Towards the Integration of Digital Avatars in Urban Digital Twins on the Cloud-to-Thing Continuum*

Lorenzo Toro-Gálvez, Rafael García-Luque, Javier Troya, Carlos Canal, and Ernesto Pimentel

ITIS Software, Universidad de Málaga, Spain
{lorenzotoro,rafagarcialuque,jtroya,carloscanal,epimentel}@uma.es

Abstract. Urban Digital Twins (UDTs) represent a powerful tool to effectively make cities smart. Over the last few years, the interest in the social aspects of smart cities is growing fast. For this reason, citizens must be considered as first-class entities in UDTs. At the same time, citizens' privacy cannot be compromised. In this paper, we propose to integrate citizens through their digital avatars (DAs) into UDTs. DAs allow to exploit citizens' information, behavioral habits and personal preferences, while allowing them to have full control of their own data. We present our envisioned architecture that makes use of the cloud-to-thing continuum for optimizing the available processing resources. We focus on a case study of the public transformation service of the city of Malaga (Spain) and describe how we are addressing its implementation.

Keywords: Urban Digital Twin · Digital Avatar · Cloud-to-thing Continuum.

1 Introduction

Smart cities use technology to try to offer solutions to global problems such as climate change or resource depletion, and aim to provide useful services to their citizens and to solve specific urban problems, such as transportation and accessibility. The concept of Digital Twin City [6], also known as Urban Digital Twin (UDT), has been defined as a way to effectively make cities smart. A UDT is a digital twin capable of modeling specific city aspects such as transportation, heat maps of population density or environmental factors. Similar to other digital twins, UDTs facilitate two-way feedback between the model and the physical entities represented in it. These digital twins cannot only enable real-time remote monitoring for cities, but also suggest adaptation policies and enhance decision-making processes in areas such as improved urban governance, smart healthcare or smart transportation.

* This work was partially funded by the Spanish Government (FEDER/Ministerio de Ciencia e Innovación–Agencia Estatal de Investigación) under projects PID2021-125527NB-I00 and TED2021-130523B-I00, as well as by UMA.

Over the last few years, the interest in the social aspects of smart cities is growing fast. For this reason, citizens must be considered as first-class entities of the UDT, since they are the fundamental key to this ecosystem. However, the role of people has often been ignored, treating them as a crowd and not as individuals. And if citizens are to be considered as individuals, then their privacy would be compromised [2]. In previous works [3,4,13,14], we presented the concept of *Digital Avatar (DA)*, which represents a virtual representation of its user and through which the latter can decide which personal information to share. The DA resides in the smartphone of its user.

In this paper we propose to integrate citizens through their DAs into UDTs. The idea is to exploit citizens' information, behavioral habits and personal preferences, at the same time allowing them to have full control of their own data, including access control and storage location. Citizens who share specific information can benefit from better services from the organizations to which they entrust their data.

We envision the development of a framework which makes use of the cloud-to-thing continuum [9,10,11] for optimizing the available processing resources. For this, we propose a distributed architecture based on four layers, namely *mist*, *edge*, *fog* and *cloud*, and define the devices and services to be placed in each layer by focusing on a case study of the public transportation service of the city of Malaga (Spain). We also present the technology stack with which we plan to implement our framework, and give a glimpse of the implementation centered on digital twins.

The remainder of this paper is structured as follows. Section 2 gives an insight into the basics needed to understand our approach. Then, Section 3 presents a case study of a UDT dealing with public transportation that motivates our proposal. Sections 4 and 5 present the architecture to be applied to our envisioned framework and the implementation centered on digital twins that we are carrying out, respectively. Finally, Section 6 concludes the paper with an insight to future work.

2 Background

2.1 Urban Digital Twins

A Digital Twin (DT) is a comprehensive digital representation of a real system, service or product (the Physical Twin, PT), synchronized at a specified frequency and fidelity [1]. The DT includes the properties, conditions, and behavior of the physical entity through models and data, and is continuously updated with real-time data about the PT's performance, maintenance, and health status throughout its entire lifetime [5,7].

One specific type of DT is the Urban Digital Twin (UDT), which is a virtual representation of a city's physical assets that utilizes data, analytics, and machine learning to create real-time, adaptable simulation models. The digital twin describes the reality (and the history of it), while it is the additional applications that bring the real intelligence and help create the common picture of reality that is the added value of an urban digital twin. UDTs enable more informed decision-making, participatory governance, and improved urban planning by offering a risk-free testing environment for long-term predictions and impact assessments.

2.2 Digital Avatars

A Digital Avatar (DA) [4,13] is a virtual entity residing in an individual's smartphone or tablet that records information about its owner and offers different services for interacting with their surroundings and other users' DAs. This concept arises from the People as a Service (PeaaS) [8] paradigm, which provides a conceptual model for application development focused on the smartphone as a representative and interface to its owner. The main purpose of a DA is to serve as a digital representation of a person, facilitating their participation in collaborative social computing applications. In this context, the DA of a user is designed to collect data about the user and their habits, integrate this information with external sensor data or open data sources, and interact with both the environment and other users' DAs.

Additionally, a DA is responsible for representing the user in responding to requests coming from other avatars or social computing applications, as well as adjusting its records and anticipated behavior as needed. Consequently, DAs empower users by enabling them to (i) take control of the information and contents they create, (ii) manage how all that information is accessed and exploited in a secure manner by third parties, and (iii) be in control of their changes and adaptations. Thus, privacy and security can be ensured [12], since users can have the control over their own data and decide whether or not they want to share it, and with whom. However, there are limitations to the capabilities of DAs, such as lack of support for global decision making and comprehensive planning, which are key elements of any smart city application.

3 Case study

The case study we are dealing with is urban transportation in smart cities, specifically in the city of Malaga (Spain). We have access to real-time information about the use of buses in the city: transport lines, stops and schedules, GPS position of the buses, traffic status, among others, provided by the City Council as open data[1].

Our main goal is to develop an urban digital twin with the ability to use all the information about the transportation network to make predictions about peak occupancy hours, recognize usage patterns and allocate the resources for optimal use. The digital twin can also monitor the current state of transport and react or adapt to unexpected peaks or incidents.

Furthermore, the UDT would be improved if it was connected to the citizens' DAs. Thus, one of our goals is also to infer information from them: how

[1] https://datosabiertos.malaga.eu/group/transporte

individuals use the transportation system, their habits and routines in relation to it, or their behavioral patterns for commuting to work or moving around the city. This will allow for more accurate predictions made on individuals. In this manner, citizens may get improved and personalized services.

Another benefit of DAs is that personal preferences could be taken into consideration to suggest alternative transportation methods more suited to citizens' likings and needs. Knowing their preferences could be very useful in deciding about the overall transport solutions to be implemented.

4 Architecture

To carry out the above case study, we rely on the Cloud-to-Thing Continuum [9,15]. This concept refers to a highly distributed, decentralized, and dynamic environment spanning from IoT devices to the cloud. The continuum is designed to address the challenges of next-generation IoT systems that involve a large number of heterogeneous devices generating massive amounts of data and require real-time processing, low service response times, and enhanced reliability and security.

Following the continuum model, we propose a four-tier architecture where each layer offloads the upper layer by taking over some of its functionalities. Next, we present the four layers starting from the lowest layer of the architecture: mist, edge, fog and cloud computing [11]. A snapshot of our architecture is shown in Figure 1.

Mist represents the layer closest to the citizens, and we place here the cizitens' smartphones that include their DAs. Therefore, all citizens' data coming from their DAs is collected at this level: positioning, information on their habits and routines, personal preferences, etc., always ensuring their privacy and data anonymization and allowing users to decide with whom they share them. Recall that citizens' information resides in their smartphones, from where they decide what to share and with whom [13]. *Mist* is also the layer through which citizens receive individualized and personalized information, coming from the processing in the upper layers, based on their virtual profiles.

The next layer is *edge* computing. This layer processes data directly coming from the mist as well as data gathered in this layer. Here, we consider that the bus stops and buses can gather and process information depending on the environment and different situations. Thus, this architecture makes it possible to run specific applications in a fixed location providing direct data transmission. Specifically, the inclusion of tracking and processing devices such as Bluetooth beacons in the different buses or bus stops makes it possible to send notifications to citizens and provide significant information, e.g., about the bus being late, overcrowded, or skipping a particular stop.

The third layer is *fog* computing, which enables ubiquitous access to a shared range of scalable computing resources. In the context of the proposed scenario, the *fog* layer will include the state of the transportation system including relevant factors organized in city areas. By using fog computing to process and

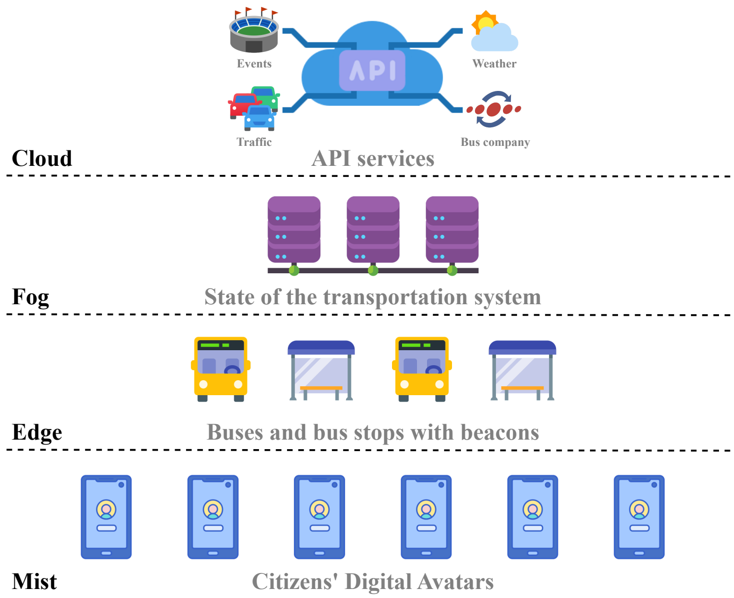

Fig. 1. Proposed 4-layer architecture based on the cloud-to-thing continuum.

analyze real-time data from the transportation system and citizens, we can gain valuable insights into factors such as traffic congestion, passenger load and route optimization, all of which can help improve the overall efficiency and reliability of the transportation system. Furthermore, this layer is essential for making predictions. By analyzing real-time and historical data, we can predict transportation costs, identify potential traffic bottlenecks and crowded areas, forecast peak demand times, and recognize popular routes and transit hubs.

The *cloud* layer, as the top tier of our proposed architecture, provides access to a range of services and real-time data that can have a direct effect in both the current and future state of the transportation system. This information is directly translated to the fog layer for analysis. Examples of services that can have an effect in the state of the transportation system include weather information and forecast, events taking place in the city (sport games, concerts or festivals) and real-time traffic information—for instance, the TomTom API offers information on the current traffic status. All these services offer information that can be crucial for optimizing transportation routes and reducing travel time for commuters.

Note that some information on the urban public transportation system is also accessible from the cloud layer, such as the route or position of the different buses. Information coming from this layer can be used to alert citizens about general incidents in the transportation system, such as accidents or cancellations, non-operational lines on special days, etc. Recall that other information related to transportation comes from the mist layer (specifically from the location of

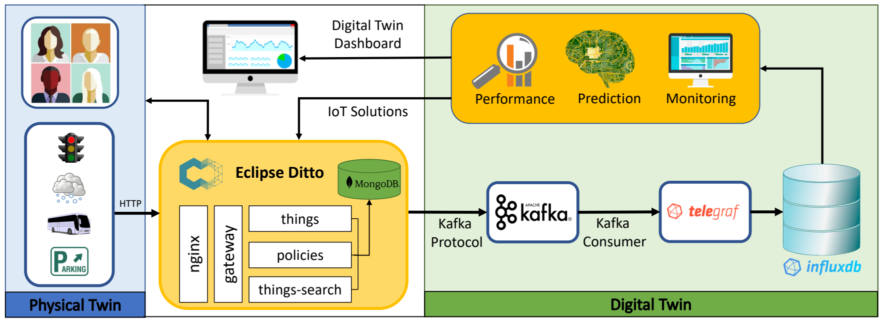

Fig. 2. An overview of the Digital Twin Architecture composed of microservices.

citizens' digital avatars) or from the edge layer (thanks to the beacons mentioned above).

5 Implementation

For materializing the architecture described before, we employ Eclipse Ditto[2], an open-source domain-agnostic framework that helps build digital twins. Figure 2 shows our envisioned architecture from a digital twins point of view. It is composed of independent microservices that communicate with one another in real time, using protocols such as AMQP or Apache Kafka. Each of these services is encapsulated in containers managed by Docker, ensuring isolation for both, which in turn enables portability and guarantees correct execution.

The physical twin in our system is comprised of real-world entities providing information to the digital twin, with the Eclipse Ditto platform situated between them. The physical twin is scalable, since more entities can be added, and so is the digital twin. We also include the citizens, through their digital avatars, as part of the physical twin, since they are an integral part in any smart city.

At the current point of the implementation, we are gathering information from the entities of the physical twin. For this, we have implemented a Java application, using Maven, that makes requests to the different APIs and collects the information through Eclipse Ditto. However, Ditto is only responsible for storing and updating the most recent state of each entity in its database, implemented in MongoDB, which prevents having historical records. For this reason, we use, as part of the digital twin, InfluxDB[3], an open-source database that enables fast storage and retrieval of time series data, with high availability. Since Ditto requires an intermediary in the connection to InfluxDB, we use Telegraf[4], which plays the role of a server-based agent responsible for collecting and sending metrics to databases, IoT systems, etc.

[2] https://www.eclipse.org/ditto/intro-overview.html
[3] https://www.influxdata.com/products/influxdb-overview/
[4] https://www.influxdata.com/time-series-platform/telegraf/

The information stored as time series data can be used for different purposes, such as obtaining performance insights, monitoring and doing predictions. All these can be used to improve or correct the service offered by the different entities of the physical twin. Some of this information can also be used to improve and create easy-to-understand recommendations to citizens through their digital avatars, such as suggesting alternative routes when commuting to work or avoiding traffic congestion in rush hours.

As for the data we are collecting so far from the services of the physical entities, we are making periodic requests to the APIs of (i) Malaga City Council, (ii) OpenWeatherMap, and (iii) Tomtom. This means we are storing information about real-time buses location, parking occupancy, urban traffic and meteorological parameters of the city of Malaga. For instance, regarding information from public buses, our framework has currently collected around 79k documents containing the geolocation of all buses in each time interval (requests are made every minute). This data-base knowledge allows to set the location of each bus with respect to its established stop and, thus, estimate the duration of all the routes.

The combination of the times taken by the buses, together with the climatic aspects, parking occupancy and traffic congestion, will pave the way for establishing patterns that infer the performance of buses and, based on this, adapt it to the preferences and needs of the citizens.

6 Conclusions and future work

This contribution describes the architecture for our envisioned framework in which we propose to integrate citizens through their DAs into UDTs, making use of the cloud-to-thing continuum to arrange different components in each of the four proposed layers: *mist*, *edge*, *fog* and *cloud*.

As a proof of concept of the proposed approach, we have developed an initial version of the UDT based on Eclipse Ditto, that takes into consideration real-time data of services such as bus transportation, traffic status and weather information.

This work can be continued in several directions. First, we plan to integrate the UDT with citizens' custom information to know how citizens use the transportation system for moving around the city, their daily routines and their preferences, among others. We also aim to expand the UDT by including other elements such as beacons and processing devices on the buses to collect more data and provide a more comprehensive analysis of the transportation system. Furthermore, we plan to incorporate data on specific events happening in the city, such as festivals or sports events, which may have an impact on the transportation system. Finally, we also plan to explore the integration of other techniques like ML and IA to further improve the system's performance and accuracy in its predictions.

Overall, this proposed framework has the potential to significantly improve the transportation system of the city of Malaga by providing real-time data and analysis that can optimize routes, reduce travel time, and enhance the efficiency and reliability of the system.

References

1. Digital Twin Consortium. 2021. Glossary of Digital Twins
2. Badii, C., Bellini, P., Difino, A., Nesi, P.: Smart City IoT Platform Respecting GDPR Privacy and Security Aspects. IEEE Access **8**, 23601–23623 (2020)
3. Bandera, D., Pozas, N., Bertoa, M.F., Álvarez, J.M., Canal, C., Pimentel, E.: Extensión de Digital Avatars para crowdsensing distribuido. In: Sistedes JCIS (2021)
4. Bertoa, M.F., Moreno, N., Pérez-Vereda, A., Bandera, D., Álvarez-Palomo, J.M., Canal, C.: Digital Avatars: Promoting Independent Living for Older Adults. Wirel. Commun. Mob. Comput. **2020**, 1–11 (2020)
5. Bordeleau, F., Combemale, B., Eramo, R., van den Brand, M., Wimmer, M.: Towards Model-Driven Digital Twin Engineering: Current Opportunities and Future Challenges. In: ICSMM Proceedings. vol. 1262, pp. 43–54 (2020)
6. Deng, T., Zhang, K., Shen, Z.J.M.: A systematic review of a digital twin city: A new pattern of urban governance toward smart cities. MSC Journal **6**(2), 125–134 (2021)
7. Grieves, M.: Digital Twin: Manufacturing excellence through virtual factory replication **1**, 1–7 (2014)
8. Guillén, J., Miranda, J., Berrocal, J., García-Alonso, J., Murillo, J.M., Canal, C.: People as a Service: A Mobile-centric Model for Providing Collective Sociological Profiles. IEEE Softw. **31**(2), 48–53 (2014)
9. Herrera, J.L., Galán-Jiménez, J., Garcia-Alonso, J., Berrocal, J., Murillo, J.M.: Joint Optimization of Response Time and Deployment Cost in Next-Gen IoT Applications. IEEE IoT Journal **10**(5), 3968–3981 (2023)
10. Laso, S., Berrocal, J., Fernandez, P., García, J.M., Garcia-Alonso, J., Murillo, J.M., Ruiz-Cortés, A., Dustdar, S.: Elastic data analytics for the cloud-to-things continuum. IEEE IC **26**(6), 42–49 (2022)
11. Lynn, T., Mooney, J.G., Lee, B., Endo, P.T.: The Cloud-to-Thing Continuum: Opportunities and Challenges in Cloud, Fog and Edge Computing (2020)
12. Muñoz, P., Pérez-Vereda, A., Moreno, N., Troya, J., Vallecillo, A.: Incorporating Trust into Collaborative Social Computing Applications. In: Proc. of EDOC'21 (Oct 2021)
13. Pérez-Vereda, A., Canal, C., Pimentel, E.: Modelling digital avatars: A tuple space approach. Science of Computer Programming **203**, 102583 (2021)
14. Pérez-Vereda, A., Murillo, J.M., Canal, C.: Dynamically programmable virtual profiles as a service. In: 2019 IEEE SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI. pp. 1789–1794 (2019)
15. Čilić, I., Žarko, I.P., Kušek, M.: Towards Service Orchestration for the Cloud-to-Thing Continuum. In: 2021 6th SpliTech. pp. 01–07 (2021)