Incorporating a Prediction Engine to a Digital Twin Simulation for Effective Decision Support in Context of Industry 4.0

Rushan Arshad[1], Paul Ton de Vrieze[1[0000-0002-9612-8951]], Lai Xu[1[0000-0003-0063-759X]]
[1] Department of Computing and Informatics, Faculty of Science and Technology
Bournemouth University, Talbot Campus,
Bournemouth, UK, BH12 5BB.
arshadr@bournemouth.ac.uk, pdvrieze@bournemouth.ac.uk, lxu@bournemouth.ac.uk

Abstract. Simulation has been widely used as a tool to enhance the manufacturing processes by effectively detecting the errors and performance gaps at an early stage. However, in context of industry 4.0, which involves increased complexity, decisions need to be made more quickly to maintain higher efficiency. In this paper, we use a prediction engine along with a Digital Twin simulation to enhance the decision-making process. We show how, based upon a simulation of a process, a prediction model can be used to determine process parameters based upon desired process outcomes that enhance the manufacturing process. To evaluate our architecture, an industrial case study based on Inventory, Storage and Distribution will be used.

Keywords: Digital Twins, Industry 4.0, Simulation, Federated Simulation, Machine Learning

1 Introduction

Simulation of complex processes is used to enhance the understanding of the working of the overall system as well as enhancing the processes through comprehensive analysis of simulation data. In context of modern manufacturing systems and industrial phase shift towards industry 4.0, the processes have become more complex and difficult to simulate. In a collaborative context, the requirement of independence presents an additional challenge for simulation of the overall process. Traditional simulation approaches (like Discrete Event, Agent Based) and industrial tools (like Simio, AnyLogic) are not capable of simulating these processes in an isolated way.

In making decisions to generate maximum business value there is value in decision support, using analysis techniques especially related to data. A key technique of decision support is through simulation of What-if scenarios which provides the data that can then be used to make decisions related to error detection, process enhancement, system improvements among other things. To generate maximum benefits, accuracy of decisions is significant because inaccurate decisions can lead to severe financial consequences for the business.

When simulating complex processes the simulation involves simulation of multiple parts. Digital twins, as digital representations of physical entities, provide such partial simulation capability. The data from the physical system is shared with the Digital Twin for updated results [8]. Using predictive modelling and Artificial Intelligence, these updated simulations can be transformed into meaningful data that can help in enhancing the decision-making process. However, in the context of collaborative processes (with different types of stakeholders and needs of confidentiality), it is not realistic to expect all digital twin simulation models to be shared with all partners, or even for the simulation models to be based on the same simulation platform.

In this paper, we present an architecture based on the Digital Twin concept which can be used for predictive analysis of processes to enhance decision making. This architecture can simulate a collaborative process involving multiple organizations and then it extracts the relevant and meaningful data. This data is then fed into a prediction engine for predictive analysis and efficient decision making. A secondary manufacturing process is used to evaluate the architecture and depict basic understanding of the working of the proposed architecture.

It is worth noting here that the optimization of simulation or predictive parameters is not the primary objective at this point for the proposed architecture. The goal is to show the working of a Digital Twin based simulation of collaborative processes and how predictive engine can be used to process that simulation data into meaningful decisions. For the remainder of this paper, section 2 provides related work for digital twins and decision support, section 3 introduces the architecture and the core working elements whereas section 4 depicts the evaluation and implementation of the architecture using an industrial case study. Section 5 concludes the paper.

2 Related Work

Simulation has been widely used in manufacturing industry as a tool to analyse, monitor and in some cases, to predict the behaviour of a system or a process in a particular scenario. Simulation has been used in each part of the manufacturing, from design, machining. production to scheduling, planning and supply chain. As the manufacturing industry is evolving, the simulation practices are also improving with the proactive usage of Virtual Reality and Augmented reality to simulate different parts of the manufacturing process [11].

The manufacturing industry have seen a significant improvement due to the integration with information technology. The complexity of modern manufacturing systems has increased many folds during the last decade with the advancements in technologies like Internet of Things (IoTs), Robotics and due to the varying (dynamic) demands of the consumers. COVID-19 pandemic led to a shift in the manufacturing industry when many industries had to make drastic changes to their supply chain and production mechanisms. For example, some cosmetics and chemical production industries started manufacturing sanitizers and related products due to a high increase in demand by making changes to their production lines. This caused companies to make highly informed decisions to remain competitive in a highly uncertain environment. This is where the importance of decision support becomes extremely relevant.

Decision Support Systems (DSS) corresponds to systems which are designed to enhance the decision-making process with the help of integrated and updated data. Simulation has been used to enhance the decision-making process in many ways. For example, simulation is used to design dispatching rules for dynamic scheduling [18], defect prevention in production systems [13], in measuring the time constraints in a semiconductor manufacturing plant [1] to name a few. However, in context of collaborative manufacturing and industry 4.0, traditional simulation approaches in isolation are generally monolithic in nature and hence have inherent limitations.

2.1 Simulation Using Digital Twins

In a complex, heterogeneous process and its simulation, digital twins are a key enabler. This is enabled through providing continuously updated simulation models representing the physical entities. A Digital Twin is essentially an updated simulation of a system and utilizes technologies like Artificial Intelligence to help detect errors at an early stage and provide predictive behavior of the system in a particular scenario. The most widely talked about concept in simulation when it comes to industry 4.0 and collaborative processes is of Digital Twins because of its usage under varying complex simulation scenarios.

Digital Twins based models and simulations present a significantly improved alternative to traditional simulations when it comes to decision support. These models provide enhanced simulations as well as analytical support for business managers to gather useful insights for better decision making. Such an architecture is capable of modelling shop floor assets, data storage and analytical support [17]. The real-time simulations and the data insights extracted from it help enhance the decision-making process.

Predictive and Preventive maintenance provide the solution to the maintenance issues that can lead to serious financial consequences. Predicting the possible errors or anomalies can help in making decisions that can save a lot of problems that otherwise can hinder the execution of time-sensitive and processes of higher priorities. A Digital Twin architecture is used to build dynamic models to implement preventive maintenance operations using updated simulations [12]. The scope, however, is limited to the scheduling and considers only a few indicators like downtime and machine idleness for preventive maintenance.

Supply chain twins [3] are also introduced to monitor the challenges of supply chain during the aftermath of COVID-19 pandemic. The data from customer behavior, online ordering, irregular shipments, and inventory is used to simulate the overall supply chain process. Disruptions such as increased demand, problems in transportation, supplier shutdown problems are introduced to analyze the response and resilience of supply chain twin. However, predictive analytics can be introduced to enhance the decision making for future supply chain disruptions.

Digital Twins are used for decision support in various applications, developed from different perspectives for example in urban agriculture and urban farming production [6], enhancing production systems [16], order management [9] etc. The accuracy of the Digital Twins based simulation remains a challenge because of the lack of approaches to determine the faults and performance of a Digital Twin and in some cases lack of sufficient data that reduces the quality of predictions.

2.2. Predictive Analytics and Digital Twins

Predictive analytics is one of the key areas which can help in realizing true potential of Digital Twins. The data from the twin of physical device or process can help in predicting key parameters that can enhance the overall manufacturing system. One solution is to use hybrid and cognitive twins along with predictive algorithms to improve the production performance and to reduce the manufacturing overheads such as over-heating of machines causing downtime. COGNITWIN [19] is a toolbox consisting of tools for simulation, data acquisition from sensors, predictive analytics and provides a modular approach for integrating hybrid and cognitive twins for process enhancement. The quality of data, computing power required, and the lack of support for collaborative processes are some of the challenges of this approach.

Predictive analytics is also utilized in Maintenance, Repair and Overhaul (MRO) operations of aircraft industry by utilizing data fusion with other operations of Digital Twin ecosystem such as sensors, physical models etc [10]. Assumptions like achieving high Signal to Noise Ratio (SNR) for better data collection and over-reliance on sensory data for predictive decision making deem this approach less reliable in context of collaborative processes.

Machine Learning and data analytic techniques are utilized for error correction [20], scheduling maintenance of production machines, transport systems [15] and preventive, corrective, and predictive maintenance of manufacturing systems [4]. There are approaches where AutomationML is used to create efficient data models to be used as a Digital Twin for enhancement of the production systems. However, challenges such as communication between twins, making efficient decisions based on integrated data of all collaborators still need to be addressed to achieve maximum benefits.

Most of the approaches based on Digital Twins use monolithic simulations and try to address a problem within a part of an organization. However, in context of collaborative processes and industry 4.0, the capabilities of such approaches are limited. Hence, to address this issue, we propose an architecture based on Digital Twins concept in the next section which is capable of simulating collaborative process which can involve multiple stakeholders.

3 Digital Twins Simulation System Architecture

Simulation of complex collaborative processes requires an architecture that not only supports effective simulation, but also allows for the diversity of components and limits to sharing of information. To meet these needs, a federated simulation architecture, as presented in Figure 1, is needed. This architecture, based on the digital twin concept, allows component simulations to be independent. The architecture consists of. of four main parts: Input, Simulation, Results and Data Processing and Prediction Engine.

The inputs consist of initial configurations that are provided to start the simulation. This can be initialization of parameters, a random seed etc. This provides a starting point to the simulation. Then, we have a Digital Twin of a process or a system. As it can be seen in Figure 1, there are multiple simulators within this Digital Twin. This simulation concept is derived from our federated simulation framework [2].

The multiple simulators are part of a federated environment. Each simulator represents a collaborator and is responsible for simulating a part of a collaborative process. The rationale behind using this framework is to make sure maximum confidentiality of each collaborator is ensure while sharing necessary data between the simulators as required and through simulation coordinator. The details of working of each component can be found in our previous work [2].

The simulation coordinator governs the data exchanges between the collaborating simulators. The synchronization of simulators to ensure simulation accuracy and to minimize redundancy is also a responsibility of the simulation coordinator. This collaborative (and federated) environment enables the collaborative partners to use existing simulators while simulating their own designated part of the process efficiently.

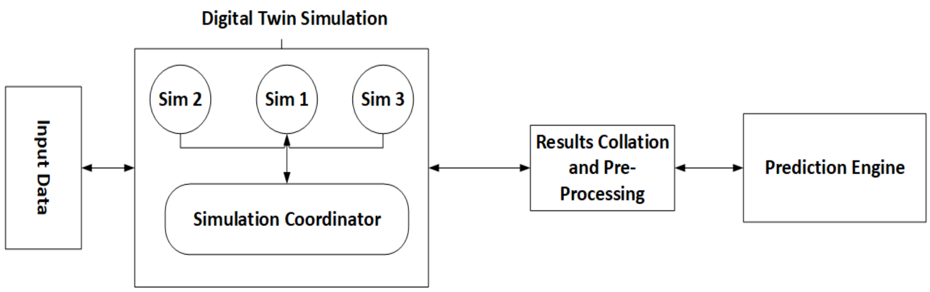

Figure 1 Digital Twins Architecture

Data sharing is an important part of this Digital Twin simulation. There are two mechanisms that are used for the communication between the simulators: 1) messaging and 2) buffer queues. All messages are communicated through the simulation coordinator between the simulators. For data sharing, a buffer queue is created. Concept of producer and consumer is used as depicted in Figure 3. The implementation details are briefly discussed in section 4.

The data during the simulation process is being generated by multiple simulators. It is important to collate them in way that redundancy is avoided, and the integrity of data is also maintained. The results collation module combines the data from the simulators as it is being generated. After data is collated (on the cloud or stored locally in logs and CSV files), it is pre-processed for prediction. Pre-processing involves gathering the data that is necessary and helpful for predictions. For example, the features that are used to predict a certain type of outcome.

The prediction engine deals with the prediction based on the input features and target variables that are selected. The prediction is used to predict the required parameter (output variable). It can be any variable from the simulation. Simulation can provide the data for certain variables in a particular scenario but prediction using Machine Learning for example can predict the value of a variable based on the inputs provided with reasonable accuracy. Hence, prediction is more suitable in this case.

Machine Learning and Deep Learning algorithms can be used according to the requirements of the simulation data. There are certain criteria that can be followed to select appropriate algorithms for example classification, forecasting etc. The data from the prediction can then be fed back for comparisons with the simulated data for further decision making. In terms of accuracy of the prediction model, various parameters can be used for example Means Squared Error or accuracy score etc. Again, this depends on the type of Algorithms that are used and the machine learning approach that is being implemented. The architecture provides flexibility in this.

The proposed architecture provides a process which can result in improvement of decision-making through comprehensive simulation, effective prediction and using a reliable data. The simulation in the architecture is also scalable to accommodate collaborative processes which is essential in context of industry 4.0.

4 Evaluation of the Proposed Architecture

To evaluate the proposed architecture, an industrial case study from the literature [5] is used (see Figure 2). There are some minor changes made to the implementation of the case study, for example instead of a monolithic simulation, different parts of secondary manufacturing are federated into multiple simulators.

The secondary manufacturing (storage, packaging, and distribution) process of medicinal products is simulated. There are three main phases of the process, simulated using 3 different simulators depicting the collaboration. The first phase is *Receipt and Inventory*, then comes *Storage and Monitoring* and finally *Distribution* is done in the last phase. There are 3 types of staff resources, including 18 technicians (G1), 2 service (G2) and 12 Quality Control (G3). In the first phase, different types of materials like cryoproducts and shippers (containers to store products) arrive and are received by the service staff.

Some of shippers are returned due to the bad quality. The remaining shippers are then stored in the pre-storage in phase 2. The shippers are the documented and their data loggers are created. After the documentation phase is completed, the inventory is updated at the last step of Storage and Monitoring phase. In the final phase, the orders that are placed are then sorted out in the order planning and then after the quality check, they are ready to be dispatched.

We applied our digital twins based architecture (Figure 1) to the manufacturing case. The implementation is done in Python programming language using a Discrete Event based simulation. Three simulators were implemented, simulating 3 phases as part of a federation concept adopted from our previous work [2]. The data sharing mechanism between the simulators is implemented using buffer queues as shown in Figure 3.

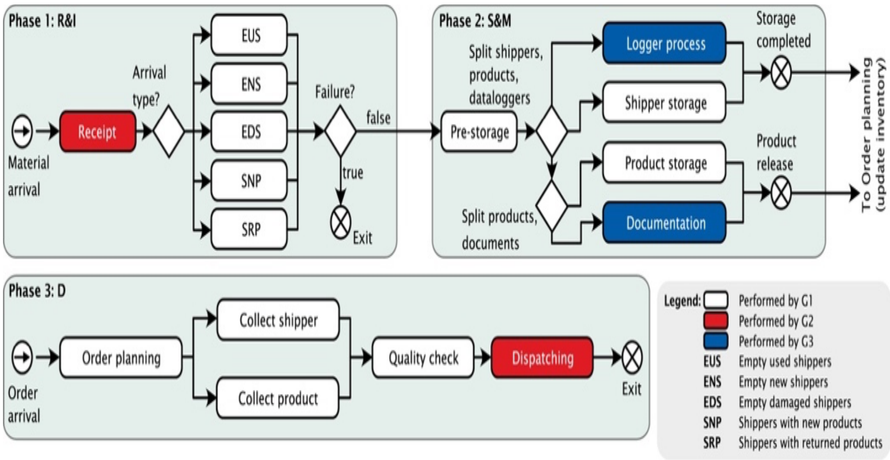

Figure 2 Secondary Manufacturing Case Study [5]

Producer is one simulator, and the consumer is another simulator which want to use the data that can be accessed from the queue. The buffer queue is implemented using the concept of First In First Out (FIFO). At the quality check process, we implemented a soft check to simulate some failures and some successes of the quality checking processes based on the quality control staff performance (quantifiable from 1-5 with 1 being lowest and 5 being the best). The simulation was run 50,000 times to generate enough data for the predictions and to expose the variation in the performance.

After gathering data from the simulation runs in form of log files (step by step execution and messaging between the simulators) and organized feature extraction in CSV files, a prediction engine is prepared. The prediction is developed to predict the number of orders that pass the quality check. Firstly, feature extraction process is done and the input features that are selected include Arrival Rate, Number of Orders Processed, Performance of the Quality Control Staff.

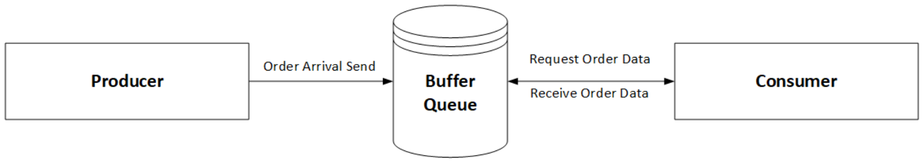

Figure 3 Data Sharing Between Simulators

The output or predicted feature is the number of orders that pass the quality check. Two different models (Logistic Regression [7] and Light Gradient Boosting Machine (LGBM) [14]) are used to predict the outcome. The data is divided into training (80 percent) and testing (20 percent). After applying the models on training and testing data, the accuracy of models is evaluated using Root Mean Square Error (RMSE) and R^2 (which means how well the regression model is fit to the data). The results of the evaluation are shown in Table 1:

The results in Table 1 show that accuracy in LGBM is improved as compared to Logistic Regression. This signifies the importance of model selection based on the available data. MSE is also minimized in LGBM as compared to Logistic Regression. The accuracy in LGBM is 67 percent which is also on the lower side which can be enhanced by improving feature selection and model tuning which will be continued in the future work. The idea behind this implementation is to show how Digital Twins concept can be efficiently utilized to enhance the decision-making process by incorporating a prediction engine. This experimentation shows that the prediction is feasible but there is a room for improvement as this is a preliminary attempt.

Table 1 Prediction Accuracy Results

Algorithm	**Training Accuracy**	**Test Accuracy**	**Mean Squared Error**	**R^2**
Logistic Regression	0.49	0.50	0.67	0.50
LGBM	0.66	0.67	0.45	0.66

The predictions gathered from the above implementation can be used to make decisions like how many products pass through the quality check and what are the other features of the product that could be responsible for failing a quality check. This implementation is preliminary but provides a basis for enhancement and core working of the proposed architecture.

5 Conclusion and Future Challenges

Simulation is an important ingredient that can help enhance the decision-making process in manufacturing industry. In context of industry 4.0 and collaborative processes, simulation can play a vital role in generating insights regarding performance of each part of the process. In this paper, we proposed a Digital Twins based architecture to enhance the decision support in context of industry 4.0 and collaborative processes. We provided a blueprint to use simulation along with machine learning to predict certain parameters based on the required objectives to enhance the decision making for improved manufacturing processes.

We used Digital Twins based simulation in a federated environment to simulate a secondary manufacturing process for the evaluation of proposed architecture. Using machine learning prediction algorithms and simulation data as input, the desired parameters were predicted. The accuracy of the models, however, can be enhanced by utilizing more correlated features and better parameters tuning. More complex case studies will be utilized to further enhance the working of the proposed architecture.

Acknowledgments. This research is partially funded by the State Key Research and Development Program of China (2017YFE0118700) and it is part of the FIRST project which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 734599

References

1. Anthouard, B., Borodin, V., Dauzere-Peres, S., Christ, Q., Roussel, R.: Management of time constraints tunnels in semiconductor manufacturing: A decision support system. In: 23ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (2022)
2. Arshad, R., Vrieze, P.T.d., Xu, L.: A federated simulation framework for cross-organisational processes. In: Working Conference on Virtual Enterprises. pp. 267–279. Springer (2021)
3. Burgos, D., Ivanov, D.: Food retail supply chain resilience and the covid-19 pandemic: A digital twin-based impact analysis and improvement directions. Transportation Research Part E: Logistics and Transportation Review 152, 102412 (2021)
4. Eirinakis, P., Kalaboukas, K., Lounis, S., Mourtos, I., Rožanec, J.M., Stojanovic, N., Zois, G.: Enhancing cognition for digital twins. In: 2020 IEEE International Conference on Engineering Technology and Innovation (ICE/ITMC). pp. 1–7. IEEE (2020)
5. Erkoyuncu, J.A., Farsi, M., Ariansyah, D., et al.: An intelligent agent-based architecture for resilient digital twins in manufacturing. CIRP annals 70(1), 349–352 (2021)
6. Ghandar, A., Ahmed, A., Zulfiqar, S., Hua, Z., Hanai, M., Theodoropoulos, G.: A decision support system for urban agriculture using digital twin: A case study with aquaponics. IEEE Access 9, 35691–35708 (2021)
7. Gladence, L.M., Karthi, M., Anu, V.M.: A statistical comparison of logistic regression and different bayes classification methods for machine learning. ARPN Journal of Engineering and Applied Sciences 10(14), 5947–5953 (2015)
8. Hou, L.,Wu, S., Zhang, G., Tan, Y.,Wang, X.: Literature review of digital twins applications in construction workforce safety. Applied Sciences 11(1), 339 (2020)
9. Kunath, M., Winkler, H.: Integrating the digital twin of the manufacturing system into a decision support system for improving the order management process. Procedia Cirp 72, 225–231 (2018)
10. Liu, Z., Meyendorf, N., Mrad, N.: The role of data fusion in predictive maintenance using

digital twin. In: AIP conference proceedings. vol. 1949, p. 020023. AIP Publishing LLC (2018)
11. Mourtzis, D.: Simulation in the design and operation of manufacturing systems: state of the art and new trends. International Journal of Production Research 58(7), 1927–1949 (2020)
12. Neto, A.A., Carrijo, B.S., Brock, J.G.R., Deschamps, F., de Lima, E.P.: Digital twin-driven decision support system for opportunistic preventive maintenance scheduling in manufacturing.Procedia Manufacturing 55, 439–446 (2021)
13. Psarommatis, F., Kiritsis, D.: A hybrid decision support system for automating decision making in the event of defects in the era of zero defect manufacturing. Journal of Industrial Information Integration 26, 100263 (2022)
14. Rai, M., Mandoria, H.L.: Network intrusion detection: A comparative study using state-of the-art machine learning methods. In: 2019 International Conference on Issues and Challenge in Intelligent Computing Techniques (ICICT). vol. 1, pp. 1–5. IEEE (2019)
15. Rudskoy, A., Ilin, I., Prokhorov, A.: Digital twins in the intelligent transport systems. Transportatio Research Procedia 54, 927–935 (2021)
16. dos Santos, C.H., Montevechi, J.A.B., de Queiroz, J.A., de Carvalho Miranda, R., Leal, F.: Decision support in productive processes through des and abs in the digital twin era: a systematic literature review. International Journal of Production Research pp. 1–20 (2021)
17. Tao, F., Zhang, M.: Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. Ieee Access 5, 20418–20427 (2017)
18. Turker, A.K., Aktepe, A., Inal, A.F., Ersoz, O.O., Das, G.S., Birgoren, B.: A decision support system for dynamic job-shop scheduling using real-time data with simulation. Mathematics 7(3), 278 (2019)
19. Unal, P., Albayrak, Ö., Jomâa, M., Berre, A.J.: Data-driven artificial intelligence and predictive analytics for the maintenance of industrial machinery with hybrid and cognitive digital twins. In: Technologies and Applications for Big Data Value, pp. 299–319. Springer (2022)
20. Vathoopan, M., Johny, M., Zoitl, A., Knoll, A.: Modular fault ascription and corrective maintenance using a digital twin. IFAC-PapersOnLine 51(11), 1041–1046 (2018)