method based on FAHP. It combines the analytic hierarchy process (AHP) with the fuzzy comprehensive evaluation method. The intention is to overcome the problem that the subjective judgment of human beings as a person in the traditional AHP will have a great impact on the results. The literature (Yang *et al.*, 2003) uses FAHP to calculate the weights of the factors affecting the credibility of the simulation, but the final credibility evaluation results only consider a set of expert scores.

5.1.2 Model validation method. Another important method of credibility research is the model credibility calculation method. It could help in verifying the credibility of a simulation. Several researchers have investigated and examined various validation approaches for different types of simulation models.

Birta and Ozmizrak (1996) proposed a method of a validation knowledge base, captured as a set of relationships between input and output variables of a simulation model. Kleijnen (1999) presented different statistical techniques to be used for simulation model validation based on the available data. Balci recommended that a validation method is the comparison of graphical outputs from simulations with experimental data and testing the degree of statistical agreement between the two (Balci and Sargent, 1982a, 1982b, 1984). Cooley and Solano (2011) describe the use of validation methods in model building. They discussed the stages of simulating an agent-based simulation model and presented six specific validation approaches. Ahn (2007) proposed a novel method for the validation of agent-based evacuation and crowd simulation. They used concepts from the field of human computation. Liu (2001) proposed the principle and fuzzy quantitative evaluation method for establishing the fuzzy comprehensive evaluation system for the credibility evaluation of large-scale simulations. However, the FAHP is not sufficient to determine the weight of each index, which reflects the ambiguity of expert judgment. Peng *et al.* (2017) believe that a similar degree method can be suitable for evaluating the credibility of a simulation system and an actual system. Tian *et al.* (2012) proposed that after applying the gray clustering method to the Integrative Avionics System, one can use this method after combining with the AHP for large-scale simulations.

A summary of a few VV&A method schemes is presented in Table III. In conclusion, there are several methods such as AHP, spectrum analysis and others that can also be applied to the simulation verification process. However, each method has its own advantages, disadvantages and scope of applications. It is necessary to select the

Method	Scheme	Application	Evaluation
Similar degree method	Peng (2017)	Suitable for evaluating the credibility of simulation systems and the actual system	Only applied to the credibility evaluation of simple systems
Gray clustering method	Tian (2012)	Integrative Avionics System; large-scale simulations	Easy idea; need combining with the analytic hierarchy process (AHP)
AHP	Shangguan *et al.* (2014)	Cooperative vehicle infrastructure simulation system; large-scale simulations	Simple and practical; many indicators easily lead to the difficult confirmation of a judgment matrix
Spectral analysis	Fishman (1967)	Time-series data generated by the simulated stochastic models	To evaluate the equivalence between the simulation model and the actual system
The software credibility test method based on behavior declaration	Lv (2016)	The verification model of the Web software	Application scenarios are too simplistic
The hybrid metric design method for the application software credibility	Han *et al.* (2013)	Information system	Lacks in practical significance
The credible evaluation method based on FAHP	Qi *et al.* (2018)	Information system	The method constraints are complex
The knowledge-based approach for the validation of simulation models	Birta and Ozmizrak (1996)	Behavioral model	Enabled the application of a variety of solution-oriented techniques
The method is based on similarity theory and evaluates confidence level with fuzzy mathematics	Liu (2001)	A large complex simulation system	Lacking consideration of the impact in-between complex subsystems
The behavior validation method based on a dynamic model	Yaman Barlas (1989)	Dynamic behavioral model	Limited toward single application type

Table III.
Summary of VV&A based on methods

appropriate method according to the requirements of different application scenarios and its characteristics. For large-scale simulations, which have the characteristics of complex computation, interactive and autonomous, one can use the gray theory (such as gray prediction method, gray decision method) and the fuzzy mathematics theory to effectively deal with complex logic problems. Furthermore, the AHP can deal with multi-objective decision-making in a simulation. By summarizing the different methods proposed and discussed by various researchers, almost every researcher believes that a single method or theory is difficult to verify the various problems associated with a simulation system without applying for medicine according to indications.

5.2 Technology

Based on these methods, a substantial amount of research has been performed to define a technique for verifying and validating simulation models.

Balci (1994) surveyed current software validation techniques and current simulation model VV&T techniques and describes how they can all be applied throughout the life cycle of a simulation study. The technology can be divided into informal, static, dynamic and formal. Moreover, a distinct difference between each classification exists (Figure 7):

- Informal techniques are among the most commonly used techniques. They are called informal because the tools and approaches used rely heavily on human reasoning and subjectivity without stringent mathematical formalism.
- Static techniques are concerned with accuracy assessment on the basis of characteristics of the static model design and source code. Static techniques do not require machine execution of the model, but mental execution may be used (Whitner and Balci, 1989).
- Dynamic techniques require model execution and are intended for evaluating the model based on its execution behavior. Most of the dynamic V&V techniques require model instrumentation.

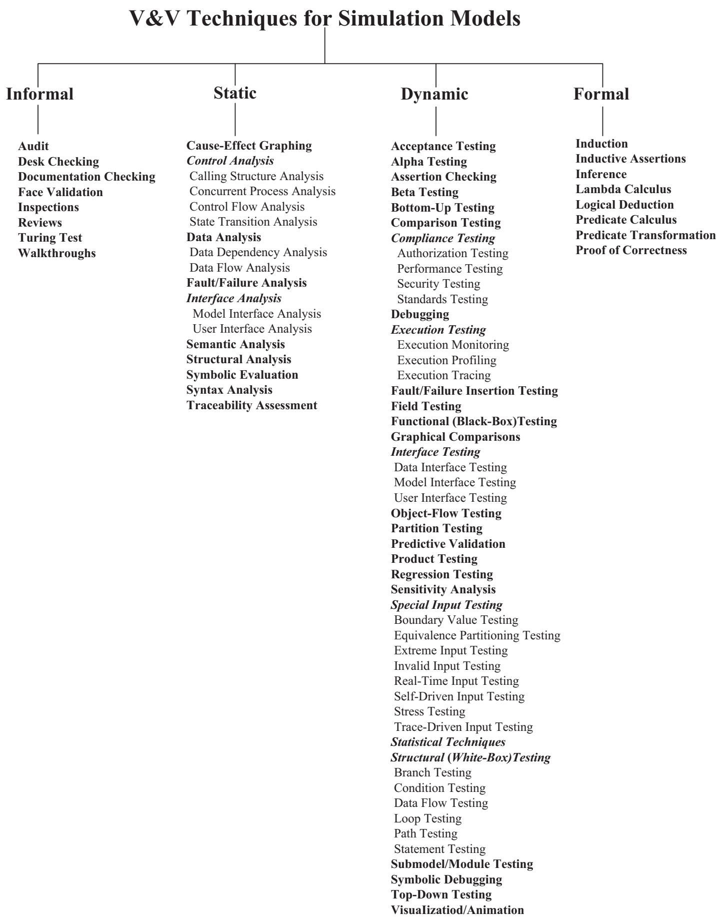

Figure 7.
A taxonomy of techniques for conventional simulation models

- Formal techniques are based on mathematical proof of correctness. If attainable, proof of correctness is the most effective means of model V&V. Current state-of-the-art proof of correctness techniques is simply not capable of being applied to even reasonably large-scale simulations.

A taxonomy of more than 77 V&V techniques for conventional simulation models is presented in Figure 7. Detailed descriptions of these techniques can be found in previous works (Balci, 1998; DoD, 1996). Balci (1994) summarized VV&T techniques that are applicable to each of the ten credibility assessment stages described in Table IV. The more of these techniques we apply, the more confidence we may gain in the credibility of a life cycle phase.

Balci (1995) refers to the validation techniques listed in the US DoD VV&A recommendations and combines them with the relevant literature (Balci, 1994; DoD, 1996; Yilmaz and Balci, 1997). As per them, VV&T techniques are classified into six distinct credibility assessment perspectives: informal, static, dynamic, symbolic, constraint and formal. The object-oriented paradigm provides numerous advantages such as maintainability and reusability over the procedural paradigm (Sargent, 1997; Yilmaz and Balci, 1997). These techniques come from the software engineering discipline and are applicable to object-oriented simulation model V&V. They further divided the V&V techniques for object-oriented simulation models into conventional, adaptive and specific.

Based on special scenarios, especially in complex agent environments, some scholars have proposed relevant verification techniques. Railsback and Grimm (2011) have studied about testing agent-based simulation models. In this study, they define 10 important techniques to find and fix software errors: i.e. syntax checking, visual testing, print statements and spot tests with “agent monitors”, stress tests, test procedures, test programs, code analysis and statistical analysis of file output and independent reimplementation of submodels. However, they found the approaches they have proposed are far from the automation of the model testing process; the primary reason is that they do not have an architectural perspective about how these solutions could be integratively constructed and conducted. Moreover, some of their solutions are not generic and completely depend on the NetLogo simulation framework (Sklar, 2007). Louloudi and Klügl (2012) proposed a new technique to validate agent-based simulation models. A novel face validation technique is presented that enables systematic plausibility checks by a human expert immersed into a fine-grained virtual reality environment that is the exact representation of the simulated multi-agent model. Tabak *et al.* (2010) proposed to use radio frequency identification technology applied for validation of an office simulation model. This technique was proved to be effective by verifying that there were no significant differences between the predicted and observed activity behavior. To better utilize the potential of the system of simulation models and simulators, industrially applicable methods for VV&UQ are crucial. Eek *et al.* (2015) presented an exploratory case study of VV&UQ techniques applied on models integrated into aircraft system simulators at Saab Aeronautics and in driving simulators at the Swedish National Road and Transport Research Institute (VTI). Results show that a large number of V&V techniques are applied, some of which are promising for further development and used in simulator credibility assessment.

We can draw the following conclusions through observing Figure 7 and the Table IV: Informal techniques subjectively do not have stringent mathematical formalism. Most of them are used in formulated problem stages and system and objectives definition stages. Static techniques are mainly concerned with accuracy assessment on the basis of characteristics of the static model design; its main application is model design stage. Dynamic techniques(like sensitivity analysis and statistical techniques) are intended to

	FP VV&T	FA of Sim	S&OD VV&T	Model Qual.	CM VV&T	PM VV&T	ED VV&T	Data VV&T	EM VV&T	Pres. VV&T
Assertion checking						✓	✓	✓	✓	
Audit	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Black-box testing						✓	✓		✓	
Bottom-up testing						✓			✓	
Boundary analysis						✓			✓	
Cause–effect graphing	✓				✓	✓			✓	
Consistency checking	✓		✓	✓	✓	✓	✓	✓	✓	✓
Data flow analysis					✓	✓	✓	✓	✓	
Debugging						✓	✓		✓	
Desk checking	✓		✓	✓	✓	✓	✓	✓	✓	✓
Execution monitoring						✓	✓		✓	
Execution profiling						✓	✓		✓	
Execution tracing						✓	✓		✓	
Face validation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Field testing									✓	
Graph-based analysis					✓	✓	✓		✓	
Graphical comparisons						✓	✓		✓	
Induction						✓			✓	
Inductive assertion						✓			✓	
Inference						✓			✓	
Inspections	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lambda calculus						✓			✓	
Logical deduction						✓			✓	
Partition analysis						✓			✓	
Path analysis						✓	✓		✓	
Predicate calculus						✓			✓	
Predicate transformation						✓			✓	
Predictive validation									✓	
Proof of correctness						✓			✓	
Regression testing						✓			✓	
Reviews	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Semantic analysis						✓	✓		✓	
Sensitivity analysis						✓	✓		✓	
Statistical techniques								✓	✓	
Stress testing						✓			✓	
Structural analysis	✓		✓		✓	✓	✓		✓	✓
Submodel testing						✓			✓	
Symbolic debugging						✓	✓		✓	
Symbolic execution						✓	✓		✓	
Syntax analysis						✓	✓		✓	
Top-down testing						✓			✓	
Turing test									✓	
Visualization						✓			✓	
Walkthroughs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
White-box testing						✓			✓	

Notes: FP VV &T = Formulated Problem VV&T; FA of Sim = Feasibility Assessment of Simulation; S&OD VV&T = System and Objectives Definition; VV&T Model Qual = Model Qualification; CM VV&T = Communicative Model VV&T; PM VV&T = Programmed Model VV&T; ED VV&T = Experiment Design VV&T; EM VV&T = Experimental Model VV&T; Pres. VV&T = Presentation VV&T

Table IV. The VV&T techniques for the credibility assessment stages evaluate the model based on its execution behavior; Formal techniques are based on mathematical proof of correctness and mainly targeted at programmed model VV&T. Actually, the dynamic technology is absolutely more suitable for facing the more complex behavioral activities in large-scale simulations.

Strengthening the research on VV&A technique and methods of large-scale simulations can not only establish the foundation for the development of VV&A software tools but will also further improve and strengthen VV&A theory. Furthermore, at the same time, it can solve the problem of single-operation methods in VV&A work and can also provide the technical support for performing VV&A work during the life cycle effectively and comprehensively. The preceding section discusses the basic tool used in VV&A process.

6. The development of VV&A validation tools of large-scale simulations

VV&A is of great significance in the simulations. In the actual VV&A activities, it is difficult to implement the complete VV&A principles due to the systematic engineering, which requires a sound program design, effective organization, efficient management, a large number of testing and data processing work (Mazhen *et al.*, 2016). Furthermore, comparing with simple simulations, most of the difficulties that can be found during the assessment of the credibility of large-scale simulations are mainly manifested in the following aspects (Chens *et al.*, 2001):

- Large-scale simulations have a huge scale, including many subsystems, and the evaluation workload is quite large. It is difficult to complete the task by manual work alone (Qin, 2009).
- The large-scale simulations have many evaluation indexes, complex index tree structure and a huge amount of expert evaluation data and need a lot of calculation in the evaluation process. Without special auxiliary tools, it is difficult to ensure the correctness of evaluation calculation (Birta and Ozmizrak, 1996; Zupan *et al.*, 2006).
- There are many evaluation methods for large-scale simulations, and it is difficult to ensure the accuracy of the evaluation results by a single evaluation method. How to select and use statistical methods correctly has become a major problem for the validators of large-scale simulations (Deslandres and Pierreval, 1991; Balci *et al.*, 2000).

These problems can be solved by using VV&A tools to some extent. Therefore, it is of great significance to study the simulation model VV&A tools, which can be shown as follows:

- To improve VV&A automation level of the simulation, the VV&A of the simulation is an extremely complicated task, which involves many links, and these links have a division of labor that works together. However, VV&A tools can not only improve the economy and rapidity of VV&A work and reduce the work intensity of VV&A personnel but also greatly improve the automation degree of VV&A.
- To improve VV&A collaboration of the simulation, the VV&A tool of simulation can support VV&A personnel to cooperate closely with project managers, design developers and model users, to share various resources in the process of modeling, exchanging various information frequently and greatly improve the collaboration degree of imitation VV&A.
- To improve the integration of the VV&A, the use of the VV&A tool integration framework to integrate all kinds of VV&A tools that have been and will be developed can not only save resources and facilitate data sharing but also significantly improve the integration degree of VV&A.

Therefore, the simulation requires a large number of model VV&A tools to assist in the VV&A process. These tools support the VV&A work throughout the life cycle of the model development with different functions at different stages of the VV&A development model. Ma Zhen divided VV&A's tools into two major types: tools for VV&A activity and function and artificial intelligence. The VV&A activity tools are shown in Figure 8.

Moreover, the simulation model VV&A tool first appeared in the early 1990s, mainly based on the development of VV&A tools in some West developed countries. By the end of the twentieth century, dozens of assistant tools have been developed abroad for one or more stages of the VV&A process (Balci, 1998; Dean, 2004; Zeigler and Sarjoughian, 2002). Some VV&A tools developed abroad and their functions are summarized, as shown in Table V.

Besides, many scholars have also designed VV&A tools for different scenarios. To reduce the workload and save the resource of evaluation, Qin (2010) designs and develops a simulation credibility evaluation assistant tool based on hierarchical evaluation. HIT-CET (Harbin Inst. of Tech. Credibility Evaluation Tool) can effectively assist to finish the evaluation work, improve the work efficiency and reduce the cost of the evaluation. Reid *et al.* (2012) created the STAT (Simulation Team Assessment Tool) to evaluate key components of all pediatric resuscitations, not only to evaluate specific scenarios. They created a valid, reliable tool for the evaluation of a team's comprehensive performance during a simulated pediatric resuscitation, which includes medical decision-making,

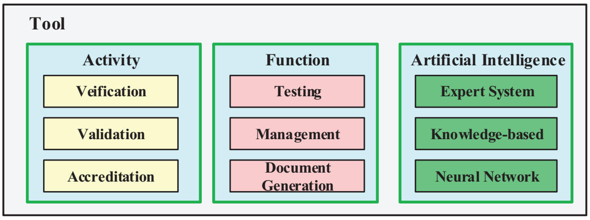

Figure 8.
The classification of VV&A tools for a simulation model

No.	Tool	R and D company	Function
1	Accreditation, Assessment, Assistant	Joint, accreditation support activity	Verification, validation and accreditation
2	V&V Managers Toolkit	TRADOC	
3	VVACET	Tecmasters	
4	Analyst-Pro	Goda Software, Inc.	
5	Caliber-RBT	Technology Builders, Inc.	Verification
6	Code Wizard	Para Soft	
7	DAVIE	DMSO Data Engineering	Validation
8	JASA	Joint Accreditation Support Activity	Accreditation
9	Authoritative Data Source Library PC Version	DMSO	
10	DAKOTA	Sandia National Laboratory	Verification, validation
11	Temporal Verification Framework	Arizona Center for Integrative M&S	
12	HLA Lab Works Suite of Tools	Aegis Technologies Group, Inc.	Verification, accreditation
13	Vertical Sky Solution 3.1	Vertical Sky	

Table V.
The part of a foreign VV&A tool for the simulation model

technical skill and human factors. Several tools are also developed proposing different ways of validating multi-agent simulations. Klugl (2008) proposes a process for validating agent-based simulation models which combines face validation, sensitivity analysis, calibration and statistical validation. Niazi *et al.* (2009) propose a validation and verification tool for agent-based simulation models for a wide variety of models. In this tool, a multi-agent overlay is created on top of the actual simulation model. The agents populating this overlay have a main task to monitor the simulation's run based on predefined constraints which when violated are logged. Montanola-Sales *et al.* (2011) present the verification and validation of an agent-based demographic simulation model implemented using a parallel demographic simulation tool (Yades) using white-box validation methods described by Pidd (2004). In this sense, Montanola-Sales et al. divide their model into smaller components and test the correctness of each component.

VV&A work of large-scale simulations is a complicated process. It is more difficult and inefficient to carry out VV&A work completely relying on the human. Therefore, it is necessary to develop VV&A tools of large-scale simulations to make the management more systematic, professional and automatic with an intention of improving the efficiency of VV&A work. This section summarizes the current VV&A tools and explains the importance and necessity of VV&A tools in the simulations. The following section will introduce the main content of this article and the plan for the next step.

7. Summary

VV&A is the most important method to assess the credibility of the simulation system. This paper presents a review of 114 papers on VV&A research of simulation and elaborates the application of VV&A in the simulation system by reviewing VV&A concept, VV&A standard, VV&A framework, VV&A technology, methods, VV&A tool, etc. The primary motivation is to present an in-depth theoretical review and analysis for the application of VV&A in large-scale simulations. We focused on the related research of VV&A in large-scale simulation; the standards of VV&A have defined the processes and procedures. The framework which facilitates the full application of techniques and methods to simulation. Moreover, we also examined the tools that can improve automation level and operation efficiency of VV&A in simulations.

Crowd network is the main mode of the modern service industry and future economy society (Chai *et al.*, 2017; Sun and Zhang, 2017), which has the characteristics of large-scaled, open-style, self-organized and ecological intelligent network (Nan *et al.*, 2017). Compared with traditional large-scale simulations, crowd science simulations have several obvious challenges as follows:

- Dynamic is the first challenge. Member attributes and states of crowd science simulations may vary at any time in an uncertain mode. Members are more loosely coupled, but member behaviors and intention variations may lead to a change of group states and intentions in extending scopes.
- Diversification is performed as a key feature; for example, time advance strategy may base on changes of slow variables, events and clock or hybrid mode. Moreover, as members are multiform and multi-disciplinary, transactions are uncertain and various, disturbances have several sources and subscriptions exist in a different layer and aspects, disturbances injection strategy and matching strategy are all needed to take diversification into consideration.
- The scale of crowd science simulation may need to achieve millions or even more trillions to uncover or verify principals and regularities of crowd science.

In a nutshell, crowd science simulation is a new development of large-scale simulations. To better improve the credibility of the crowded network and solve the challenges it creates, it is very essential, urgent and imperative to make standards and frameworks on VV&A working of crowd science simulations. Moreover, this research can help researchers to provide support of a recommendation for formulating VV&A in large-scale simulations.

References

A, R.E. (1967), "Pattern recognition: theory, experiment, computer simulations, and dynamic models of form perception and discovery", *Brain Research*, Vol. 2 No. 3, p. 299.

Abrahamson, G. (1980), "Terminology for model credibility", *Simulation*, Vol. 35 No. 6, pp. 206-207.

Ahn, V.L. (2007), "Human computation", *Proceedings of the 4th International Conference on Knowledge Capture, K-CAP 2007, ACM, New York, NY*, pp. 5-6.

Andrews, J. and Graef, S. (1970), "Dynamic modeling and simulation of the anaerobic digestion process", Anaerobic Biological Treatment Processes, pp. 126-162.

Balci, O. (1994), "Validation, verification and testing through the life cycle of a simulation study", *Annals of Operations Research*, Vol. 53 No. 1, pp. 121-173.

Balci, O. (1995), "Principles and techniques of simulation validation, verification, and testing", *Proceedings of the 27th Conference on Winter Simulation*, *IEEE Computer Society*, pp. 147-154.

Balci, O. (1997), "Verification validation and accreditation of simulation models", in *Proceedings of the 29th Conference on Winter Simulation*, *IEEE Computer Society*, pp. 135-141.

Balci, O. (1998), "Verification, validation, and testing", *Handbook of Simulation*, Vol. 10, pp. 335-393.

Balci, O. and Nance, R.E. (1985), "Formulated problem verification as an explicit requirement of model credibility", *Simulation*, Vol. 45 No. 2, pp. 76-86.

Balci, O. and Sargent, R.G. (1982a), "Some examples of simulation model validation using hypothesis testing", in *Proceedings of the 14th Conference on Winter Simulation*, Vol. 2, pp. 621-629.

Balci, O. and Sargent, R.G. (1982b), "Validation of multivariate response models using hoteling two-sample t2 test", *Simulation*, Vol. 39 No. 6, pp. 185-192.

Balci, O. and Sargent, R.G. (1984), "Validation of simulation models via simultaneous confidence intervals", *American Journal of Mathematical and Management Sciences*, Vol. 4 Nos 3/4, pp. 375-406.

Balci, O., Saadi, S.D. and Ormby, W.F. (2000), "Planning for verification, validation, and accreditation of modeling and simulation applications", *Simulation Conference*, *IEEE*, pp. 829-839.

Barlas, Y. (1989), "Multiple tests for validation of system dynamics type of simulation models", *European Journal of Operational Research*, Vol. 42 No. 1, pp. 59-87.

Barnes, C., Verdier, F., Pegatoquet, A., Gaffe, D. and Cottin, J.M. (2016), "Wireless sensor network protocol property validation through the system's simulation in a dedicated framework", *10th International Conference on Signal Processing and Communication Systems, ICSPCS 2016 – Proceedings*, available at: https://doi.org/10.1109/ICSPCS.2016.7843355

Bayarri, M.J., Berger, J.O., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C.-H. and Tu, J. (2007), "A framework for validation of computer models", *Technometrics*, Vol. 49 No. 2, pp. 138-154.

Birta, L.G. and Ozmizrak, F.N.J. (1996), "A knowledge-based approach for the validation of simulation models: the foundation", *ACM Transactions on Modeling and Computer Simulation*, Vol. 6 No. 1, pp. 76-98.

Borko, H. (1962), *Computer Applications in the Behavioral Sciences*, Prentice Hall, Englewood.

Chai, Y., Miao, C., Sun, B., Zheng, Y. and Li, Q. (2017), "Crowd science and engineering: concept and research framework", *International Journal of Crowd Science*, Vol. 1 No. 1, pp. 2-8.

Cooley, P. and Solano, E. (2011), "Agent-based model (ABM) validation considerations", *Proceedings of the Third International Conference on Advances in System Simulation (SIMUL 2011), IARIA, Barcelona*, pp. 134-139.

Dean, S. (2004), "Verification and validation in military simulations", *Proceedings of the 2004 Winter Simulation Conference*, pp. 925-932.

Department of Defense (1996), "Department of defense verification, validation, and accreditation (W and A) recommended practices guide", Defense Modeling and Simulation Office, Alexandria, VA, November, in Balci, P.A., Glasow, P., Muessig, E.H., Page, J., Sikora, S., Solick, S. and Youngblood, S. (Eds), available at: http://triton.dmso.mil/docslib/mspolicy/vva/r

Department of Defense Instruction (DoDI) (2003), "DoD modeling and simulation verification, validation and accreditation", available at: www.docin.com/p-116071952.html

Deslandres, V. and Pierreval, H. (1991), "An expert system prototype assisting the statistical validation of simulation models", *Simulation-Transactions of the Society for Modeling and Simulation International*, Vol. 56 No. 2, pp. 79-89.

Drchal, J., Certický, M. and Jakob, M. (2016), "VALFRAM: validation framework for activity-based models", *Jasss*, Vol. 19 No. 3, available at: https://doi.org/10.18564/jasss.3127

Easterling, R.G. (2001), "Measuring the predictive capability of computational methods: principles and methods, issues and illustrations", Office of Scientific and Technical Information Technical Reports.

Eek, M., Kharrazi, S., Gavel, H. and Olvander, J. (2015), "Study of industrially applied methods for verification, validation and uncertainty quantification of simulator models", *International Journal of Modeling, Simulation, and Scientific Computing*, Vol. 6 No. 2, p. 1550014.

Evans, G.W., Wallace, G.F. and Sutherland, G.L. (1967), "Simulation using digital computers", Digital Simulation.

Fishman, G.S. and Kiviat, P.J. (1967), "The analysis of simulation–generated time series", *Management Science*, Vol. 13 No. 7, pp. 525-557.

Glasow, P., Muessig, P., Sikora, J., Youngblood, S., Balci, O., Solick, S. and Page, E. (1996), "Verification, validation, and accreditation (VV&A) recommended practice guide", Office of the Director of Defense Research and Engineering, Defense Modeling and Simulation Office.

Glasow, P., Muessig, P., Sikora, J., Youngblood, S., Balci, O., Solick, S. and Page, E. (2000), "Verification, validation, and accreditation (VV&A) recommended practice guide 2", Office of the Director of Defense Research and Engineering, Defense Modeling and Simulation Office.

Glasow, P., Muessig, P., Sikora, J., Youngblood, S., Balci, O., Solick, S. and Page, E. (2005), "Verification, Validation, and Accreditation (VV&A) Recommended Practice Guide 2.5", Office of the Director of Defense Research and Engineering, Defense Modeling and Simulation Office.

Goodman, C.J., Mellitt, B. and Rambukwella, N.B. (1987), "CAE for the electrical design of urban rail transit systems", Computers in Railway Operations, pp. 173-193.

Gurcan, O., Dikenelli, O. and Bernon, C. (2011), "Towards a generic testing framework for agent-based simulation models", Computer Science and Information Systems, IEEE, pp. 635-642.

Han, D. (2013), "Research and implementation on application software hybrid trust measurement", working paper, Beijing University of Technology, Beijing, 15 June.

Howard, C.E. (2011), "Simulation and training: expecting the unexpected", *Military and Aerospace Electronics*, Vol. 22 No. 11, pp. 12-14.

Interactive, D. and Committee, S. (1998), "IEEE Trial-Use recommended practice for distributed interactive simulation verification, validation, accreditation, IEEE trial", IEEE.

Interoperability, S., Committee, S. and Computer, I. (2007), "IEEE recommended practice for verification, validation, and accreditation of a federation – an overlay to the High-Level architecture federation development and execution process".

Kheir, N.A. and Holmes, W.M. (1978), "On validating simulation models of missile systems", *SIMULATION*, Vol. 30 No. 4, pp. 117-128.

Kheir, W.M. *et al.* (1986), "Flight mechanics panel working group WG-12 on validation of missile system simulation", Advisory Group for Aerospace Research and Development Neuilly-Sur-Seine (France), available at: https://apps.dtic.mil/docs/citations/ADA166617

Kleijnen, J.P.C. (1999), "Validation of models, statistical techniques and data availability", *Simulation Conference IEEE*.

Klügl, F. (2008), "A validation methodology for agent-based simulations", *Proceedings of the 2008 ACM Symposium on Applied Computing*, *ACM*, pp. 39-43.

Klügl, F. (2009), "Agent-based simulation engineering", PhD thesis, Habilitation Thesis, University of Wurzburg.

Law, A.M. (2007), *Simulation Modeling and Analysis*, 4th ed., McGraw-Hill, New York, NY.

Lent, M.V., Hill, R., McAlinden, R. and Brobst, P. (2003), "2002 Defense modeling and simulation office (DMSO) laboratory for human behavior model interchange standards", University of Southern California Marina Del Rey CA Inst for Creative Technologies.

Li, H. and Sun, F. (2007), "A parallel multi-agent simulation planning approach to complex logistics system with genetic optimization", *International Conference on Wireless Communications*, *IEEE*.

Li, L. (2012), "Research on simulation credibility assessment of marine electric propulsion system", Working paper, Harbin Engineering University, Harbin.

Li-Ping, Z. and Xiao-Ping, L. (2007), "Research on operational validity evaluation of man-in-the-loop simulation system", *Journal of System Simulation*, Vol. 19 No. 7, pp. 1417-1421.

Liu, M. (2017), "Research on restful software credibility testing based on behavior declaration", Working paper, Beijing University of Technology, Beijing, 25 May.

Louloudi, A. and Klügl, F. (2012), "Immersive face validation: a new validation technique for agent-based simulation", Computer Science and Information Systems IEEE.

Lv, H. (2016), "Research on software credibility testing method based on behavior declaration", Working paper, Beijing University of Technology, Beijing, 12 June.

McLeod, J. (1968), *Simulation: The Dynamic Modeling of Ideas and Systems with Computers*, McGraw-Hill.

Machlup, F. (1955), "The problem of verification in economics", *Southern Economic Journal*, Vol. 22 No. 1, pp. 1-21.

Martin, F.F. (1968), *Computer Modeling and Simulation*, Wiley, New York, NY.

Mehrabadi, N.R., Wen, B., Burgos, R., Boroyevich, D. and Roy, C. (2014), "Verification, validation and uncertainty quantification (VV&UQ) framework applicable to power electronics systems", SAE Technical Paper 2014-01-2176, available at: https://doi.org/10.4271/2014-01-2176

Miller, R.L. (1981), "Computer simulation in Geology", *Earth Science Reviews*, Vol. 7 No. 3, pp. A130-A131.

Montanola-Sales, C., Onqqo, S. and Casanovas-Garcia, J. (2011), "Agent-based simulation validation: a case study in a demographic simulation", *Proceedings of the Third International Conference on Advances in System Simulation (SIMUL 2011)*, *Barcelona*, pp. 109-115.

Montgomery, D.C. and Conard, R.G. (1980), "Comparison of simulation and flight test data for missile systems", *SIMULATION: Transactions of the Society for Modeling and Simulation International*, Vol. 34 No. 2, pp. 63-72.

Muessig, P.R. (2001), "A 'smart' approach to VV&A".

Nan, Y., Liu, Y., Shen, J. and Chai, Y. (2017), "A study on the MCIN model in intelligent clothing industry", *International Journal of Crowd Science*, Vol. 1 No. 2, pp. 133-145.

Niazi, M.A., Hussain, A. and Kolberg, M. (2009), "Verification and validation of agent-based simulations using the VOMAS (virtual overlay multi-agent system) approach", *CEUR Workshop Proceedings*, Vol. 494, pp. 340-346.

Oberkampf, W.L. (1994), "A proposed framework for computational fluid dynamics code calibration/validation", Nasa Sti/recon Technical Report N, 95.

Oberkampf, W.L. and Barone, M.F. (2006), "Measures of agreement between computation and experiment: validation metrics", *Journal of Computational Physics*, Vol. 217 No. 1, pp. 5-36.

Peng, P., Jian-Bing, T. and Ya-Bing, Z. (2017), "Amelioration and application of similar degree method for simulation credibility evaluation", *Journal of System Simulation*, Vol. 19 No. 12, pp. 2658-2660.

Pengfei, X., Lees, M., Nan, H. and Viswanthatn, T.V. (2011), "Validation of agent-based simulation through human computation: an example of crowd simulation", *International Conference on Multi-Agent-Based Simulation*, Springer-Verlag, pp. 1-13.

Pidd, M. (2004), *Computer Simulation in Management Science*, John Wiley and Sons Inc, Indianapolis, IN.

Railsback, S.F. and Grimm, V. (2011), *Agent-Based and Individual-Based Modeling: A Practical Introduction*, Princeton University Press, Princeton, NJ.

Reid, J., Stone, K., Brown, J., Caglar, D., Kobayashi, A., Lewis-Newby, M. and Partridge, R. *et al.* (2012), "The Simulation Team Assessment Tool (STAT): development, reliability and validation", *Resuscitation*, Vol. 83 No. 7, pp. 879-886. European Resuscitation Council, American Heart Association, Inc., and International Liaison Committee on Resuscitation. Elsevier Ireland Ltd.

Rothenberg, J. (1989), "The nature of modeling", Artificial intelligence, simulation and modeling.

Sanders, P. (1996), "DoD modeling and simulation (M&S) verification, validation, and accreditation (VV&A) (no. DODI-5000.61)", Office of the Under Secretary of Defense (Acquisition And Technology), Washington, DC.

Sargent, R., Glasow, P., Kleijnen, J., Law, A., McGregor, I. and Youngblood, S. (2000), "Strategic directions in verification, validation, and accreditation research", *Proceedings of the 2000 Winter Simulation Conference*, pp. 909-916.

Sargent, R.G. (1997), "Verification, validation, and accreditation of simulation models", *Applied System Simulation*, Vol. 1 No. 4, pp. 487-506.

Sargent, R.G. (2001), "Some approaches and paradigms for verifying and validating simulation models", *Proceedings of the 2001 Winter Simulation Conference*, *IEEE*.

Sargent, R.G. (2005), "Verification and validation of simulation models", *Proceedings of the 37th Conference on Winter Simulation, WSC'05, Winter Simulation Conference*, 4-7 December, *Orlando, FL*, pp. 130-143.

Shangguan, W., Yuan, C.Y., Zhang, J.M., Cai, B.G. and Wang, J. (2014), "Verification and evaluation method of cooperative vehicle-infrastructure simulation system based on AHP-GRAP", *International Conference on Electromagnetics in Advanced Applications*, *IEEE*.

Shannon, R. (1975), *The Art and Science*, Prentice Hall.

Shichman, H. and Hodges, D. (2003), "Modeling and simulation of insulated-gate field-effect transistor switching circuits", *IEEE Journal of Solid-State Circuits*, Vol. 3 No. 3, pp. 285-289.

Sklar, E. (2007), "NetLogo, a multi-agent simulation environment", *Artificial Life*, Vol. 13 No. 3, pp. 303-311.

Snyder, M.F., Rideout, V.C. and Hillestad, R.J. (1968), "Computer modeling of the human systemic arterial tree", *Journal of Biomechanics*, Vol. 1 No. 4, pp. 341-353.

Sornette, D., Davis, A.B., Ide, K., Vixie, K.R., Pisarenko, V. and Kamm, J.R. (2007), "Algorithm for model validation: theory and applications", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 104 No. 16, pp. 6562-6567.

Sun, H. and Zhang, M. (2017), "An HLA based simulation framework for crowd science", *2017 International Conference on Mathematics*, Modelling and Simulation Technologies and Applications (MMSTA 2017), pp. 714-718.

Tabak, V.V., Vries, D.B.B. and Dijkstra, J.J. (2010), “RFID technology applied for validation of an office simulation model”, Pedestrian and Evacuation Dynamics 2008, pp. 269-275.

Tang, J.B., Zha, Y.B. and Li, G. (2006), “An overview of the research on VV&A in simulation”, *Computer Simulation*, Vol. 11 No. 4, pp. 82-86.

Terano, T. (2006), “Exploring the vast parameter space of multi-agent based simulation”, *International Conference on Multi-Agent-Based Simulation VII*, *Springer-Verlag*, pp. 1-14.

Terano, T. (2007), “Exploring the vast parameter space of multi-agent based simulation”, in Antunes, L. and Takadama, K. (Eds), *Multiagent-Based Simulation VII. Lecture Notes in Computer Science*, Springer, Berlin Heidelberg, Vol. 4442, pp. 1-14.

Tian, G.W., Yong, W.U., Dong, Z.J., Bao, T.H. and Tao, J.J. (2012), “A grey theory model using in the simulation credibility evaluation of the integrative avionics system”, Fire Control and Command Control.

Wang, Z., Li, Z., Feng, Y. and Rong, G. (2016), “Crude oil operations under uncertainty: a continuous-time rescheduling framework and a simulation environment for validation”, *Industrial and Engineering Chemistry Research*, available at: acs.iecr.6b01108

Weisel, E.W. (2004), “Models, composability, and validity”, Doctoral dissertation, Old Dominion University, Norfolk, VA.

Weisel, E.W., Petty, M.D. and Mielke, R.R. (2003), “Validity of models and classes of models in semantic composability”, *Proceeding Fall Simulation Interoperability Workshop 2003*, Vol. 9, p. 68.

Whitner, R.B. and Balci, O. (1989), “Guidelines for selecting and using simulation model verification techniques”, in MacNair, E.A., Musselman, K.J. and Heidelberger, P. (Eds), *Proceedings of the 1989 Winter Simulation Conference, IEEE, Piscataway, NJ*, pp. 559-568.

Wigan, M.R. (1972), “Fitting, calibration, and validation of simulation models, simulation”, Vol. 18 No. 5, pp. 188-192.

Yilmaz, L. and Balci, O. (1997), “Object-oriented simulation model verification and validation”, *Proceedings of the 1997 Summer Computer Simulation Conference*, *SCS*, *San Diego, CA*.

Yu, X.J. and Xiao, R. (2018), “Credibility verification method and calculation based on application behavior declaration”, *Computer Systems and Applications*, Vol. 27 No. 11, pp. 17-26.

Zeigler, B.P. and Sarjoughian, H.S. (2002), “Implications of M&S foundations for the V&V of large scale complex simulation models”, *Proceedings of the Foundations for V&V in the 21st Century Workshop*, *Laurel, MD*.

Zha, Y. and Kedi, H. (1997), “A survey on the credibility of system simulation”, *Journal of System Simulation*, Vol. 9 No. 1, pp. 4-9.

Zhang, S., Geng, H., Tong, J. and Du, M. (2012), “The research review on VV&A working system of complex simulation system”, AsiaSim 2012, pp. 534-539.

Zhang, W. and Wang, X.R. (2001), “Simulation credibility”, *Xitong Fangzhen Xuebao/Acta Simulata Systematica Sinica*, Vol. 13 No. 3, pp. 312-314.

Zoumpoulaki, A., Avradinis, N. and Vosinakis, S. (2010), “A multi-agent simulation framework for emergency evacuations incorporating personality and emotions”, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 6040 LNAI, pp. 423-428.

Zupan, B., Holmes, J.H. and Bellazzi, R. (2006), “Knowledge-based data analysis and interpretation”, *Artificial Intelligence in Medicine*, Vol. 37 No. 3, pp. 163-165.

Further reading

Brade, D., Jacquart, R., Voogd, J.M. and Yi, C.H. (2005), “Weag Thales jp11.20 – Final state of the REVVA methodology”.

Department of Defense (2000), "Verification validation and accreditation (VV&A) recommended practice guides BUILD 2", available at: www.docin.com/p-116071952.html

Gang, X., Ju-Long, L., Yu-Xiang, H.U. and Shi-Ran, L. (2016), "Evaluation approach for network components performance using trustworthiness measurement", *Journal on Communications*, Vol. 37 No. 3, pp. 118-128.

Hui-Zhen, Y., Feng-Ju, K. and Jun, L. (2003), "A fuzzy AHP method for credibility evaluation of system simulation", *Computer Simulation*, Vol. 20 No. 8, pp. 43-46.

Jacquart, R., Bouc, P. and Brade, D. (2004), "A common verification, validation, and accreditation framework for simulations: Project JP11.20, REVVA", *Spring Simulation Interoperability Workshop*.

Lei-Sheng, Q., Xiao-Na, Z., Xing-Huo, M. and Bo, L. (2018), "Credibility evaluation of the C4ISR system simulation test based on the fuzzy AHP method", *Software Engineering*, Vol. 21 No. 1, pp. 8-11.

Li-Ge, Q. (2009), "Research on the credibility evaluation and its tool of complex simulation system", Working paper, Harbin Institute of Technology, Harbin, 3 July.

Li-Ge, Q., Ming, Y. and Ke, F. (2010), "Research on the simulation credibility evaluation assistant tool based on hierarchical evaluation", *Computer Simulation*, Vol. 27 No. 6, pp. 118-119.

Oberkampf, W.L. and Trucano, T.G. (2007), "Verification and validation in computational fluid dynamics", *Advances in Mechanics*, Vol. 38 No. 3, pp. 209-272.

Shi-Jie, C., Chen, L.C. and Lin, L. (2001), "Knowledge-based support for simulation analysis of manufacturing cells", *Computers in Industry*, Vol. 44 No. 1, pp. 33-49.

Shi-Kao, L. and Xing-Tang, L. (2001), "A new method of elevation of confidence level of large-scale perplexing simulation system", *Acta Simulata Systematic Sinica*, Vol. 13 No. 5, pp. 666-670.

Xue-Rong, Y., Ying, L. and Xiang-Jun, F. (2006), "Design of the M&S and VV&A management system", *Computer Simulation*, Vol. 23 No. 10, pp. 284-288.

Yu, X. and Wang, J. (2018), "A preliminary study of the automatic generation of credibility test cases based on the immune algorithm", Cluster Computing.

Yu-Jie, Z. and Xin-Xing, L. (2014), "Fuzzy evaluation method for Web software trustworthiness", *Application Research of Computers*, Vol. 31 No. 4, pp. 1072-1076.

Zhen, M., Xiao-Yan, W., Xiang-Wei, B. and Rui, Z. (2016), "VV&A tools for simulation model", *Modern Defence Technology*, Vol. 44 No. 1, pp. 224-229.

Corresponding author
Sun Hongbo can be contacted at: hsun@ytu.edu.cn