



Exploiting digital twin as enablers for synthetic sensing

Roberto Minerva, Faraz Malik Awan, Noel Crespi

► To cite this version:

Roberto Minerva, Faraz Malik Awan, Noel Crespi. Exploiting digital twin as enablers for synthetic sensing. IEEE Internet Computing, 2022, 26 (05), 10.1109/MIC.2021.3051674 . hal-03100719

HAL Id: hal-03100719

<https://hal.science/hal-03100719>

Submitted on 6 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Department: Head
Editor: Name, xxxx@email

Exploiting Digital Twin as enablers for Synthetic Sensing

Roberto Minerva

Telecom SudParis, Institut Mines-Telecom, Institut Polytechnique de Paris

Noel Crespi

Telecom SudParis, Institut Mines-Telecom, Institut Polytechnique de Paris

Faraz M. Awan.

Telecom SudParis, Institut Mines-Telecom, Institut Polytechnique de Paris

Abstract— The Digital Twin is increasingly used in several problem domains and it is an important enabler for Internet of Things applications. It is a useful framework for representing and simulating how physical objects interact in target environments. The Digital Twin has the ability to store and maintain historical and current data about the physical object. In the field of IoT, the concept of general-purpose sensing is aiming at determining a set of basic sensing capabilities from which deriving, by means of Artificial Intelligence algorithms and techniques (termed Synthetic Sensing), relevant information for describing an environment. This paper explores the possible relationships and the feasibility of an integration of Synthetic Sensing within a Digital Twin framework. The relevant concepts and associated technologies, the challenges and some of the enabled scenarios that this integration can bring are presented. This approach is in its infancy and there is a strong need to validate and assess its viability, feasibility and benefits. The paper identifies some challenges and validation steps that can lead to a consolidation and broad adoption of this approach. Finally, the paper presents some future work that will be carried out with the goal to prove and promote the approach.

■ **THE APPLICATIONS** of the Digital Twin are increasing and they span beyond the original problem domain of manufacturing (i.e., the Product Lifecycle Management, PLM [1]) to new applications like e-health, smart cities, transportation, and many others [2] [3] [4]. In this paper, the Digital Twin is intended as the constant association between a physical object and its software representations within a virtualized environment. Events, actions, features, status, data and

all the other information characterizing the (physical) object are to be timely reflected into the logical (software) object (and vice versa). The Digital Twin refers to the relationship between a physical object, the PO, and its software representations, Logical Objects, LOs. This relationship, as pointed out in [5], is characterized by several properties that fully qualify the Digital Twin. Foundational properties are Identity; Representativeness (with Contextualization and

Modelling); Reflection; and Entanglement. They ensure that the LOs are identifiable and associated to the PO. They represent relevant features and aspects of the physical object respect to a specific application context. These properties ensure that changes in the PO are reflected in the LOs and that a strong link is established between the PO and its associated LOs. Software properties are features that a platform should offer to support the lifecycle of several DTs, namely replication, persistency, memorization, composability, accountability / manageability; augmentation; and predictability. They ensure that LOs representing a PO can be used and programmed for providing applications. In particular, LOs can be replicated (and distributed) in the software environment. They are persistent in the software space (at least a replica of the PO must always be available to applications). LOs must have memory capabilities to store changes and attributes of the DT over time. LOs can be aggregated and composed to describe complex POs. They can be managed and accounted for their usage. LOs can be extended and augmented by means of new functions and interaction modes and related Application Programming Interfaces, APIs. Their behavior can be simulated and predicted over time. Business properties refer to servitization, i.e., the ability to transform a product/object into a set of services and functions; and ownership, i.e., the possibility to identify the owner of each PO, and the individual owners of associated LOs. The DT is particularly appealing in relationship with Internet of Things and Smart City applications. The ability to represent features and aspects of sensors/actuators or complex aggregation of physical objects (e.g., a building) offers the possibility to track the behavior of objects and their evolution over time [6] [7]. Figure 1 represents some of the Digital Twin properties [5].

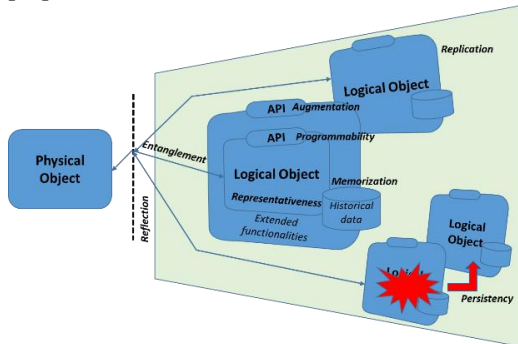


Figure 1: some Digital Twin properties

GENERAL-PURPOSE AND SYNTHETIC SENSING

In the Internet of Things, an interesting trend towards the Synthetic Sensing [8] is emerging, i.e., the possibility of determining complex events by means of simple sensed data like noise, temperature, light and other basic measureable features (general-purpose sensing). The application of Artificial Intelligence techniques to these basic data yields to the identification of complex events and information that can fully describe the situation occurring in an environment. This capability is termed Synthetic sensing. This approach is appealing because it supports the standardization and consolidation of sensing capabilities and it moves the complexity to the software level. However, its applicability has to be proved in large and complex contexts, e.g., a factory and in perspective to a smart city.

There are two major research challenges in this approach: the first is to validate the approach in large and complex environments; and the second is the identification of a minimal set of sensing capabilities that can be used to support the largest number of applications. In particular, it is important to determine if video sensing, i.e., the usage of video cameras for collecting data, is a basic and needed capability. The problem with video is the possibility to collect information of individuals and then infringing their privacy. Figure 2 represents the relationship between the general-purpose sensing (i.e., the capture of basic measures of the environment), the collection of data and a software functionality that, by means of AI techniques, can infer complex data and information about the environment.

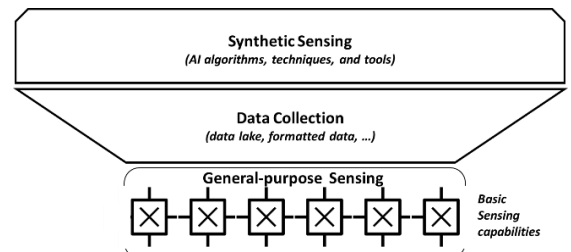


Figure 2: General-purpose and synthetic sensing

DIGITAL TWINS AND SIGNATURES

The DT of a product can represent the visual, audio, electrical, chemical and other characteristics as well as the “signatures” of the product in different states (e.g., normal or under stress or in malfunctioning states). The

DT can then be seen as a repository of signatures that can describe or uniquely identify the (class of a) product.

How the DT, the Synthetic Sensing and the Signature recognition fit together? Let us consider an example. Electric cars are less noisy than traditional vehicles. However, for safety and business reasons [9], different types of cars (even of the same maker) will have distinctive engine sounds. This feature is as a sort of “signature” that characterize a specific model. A digital twin representing that car model should store the “engine sound” and make it publicly available. The sound signature of the car can be used by synthetic sensing applications to identify the car and activate specific functions. The signature becomes a synthetic identifier of the car / object (a sort of advanced RFID). There is a vast literature on signatures ranging from acoustic studies (e.g., [10]) to electric ones [11] and more. The signature under certain circumstances can be useful even to predict unintended functioning or malfunctioning of products and devices [12].

The Digital Twin can store and make available a set of signatures that characterize the physical object under normal operation or during critical conditions. The Digital Twin is a “signature” repository to be used to identify an object, to recognize some events and for determining some of behaviors of the PO (see Figure 3).

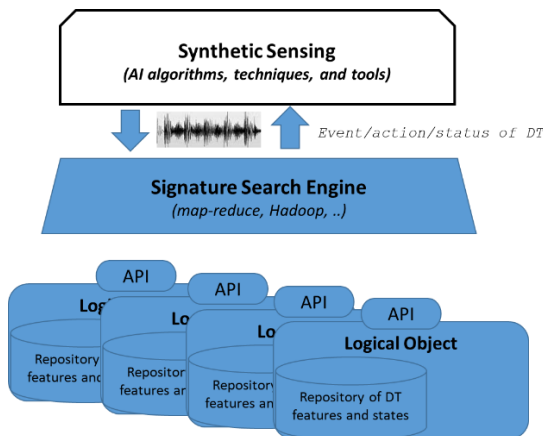


Figure 3: Digital Twin and Signatures

The identification can be generic, i.e., the identification of the class of physical object (e.g., the model of a vehicle of a certain brand), or specific, i.e., the univocal identification of a specific physical object (e.g., the specific car with a univocal identifier, e.g., the plate). The specific identification is very important, but it may introduce some privacy and security issues. The

identification of specific or generic objects is not always easy in large complex systems. For instance, the recognition of generic vehicles at crossroads of a jammed smart city can be difficult or impossible. In this case, it would be important to be able to “measure” the level of congestion and traffic and associate to it a set of acoustic, pollution and visual signatures (a characteristic of the DT of the specific crossroad). The traffic level signature can be considered as the “sum” of the different signatures of the actual vehicles at the specific crossroad (plus a certain quantity due to other acoustic, pollution, visual elements).

A NEW APPROACH

Large IoT applications are coping with complex environments measurements. They use specialized sensors for detecting, acquiring, and collecting data related to specific events (e.g., traffic density measurement). It is possible to exploit the DT features and capabilities to understand what is happening in a complex environment and how the physical objects influence and determine the context. The DT is representing either a physical object capable of measuring its changes and characteristics (an instrumented physical object); or the logical objects are able to continuously collect and process data that represent the variation of the physical object. The DT becomes a proactive element capable of acting on the environment and directly changing the situation. Inferring and representing the relationships between the single DT and the others in the specific environment provides a more granular knowledge of the phenomena. By identifying classes and (if possible) instances of physical actors responsible of the modifications, their individual contribution to changes of status of the environments and the interdependencies between the objects themselves, make more precise the understanding of the situation. Using the DTs in this manner can provide an explicable and understandable view about the current situation. Also simulations can provide a better understanding and insight of the dynamics that contributed to the current status of the environment. The DTs can be used to calculate and predict additional measurements that are not necessarily provided by deployed sensors. It is important to derive models and algorithms capable of supporting these calculations. This may require a continuous calibrations and tuning up of inferred data based on processing of the basic sensed values and measures collected in the field. Many application domains could benefit from a change of paradigm from

LO) and get a description of its typical functioning and the set of signatures associated to it. Additional information could also be provided by the DT, such as typical processes needed to use it, the association with other objects and more. Figure 5 sketches a simple scenario for the home environment.

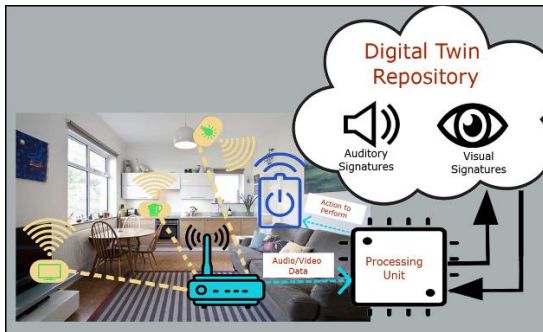


Figure 5: the use of signatures and DT in a home environment.

When the home system will detect a new event (for instance a new sound) it compares it with the set of known signatures (in this case a sound) associated to the devices present in the home. The set of possibilities is, in most of the cases, well defined and easily and quickly recognized. In addition, if the recognized event is classified as a potential step within a process, the home system could be “expecting” the next event and prepare in advance for accommodate the action. If the signature is not recognized in this “most likely set”, then the search for a mapping of the event and the signature of DT can be extended to malfunctioning cases, or to exceptions or to a general repository of accessible signatures. If the event is not recognized, a supervised training can take place and the human intervention could help in identify, classifying and assigning the event to the signature of a DT. Another important aspect of this approach is its ability to provide an explanation of the event within a process. This may help humans to understand why the system is behaving in a particular way. For instance, in case of malfunctioning, the system could associate a particular event with a specific signature (a spike in electric consumption, for instance) and act in such a way to recognize the malfunctioning and put in place actions for reducing the impact of the damage. These actions can be suggested by the description of the DT itself (much like an instruction manual) or to be determined by some rule based logic.

In this example we have considered devices (POs) that have a weak entanglement with the corresponding

virtual representation (LOs), i.e., the physical device is not able to determine the changes of status and communicate them immediately to the LOs. The DT representation is a static one: the signatures are those determined in the factory and only a few new ones are introduced in case of issues and malfunctions. In case of strong entanglement, i.e., when the device has sensing capabilities that can provide a full description of the status of the PO, then the home scenario DT-based can be simpler. The device is able to understand the events and it will simply communicate them (and the process and expected future actions) to the LOs that will forward the information to the control system of the Home. This DT based approach can be useful for helping in many situations by reducing the complexity of understanding the behavior of the physical objects and their interactions in specific contexts.

Scenario 2: Smart City and traffic – pollution relationship.

In this scenario, the following situation is considered: a limited set of traffic density sensors has been deployed in a smart city, while general-purpose sensing capabilities (comprising pollution sensors) are widely used to monitoring and measuring the functioning of the city. Noise and pollution sensing are considered “basic” capabilities. Measuring the noise and the pollution levels often can provide a measure of the “density” of traffic. The assumption is that a certain number of vehicles in a specific area of the city may have a signature in terms of pollution and noise levels. Ideally, from noise and sound, it could be possible to capture the typical signatures of vehicles producing them and to aggregate their expected pollution contribution. It can be compared with the one measured by the sensors. Cars have an individual “sound” signature (per model and engine used). If the general-purpose sensory is able to detect the major part of the signatures of nearby cars, it would then be possible to “calculate” the expected pollution level of the aggregation of the detected cars. This allows the creation of a “composite signature” for that level of traffic. Other sources of pollution can be considered if they are represented by DT: nearby smart buildings and factories could contribute information for approximating the total pollution present in that area of the city at the time. Identifying many vehicles at once by means of their sound signatures can be impossible, so other basic measurements should be considered. Video footage or even still pictures of the congested areas in conjunction with visual signatures can be used

to dynamically determine the number of vehicles, their types and models. From individual signatures, specific contributions to the pollution production can be determined and the “pollution and noise” traffic signature can be generated. This signature using cameras or video can be calculated and verified in a subset of the main crossroads of the city. The derived signature is then used in other points of the city with a good approximation and accuracy. The very specific-purpose traffic sensors could be limited or eliminated in favor of a sensory infrastructure exploiting the synthetic sensing and DT representations. Also in this example, the considered DTs are static, i.e., they provide a set of signatures and related general information of the class of physical objects. More proactive DTs can dynamically provide information such as location, speed, density of cars around the physical object, their detailed consumption and pollution produced. The same scenario could be implemented, but with less sensory and software complexity. The DTs could provide all the needed information and the smart city system should be designed in such a way to collect the information and create a representation of the current situation. These DTs provide crowdsensing capabilities designed to proactively contribute to the measurement and optimization of the environment by directly and timely providing general-purpose data.

A mixed approach, i.e., general-purpose sensing in the public infrastructure and fully-fledged sensing in the POs, is a promising possibility for the coverage and monitoring of a city.

CONCLUSIONS

The paper has introduced a vision for the integration of Digital Twin and the Synthetic Sensing. This approach can have an impact on the capability of representing and programming complex. The depicted approach is a long-term vision that not necessarily will be implemented and realized as described. However, the Digital Twin will pervade many applications domains and its relationship with data fusion, artificial intelligence tools and algorithms should be further investigated to exploit its possibilities and features. A general remark respect to the approach is related to security and privacy. The identification of objects by signature may be quite precise in small confined environments, while in larger one can be difficult and can introduce privacy and security concerns for

individuals. Security breaches or the easiness of infringing privacy or security of individuals is a major challenge for the approach. To validate and assess the approach, its feasibility and its benefits, some incremental steps are envisaged in terms of research challenges:

1. General-purpose sensing: the challenge is the identification of a minimal set of basic sensing capabilities that can cover the largest part of needs of smart environments. Experimentations in this field could be helpful to determine these capabilities and to refine techniques, algorithms and tools for implementing functionalities of synthetic sensing.
2. DT repository usage: the static description of DT requires the ability of providing well-organized collections of information for describing the intended usage and behavior of the object. Also indications, instructions about the expected behavior in case of malfunctions or critical situations must be formatted and presented for automatic treatment. The validity, efficiency and usability of these solutions have to be proved by means of experimentations.
3. DT for fusion: the DT collects and represents different types of data describing the physical objects under different facets. The ability of DTs to store, mix and extract new information has to be extensively verified to identify a set of best practices for helping developers and practitioners in the large scale usage of this approach.
4. DT for Situation Awareness and explainability: the capability of the DT in helping the understanding and explainability of the behavior of products/objects in specific situations has to be proved and consolidated from simple to more complex cases. The differences between static versus dynamic DT in several contexts has to be evaluated in terms of benefits and complexity of the solutions.
5. Fully-fledged DT systems: large systems of DTs have to be created, operated and managed to optimize their usage and capabilities. Large use cases have to be first prototyped and then industrialized. The progression towards “horizontal” platforms of DTs has to be proved and justified by the easiness to program and the

possibility of representing the complexity of interacting vertical use cases.

The focus of our next activities in this field will concentrate on these steps and challenges by means of the prototyping of increasingly more complex use cases that will involve and use several AI technologies as well as basic data collection.

REFERENCES

1. K.Y.H Lim, P. Zheng, P., & C.H. Chen, (2019). A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, 1-25.
2. R. Martinez-Velazquez, R. Gamez and A. E. Saddik, "Cardio Twin: A Digital Twin of the human heart running on the edge," 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Istanbul, Turkey, 2019, pp. 1-6, doi: 10.1109/MeMeA.2019.8802162.
3. M. González, O. Salgado, J. Croes, B. Pluymers and W. Desmet, "A Digital Twin for Operational Evaluation of Vertical Transportation Systems," in *IEEE Access*, vol. 8, pp. 114389-114400, 2020, doi: 10.1109/ACCESS.2020.3001686.
4. W. Kinsner and R. Saracco, "Towards Evolving Symbiotic Cognitive Education Based on Digital Twins," 2019 IEEE 18th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), Milan, Italy, 2019, pp. 13-21, doi: 10.1109/ICCI*CC46617.2019.9146095.
5. R. Minerva, G. M. Lee and N. Crespi, "Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models," in *Proceedings of the IEEE*, vol. 108, no. 10, pp. 1785-1824, Oct. 2020, doi: 10.1109/JPROC.2020.2998530.
6. T. Ruohomäki, E. Airaksinen, P. Huuska, O. Kesäniemi, M. Martikka and J. Suomisto, "Smart City Platform Enabling Digital Twin," 2018 International Conference on Intelligent Systems (IS), Funchal - Madeira, Portugal, 2018, pp. 155-161, doi: 10.1109/IS.2018.8710517.
7. R. Barthel, K. Leder, A. Hudson-Smith, A. Karpovich, M. De Jode, and C. Speed "An Internet of Old Things as an Augmented Memory System", *Journal Of Personal and Ubiquitous Computing*, Springer, DOI: 10.1007/s00779-011-0496-8.
8. G. Laput, Y. Zhang, and C. Harrison. "Synthetic Sensors: Towards General-Purpose Sensing". In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, pp. 3986-3999. DOI: <https://doi.org/10.1145/3025453.3025773>
9. D. Miljković, "Sample Based Synthesis of Car Engine Noise," 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2020, pp. 1012-1017, doi: 10.23919/MIPRO48935.2020.9245323.
10. A. Alkilani, A. Shirkhodaie. "Acoustic signature recognition technique for Human-Object Interactions (HOI) in persistent surveillance systems." In *Signal Processing, Sensor Fusion, and Target Recognition XXII*, vol. 8745, p. 87451C. International Society for Optics and Photonics, 2013.
11. P.L. Mendonça, E.L. Bonaldi, L.E.L. de Oliveira, G. Lambert-Torres, J.G. Borges da Silva, L.E. Borges da Silva, C.P. Salomon, W.C. Santana, A.H. Shinohara, "Detection and modelling of incipient failures in internal combustion engine driven generators using Electrical Signature Analysis", *Electric Power Systems Research*, Volume 149, 2017, Pages 30-45, ISSN 0378-7796, <https://doi.org/10.1016/j.epsr.2017.04.007>
12. S. Ierace, M. Garetti, L. Cristaldi. "Electric Signature Analysis as a cheap diagnostic and prognostic tool." In *Engineering Asset Lifecycle Management*, pp. 750-757. Springer, London, 2010.
13. J. Wang, Y. Chen, S. Hao, X. Peng, L. Hu, "Deep learning for sensor-based activity recognition: A survey", *Pattern Recognition Letters*, Volume 119, 2019, Pages 3-11, ISSN 0167-8655, <https://doi.org/10.1016/j.patrec.2018.02.010>.
14. W. Ding, X. Jing, Z. Yan, L. T. Yang, "A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion, *Information Fusion*", Volume 51, 2019, Pages 129-144, ISSN 1566-2535, <https://doi.org/10.1016/j.inffus.2018.12.001>.
15. Z. Liu, N. Meyendorf, N. Mrad. "The role of data fusion in predictive maintenance using digital twin." In *AIP Conference Proceedings*, vol. 1949, no. 1, p. 020023. AIP Publishing LLC, 2018.
16. S. Nosratabadi, A. Mosavi, R. Keivani, S. Ardabili, F. Aram. "State of the art survey of deep learning and machine learning models for smart cities and urban sustainability." In *International Conference on Global Research and Education*, pp. 228-238. Springer, Cham, 2019.
17. A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," in *IEEE Access*, vol. 6, pp. 52138-52160, 2018, doi: 10.1109/ACCESS.2018.2870052.
18. A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI", *Information Fusion*, Volume 58, 2020, Pages 82-115, ISSN 1566-2535, <https://doi.org/10.1016/j.inffus.2019.12.012>.
19. F. M. Cau, "Explaining Black Box Models Through Twin Systems." In *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion*, pp. 27-28. 2020.
20. D.J. Rao, S. Mane. "Digital Twin approach to Clinical DSS with Explainable AI." *arXiv preprint arXiv:1910.13520* (2019).

Roberto Minerva (M'92, SM'19) received the M.S. degree (summa cum laude) in computer science from the University of Bari, Italy, in 1987 and doctoral degree in "computer Science and Telecommunications" from the Pierre and Marie Curie University – Sorbonne University, Paris, France in 2013. From 1987 to 1996, he was a researcher in the area of Service Architectures and Network Intelligence within Telecom Italia Research Center. In the following years, he was appointed responsible of several research groups related to Network

Department Head

Intelligence and evolution to next generation networks. From 2013 up to 2016 he was appointed to the Strategic Initiatives of TIM. From 2016 he was the Technical Project Leader of SoftFIRE, an European Project devoted to the experimentation of NFV, SDN and edge computing. From 2018, he is an Associate Professor in Softwarisation within the Service Architecture Lab of the Wireless Networks and Multimedia Services Department of Institut Mines Telecom Telecom Sud Paris part of Institut Polytechnique de Paris, Evry, Paris 91000, France. He is author of the book “Networks and new services: a complete story”. He is author of more than 50 articles in journals and international conferences. Dr. Minerva has been the chairperson of the IEEE IoT Initiative in the period 2014-2016. He is currently member of the Scientific Committee of the Fondazione Bruno Kessler in Trento, Italy. Contact him at Roberto.minerva@telecom-sudparis.eu.

Faraz Malik Awan is a research engineer and doing his Ph.D. at Telecom SudParis (Institute Polytechnique de Paris). He received BS Computer Science degree from COMSATS Institute of Information Technology, Pakistan (now known as COMSATS University Islamabad, Pakistan) in 2015. After graduation, he joined a German development agency, called Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) in 2015 and started working as Software Engineer. He completed MS degree from Chung-Ang University, Seoul, South Korea majoring in System Software in 2018. Along with MS degree, he worked as Research Assistant in Real-Time Software Engineering Laboratory. His research interest includes Deep/Machine Learning focusing on Smart Cities and Natural Language Processing. Contact him at: faraz_malik.awan@telecom-sudparis.eu

Prof. **Noël Crespi** holds Masters degrees from the Universities of Orsay (Paris 11) and Kent (UK), a diplôme d’ingénieur from Telecom ParisTech, a Ph.D and an Habilitation from UPMC (Paris-Sorbonne University). From 1993 he worked at CLIP, Bouygues Telecom and then at Orange Labs in 1995. He took leading roles in the creation of new services with the successful conception and launch of Orange prepaid service, and in standardisation (from rapporteurship of IN standard to coordination of all mobile standards activities for Orange). In 1999, he joined Nortel Networks as telephony program manager, architecting core network products for EMEA region. He joined Institut Mines-Telecom in 2002 and is currently professor and Program Director, leading the

Service Architecture Lab. He coordinates the standardisation activities for Institut Mines-Telecom at ITU-T and ETSI. He is also an adjunct professor at KAIST (South Korea), an affiliate professor at Concordia University (Canada), and guest researcher at the University of Goettingen (Germany). He is the scientific director the French-Korean laboratory ILLUMINE. His current research interests are in Data Analytics, Internet of Things and Softwarisation. Contact him at noel.crespi@mines-telecom.fr.