Effect-Driven Selection of Web of Things Services in Cyber-Physical Systems Using Reinforcement Learning

KyeongDeok Baek[✉] and In-Young Ko

School of Computing, Korea Advanced Institute of Science and Technology,
Daejeon, Republic of Korea
{kyeongdeok.baek,iko}@kaist.ac.kr

Abstract. Recently, Web of Things (WoT) expands its boundary to
Cyber-physical Systems (CPS) that actuate or sense physical environ-
ments. However, there is no quantitative metric to measure the quality
of physical effects generated by WoT services. Furthermore, there is no
dynamic service selection algorithm that can be used to replace services
with alternative ones to manage the quality of service provisioning. In
this work, we study how to measure the effectiveness of delivering vari-
ous types of WoT service effects to users, and develop a dynamic service
handover algorithm using reinforcement learning to ensure the consistent
provision of WoT services under dynamically changing conditions due to
user mobility and changing availability of WoT media to deliver ser-
vice effects. The preliminary results show that the simple distance-based
metric is insufficient to select appropriate WoT services in terms of the
effectiveness of delivering service effects to users, and the reinforcement-
learning-based algorithm performs well with learning the optimal selec-
tion policy from simulated experiences in WoT environments.

Keywords: Service effectiveness ·
Effect-driven WoT service selection · Reinforcement learning ·
Cyber-physical System

1 Introduction

Cyber-physical systems (CPS) are the systems in which computational resources
lie on abstract cyberspace and physical devices lie on physical spaces are con-
nected and coordinated with each other to provide complex services that are nec-
essary to accomplish users' goals [5]. Already there are many types of CPS that
have been deployed in our urban environments such as smart homes, vehicle-
to-everything (V2X), and smart factories. In particular, CPS has become an
important part of Web of Things (WoT) because their key components are con-
nected with each other via the Web, and it is essential to effectively find, access
and utilize physical WoT resources that are necessary to accomplish users' goals.

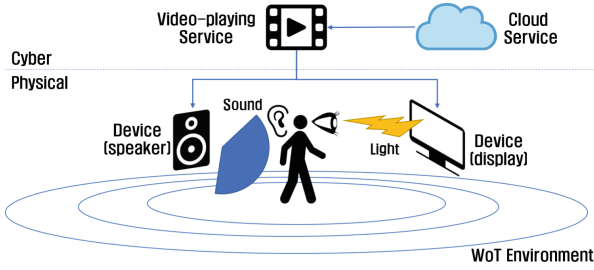
I.-Y. Ko—Ph.D. Supervisor.

Fig. 1. WoT environments based on Cyber-physical Systems

Figure 1 shows an example of CPS-based WoT environment, that is divided in two layers, namely, cyber and physical layer, where traditional Web and WoT services are lied on the cyber layer, and physical devices and users are lied on the physical layer. Via actuating devices such as displays and speakers, a video-playing service in the cyber layer can deliver video contents to users by generating light and sound effects to the physical layer. Obviously, it is necessary to define the metrics to measure the quality of delivering service effects by WoT services to support users accomplishing their goals by providing services in required quality. Moreover, service selection problem, which is to select the most appropriate services among available candidates, becomes more challenging in WoT environments because of its physical-aware and highly dynamic nature.

In this work, we identify the essential characteristic of CPS that need to be considered to make WoT services to effectively interact with physical environments and human users while generating or sensing physical effects such as lights and sounds via physical media that are deployed over the physical environments. Especially, there are effect-generating services that produce and deliver physical effects to users, such as news-delivery and music-playing services as shown in Fig. 1. The quality of such effect-generating services affects users' satisfaction, so the selection of services should be done in a user-centric manner by evaluating how well the generated effects are delivered to the user. However, existing works on Web service selection only considers network-level quality of services (QoS) attributes, such as latency that affect the general quality perceived by users, but cannot reflect the quality of physical effects of the effect-generating services.

2 Research Issues

2.1 Service Effectiveness

Figure 2 shows an example categorization of WoT services, where solid boxes indicate categories and dashed boxes indicate an example service for each category. Most of the WoT services that interact with physical environments can be categorized as actuating or sensing services. In this work, we mainly focus on actuating services because actuating services can contribute directly to the accomplishment of users' goals by generating effects in physical environments,
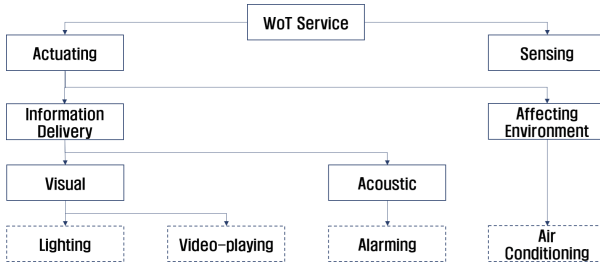
Fig. 2. Example categorization of WoT services

while the role of sensing services is simply about collecting information. Obviously, the effectiveness of such actuating services need to be evaluated differently according to their physical effects. However, to the best of our knowledge, there is no quantitative measure proposed to evaluate the quality of physical effects generated by WoT services, which we call *service effectiveness*. Therefore, it is necessary to model a specific effectiveness metric for each type of physical effects.

In addition, the physical effects generated by the actuating services may cause constructive or destructive interference when there are more than one effect generated in the same space. Moreover, in users' perspective, there can be service-level interference. For instance, the effectiveness of a movie-playing service increases if the associated display and speaker devices are located cohesively to each other in a space [1]. Another example is that if there is a service that generates bright illumination, it may cause glare and degrades the user's satisfaction on watching movies. Although there are some work done on analyzing the correlations among QoS attributes [4], there have been no efforts on modeling and measuring service-level interference in terms of delivering physical effects.

2.2   Predictive Service Selection and Dynamic Handover

Service provisioning in CPS environments needs to be done usually for a long time, and therefore, it is essential to ensure the required quality of services for a user task for a long period of time in a continuous and consistent manner. However, most of the existing dynamic service selection algorithms consider the quality of the candidate services at the time when they choose the services rather than considering the future quality of the services [11]. Especially in dynamic CPS environments, we cannot assume that the quality of a service that is monitored at a time when the service is selected will be remained the same throughout the service provisioning period. For instance, while a graphical content is shown to a user by using a nearby display device, if the user moves far from the device or the display suddenly blacked out, the content cannot be perceived by the user effectively anymore.

To deal with the above problem, we have identified two research directions. First, to maintain a certain level of service quality during a service provisioning

period, dynamic service selection needs to be done in an iterative manner to replace some of the services that show degradation of their quality with alternative ones. We call this process as *dynamic service handover* [1,2]. Second, service selection should be done in a predictive manner, so that not only considering the current quality of services but also we can predict the future quality of services and make the service provisioning more stable. By performing predictive service selection, the number of handovers, which may cause service-migration overheads and service interruptions, can be minimized.

3   Previous Works

3.1   Service Effectiveness

In our previous works, we considered physical locations of mobile users and devices, and selected services that are located in a spatially cohesive manner centered by the user [1,2]. We defined a metric named *spatio-cohesiveness* to measure how the user and the selected services are located cohesively in terms of the devices associated with the services. However, one limitation of this method is that the services that are located cohesively cannot guarantee the effectiveness of delivering physical effects to users and improve the perceived service quality. As a counterexample, if we consider only spatio-cohesiveness, the service selection algorithm selects services based on the Euclidean distance between available candidates and the user, so the WoT devices that are associated with the selected services may be located behind a wall, and the user cannot perceive the effects that are generated by the devices.

In our on-going work, we define a rule-based model of visual service effectiveness, which evaluates whether the generated content can be perceived successfully by a user or not. The model was designed based on domain knowledge of the human vision system and simple physics of light, and contains three constraints. First, if the device is to far from the user, then the effectiveness is zero because the user cannot recognize the content correctly. Second, if the device is not in the Field of View (FoV) of the user, then effectiveness is zero because the user cannot perceive the light from the device at all. Third, if the device is not facing the user, then effectiveness is zero because the user would only see the back of the device. Finally, service effectiveness is 1 if all constraints are passed.

3.2   Predictive Service Selection and Dynamic Handover

In our previous works, we adopted a reinforcement learning algorithm to effectively select and dynamically handover services in a predictive manner [2]. Specifically, we developed a service selection agent that makes decisions of selecting services and trained the agent by using a reinforcement learning algorithm in a simulated WoT environments. We found that the agent could learn the optimal policy of selecting services in terms of spatio-cohesiveness. Our service selection agent is designed based on the Actor-Critic algorithm [7], Deep-Q Network (DQN) [10], and Deep Reinforcement Relevance Network (DRRN) [6].
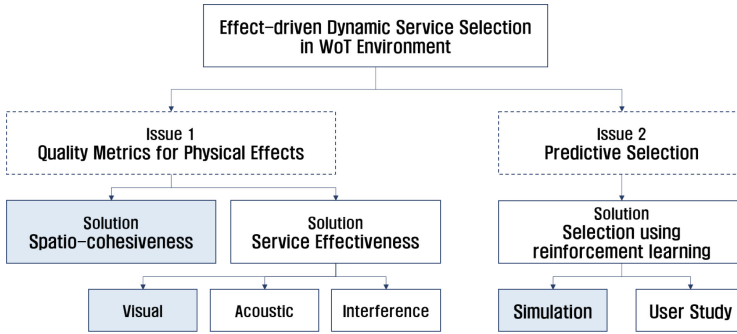
Fig. 3. Research road map

4   Research Plans

Figure 3 shows the research road map of this work, and the shaded boxes indicate the research issues that have been dealt in our previous works.

4.1   Service Effectiveness

Type-Specific Service Effectiveness Model. We have studied only the visual service effects, and we plan to investigate the ways of measuring the effectiveness of delivering acoustic effects. Furthermore, our current model of visual service effectiveness is a simple rule-based model, so we plan to evaluate and improve the practicality of the model by performing user-studies.

Service Interference. We plan to analyze service-level interference among the services that generate similar or different types of physical effects, and develop a service selection algorithm to choose cooperating services that have constructive interference and avoid destructive interference.

4.2   Predictive Service Selection and Dynamic Handover

Ideally, the training of our service selection agent should be done in real-world WoT environments, but we performed the training in simulated WoT environments. Training in real-world environments is known to be a challenging problem for reinforcement learning researchers because collecting real-world samples costs too much and difficult to make the agent experience the world in an iterative manner. We have two research directions regarding to this issue.

Virtual Reality-Powered User Study. First, we will perform user-studies in virtual WoT environments using Virtual Reality (VR) technologies. In some recent works, VR technologies are used to mimic psychological experiments through Web-based crowd sourcing platforms [9], and to let users experience

elderly peoples' sight by reducing visual acuity virtually [8]. Currently, we are implementing virtual WoT environments using VR technologies to evaluate and improve our visual service effectiveness model.

Learn from Human Preferences. Second, in a recent work, the researchers studied how reinforcement learning agents can learn policies from guidance based on human preferences rather than from reward signals [3]. We plan to adopt this technique and conduct user studies to train our service selection agent following human preferences data examined by real users.

References

1. Baek, K.-D., Ko, I.-Y.: Spatially cohesive service discovery and dynamic service handover for distributed IoT environments. In: Cabot, J., De Virgilio, R., Torlone, R. (eds.) ICWE 2017. LNCS, vol. 10360, pp. 60–78. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-60131-1_4

2. Baek, K.D., Ko, I.-Y.: Spatio-cohesive service selection using machine learning in dynamic IoT environments. In: Mikkonen, T., Klamma, R., Hernández, J. (eds.) ICWE 2018. LNCS, vol. 10845, pp. 366–374. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91662-0_30

3. Christiano, P.F., Leike, J., Brown, T., Martic, M., Legg, S., Amodei, D.: Deep reinforcement learning from human preferences. In: Advances in Neural Information Processing Systems, pp. 4299–4307 (2017)

4. Deng, S., Wu, H., Hu, D., Zhao, J.L.: Service selection for composition with QoS correlations. IEEE Trans. Serv. Comput. **9**(2), 291–303 (2016)

5. Gill, H., Midkiff, S.F.: Cyber-physical systems program solicitation (2009). https://www.nsf.gov/pubs/2008/nsf08611/nsf08611.htm

6. He, J., et al.: Deep reinforcement learning with an action space defined by natural language. In: Proceedings of the 2016 Workshop Tracks of International Conference for Learning Representations (ICLR) (2016)

7. Konda, V.R., Tsitsiklis, J.N.: Actor-critic algorithms. In: Advances in Neural Information Processing Systems, pp. 1008–1014 (2000)

8. Krösl, K., Bauer, D., Schwärzler, M., Fuchs, H., Suter, G., Wimmer, M.: A VR-based user study on the effects of vision impairments on recognition distances of escape-route signs in buildings. Vis. Comput. **34**, 911–923 (2018)

9. Ma, X., Cackett, M., Park, L., Chien, E., Naaman, M.: Web-based VR experiments powered by the crowd. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web, pp. 33–43. International World Wide Web Conferences Steering Committee (2018)

10. Mnih, V., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529 (2015)

11. Moghaddam, M., Davis, J.G.: Service selection in web service composition: a comparative review of existing approaches. In: Bouguettaya, A., Sheng, Q., Daniel, F. (eds.) Web Services Foundations, pp. 321–346. Springer, New York (2014). https://doi.org/10.1007/978-1-4614-7518-7_13