Digital Twin Federation and Data Validation Method

Myung-Sun Baek [1, 2]
[1]Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea
[2] University of Science and Technology (UST), Daejeon, South Korea
sabman@etri.re.kr

***Abstract*—Digital twin is a technology developed to solve problems in the real world or to optimize processes and systems in the real world. Digital twins virtually replicate various objects in the real world, and based on this, simulations in the digital world can be performed at low cost to solve or optimize problems in the real world. Until now, these digital twin technologies have been designed and developed to target a single system or single process. Recently, some researchers and developers have tried to optimize a more complex world or to solve complex problems by using digital twin. The concept that emerged based on this is the federated digital twin, which organically unites multiple digital twins. However, the first procedure for effective operation of multiple digital twins is to verify the integrity and validity of data. In this paper, a data validation method for constructing the federated digital twin is introduced.**

Keywords— Digital twin, federation, federated digital twin

I. INTRODUCTION

Digital twin is a technology developed to solve or optimize problems in the real world by digitally replicating objects (people, things, and spaces) in the real world and performing various simulations through them [1]. In order to convert objects to digital or to perform predictive simulations through digitally converted objects, the latest technologies such as advanced machine learning, data analysis, data and object visualization, and advanced simulation are required [2]. This digital twin concept was presented by Michael Grieves in 2002. At that time, he introduced the concept of digital twin through a presentation called Product Lifecycle and Management (PLM) [3]-[4]. Since then, a digital twin has been attempted by NASA to simulate a physical model of a spacecraft. Digital twin is mainly used to succeed in projects that can waste a lot of money and time if a mistake is generated. The concept of the digital twin continues to evolve, and is mainly developed and utilized for optimizing industrial processes or equipment.

Recently, various projects at the national level using digital twins have been introduced. Singapore has developed and operated 'Virtual Singapore' using digital twin technology at the national level. Furthermore various digital twins are being developed in Korea, mainly in Busan, Sejong, and Jeju. It is very difficult to build a digital twin at the city or national level based on a single digital twin concept. Therefore, in order to build a digital twin of a large system or to solve complex problems and optimize the system, a federated digital twin technology that can operate as one large-scale digital twin by uniting multiple digital twins is required. By using the federated digital twin technology, it is possible to select multiple digital twins according to the purpose, and to federate and operate them [5]. However, since the digital twin is a technology implemented by combining various latest technologies, it is very sensitive to errors. Therefore verification of the operation and data validity of each technology is very important.

In this paper, we study the federated digital twin data validation method to effectively operate the federated digital twin. We study the validation method according to the various data types to constitute the federated digital twin and the method of verifying whether a valid output is derived by verifying the operation of a single digital twin.

II. DIGITAL TWIND FEDERATION MODEL

As shown in Fig. 1, we consider a federated digital twin in which multiple digital twins work together. As shown in the figure, a number of various single digital twins are placed at the bottom of the platform, and a federated digital twin is built by selecting multiple single digital twins according to the purpose of the federated digital twin operation. By operating the federated digital twin constructed in this way, large-scale system optimization can be performed. If necessary, not only a single digital twin but also individual objects of the physical world are replicated as digital objects and added to the federated digital twin. In this paper, we consider the part of constructing a federated digital twin and verifying the validity of the corresponding digital twin.

III. DATA VALIDATION METHOD

In order to optimize a large-scale system by combining multiple digital twins, it is very important to determine whether single digital twins are operating normally and whether the output data is valid. A single digital twin is also a complex system of various digital technologies and data. Therefore, determining whether each of these systems is operating normally is an essential element in the operation of a federated digital twin. This section describes how to determine whether a single digital twin operates normally and verify the digital twin data validity. Fig. 2 shows the digital twin data validation method for federated digital twins.

In the case of digital twin data validation, analysis of characteristics and patterns by data properties and types should be firstly performed. And the data validation platform then determines the type of input data. Therefore, it is determined whether the data is time-series characteristic data, sporadic event data, anomaly sensing data such as traffic accidents in the traffic digital twin or river overflow in the river management digital twin. And the data validation

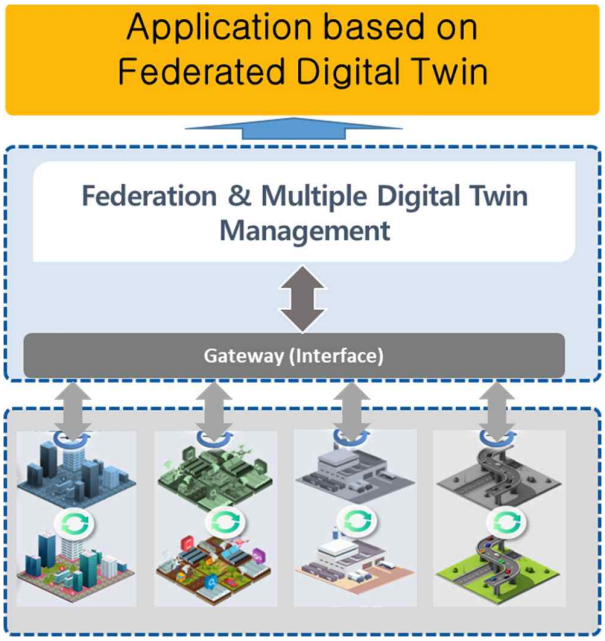

Fig. 1. Conceptual diagram of federated digital twin

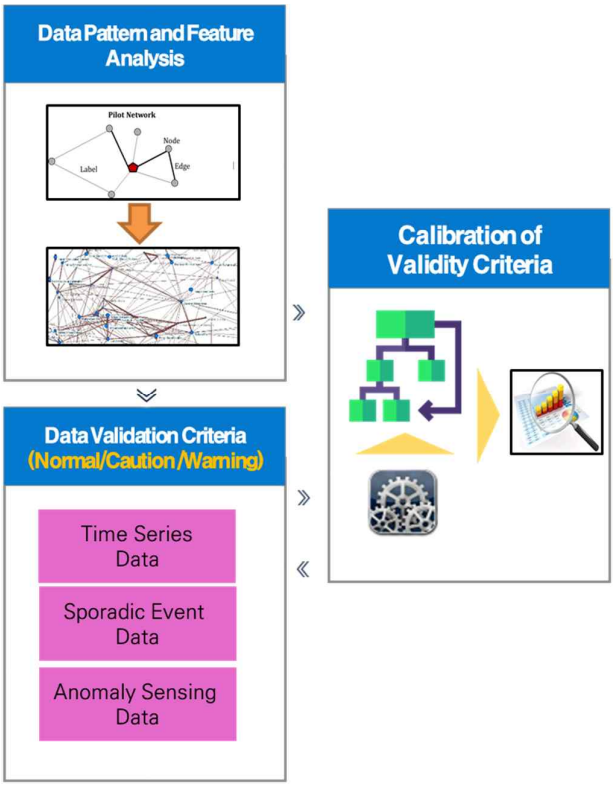

Fig. 2. Digital twin data validation process.

platform analyzes the characteristics and patterns of the input data. Through this analysis, efficacy criteria or thresholds are established. In this case, as for the validity criteria, criteria such as normal, caution, and warning can be set. Although it is possible to set the validity criteria based on statistical characteristics, it is more effective to set a rough reference value through an expert with domain knowledge and to upgrade the reference value primarily. When correcting the validity criterion for each attribute, the validity criterion value can be corrected based on data causality analysis such as RNN (Recurrent Neural Network). Through this, it is possible to check whether the data of the digital twin is output within the effective range. Real-time analysis is required when validating time series data. This is because time series data output in real time should be immediately input and validated.

Thereafter, a procedure for determining whether a single digital twin operates normally is performed. The operation of a single digital twin can be expressed as a combination of various data. For example, in the case of an operation of a traffic digital twin, weather data, temperature data, and vehicle movement data are combined to output one operation. Therefore, the operation of a single digital twin can be verified based on a validation tool in which various data properties are combined. The validation tool can be implemented with various data attribute verifiers, and in order to improve implementation efficiency, it is necessary to design each data attribute device so that it can be reused in various validation tools.

In addition, the results of the single digital twin operation validation module should be able to be displayed to the administrator through the user GUI.

IV. Conclusions

In order to optimize the large-scale real world and complex systems, it is necessary to implement a federated digital twin in which multiple single digital twins are federated. For the smooth operation of these joint digital twins, validation to determine whether digital twin data and operation are normal must be ensured. In this paper, data and single twin operation validation methods for joint digital twin were studied. If the validity of data and individual twins is verified through the proposed method, high-reliability/high-efficiency operation of the federated digital twin is possible.

Acknowledgment

This paper was supported 100% by IITP grant funded by ICT R&D program of MSIT in 2022 year (No. 2022-0-00545, Development of Intelligent digital twin object federation and data processing technology)

References

[1] D. -Y. Jeong et al., "Digital Twin: Technology Evolution Stages and Implementation Layers With Technology Elements," *IEEE Access*, vol. 10, pp. 52609-52620, May, 2022.

[2] E. K. Maskooni, S. A. Naghibi, H. Hashemi, and R. Berndtsson, "Application of advanced machine learning algorithms to assess groundwater potential using remote sensing-derived data," *Remote Sensing*, vol. 12, no. 17, pp. 1–25, Aug. 2020.

[3] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," *Transdisciplinary Perspectives on Complex Systems. Cham, Switzerland: Springer*, pp. 85–113, 2017.

[4] B. R. Barricelli, E. Casiraghi, and D. Fogli, "A survey on digital twin: Definitions, characteristics, applications, and design implications," *IEEE Access*, vol. 7, pp. 167653–167671, 2019.

[5] D.-Y. Jeong, et. al., "Korea digital twin technology roadmap ver. 1," *IITP*, Decmber 2021