A Time-Series Process Event Log Preprocessing Approach for Data-Intensive and Predictive Operationalization of Smart Factories

Minjae Park*, Sohyang Park**, Thanh-Hai Nguyen**, Dinh-Lam Pham***, Dongkeun Oh****,
Kyoung-Sook Kim****, In-Kyoo Chun****, Kwanghoon Pio Kim**
*Div. of Computer Information, DAELIM UNIVERSITY COLLEGE, Anyang Gyeonggi, 13916, South Korea
**Div. of AI Comp. Science & Engineering, KYONGGI UNIVERSITY, Suwon Gyeonggi, 16227, South Korea
***Contents Convergence SW Res. Inst., KYONGGI UNIVERSITY, Suwon Gyeonggi, 16227, South Korea
****Smart Factory Research Institute, IC&IT Co., Guro Seoul, 08377, South Korea
mjpark@daelim.ac.kr, fsa7291, nguyenthanhhai, phamdinhlam, kwang,g@kgu.ac.kr, fdkoh, kskim, tazg@icnit.co.kr

Abstract—This paper proposes a time-series process event log preprocessing approach applied to realize data-intensive and predictive operationalization of IoT-supported smart factories. The advanced conceptual approach of the data-intensive and predictive operationalization processes is newly proposed for effectively and predictively operating very large-scale smart factories that are supported by the Internet of Things platforms with temporal operationalization processes and their time-series event log datasets. The authors' research group is trying to develop a predictive smart factory operationalization system, which is characterized by the advanced conceptual approach, and the starting point of which is at acquiring the time-series process event log datasets with a proper data preprocessing approach. Hence, this paper devises a time-series process event log preprocessing approach with its algorithmic details to arrange the proper datasets to be used for concretizing the data-intensive and predictive operationalization in the very large-scale smart factories. Finally, we develop an XES-formatted event log preprocessing system based upon the devised preprocessing algorithms and apply to a real and nation-wide production smart factory in order to validate the operationality and feasibility of the developed system.

Keyword—Smart Factory, Internet of Things, Data-Intensive Smart Factory, Operationalization Process Automation, Digital Twin Platform, IoT-Sensed Dataset, Time-Series Dataset

Minjae Park is an assistant professor in the division of computer and information at Daelim University College, South Korea. He received B.S. degree, M.S. degree, and Ph. D. degree in computer science from Kyonggi University in 2004, 2006, and 2009, respectively. He had worked as a researcher and developer at BISTel, Inc. in South Korea. His research interests include semiconductor manufacturing and fault-detection systems, large-scale database systems and applications, enterprise information systems, the Internet of Things collaboration architectures and services, process-aware knowledge discovery, process mining, and process monitoring and predictive monitoring.

Dinh-Lam Pham received the B.S. and M.S. degrees in computer science from Thai Nguyen University, Vietnam, in 2008 and 2010, respectively, and the Ph.D. degree in computer science from Kyonggi University, South Korea, in 2021. He is currently a Research Professor with the Contents Convergence Software Research Institute and a Research Member with the Data and Process Engineering Research Laboratory, Kyonggi University. His research interests include process mining, deep learning-based process predicting, workflow systems, workflow-supported social and affiliation networks discovery and analysis, and process-aware Internet of Things architecture and services.

Kyoung-Sook Kim received the B.S. degree in computer science from Kyonggi University, South Korea, in 1984. She also received the M.S. degree and the Ph.D. degree in computer science and engineering from Kyunghee University, South Korea, in 2004 and 2019, respectively. She is currently a lecturer with the Division of AI Computer Science and Engineering, Kyonggi University. She is also a research member of the Contents Convergence Software Research Institute in Kyonggi University. Her research interests include data and knowledge engineering, process-aware enterprise information systems, business process intelligence, process mining, and deep learning and knowledge engineering using active contents big data.

Kwanghoon Pio Kim (M'21) received the B.S. degree in computer science from Kyonggi University, South Korea, in 1984, the M.S. degree in computer science from Chungang University, South Korea, in 1986, and the M.S. and Ph.D. degrees in computer science from the University of Colorado Boulder, USA, in 1994 and 1998, respectively. He had been the Dean of the Computerization and Informatics Institute, Kyonggi University, from 2017 to 2021. Since 2007, he has been the Founder and the Director in charge of the Contents Convergence Software Research Institute, Kyonggi University, where he has been leading and fulfilling a multi-million dollars research project that will be continuously supported and funded by the National Research Foundation of Korea, from 2020 to 2029, ever since the institute was officially designated as the National Science and Engineering Research Institute by the Ministry of Education, South Korea, in 2020. He had worked as a Researcher and a Developer at Aztek Engineering, American Educational Products Inc., and IBM, USA, and a Research Member of Staff at the Electronics and Telecommunications Research Institute (ETRI), South Korea. He is currently a Full Professor with the Division of AI Computer Science and Engineering and the Founder and the Supervisor of the Data and Process Engineering Research Laboratory, Kyonggi University, South Korea. His research interests include groupware, workflow and business process management systems, process warehousing and mining, predictive process monitoring, workflow-supported social networks discovery and analysis, process-aware Internet of Things, predictive process modeling, process deep-learning.