FengHuoLun: A Federated Learning based Edge Computing Platform for Cyber-Physical Systems

Chong Zhang
School of Information Technology
Deakin University
Geelong, Australia
zhangcho@deakin.edu.au

Xiao Liu
School of Information Technology
Deakin University
Geelong, Australia
xiao.liu@deakin.edu.au

Xi Zheng
Department of Computing Macquarie University
Sydney, Australia
james.zheng@mq.edu.au

Rui Li
School of Information Technology
Deakin University
Geelong, Australia
rui@deakin.edu.au

Huai Liu
Department of Computer Science and Software Engineering
Swinburne University of Technology
Melbourne, Australia
hliu@swin.edu.au

***Abstract*— Cyber-Physical Systems (CPS) such as intelligent connected vehicles, smart farming and smart logistics are constantly generating tons of data and requiring real-time data processing capabilities. Therefore, Edge Computing which provisions computing resources close to the End Devices from the network edge is becoming the ideal platform for CPS. However, it also brings many issues and one of the most prominent challenges is how to ensure the development of trustworthy smart services given the dynamic and distributed nature of Edge Computing. To tackle this challenge, this paper proposes a novel Federated Learning based Edge Computing platform for CPS, named "FengHuoLun". Specifically, based on FengHuoLun, we can: 1) implement smart services where machine learning models are trained in a trusted Federated Learning framework; 2) assure the trustworthiness of smart services where CPS behaviours are tested and monitored using the Federated Learning framework. As a work in progress, we have presented an overview of the FengHuoLun platform and also some preliminary studies on its key components, and finally discussed some important future research directions.**

Keywords*— *Federated Learning, Edge Computing, Cyber-Physical Systems, Trustworthy, Microservices

I. INTRODUCTION AND RELATED WORK

With the rapid growth of IoT (Internet of Things) and smart services, many CPS (Cyber-Physical Systems) such as intelligent connected vehicles, smart farming and smart logistics are being developed in the market [1]. In the meantime, billions of mobile phones and IoT devices are constantly generating tons of data at the network edge. Conventional Cloud Computing architecture cannot keep up with the increasing demand of massive data processing and the requirement of real-time response. As a result, Edge Computing (also known as Fog Computing) is emerging the next generation IT infrastructure for smart services. Edge Computing provisions computing resources close to the End Devices from the network edge and it can fully utilise three layers of computing resources including End Devices, Edge Nodes and Cloud Servers. Therefore, Edge Computing becomes the ideal platform for CPS [2].

Typical application domains for Edge Computing based CPS include smart vehicles, smart farming and smart logistics. Specifically, there are many on-board smart services equipped for autonomous driving vehicles, for example, voice and speech recognition, gesture controls, eye tracking, mapping and safety systems, and so on. Meanwhile, Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) communications further extends the capability of individual vehicles through data exchange with Roadside Units (RUs) and nearby vehicles to build large-scale and collaborative smart services such as traffic control, collision prevention, and smart parking [3]. There are many smart farming applications [4] such as Precision Farming, Precision Livestock Farming, Agricultural Drones, Internet of Food and Farm, and so on. Typical smart technologies used in smart farming include, for example, sensing, positioning, driverless tractors and drones, robotics and automation, automatic watering and irrigation, big data analytics, and so on. Smart logistics makes supply chains more effective and efficient through connected devices and data analytics in the supply chain [5]. Huawei intelligent logistics solutions [6] are based on Huawei OceanConnect IoT platform, NB-IoT and RFID technologies. The whole process of cargo transportation is managed in a visual way, and the monitoring of cold chain transportation improves safety and quality. Currently, most giant logistics companies have been conducting pioneering projects on UAV (Unmanned Aerial Vehicle) especially smart drones to solve the last mile delivery problem including DHL and Amazon.

It can be easily seen from the above that Edge Computing can play a significant role in many CPS. However, it also brings many issues and one of the most prominent challenges is how to develop trustworthy smart services with a large number of distributed resource-constrained End Devices in Edge Computing. Specifically, there are two major issues need to be solved:

1) How to implement distributed learning models. Machine learning models are the core for most smart services. In Edge Computing, the data required for training the models are often distributed among many resource-constrained End Devices. Therefore, instead of centralised learning, a distributed learning framework is required where End Devises can be coordinated and locally learned knowledge can be aggregated to form a global view.

2) How to ensure the trustworthiness of smart services. Trustworthiness is the essential requirement for any CPS. However, given the complex and distributed nature of both Edge Computing and CPS components, testing the trustworthiness of smart services is a challenging issue. Clearly, instead of a centralised testing framework, a distributed testing framework for monitoring the runtime behaviours of smart services is required.

Currently, there are a few Edge Computing development platforms/SDKs released by some public Cloud service providers to support the development of smart services, but they are usually designed in a top-down fashion and their key motive is to leverage their business Cloud services. For example, Microsoft Azure IoT Edge [7] allows developers to employ Azure AI services and deploy their own codes on the Edge Nodes. Specifically, the IoT Edge runtime runs at the Edge Nodes and manages the communication between the End Devices, the Edge Nodes and the Cloud Servers. An IoT Edge simulator is also provided for debugging modules running at the Edge. Huawei Intelligent EdgeFabric (IEF) [8] focuses on the management of Edge Nodes and the push-down of Huawei Cloud services in FunctionGraph to the Edge. KubeEdge [9] is an open source system built upon Kubernetes to extend native containerised application orchestration capabilities to the Edge.

Unfortunately, existing Edge Computing development platforms cannot effectively solve the two issues mentioned above. In this paper, we propose the idea of using Federated Learning as a solution. Federated Learning is a latest technique for machine learning that can train a global model across multiple decentralized end devices holding local data samples without exchanging the data. Therefore, strong data security and privacy protection can be achieved by design [10]. Edge Computing systems can adopt the Federated Learning technique to achieve better learning results for smart services while protecting data security and privacy. For example, Google applied FL to the Gboard data stored on hundreds of millions of devices to collaboratively learn a shared prediction model for people's interactions with mobile devices. Federated Learning has also been intensively applied in the financial industry for such as anti-money laundering, risk modelling and insurance pricing. Some open-source Federated Learning frameworks have been released such as TensorFlow Federated [11] and PySyft [12].

Given the same distributed nature of Federated Learning and Edge Computing, to adopt the Federated Learning framework to achieve better learning results and data protection for Edge Computing based CPS is a very promising direction. In this paper, we propose the idea of "FengHuoLun", a Federated Learning based Edge Computing platform for CPS. The platform overview and some preliminary results for its key components are presented later.

II. FEDERATED LEARNING BASED EDGE COMPUTING PLATFORM FOR CYBER-PHYSICAL SYSTEMS

A. Platform Overview

Figure 1 presents the overview of the FengHuoLun Platform. The design of FengHuoLun considers three different views from macro to micro including Global view, Edge view and Entity view. Specifically, Global view represents the overview of the whole CPS and the business requirements of stakeholders; Edge view implements the business requirements with smart service and business processes; Entity view defines the physical objects including the end devices, computing nodes and networks. As shown in the middle of the figure, FengHuoLun adopts the three-layer system architecture for Edge Computing which includes the End Devices layer (consists of such as smart phones, vehicles, drones and other smart devices) at the bottom, the Edge Node layer (consists of a large number of computing nodes at the network edge) at the middle, and the Cloud Server layer (consists of unlimited resources with powerful computation and massive storage) at the top. The End Devices are physically close and connected to the Edge Nodes via fast wireless connections such as the 5G network, while the Edge Nodes are often far away from but connected with the Cloud Servers via broadband. Simulators for resources at different layers are also available in FengHuoLun to support the software development and testing. As shown in the right of the figure, Federated Learning has been deployed at all three layers. Specifically, many smart services are implemented at the End Device layer and Edge Node layer where End Devices are training their local model with their own data and then aggregated by the central server at the Edge Node layer to construct the global model. Similarly, some smart services including our testing utilities can be deployed at the Edge Node layer and the Cloud Server layer. More examples are presented below.

B. Federated Learning based Smart Services

FengHuoLun can be used in many large-scale CPS. For instance, in smart agriculture applications, each plant can be equipped with a sensor box (e.g., with humidity and temperature sensors) and light-weight pre-trained models can be deployed locally to detect anomaly (Entity view) efficiently but with some acceptable error range. According to the idea of a Federated Learning framework, all the plants within a specific region (e.g., inside 2km range) are equipped with an Edge Node with pre-trained regional models along with pre-trained local models to be downloaded for each plant's end node within its service range. The regional model is more complex than the local models (and slower in giving prediction results) but is able to provide a more comprehensive view (Edge view) over the plants within its service range (e.g., anomaly detection). The global model is deployed at the Cloud Server, which is as complex as required to handle vast amount of data generated by all plants covered by the system. The global model is able to give an overarching view (Global view) of the entire plant farm. FengHuoLun also aims to play an important role in highly mobile large-scale CPS such as intelligent connected vehicles where vehicles can share important context-sensitive information among themselves. FengHuoLun can be used to learn the context information for each vehicle (Entity view), learn the context information for each road side unit (Edge view), and the large area traffic information (Global view) for the city so that the best informed decision can be made for each connected vehicle to guarantee safety and optimal usage of the road infrastructure.

C. Federated Learning based Testing Utilities

In FengHuoLun, we are particularly focused on the trustworthiness and robustness of smart services and the ecosystem they compose. More specifically, a variety of testing utilities will be applied to validate and verify the services and the whole system against expected behaviours from different views. Metamorphic Testing (MT), for example, can be used to identify necessary properties of system behaviours, represented by the so-called metamorphic relations among multiple system inputs and associated outputs. Note that due to different views of our FengHuoLun platform and the distributed nature of Federated Learning, the metamorphic relations may be presented in a hierarchical structure.

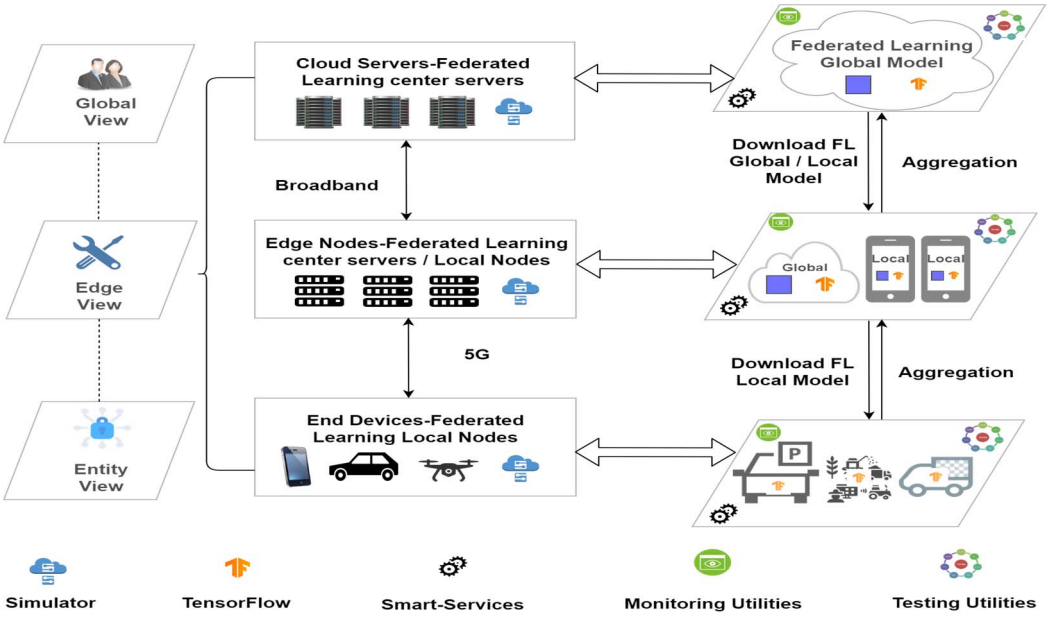

Figure 1: Overview of the FengHuoLun Platform

On every Edge Node, a set of metamorphic relations can be constructed (Edge view), which are particularly appropriate for the testing and monitoring of the local models used in the smart services (Entity view). Hereby, the major purpose is to guarantee that the local services and models can deliver the expected behaviours (in other words, to assure the trustworthiness of these models and corresponding services). The relations on different models do not need to be strongly consistent. An "eventual consistency" will be pursued one a higher level: A unified family of more general metamorphic relations will be constructed and used for the testing and monitoring of the global model deployed in the central cloud server (Global view). Analogous to the aggregation of local models to the global model, the metamorphic relations for each different local model will be aggregated and generalized to construct some global-level relations such that different aspects of the whole system can be combined to represent the overall functionalities. The globalized relations are particularly useful to validate whether and to what extent the whole ecosystem can achieve the expected system behaviours under a wide variety of scenarios (in other words, to check the robustness of the system). For illustration, let us look at autonomous driving as the concrete example. Different models on local servers may be focused on distinct perspectives of driving behaviours, such as speeding and steering angles. Correspondingly, the metamorphic testing for these local models could use the relations about whether and to what extent the change in driving condition would cause the driverless vehicle to increase/reduce speed and take a turn with a certain degree, respectively. Those relations should then be combined to compose general MRs for the global model, to justify the correctness of the overall behaviours of the system.

III. PRELIMINARY RESULTS

In this section, we present some preliminary results for the key components of FengHuoLun, including Microservice which is the major software architecture for Edge Computing based CPS, the simulation tool for an Edge Computing environment which is a key component for the CPS development and testing, and anomaly detection algorithm for faulty sensor nodes which is a key component for monitoring CPS system behaviours.

A. Testing and Deployment of Microservices

Microservice is the default software architecture for Edge Computing based systems [7, 9]. However, due to unique characteristics of microservices as well as various machine learning models attached to them, it is extremely difficult, if not impossible, to verify test results given any possible input to a certain microservice. This so-called oracle problem is a fundamental problem in software testing and verification. MT has been demonstrated as a simple yet effective approach to the oracle problem in the community of software testing. It has helped detect different types of real-life defects in various application domains. In our preliminary study [13], Metamorphic Testing was applied into the verification of microservices. An experimental study was conducted based on three typical microservices, for which seven simple metamorphic relations (MRs) were constructed. Despite the simplicity of MRs, MT was justified to have a very high effectiveness in detecting various faults without the need of complete oracles. As discussed in Section II, MT can be integrated with Federated Learning to play a significant role in FengHuoLun.

B. Simulation Tool for Edge Computing

Due to the complicated environment of Edge Computing, simulation tools are often required during the development and testing of CPS before deployment in the real-world environment. In our previous work [14], we have created an open-source simulation tool to evaluate the runtime performance of Edge/Fog Computing systems. The proposed simulation tool is an efficient and extensible toolkit for automatically evaluating resource and task management strategies in Edge Computing with simulated user applications defined in DAG (Directed Acyclic Graph), namely workflows. The simulation tool is able to: 1) automatically set up a simulated Edge Computing environment; 2) automatically simulate the execution of user

submitted applications; and 3) automatically evaluate and compare the performance of different computation offloading and task scheduling strategies with three basic performance metrics, including time, energy and cost. As shown in Figure 1, we are planning to extend the simulation tool to include Federated Learning models and integrate into FengHuoLun to simulate objects at different layers.

C. *Anomaly Detection in Complex WSNs*

Anomaly detection is a very important yet challenging task in Edge Computing based CPS. In our previous work [15], we have investigated dictionary learning based on a non-negative constraint to detect anomaly nodes in wireless sensor networks with sparse representation. Through experiment on a specific thermal power plant in China, we evaluate our proposed method in detecting abnormal nodes against four state of the art approaches and prove the method is more robust. Based on this work, we can extend this anomaly detection algorithm into the monitoring utilities for FengHuoLun where the monitoring will be instead based on Federated Learning and distributed not only across sensor nodes (as in our previous work) but also across three different layers with different views. The monitoring utilities are themselves Deep Learning models created and deployed with Federated Learning framework. The monitoring models on entity view are able to detect anomaly promptly but with limited layers and acceptable accuracy. The monitoring models on Edge and Global view are getting more complex with improved accuracy but increased latency. Such hierarchical monitoring architecture enables FengHuoLun to detect anomaly at runtime for various CPS with different requirement for latency and accuracy.

IV. CONCLUSIONS AND FUTURE WORK

Given its unique advantages, Edge Computing is becoming the mainstream architecture for developing Cyber-Physical systems. As most CPS are time- and safety-critical, how to ensure the development of trustworthy Edge Computing based smart services is an essential challenge. In this paper, we propose the idea of a Federated Learning based Edge Computing platform for CPS, named "FengHuoLun". Specifically, based on FengHuoLun, we can implement smart services, testing utilities, and runtime monitoring utilities using the Federated Learning framework to ensure the trustworthiness of CPS. The design of FengHuoLun and some of the preliminary results for its key components have been presented in this paper.

The next step is to complete the development of FengHuoLun and evaluate its effectiveness in the development of real-world CPS. There are few major research directions that we aim to investigate in the future based on FengHuoLun:

1) The extension of the preliminary work [13] on Metamorphic Testing to the more general test-driven software development to extend the support to software developers.

2) The extension of the preliminary work [14] on Edge simulation to include the simulation of Federated Learning models.

3) The extension of the preliminary work [15] for the detection of anomaly sensors at one layer to the hierarchical runtime anomaly detection at different layers for the underlying CPS.

REFERENCES

[1] A. Ahmada, A. Paula, M. Rathore and H. Chang. Smart cyber society: Integration of capillary devices with high usability based on cyber–physical system. Future Generation Computer Systems, vol. 56, 493-503, 2016.

[2] B. Bordel, R. Alcarria, T. Robles and D. Martin. Cyber–physical systems: Extending pervasive sensing from control theory to the Internet of Things. Pervasive and mobile computing, vol. 40, 156-184, 2017.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[3] J. Zhao, Y. Chen and Y. Gong. Study of connectivity probability of vehicle-to-vehicle and vehicle-to-infrastructure communication systems. Proc. IEEE 83rd Vehicular Technology Conference, Nanjing, China, 2016.

[4] Sciforce. Smart Farming, or the Future of Agriculture, Available: https://medium.com/sciforce/smart-farming-or-the-future-of-agriculture-359f0089df69. Last Accessed on 01/12/2019.

[5] L. Barreto, A. Amaral, and T. Pereira. Industry 4.0 implications in logistics: an overview. Procedia Manufacturing, vol. 13, 1245-1252, 2017.

[6] Huawei. Making Connections with smart logistics - Huawei Industry Insights, Available: https://www.huawei.com/en/industry-insights/digital-huawei/cases/smart-logistics. Last Accessed on 01/12/2019.

[7] IoT Edge, https://azure.microsoft.com/en-au/services/iot-edge/. Last Accessed on 01/12/2019.

[8] Intelligent EdgeFabric, https://support.huaweicloud.com/ief/index.html. Last Accessed on 01/12/2019.

[9] KubeEdge, https://kubeedge.io/en/. Last Accessed on 01/12/2019.

[10] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konecny, S. Mazzocchi and H. B. McMahan, Towards federated learning at scale: System design, arXiv preprint, arXiv:1902.01046, 2019.

[11] TensorFlow Federated, https://www.tensorflow.org/federated. Last Accessed on 01/12/2019.

[12] PySyft, https://github.com/OpenMined/PySyft. Last Accessed on 01/12/2019.

[13] G. Luo, X. Z., H. Liu, R. Xu, D. Nagumothu, R. Janapareddi, E. Zhuang and X. Liu. Verification of microservices using metamorphic testing. Proc. 19th International Conference on Algorithms and Architectures for Parallel Processing, Melbourne, Australia, 2019.

[14] X. Liu, L. F., J. Xu, X. Li, L. Gong, J. Grundy, Y. Yang. FogWorkflowSim: An automated simulation toolkit for workflow performance evaluation in fog computing. Proc. 34th IEEE/ACM International Conference on Automated Software Engineering, San Diego, California, United States, 2019.

[15] X. Li, G. Xu, X. Zheng, K. Liang, E. Panaousis, T. Li, W. Wang and C. Shen. Using sparse representation to detect anomalies in complex WSNs, ACM Transactions on Intelligent Systems and Technology, 10(6), Article No. 64, December 2019.