Challenges in Using Digital Twins for Modelling Human Behaviour in Environments

Hadi Banaee, **Franziska Klügl** and **Amy Loutfi**
Centre for Applied Autonomous Sensor Systems (AASS), Örebro, Sweden
{hadi.banaee, franziska.klugl, amy.loutfi}@oru.se

Abstract

Digital Twins are powerful tools focusing on scene representation and environmental sensing. Still these tools are not yet used to uncover the complexities of human behaviour and their interaction with the environment. In particular deriving activities, intentions and casualties from digital twins in order to better capture human environment interaction can provide useful tools for data-driven decision support in a number of applications. This short paper lifts some of the challenges, and opportunities in augmenting current digital twins with a cognitive (AI) layer.

Motivation

To understand and predict human behaviour in space forms an essential ingredient to many areas of research and design, ranging from urban planning to advertisement. The need for understanding goes beyond knowing what are relevant points of interest for an (individual) human or what is the next step that a human would probably make given her previous movement. This work is about discussing the need and potential for new data-driven approaches in modelling human activity. Yet, simply generating information from trajectories or IoT sensor data just using innovative data-driven methods, such as classifying or predicting movement is not enough [Wang *et al.*, 2019]. Human behaviour captured by sensors may be unexpected, which itself is challenging. Beyond that, it is particularly important to be able to explain why the observed human decided to e.g. sit on the stairs instead of using the benches. Maybe all benches were occupied, dirty or simply too exposed to direct sun on a hot day. Environmental context and how it impacts on the agent must be explicitly considered, in addition to isolated position sequences.

A concrete example of a Digital Twin which we are working on is an urban square in Örebro, Sweden. The place is for pedestrians, equipped with a walk-able fountain, benches, rows of trees at two sides and bicycle stands, public toilet, etc. There are restaurants and bars at two of its sides, offices and the river at the other two, all separated by roads from the actual place. As the place connects a public transportation hub to a road towards the railway station, many people simply cross the place, but there are a variety of other activities during day and night time. This makes the place an interesting case for an urban digital twin. We currently plan to install a camera system that will continuously produce position trace information from individual, anonymous humans. There are mainly the following motivations for creating this digital twin, there: 1) Answer questions on what is going on now in a fast accessible way: Is there a critical situation emerging or (a robbery at the toilet, a brawl at the ice cream shop) or has there been a change in usage recently (parts of the area are more and more ignored by pedestrians)? 2) Provide general information that urban planning can use to improve the infrastructure for supporting desired activities and demotivating undesired ones?

Numerous research has been done to simulate human behaviour in a crowd [Duives *et al.*, 2013; Karamouzas *et al.*, 2017; Haghani and Sarvi, 2018]. Most of these studies have focused on developing models to imitate expected human behaviours by involving prior expert knowledge about the suitable interactions between humans and environments. However, a challenging question is how analysing human behaviour in an actual environment can be grounded using data-driven models. This modelling is not a trivial task due to the complexity of the human behaviours in real-world environments. Hereby, a proposed approach should combine a number of tasks: observe the environment, extract humans' mobility and behaviour from sensor data, infer meaningful interactions with their environment, understand the relationships between the human and environment, and finally describe the derived information in a human-understandable way.

Besides data-driven approachs to analysing human behaviour, the cognitive principles of such behaviour should be in focus as well. Several computational models have been developed to capture the behaviour of human pedestrians for the sake of prediction or optimisation [Zou *et al.*, 2018]. These models are not sufficient to interpret human behaviour due to the lack of coupling data analysis and learning approaches with human cognition. Moreover, another shortcoming in the current frameworks to model human mobility is to generate information out of repetitive actions as patterns [Pappalardo and Simini, 2018; Rudenko *et al.*, 2019]. To reason how and why an activity is performed, it is necessary to analyse behaviour in a long-term observation to extract meaningful patterns (e.g., rules) in the data.

In our ongoing work, we aim to address the following is-

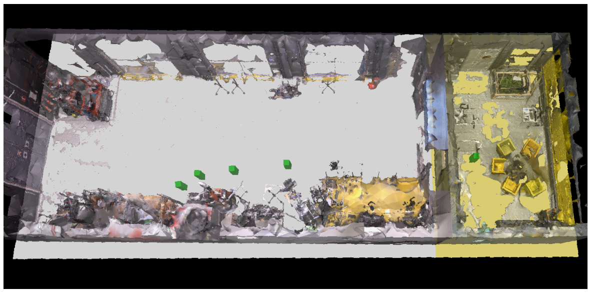

Figure 1: A screenshot of the developed digital twin for two adjacent workrooms, depicting the positions of human objects (green cubes).

sues: 1) We assume that data-driven analysis needs to explicitly account for its spatial context, thus we need to contextualise data-driven approaches into a digital twin of its environment. 2) Perception and spatial reasoning form essential ingredients into how humans move and use space. Thus, a data-driven approach needs to be connected to this type of human cognition beyond individual points of interests or movement parameters. The existing challenges and the feasible solutions to achieve such a model is discussed within a framework that we call *AI-enabled digital twins*.

Digital Twins

A digital twin (DT) is a virtual representation of physical environments including the relevant information from static structures and objects [Haag and Anderl, 2018]. Digital twins have been mainly designed for IoT purposes and are used in various applications such as manufacturing, healthcare, and smart cities [El Saddik, 2018; Kritzinger *et al.*, 2018].

Utilising digital twins to capture human behaviour has the advantage of observing actual activities of humans in relation to the digitalised environment with all its limitations. For instance, a digital twin of a university corridor should contain information about where students sit (which is not necessarily the placed benches), or a digital twin of a kindergarten should capture where kids play (which is not necessarily the playground). Thus, for enabling this, a digital twin needs to capture human activity beyond trajectories, but represent human activities – ongoing and in an abstracted pattern-like form. In this way, the generated representation may contain emergent behaviour and interaction that are not visible by looking at the current state of the system. Figure 1 shows an example of our developed digital twin, along with capturing the positions and the mobility of human objects within the environment.

AI-enabled Digital Twins

AI approaches can be employed to perform a variety of tasks upon the digital twin's data. This is more crucial when more complicated tasks such as detecting human behaviour patterns are needed to be done. Moreover, there are several cognitive models/approaches for improving our understanding on dynamic human motion and behaviour [Fridman and Kaminka, 2010; Manley and Cheng, 2018; Zou *et al.*, 2018]. Another reason to involve AI in such models is the capability of recently developed *Explainable AI* approaches. A variety of these approaches (from transparent neural networks [Strannegård *et al.*, 2012] to natural language generation [Gatt and Krahmer, 2018]) can enable a digital twin to emerge descriptive results either to generate interpretable information to the end-user of the application, or to enhance further scientific investigations on the entire environment (e.g., building practical multi-agent simulations).

Challenges

There are a number of challenges while modelling human behaviour patterns in an AI-enabled digital twin, as follows:

Learning aspect: Within machine learning approaches, learning the historical patterns is mainly employed for tasks such as optimisation and prediction. A challenging task, however, is to model the learnt patterns in a semantic representation as an interpretable set of knowledge. Moreover, human behaviour patterns can include determining interactions, correlations, rules, and causality between the occurred behaviours and the environment. Determining such pattern-based information will be more complicated than event-based tasks such as anomaly detection.

Social and contextual aspects: An important aspect of modelling human spatial activity patterns is to consider the context of environmental and social behaviours. Abstract analysis of human mobility in an environment may lead to a naive model for individuals without explaining the reasons behind the behaviours [Rudenko *et al.*, 2019]. The main challenge is how social behaviours such as force, attraction or repelling each other can be augmented to a data-driven model in order to capture social factors (e.g., psychological security) of human mobility. This issue is interesting as social and contextual factors, such as culture, type of environment, time of day can and should be used for augmenting the models learnt from sensing how humans use space, traverse a place, etc. An example of this challenge is to compare modelling human behaviour with robot navigation. In the former case, it is crucial to consider e.g., the difference between the crowd in a university corridor versus a kindergarten yard, or the trajectory of humans in an empty street during lunchtime or at midnight. However, in the latter case, there might be no difference for a robot to move around at any time of the day.

Feasible Solutions

A number of AI solutions are feasible to address the problem of modelling human behaviour patterns with respect to the discussed challenges. Digital twins typically equipped with various sensors such as camera, motion, and light to measure human mobility (e.g., entering or leaving the space). Considering the data collected in a digital twin, the tasks and their corresponding solutions can be categorised into the three following perspectives:

Historical data: To achieve pattern recognition on the historical data of human mobility, a solution is to use automatic rule generation approaches to extract meaningful relations. Rule mining aims to find (unseen) temporal and spatial relations between the actions, and present them in an interpretable model. For instance, one possible output of such approach can be: "*most of the times, when the students come to the corridor after lunchtime, they sit on the stairs near the classroom.*" This kind of results may lead to understanding the behaviour of students before starting a course lecture.

Another solution would be using learning techniques (e.g., neural network) to analyse the mobility of the crowd in a long-term and large-scale data set. The extracted information may emerge to explain the interactions between humans in a long period as a repetitive behavioural pattern.

Real-time stream of data: Analysing the stream of data a digital twin will be influential for tasks such as interpolation, anomaly and change detection. This analysis should be presented with proper visualisation to the end-user as well.

Upcoming data: Following the extracted information from the past and present, it is beneficial to also involve AI solutions to predict the upcoming activities. Knowing human behaviour patterns will ease us to understand and explain what will happen next. If we can sufficiently capture the underlying rules and context information, also what-if scenarios testing unseen environmental setups may become possible.

Final Remarks

This paper has addressed the challenges of modelling human behaviour patterns in a digital twin framework, along with feasible solutions to achieve such a model in data-driven manners. An instant step to develop the AI-enabled digital twin will be characterising appropriate features to identify human behaviours, and then, to apply suitable data-driven approaches to capture patterns in such behaviours.

References

[Duives *et al.*, 2013] Dorine C Duives, Winnie Daamen, and Serge P Hoogendoorn. State-of-the-art crowd motion simulation models. *Transportation research part C: emerging technologies*, 37:193–209, 2013.

[El Saddik, 2018] Abdulmotaleb El Saddik. Digital twins: The convergence of multimedia technologies. *IEEE MultiMedia*, 25(2):87–92, 2018.

[Fridman and Kaminka, 2010] Natalie Fridman and Gal A. Kaminka. Modeling pedestrian crowd behavior based on a cognitive model of social comparison theory. *Computational and Mathematical Organization Theory*, 16:348–372, 2010.

[Gatt and Krahmer, 2018] Albert Gatt and Emiel Krahmer. Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, 61:65–170, 2018.

[Haag and Anderl, 2018] Sebastian Haag and Reiner Anderl. Digital twin–proof of concept. *Manufacturing Letters*, 15:64–66, 2018.

[Haghani and Sarvi, 2018] Milad Haghani and Majid Sarvi. Crowd behaviour and motion: Empirical methods. *Transportation research part B: methodological*, 107:253–294, 2018.

[Karamouzas *et al.*, 2017] Ioannis Karamouzas, Nick Sohre, Rahul Narain, and Stephen J Guy. Implicit crowds: Optimization integrator for robust crowd simulation. *ACM Transactions on Graphics (TOG)*, 36(4):1–13, 2017.

[Kritzinger *et al.*, 2018] Werner Kritzinger, Matthias Karner, Georg Traar, Jan Henjes, and Wilfried Sihn. Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11):1016–1022, 2018.

[Manley and Cheng, 2018] Ed Manley and Tao Cheng. Exploring the role of spatial cognition in predicting urban traffic flow through agent-based modelling. *Transportation Research Part A: Policy and Practice*, 109:14 – 23, 2018.

[Pappalardo and Simini, 2018] Luca Pappalardo and Filippo Simini. Data-driven generation of spatio-temporal routines in human mobility. *Data Mining and Knowledge Discovery*, 32(3):787–829, 2018.

[Rudenko *et al.*, 2019] Andrey Rudenko, Luigi Palmieri, Michael Herman, Kris M Kitani, Dariu M Gavrila, and Kai O Arras. Human motion trajectory prediction: A survey. *arXiv preprint arXiv:1905.06113*, 2019.

[Strannegård *et al.*, 2012] Claes Strannegård, Olle Häggström, Johan Wessberg, and Christian Balkenius. Transparent neural networks. In *International Conference on Artificial General Intelligence*, pages 302–311. Springer, 2012.

[Wang *et al.*, 2019] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11, 2019.

[Zou *et al.*, 2018] Haosheng Zou, Hang Su, Shihong Song, and Jun Zhu. Understanding human behaviors in crowds by imitating the decision-making process. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.