A Data-Centric Approach for Model-Based Systems Engineering

Guoxiong ZHAN

*College of Information System and Management, National University of Defense Technology,
Changsha 410073, China
E-mail: zhan_guoxiong@163.com*

Bingfeng GE

*College of Information System and Management, National University of Defense Technology,
Changsha 410073, China
E-mail: bingfengge@nudt.edu.cn*

Minghao LI

*College of Information System and Management, National University of Defense Technology,
Changsha 410073, China
E-mail: liminghao4869@sina.com*

Kewei YANG

*College of Information System and Management, National University of Defense Technology,
Changsha 410073, China
E-mail: kayyang27@nudt.edu.cn*

Abstract A data-centric approach is proposed to facilitate the design and analysis of challenging complex systems and address the problems of currently existing model-based systems engineering (MBSE) methodologies. More specifically, based on three core steps of current MBSE methodologies, a high-level data meta-model, depicting the semantic relationships of high-level data concepts, is first presented to guide the data modeling for systems engineering (SE). Next, with respect to the six high-level data concepts, the data elements are collected as the modeling primitives to construct static and/or executable models, which can also act as a common and consistent data dictionary for SE. Then, the mapping associations amongst core data elements are established to associate the model elements in different steps and achieve the requirement traceability matrix. Finally, the feasibility of the proposed approach is demonstrated with an illustrative example.

Keywords data-centric; model-based systems engineering (MBSE); meta-model; data modeling; requirement traceability matrix

1 Introduction

With information technologies continuing to evolve, the size and complexity of today's systems are increasing rapidly and systems engineering (SE) is undergoing major changes. Obviously, traditional document-centric systems engineering suffers from limitations arising from

Received June 23, 2015, accepted August 15, 2015

Supported by National Natural Science Foundation of China (71201168)

manual updates, ambiguity, unstructured representation of information, etc. This has brought lots of trouble to the system design process^[1]. In this respect, model-based systems engineering (MBSE) emerges to formalize the application of models^[2], moving engineering and management from paper documentation as a communication medium to a paperless environment with model-based visualization^[3]. MBSE centers the model construction and analysis in the whole systems engineering process (SEP), and achieves the system design by continuous evolvement and iterative improvement of models^[4]. Compared with the document-centric systems engineering, MBSE possesses many advantages, such as enhancing the communication, strengthening the information traceability, improving the reusability of elements, etc^[5-7]. At present, there are many MBSE methodologies including the IBM/Telelogic/i-Logix Harmony SE, INCOSE object oriented systems engineering method (OOSEM), IBM rational unified process for systems engineering (RUP SE), Vitech model-based system engineering methodology, and JPL state analysis (SA)^[8]. These methodologies not only deepen the understanding of complex systems to a large extent, but also facilitate the management of the complexity effectively. No doubt, MBSE has grown in popularity as a way to deal with the design of complex systems. However, the foregoing MBSE methodologies place so much emphasis on the representation form of models and overlook the importance of the underlying data for the verification and analysis of models, which may easily lead to some serious problems, such as the data inconsistency and duplicate definition. Moreover, without the complete definition of data elements and their associations, the modification of elements in one model is difficult to trace and update in other models, and thus the correctness and effectiveness of models are not guaranteed. In addition, there always exist different modeling standards and tools corresponding to different MBSE methodologies^[9], thus it is quite difficult to compare and integrate models across these methodologies. Therefore, in this paper, a data-centric approach is proposed to address the above problems. The remainder of this paper is structured as follows. Section 2 presents the analysis and summary of three core steps of MBSE methodologies, with an outline of the collective shortcomings. The proposed data-centric approach is discussed in Section 3, including the high-level data meta-model, data elements in three core steps, and mapping associations among core entities. An illustrative example is then used to demonstrate the feasibility of the proposed approach in Section 4. Finally, Section 5 includes the conclusion and points to future work.

2 Core Steps of MBSE

In general, each of these MBSE methodologies provides a model-based approach for performing the three core steps of SEP: Requirements analysis, system functional/behavioral analysis, and architectural design^[10]. They constitute a comprehensive, iterative and recursive problem solving process, transforming stakeholder requirements into the system architecture, generating information for decision makers, and providing inputs for the next level of development. As shown in Figure 1, a typical SEP begins with analyzing the initial stakeholder requirements, missions and environment constraints. Subsequently, a set of functional requirements and performance requirements are developed in step of requirements analysis. In next step, the system functional architecture is designed, and the behavioral modes of systems are analyzed to ensure

that all requirements are satisfied. In step of architectural design, the physical architecture solutions of systems are defined and selected through trade-off analysis. Finally, the requirements verification process is performed to ensure that all requirements are met by the proposed solution. The final outputs of the entire SEP include the decision database, system architecture, etc.

Figure 1 Three core steps of MBSE

As shown in Figure 1, all the core steps are performed around the models, including requirement models, function/behavior models, and system models. But the consistency of the underlying data is hardly assured and thereby the associations between data elements are difficult to construct and analyze. Therefore, to support the data-related check and analysis of models, the proposed data-centric approach captures the core data elements for three core steps respectively. The static or dynamic executable models in different steps can be produced by organizing these data elements regardless of what modeling language or tool is employed in the development. In other words, the models are merely the visualized representations of them, while these data elements are the true driver of the SEP.

3 Data-Centric Approach

3.1 High-Level Data Meta-Model

Figure 2 High-level data meta-model based on 5W1H

According to the model driven architecture with the meta object facility designed by the object management group^[11], systems models can be organized within a four-layered architecture. The meta-model (model of model or model of metadata) of this four-layered architecture typically defines the abstract syntax of models and the interrelationships between model elements. That is, it is usually defined as a set of concepts or model elements, as well as the constraints and rules for how they may be arranged and related to build models without necessarily providing the concrete syntax of the language^[12]. Accordingly, the meta-model is independent of modeling languages and tools, providing the basis for comparing and integrating the models across different MBSE methodologies. It thus ensures data consistency, and facilitates the automatic construction of some models from previously defined data^[13, 14]. In the three core steps of SEP, there always exists a need to answer a series of questions to collect information on the requirements, functions/behavior, and architecture (RFBA) of the system. These questions can be grouped into six semantically complete communication interrogatives, i.e., WHO, WHERE, WHAT, WHEN, WHY, and HOW (5W1H)^[15]. Clearly, 5W1H, as the high-level data concepts representing the six different aspects of the RFBA respectively, can be used to collect all the key information and thereby capture the core data elements for all steps of SEP. Moreover, as the data taxonomy of models at the highest level, 5W1H can ensure the consistency in the

meaning of each data elements, in order to eliminate the aforementioned problems in current MBSE methodologies. As shown in Figure 2, 5W1H, as well as their logical relationships of how they may be arranged and related to build models, constitute the high-level data meta-model. It can be summarized as an integrated dynamic model of sequenced Activities (HOW) and their Events (WHEN) performed by Performers (WHO) to produce and consume Resources (WHAT) in Locations (WHERE) under specified Rules and Conditions (WHY)^[16]. If performance parameters and measures are incorporated into a basic executable model constructed from core data elements and their associations, quantitative decision-making analysis in the practice of SEP can be supported effectively. The data meta-model formalizes general or high-level data concepts and their associations with standard terminology and formal semantics. Thus, in the whole SEP, all the necessary data elements in accordance with the data meta-model can be collected, maintained, and shared among various organizations as a common, consistent, and integrated data dictionary.

3.2 Data Modeling

Based on the above high-level data meta-model, all the necessary data elements in the whole SEP can be captured around 5W1H. These data elements can be divided into three object classes: entities, relationships among entities, and attributes of entities and their relationships^[17]. Figure 3 depicts the whole process of the proposed approach, and shows the data entities in different steps. In order to reduce the complexity of the picture, the relationships among entities and their attributes will be discussed later. Note that there exist corresponding relations between the data elements in different steps. For example, the Capability, Operational Activity, and System Function are connected by the relation of HOW, indicating that the three data entities are all derived from HOW that is one of the high-level data concepts. Thus, they are just different representations for the same semantic content in different steps of SEP. In particular, the elements of the functional/behavioral analysis and architectural design steps are symmetrically aligned with each other within each of the 5W1H. As shaded in Figure 3, these data elements representing the semantic concepts of 3W1H (i.e., HOW, WHO, WHERE, and WHAT) are the core entities for SE, which are also considered as the core modeling primitives with the meaning of HOW performed by WHO at WHERE to produce/consume WHAT. As shown in Figure 3, the proposed approach includes the following three steps.

Step 1 As discussed earlier, requirements analysis starts with the import of the stakeholder requirements that typically are problem-oriented statements and focus on required capabilities^[18]. That is, capabilities are the very essence of the stakeholder requirements. Thus, if the developed system possesses the desired capabilities, it will meet the requirements completely. In this sense, the major data entities include Vision, Capability, and Capability Phasing. And there exist affiliations between the capabilities, denoting the relationships. With these relationships, the general or high-level capabilities can be decomposed into more concrete capability units, namely the system functional requirements which are the inputs of next step.

Figure 3 Data elements of the three core steps

Step 2 In response to the developed functional requirements in Step 1, the behavioral modes are designed and analyzed based on a series of sequenced activities in specific scenarios. And then the operational activities are performed to check whether the desired effects are achieved to meet the requirements or not. Here the iterative and incremental development is carried out to ensure that all the functional requirements are supported by corresponding activities. Consequently, the major data entities consist of Desired Effect, Resource, Operational Activity, Logical Location, Operational Node, and Process, which are just the concretization of 5W1H. The relationships among them are aligned well with the associations of 5W1H. For example, the relationship between Resource and Operational Activity is defined as “Operational Activity produces/consumes Resource”, which is just derived from the association between HOW and WHAT. The operational resource flows are the attributes of Activity, while Operational Node contains several attributes, such as organizations, measures, and so on.

Step 3 The system architecture, including the set of sub-systems and how they connect each other, is designed to provide required system functions, in order to support the behavior modes and thereby satisfy the functional requirements. Therefore, Goal, Resource, System Function, Physical Location, System, and Process are the major data entities in the step of system architectural design. The interfaces are used to denote the relationships among Resource, System, and Logical Location, while the resource flows offer the attributes for the interfaces. In addition, the performance parameters and measures of the system can be utilized to conduct trade-off analysis for alternatives of the system architecture, in order to obtain the best solution.

3.3 Mapping Associations among Core Entities

Around 5W1H, all the necessary data elements in accordance with the aforementioned high-level data meta-model are captured for different steps of SEP in a semantically complete and consistent way. No doubt, these data elements are considered as the center and basis of the whole SEP. More specifically, Capability as the core data entity is really the foundation of the collection, modeling and analysis of requirement-related data in Step 1. And the functional/behavioral analysis is carried out around Operational Activity in Step 2. With regard

to the step of architectural design, it focuses on both of System and System Function, each of which is the core data entity. According to the guidelines on data meta-models and view products provided by the latest U.S. Department of Defense Architecture Framework (DoDAF 2.0), the core data entity Capability can be successively mapped to System Function and System with the products Capability to Operational Activities Mapping (CV-6) and Operational Activity to Systems Function Traceability Matrix (SV-5a), and the view products in Systems Viewpoint (SV)^[19]. Therefore, the mapping associations among the data entities in Figure 4 and the requirements traceability matrix in Figure 5 are established.

Figure 4 Capability mapping process

Requirements		Capability 1	Capability 2	Capability 3	Capability 4	Capability 5	Capability 6	Capability 7	Capability 8	Capability 9
Systems, System Functions										
System 1	System Function A	√							√	
	System Function B					√	√			
	System Function C		√							√
System 2	System Function A	√			√				√	
	System Function D	√				√				
System 3	System Function E									
	System Function F							√		
	System Function G			√						
	System Function H									√

Figure 5 Requirement traceability matrix

On one hand, the mapping associations among the data elements effectively solve the aforementioned problem that the elements of models are hardly associated in current MBSE methodologies. On the other hand, with the requirements traceability matrix, it becomes quite easy to identify whether the capabilities are supported by corresponding system functions and realized by the systems or not, thus the requirements verification is no longer a big headache. For example, as shown in Figure 5, there exist mapping associations between Capability 5 and System Function B and D, while the two system functions are accomplished by Systems 1 and

2 respectively. Clearly, this capability can be realized or the requirement be satisfied if only the two systems exist simultaneously.

According to the above approach, the data elements are the center of the entire SEP, while the models are just the visualized representations of them. That is, all kinds of view products can be utilized to organize these data elements to form fit-for-purpose models and thereby support more powerful description and analysis. Thus, the forms of the model should/can be selected or customized according to specific problems, not limited to the DoDAF-described view products.

4 Illustrative Example

This paper illustrates the feasibility of the proposed approach through the design of a notional air defense system. The mission of the system is to protect a restricted key area against enemy targets, that is, when any possible enemy target is detected, an interceptor is sent to intercept it. The environment constraints include the changing weather, battlefield situation, personnel, and so on. With regard to the stakeholder requirements, the system is expected to provide the capabilities of protecting the specified area.

According to the proposed approach, based on the high-level data meta-model, all the necessary data elements and their associations in the three core steps of SEP can be captured around 5W1H, as follows:

Step 1 Capabilities (HOW), as the core data entities, include the air defense capability and its sub-capabilities. The de-composition is the major relationship among them. Therefore, all kinds of visualized models can be used to organize these data elements to carry out complete requirements description and analysis. Here the view product Capability taxonomy (CV-2) is utilized to describe these data elements, as shown in Figure 6. Note that, the leaf-level capabilities are considered as the functional requirements, which are the inputs of Step 2.

Figure 6 Capability taxonomy (CV-2)

Step 2 Based on the functional requirements in Step 1, the data elements are captured around 5W1H, as follows: Activities (HOW) include Fly In, Issue Surveillance Directive (Surv_Dir), Issue Interception Directive (Int_Dir), Issue Locating Directive (Loc_Dir), Issue Assessing Directive (Ass_Dir), Track Target, Control, Intercept, Locate Target, Assess Kill, etc. Performers (WHO) include Intelligence, Surveillance, Reconnaissance (ISR), Command, Control, Interceptor, and Enemy Target, each of which is considered as the Operational Node. Locations (Where) correspond to Performers in this example. Resources (WHAT) include Enemy Targets, Interceptors, and other command and control (C2) information. Desired Effects (WHY) is to expel or destroy the Enemy Targets. Events (WHEN) include: 1) As soon as the enemy targets are detected and identified, the ISR will track them, collect relevant information, and send information to the Command; 2) According to the received information, the Command makes decisions and issues orders after a series of analysis; 3) After receiving the orders from the Command, the Control schedules the interceptors to intercept the enemy targets. 4) Finally, the ISR collects relevant battlefield information again, in order to assess the effect. Similarly, according to specific problems, the fit-for-purpose models can be constructed based on the captured data with the support of specific methodologies and presentations. Also, some DoDAF-described view products in OV can be selected to organize these data elements to support visualized analysis. Constrained by the page limitation, this example only gives an activity diagram with swim lanes, namely Operational activity model (OV-5b), to illustrate the sequenced activities needed to provide the desired capabilities, as shown in Figure 7.

Figure 7 Operational activity model (OV-5b)

Step 3 As discussed earlier, the data elements of Step 2 and 3 are symmetrically aligned with each other within each of the 5W1H, which means that they are just different representations for the same semantic content in different steps of SEP. Thus, here only gives two kinds of data elements, namely System Function (HOW) and System (WHO), while the others correspond to those of Step 2, respectively. System Functions (HOW) include Surveillance and Reconnaissance Function, Intelligence Gathering Function, Target Locating Function, Information Transmission Function, Command and Decision-making Function, Control Function, Engagement Function. Systems (WHO) include ISR Sub-system, Command Sub-system, Control Sub-system, and Interceptor. With the above data elements, all kinds of visualized models can be constructed by organizing them. This means that these elements act as the common foundation of all models in different steps of SEP, and thereby the consistency of models is ensured. In addition, there exist lots of relationships among these data elements, such as the interfaces among the sub-systems, the resource flows, etc. If performance parameters and measures are incorporated into a model constructed from these data entities and their relationships, trade-off analysis can be carried out to select the best system architectural solution. Moreover, with the mapping associations among Capability, Operational Activity, System Function, and System, the requirement traceability matrix can be achieved to facilitate the requirements verification, as shown in Figure 8. For example, with respect to the information transmission capability of the air defense system, this capability requirement can be satisfied if only all sub-systems own the information transmission function simultaneously.

Requirements Subsystems and System Functions		Leaf-level Capabilities						
		Situation Awareness	Surveillance and Reconnaissance	Information Transmission	Command and Decision-making	Real-time Control	Expulsion	Fire Strike
ISR	Surveillance and Reconnaissance	√	√					
	Intelligence Gathering	√						
	Target Locating	√						
	Information Transmission			√				
Command	Command and Decision-making				√			
	Information Transmission			√				
Control	Control Function					√		
	Information Transmission			√				
Interceptor	Engagement Function						√	√
	Information Transmission			√				

Figure 8 Requirement traceability matrix

5 Conclusions

In this paper, a data-centric approach and its corresponding process to enable SE is proposed to cope with the collective shortcomings of current MBSE methodologies. This approach abstracts essential information from the underlying complexity with the system modeling in a data-centric and semantically consistent fashion, which keeps the consistency and correctness of model elements and provides the best basis for comparing and integrating the models among different MBSE methodologies due to its independence of methodologies and modeling tools. Through the analysis and summary of three core steps of MBSE methodologies, a high-level data meta-model is first presented based on 5W1H to guide the data modeling to capture all the necessary data elements and their relationships for all steps of the entire SEP. Next, semantically complete and consistent data elements are collected around 5W1H, which are independent of modeling languages and tools and thereby can act as a common and integrated data dictionary for various organizations. In particular, as the modeling primitives, these data elements can be organized to form all kinds of fit-for-purpose models to support more powerful analysis. And then, the mapping associations among the core data entities, including Capability, Operational Activity, System Function, and System, are established to associate the model elements in different steps of SEP and achieve the requirement traceability matrix. With the matrix, it becomes quite easy to carry out requirement verification. Finally, the feasibility of the proposed approach is demonstrated with an illustrative example of a notional air defense system. Although the research in this paper is mainly based on DoDAF, the proposed approach can also be applied for the design of civil complex systems by making some revisions to the terminology of data elements. The ongoing and future work is aimed at addressing in detail the automated transformation of executable models from these data elements to facilitate the quantified analysis and optimization of the system architecture.

Acknowledgements The authors would like to thank our co-workers for their kind assistance in improving the presentation quality of this paper. The authors are also grateful to the Editor-in-Chief, Editor, and Editorial Manager for handling the processing of their paper.

References

- [1] Kalawsky R S, O'Brien J, Chong S, et al. Bridging the gaps in a model-based system engineering workflow by encompassing hardware-in-the-loop simulation. *IEEE Systems Journal*, 2013, 7(4): 593–604.
- [2] Dickerson C E, Mavris D. A brief history of models and model based systems engineering and the case for relational orientation. *IEEE Systems Journal*, 2013, 7(4): 581–592.
- [3] INCOSE. INCOSE Systems Engineering Vision 2025. San Diego, CA, USA, Jun. 2014.
- [4] Ramos A L, Ferreira J V, Barceló J. Model-based systems engineering: An emerging approach for modern systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 2012, 42(1): 101–111.
- [5] NDIA. Final report of the model based engineering (MBE) subcommittee. Systems Engineering Division M&S Committee, National Defense Industrial Association, Arlington, VA, Feb. 2011.
- [6] Fieber F, Regnat N, Rumpe B. Assessing usability of model driven development in industrial projects. *Proc. 4th Eur. Workshop: From Code Centric to Model Centric Software Engineering: Practices, Implications and ROI*, Jun. 2009: 1–9.
- [7] Nolan B, Brown B, Balmelli L, et al. Model driven systems development with rational products. New York: IBM International Technical Support Organization, 2008.
- [8] Estefan J A. Survey of model-based systems engineering (MBSE) methodologies. INCOSE INCOSE-TD-2007-003-02, 2008: 1–80.

- [9] Rao B H, Padmaja K, Gurulingam P. A brief view of model based systems engineering methodologies. *International Journal of Engineering Trends and Technology*, 2013, 4(8): 3266–3270.
- [10] Piaszczyk C. Model based systems engineering with department of defense architectural framework. *Systems Engineering*, 2011, 14(3): 305–326.
- [11] Object Management Group. Meta object facility (MOF) core specification. Needham, MA, USA, Jan. 2006.
- [12] Ge B, Hipel K W, Yang K, et al. A novel executable modeling approach for system-of-systems architecture. *IEEE Systems Journal*, 2014, 8(1): 4–13.
- [13] Piriou P Y, Faure J M, Deleuze G. A meta-model to support the integration of dependability concerns into systems engineering processes: An example from power production. *IEEE Systems Journal*, to be published.
- [14] Pfister F, Chapurlat V, Huchard M, et al. A proposed meta-model for formalizing systems engineering knowledge, based on functional architectural patterns. *Systems Engineering*, 2012, 15(3): 321–331.
- [15] Ge B, Hipel K W, Yang K, et al. A data-centric capability-focused approach for system-of-systems architecture modeling and analysis. *Systems Engineering*, 2011, 16(3): 363–375.
- [16] DoD Architecture Framework Working Group. DoD architecture framework. Department of Defense, USA, 2010.
- [17] Ring S J, Nicholson D. Activity-based methodology for development and analysis of integrated DoD architectures. *Handbook of Enterprise Systems Architecture in Practice*, Saha P, Ed. Hershey, PA, USA: Information Science Reference, 2007.
- [18] Hoffmann H P. Systems Engineering best practices with the rational solution for systems and software engineering deskbook release 3.1.2. [Online]. Available: <http://ibm.co/HarmonyDevWorks>.
- [19] Yang K W, Ge B F, Zhao Q S. An architectural approach for capability mapping and gap analysis. *Journal of Systems Science and Information*, 2013, 1(1): 86–96.