Digital Twin-empowered Integrated Satellite-Terrestrial Networks towards 6G Internet of Things

Tinh T. Bui, *Student Member, IEEE,* Long D. Nguyen, *Member, IEEE,* Berk Canberk, *Senior Member, IEEE,* Vishal Sharma, *Senior Member, IEEE,* Octavia A. Dobre, *Fellow, IEEE,* Hyundong Shin, *Fellow, IEEE,* and Trung Q. Duong, *Fellow, IEEE*

***Abstract*—Integrated satellite-terrestrial networks (ISTNs) technology in the sixth generation (6G) wireless networks has been considered as a promising candidate for global coverage and seamless connectivity for the Internet of Things (IoT). However, integrating these two complex systems poses many deployment, management, and maintenance issues. With the fundamental principle of building a virtual live representation of the networks, the digital twin (DT) technology can be used in complex ISTNs to provide a reliable environment for designing or testing, reduce risk and latency, recover the networks quickly, and optimise resource allocation for IoT devices in real-time. In this paper, we propose an ISTN framework empowered by DT technology and discuss its promising models, benefits, potential technologies, research challenges, and future research direction for 6G IoT.**

***Index Terms*—Integrated satellite-terrestrial networks (ISTNs), digital twin (DT), 6G, satellite communications.**

I. INTRODUCTION

The fifth generation (5G) wireless network is rolling out with significantly high key performance indicators in terms of the peak data rate (20 Gbit/s), low latency (1 ms) and ultra-dense connection density (10^6 devices/km^2), corresponding to its three major categories, namely enhanced mobile broadband, ultra-reliable and low latency communications (URLLC), and massive machine-type communications [1], [2].

However, the majority of the planet is not covered by cellular networks due to geographic and economic constraints on network infrastructure construction [3]. Until 2020, around 43.8% of the world's population is living outside urban areas, without the necessary infrastructure to access the internet. The sixth generation (6G) is expected to be the first to achieve global coverage [4] in addition to providing even higher data rates, latency, and network capacity compared to 5G. To this end, satellite networks complement terrestrial networks, providing ubiquitous wireless communication services to devices anytime and anywhere. Integrated satellite-terrestrial networks (ISTNs) have recently been considered a key enabler for global wireless services in 6G [5].

The ISTN architecture leverages and combines the different, featured strengths of satellite and terrestrial networks to provide universal connections for various services with diverse requirements. Satellite communication is comprised of different satellites such as geostationary equatorial orbit (GEO), medium earth orbit (MEO), and low earth orbit (LEO) satellites at the attitudes of 36000 km, 2000-36000 km, and 500-2000 km, respectively. Among them, LEO satellites require less power to transmit signals at high data rates, and their communication latency is much lower. Thanks to these advantages, LEO satellites are expected to provide global 6G communications in projects such as Starlink, OneWeb, and Telesat. However, it requires hundreds or thousands of LEO satellites to cover the Earth's surface. For example, the estimated numbers of LEO satellites to build mega constellations are 42000 with the Starlink network, 648 with the OneWeb network, and 300 with the Telesat network [3]. Unlike terrestrial networks, replacing and repairing the hardware on satellites within their life cycles after launch is challenging. Moreover, the huge number of satellites operating concurrently within the strict requirements of 6G communications poses many challenges to deployment, management, and maintenance. Integrating satellite networks into multi-layer and heterogeneous terrestrial networks remains a complex task in terms of routing, resource management, offloading, and evaluation to guarantee service continuity, reliability, and quality. Therefore, it is essential to build real-time replicas of ISTNs that can reflect the actual status of the whole network by using updated data.

The digital twin (DT) technology, which is widely used for modelling complex systems for the simulation, comprehensive analysis, monitoring, real-time decision-making, and

T. T. Bui was with Queen's University Belfast, UK and now is with Memorial University of Newfoundland, Canada (e-mail: ttbui@mun.ca).

L. D. Nguyen is with Duy Tan University, Vietnam (email: nguyendinhlong1@duytan.edu.vn)

B. Canberk is with Edinburgh Napier University, United Kingdom and Istanbul Technical University, Turkey (e-mail: b.canberk@napier.ac.uk).

V. Sharma is with Queen's University Belfast, UK (e-mail: v.sharma@qub.ac.uk).

O. A. Dobre is with Memorial University, Canada (e-mail: odobre@mun.ca).

H. Shin is with Kyung Hee University, South Korea (e-mail: hshin@khu.ac.kr).

T. Q. Duong is with Memorial University, Canada and also Queen's University Belfast, UK (e-mail: tduong@mun.ca).

The work of V. Sharma was supported by the UK Department for Science, Innovation and Technology under the Future Open Networks Research Challenge project TUDOR (Towards Ubiquitous 3D Open Resilient Network).

The work of H. Shin was supported by the National Research Foundation of Korea Grant funded by the Korean Government (MSIT) under Grant 2022R1A4A3033401.

The work of O. A. Dobre was supported in part by the Canada Research Chairs Program CRC-2022-00187.

The work of T. Q. Duong was supported in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109.

Corresponding authors are Trung Q. Duong and Hyundong Shin.

prediction, is a promising approach to tackling challenges in ISTNs [6]. Unlike terrestrial networks, satellite networks can not be built in a short period of time since the mega-constellation of LEO satellites is developed by gradually adding satellites one by one. In this progression, the optimal decision for choosing the right launching time and designing the satellites' orbits is essential to guarantee the expansion of satellite networks ensure not only the efficient cooperation with terrestrial networks but also adaptation to the increasing demand of users. Thus, digital replicas of ISTNs with real-time updating status using DT are needed for testing and optimising the new changes before actual deployment. Also, the DT technology in ISTNs provides a safe and reliable environment for validating new protocols, simulations, and applications. These new technologies imposed by strict constraints can be designed, evaluated and adjusted in virtual representations which have the same status as the real networks, subsequently limiting the risk during actual implementation. Additionally, with the support of the DT technology, the operators can also apply other emerging technologies such as real-time optimisation, artificial intelligence (AI), and mobile edge computing to improve network performance in ISTNs. Despite the necessity of DT in ISTNs, the research is still in its early stages with a few related works. In terms of network routing, in [7], an inter-satellite routing method was proposed with the support of DT to improve the quality of signalling between satellites while a DT network was considered as an environment to establish models for predicting switching and congestion in [8]. Regarding security, the authors in [9] introduced the space DT that is used for constructing a virtual ISTN to enable operators to assess security and evaluate countermeasures.

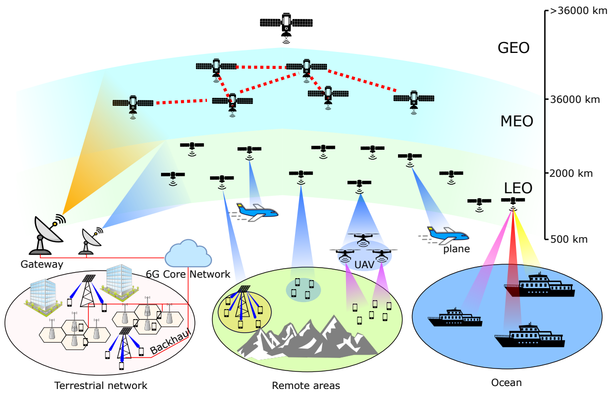

Fig. 1. An overview of ISTNs in the context of 6G global coverage.

In this paper, we propose a novel DT for ISTNs to provide ubiquitous connectivity with the integration between satellite networks and terrestrial networks. The remaining sections are organised as follows: Section II presents the general framework and the benefits of integrating DT into ISTNs. Based on this framework, in Section III, we propose potential technologies to leverage the unique characteristics of DT to tackle limitations in ISTNs. Then, Section IV introduces a simple case study to prove the efficiency of using DT for a mobile edge computing system in ISTNs, followed by the challenges and open research topics described in Section V. Finally, Section VI summarises key contributions of this study.

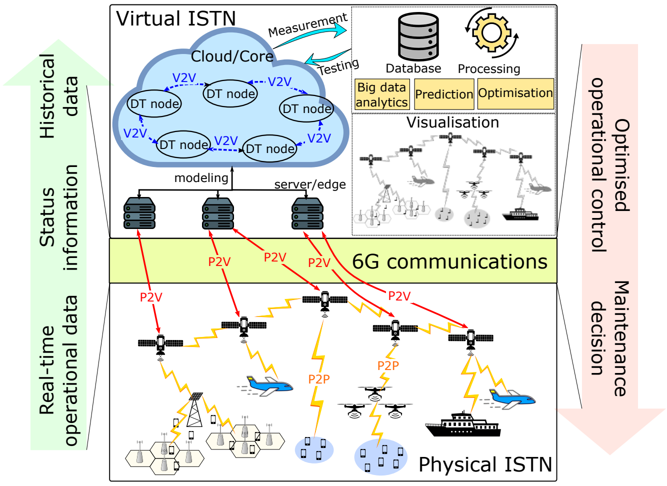

Fig. 2. A general framework of DT-ISTN.

II. DT-ENABLED 6G ISTNs

A. *Conceptual model of DT-enabled 6G ISTNs*

A typical model of ISTNs is proposed to leverage the positives of both satellite and terrestrial networks as in Fig. 1. Specifically, the ground base stations (BSs) can serve as relays to improve network reliability, as Internet of Things (IoT) devices can receive signals indirectly from the ground BSs [3]. As an extension of this architecture, the satellites provide backhaul links to the BSs not connected to the core network through any optical fibre link. In addition, the satellite networks can be complementary for offloading data traffic when the terrestrial networks are exhausted by a significant increase in the number of devices and data traffic, especially during peak time or in rural areas where the infrastructure is less developed. Finally, to tackle significant signal attenuation in the satellite-to-user links, satellite-air-terrestrial networks can be employed with the support of aerial networks, such as balloons, aeroplanes, and unmanned aerial vehicles (UAVs). Particularly, the balloons can fly continuously for several years to provide seamless connection. Dozens of million flights worldwide in a year allow aeroplanes to serve user devices in areas that they cross by. Meanwhile, despite short flying duration, UAVs still have the advantages of flexibility, optimisation, and quick deployment to provide temporary connections in hot spots or areas in the aftermath of a disaster. With the architecture of a combination of multiple layers, i.e., satellites, aerial vehicles, and ground BSs, ISTNs can leverage the unique positives of each one to provide global connectivity. Satellites with large coverage are used to provide ubiquitous connections, and the flexibility of UAVs is exploited to provide quickly high-quality line-of-sight links, while reliable and low-latency communication to users is served by ground BSs.

Although traditional simulation tools such as HYPATIA in [10] offer functionalities for network simulation and visualisation, the DT technology provides a more robust and comprehensive approach. Firstly, DT is a multi-scale model consisting of detailed component models to high-level simulations such as communication, and environments. This allows operators to zoom in on specific points or zoom out for a broader view when necessary. Secondly, real-time status updating facilitates the evolution of digital networks continually together with real networks. Thirdly, DT provides flexible models which can be adjusted in any part to develop new technologies, and protocols or effectively simulate abnormal scenarios that can happen in the future. Fig. 2 shows an overall model of DT-ISTN in 6G for global coverage with three distinct components:

- Physical ISTN consists of network elements such as satellites, balloons, aeroplanes, UAVs, ships, terrestrial BSs, and IoT devices. They are connected by physical-to-physical communication (P2P) to exchange data over various communications protocols.
- Virtual ISTN, where the DT technology is integrated into ISTN, is the core component and consists of DT networks. The connections between elements in DT networks are virtual-to-virtual communication (V2V) that reflects the P2P and allows the DT nodes to share their information and status [11]. Different from the P2P which depends on radio power, spectrum, latency, and channels, the V2V takes place in the digital world and depends on the computing capacity of DT servers and the given information of the model. Therefore, simulations in DT networks do not suffer from long latency as in physical ISTN and can be done quickly with powerful computing servers. Indeed, assume that a large amount of data needs to be transmitted from satellites to ground users, the latency in the DT network consists of copying that data from digital satellite twins to user twins with the changes caused by channels which are modelled by mathematical equations. This is useful when applying real-time optimisation or AI for testing the scheduling or resource allocation solutions multiple times to get different measurements and choose the best one, which otherwise would take considerably longer in a physical ISTN. Additionally, visualisation technologies can be integrated to enable network operators to observe and manage the networks more efficiently.
- Physical-to-virtual communication (P2V) is responsible for real-time feedback by updating information from physical objects of the ISTN to DT nodes, including real-time operational data, status information from sensors, and historical data, while the reverse way from the digital world to the physical world is used for transmitting optimised operational control and maintenance decisions. However, it is a non-trivial task to connect all network elements in ISTNs due to physical limitations. Therefore, satellites can be responsible for communicating with servers or edge servers where the updated data is synthesised for modelling DT networks, as well as where feedback data is transmitted to satellites. These servers and edge servers are placed at ground core networks and gateways to communicate with satellites, or they can be placed at high-orbit satellites such as GEO and MEO satellites, and use inter-satellite links to create distributed DT networks. Although real-time feedback through P2V is required in both software-defined networking and DT, there are key differences making DT more suitable for ISTNs: (i) DT technology allows the modelling of ISTNs from communication level to network element level where devices in a satellite have their replica in digital world; and (ii) DT technology together with cutting-edge technologies such as AI and real-time optimisation enables prediction and proactive response to issues.

B. *Outstanding benefits of DT-enabled 6G ISTNs*

Compared with traditional terrestrial networks, ISTNs are much harder to deploy, control, and especially make changes in hardware since satellites orbit in space at ultra-high speed (up to 7.65 km/s for LEO satellites). The integration of DT into ISTNs can bring several profound advantages:

- Deployment: As a complex and costly task, launching satellites into space requires detailed and thorough plans as well as a prediction of their duty during the whole life cycle. Additionally, due to lower active altitudes compared to the other types of satellites, the orbits of LEO satellites are more easily decayed by the friction in the atmosphere. In ISTNs, the terrestrial networks are usually built before the satellite ones. Thus, efficient cooperation between them is essential to be guaranteed when adding new satellites to the networks. With two-way communication between physical entities and DT, the network operators can build a digital network consisting of twins, which fully describes the current status of the physical satellites, terrestrial BSs, and communication environments, i.e. space and ground. In this way, before launching a new satellite to the existing ISTN, its new replica can be used for testing, evaluating performance, and simulating in the DT-ISTN to choose the optimal orbit and optimal device parameters, such as the number of antennas, computing capacity, energy battery.
- Management: To achieve global coverage, dozen thousands of satellites are predicted to be launched into space shortly with the orbits assigned by the International Telecommunication Union, yet they are authorised by distinct companies and nations. Together with terrestrial networks, it is thus a challenge to manage such networks with a massive number of elements, huge data traffic, and plenty of computational requirements. Additionally, due to large coverage and quick movement, LEO satellites can be used to serve many areas with different sub-terrestrial networks and varying densities of users. With the support of DT technology, operators can use digital representations, which can fully describe the status of terrestrial networks, satellites' standard information, locations, and connections to user devices to determine optimal resource allocation, unexpected malfunctions or predict and pre-solve conflicts. For example, the orbits of satellites from different vendors can change over time unexpectedly, causing the possibility of collisions, especially with the increasing number of satellites. Currently, thousands of satellites have been operating simultaneously in space where there are thousands of large debris objects and

millions of medium and small ones. It would be very dangerous and uncontrollable since space junk from only one satellite collision can cause a domino effect and damage the whole space. Thus, it is essential to build a digital replica of the ISTNs that not only manages the physical networks in real time but also predicts potential problems depending on the received information of location and sensing, and provides the optimal solution to avoid them. DT additionally offers a safe environment for simulating worst cases of collisions and evaluating the effects on the risk of other next collisions and the network performance from a communication perspective.

- Maintenance: The employment of DTs can improve the quality of the maintenance process in ISTNs for several applications. The maintenance requirements of LEO satellites are much stricter than those in current terrestrial networks since all maintenance activities have to be done remotely after launch. Therefore, a digital network replicating real ISTNs is essential for predicting possible failures and evaluating new maintenance actions before actual employment. Firstly, when there is an error in the operating process, a full-characteristic digital model and real-time updated information through the communications between physical elements and virtual ones will enable operators to quickly monitor and determine exactly where is the error. Then, multiple different methods and maintenance scenarios are tested and evaluated for their impact on the virtual model so that the optimal one is chosen to minimise risk and response time. Onsite maintenance is impossible with satellites if there is a failure in hardware, in which case software updates can be sent to satellites to activate built-in alternative parts. Secondly, DTs can use machine learning (ML) to analyse and learn historical data of the networks and the status of equipment as well. Then, the models are used to predict when equipment failures are likely to happen to make scheduled maintenance proactively so that the probability of interruption in service decreases. Finally, DTs provide a real-time, precise view of ISTNs, supporting network operators to use optimisation algorithms to adjust power, channel allocation, routing, handover process, precoding design, and many other parameters.

III. POTENTIAL TECHNOLOGIES APPLICABLE TO DIGITAL TWIN-ENABLED ISTNS

A. *Real-time optimisation in DT-enabled 6G ISTNs*

In ISTNs, high latency due to long-distance transmission is one of the bottlenecks to fulfilling the stringent requirements in 6G. As aforementioned, V2V in DT networks, which replicates real ISTNs, can be simulated at a low time cost. Additionally, real-time optimisation in DT networks helps reduce the processing time in ISTNs significantly while the decisions are guaranteed to be optimal. This combination of real-time optimisation and DT can be used to improve performance. For example, in multi-beam design for satellites, narrow beams from several arrays of thousand antennas need to be adjusted in real time to communicate with IoT devices, gateways, and other satellites simultaneously since the positions of both satellites and user devices change quickly. In DT networks, real-time optimisation can be used for pre-optimising the beamforming design according to predicted data before the satellites move to a position. Then the satellites can employ this solution as a reference for the current status and make minor changes to achieve an optimal solution with current parameters.

Moreover, as an extension of terrestrial networks, the signals from satellites or flying machines can interfere with cellular devices, while ground BS operations interfere with satellite devices. However, the majority of resource allocation problems aiming to mitigate the interference in ISTNs are usually non-convex and large-scale with many variables due to the massive number of IoT devices, satellites, BSs, and flying machines. Therefore, achieving optimal or near-optimal solutions in real-time is challenging but essential in this context. Another application is to optimise user association and handover processes, which happen frequently in ISTNs. It takes a LEO satellite at the altitude of 1000 km approximately 105 minutes to complete a full orbit, so with a gateway that can steer beams within a very wide range, the handover process happens in approximately 30 minutes. Additionally, in case of a high number of LEO satellites operating, the LEO satellite which provides the best quality of signal to a gateway changes quickly, leading even multiple handover processes happening within 30 minutes.

To achieve real-time optimisation, especially for large-scale non-convex problems with a massive number of integer and real variables as well as different structures, i.e., scalar, vectors, matrices, or even tensors, we propose the following approaches.

- Distributed computing: Most recent computers are multi-processor computing devices, powerful enough to break a complex problem into multiple simple ones to compute in parallel. In ISTNs, optimal radio resource allocation is a large-scale problem, where the computation complexity increases exponentially with respect to the variables and constraints. Therefore, splitting them into small-scale problems is one of the efficient approaches to achieving real-time computing. Although distributed approaches cannot always guarantee global optimality, especially for NP-hard problems which require specific methods for specific problems, the positives compared to centralised methods are lower complexity, allowing different suitable methods for sub-problems, and continue functioning even if some individual nodes suffer damages. In addition, game theory can also be used as a distributed tool for solving optimisation problems where players (i.e. multiple network elements) interact in compliance with game rules. As an example of distributed computing in ISTNs, in multi-beam design, the large-scale antenna array can be divided into multiple small sub-arrays to focus the narrow beams in desired directions.
- Approximation: Most optimisation problems in ISTNs are non-convex and extremely complex. Approximating objective functions and constraints by convex functions to relax the problems is essential to reduce the com-

putational complexity. There are many tools for efficiently solving convex optimisation problems, e.g., CVX, CVXPY, JuliaOpt, and GUROBI. However, the complexity of a problem also depends on the forms of objectives and constraints. Thus, to achieve real-time computing in large-scale problems, the approximation of the constraints and objective functions should be given in the form of affine functions, analytical solutions, or evolutional methods, which do not require calculating the derivative of functions.
- ML: Solutions to convex problems should be learned by ML algorithms to guarantee the efficiency of learning instead of non-convex problems since local optima in non-convex ones can probably give a low performance which should be avoided in the learning process. Moreover, choosing the optimal initial point in an iterative algorithm is very important to localise the zone of global optimum, but usually ignored since it is really hard to prove the efficiency of this initial point. With historical information, ML can be used to choose potential initial points for an iterative optimisation algorithm, resulting in the solution after convergence being global optimum or much better compared with the random initial points.

These approaches can be used cooperatively for forming amalgamations to find optimal solutions more efficiently.

B. *AI and ML-enabled DT-ISTNs*

In DT-ISTNs, AI algorithms can be used to analyse historical data and data generated by DT networks to identify future behaviour, optimise performance, and improve reliability. For example, data from network traffic or data from sensors that are transmitted to the DT networks can be used for training AI models to extract featured characteristics and cyclical changes. In this regard, AI models are able to monitor the networks, predict areas and times of high traffic congestion, and then adjust resource allocation strategies accordingly to optimise resource efficiency. Besides, the DT networks can be used as an environment for testing abnormal scenarios (e.g. security risks), and these solutions can be used for training AI models to make networks more robust to the same situations in the future. Moreover, thanks to the characteristics of intelligence, low online computing time (testing time), high precision, and prediction, DL algorithms can be used for reducing delay and optimising resource allocation in ISTNs [12].

- Supervised learning: Huge labelling data sets are required in order to predict more accurately, which can be costly, time-consuming, and even prone to errors. With the support of the DT technology, automatically labelling data sets captured by sensors and other data sources can be generated precisely by DT networks to tackle this issue. For example, in sub-channel allocation, according to the characteristic of virtual channels in DT, the packets of channel gain are generated, and an optimisation algorithm can be used to find an optimal decision as a label for supervised learning.
- Semi-supervised learning: The historical data stored in DT database can be used for training as labelled data while the remaining unlabelled amount is generated by DT networks from futuristic simulations.
- Reinforcement learning (RL): Reinforcement learning is suitable to be used in DT networks. Instead of building complex mathematical equations for describing the environment, the actions of agents in RL can be tested directly in DT networks which are the real-time estimated version of actual networks.

Aside from the aforementioned learning methods, two advanced breakthrough approaches can be applied to DT-ISTNs, namely federated learning (FL) and transformer architecture used in generative AI. Regarding FL, privacy is the most important factor since FL allows models to be trained separately at network nodes, for example, in satellites. This decentralised method also reduces the complexity of the training phase with low energy consumption and avoids data backhauling overhead to core networks through satellites. DT networks can be used to allocate the computing capacity of network nodes depending on the status of DT nodes which are continually updated by data from physical ISTNs. In terms of transformer architecture which is currently well-known for ChatGPT of OpenAI, the relationships in serial data can be analysed depending on self-attention and feed-forward networks, and the generative AI model is used to generate possible changes in the future. In the scenarios of ISTNs, data that have the serial characteristic can be used for training transformer models from measuring sensors in IoT devices, user requests, path loss, and beamforming design for the whole orbit of LEO satellites (around 90-120 minutes to complete a full orbit). These trained generative AI models can be used to create possible parameters for applications such as testing, pre-optimisation, or training other AI models.

C. *URLLC-aided mobile edge servers*

The use of URLLC-aided mobile edge servers (URLLC-MES) can be an efficient solution for supporting computing-intensive applications such as games, autonomous driving, and face or speed recognition in ISTNs [13]. In conventional ISTNs without MES, for offloading heavy tasks, user devices have to transmit a part of commands and data to satellites, and then they are forwarded to core networks for processing, resulting in long latency. Therefore, MES should be deployed closer to end devices, for example, in satellites and ground BSs; in order to reduce latency, it is required to enable real-time offloading and real-time computing. Offloading in ISTNs is a mixed-integer non-convex optimisation problem that is not only complex but also large-scale in the scenario of ISTNs. In this case, the DT technology can be used for modelling the computing capacity of network elements and, depending on the current status from real-time updating of edge servers and user devices, to provide the optimal offloading decision. Proactive AI models in DT can be trained by historical data of users' demand and channel status which is saved in the database of DT, and a massive amount of data generated quickly by DT. Thus, depending on the current demand, AI models can predict the demand and the channel quality for the next period and pre-optimisation methods can be applied to give approximate solutions. With the support of DT, offloading in ISTNs can avoid data and computing overhead at edge servers which have poor connections and large waiting queues for unprocessed requests since the up-to-date status of servers is considered. Subsequently, employing URLLC-MES to DT-ISTNs can help to optimise energy efficiency and improve the quality of service for end devices.

IV. DT-ENABLED 6G ISTNS: IMPACT OF IMPERFECT ESTIMATION OF PHYSICAL PARAMENTERS

To demonstrate the efficiency of the DT technology on 6G ISTNs, we evaluate the impact of imperfect estimations of physical parameters at ISTNs on the performance of DT-ISTNs. Here, we conduct simulations for a DT-enabled ISTN mobile edge computing (MEC) system, where satellites serve as remote edge servers (ESs) to partially execute offloaded tasks from IoT devices. This system allows computational tasks from IoT devices to be executed quickly with respect to the constrained resource budget of IoT devices, i.e., limited processing rate and small storage battery. In the ISTN, 4 satellites orbiting at the altitude of 1000 km above the ground occupy equal coverage area in the considered region of 2000×2000 km with 50 ground users in total. Each user has one task with a size of 5 MB and the required computation resource from 100 to 1400 Mega cycles. The processing rate of each satellite is 3 GHz while that of each user varies from 1 to 1.4 GHz. The other simulation parameters are established based on the settings in [14], [15]. The latency metric is chosen as the optimisation objective for minimisation. The simulation results provide a general reference for choosing suitable DT with different accuracy in ISTNs while guaranteeing the requirement of minimum latency. The error value of DT models denotes the difference between the real value and the estimated value of the processing rate of MEC servers at the satellites.

Fig. 3 demonstrates how DT estimation and the required computation resource of tasks affect the end-to-end (e2e) latency of the IoT devices. It is seen that the more accurately the DT estimates the processing rate of ESs and IoT devices, the lower latency is obtained. This reflects the implementation of DT in real-world systems, where DT services demand accurate estimation. Additionally, Fig. 3 indicates that the required computation resource significantly impacts the average latency of IoT devices. Complicated tasks naturally request more computation resources to execute, which leads to an increase in processing latency as well as the overall e2e latency. Therefore, it is crucial to implement an effective MEC architecture for time-sensitive applications by applying offloading solutions.

To illustrate the task offloading behaviour with different settings of the UE processing rate, Fig. 4 displays the average offloading portions versus different processing rates of IoT devices. In particular, when the processing rate of IoT devices increases, the average offloading portion experiences a sustainable decrease. This behaviour confirms the effectiveness of the implemented offloading solution. When the IoT devices are more powerful, they can locally execute their tasks more quickly; therefore, a lower portion of tasks is offloaded to ESs. Importantly, the impacts of DT estimation are also clearly demonstrated in Fig. 4. Again, the more accurate the DT estimates are, the better the performance is.

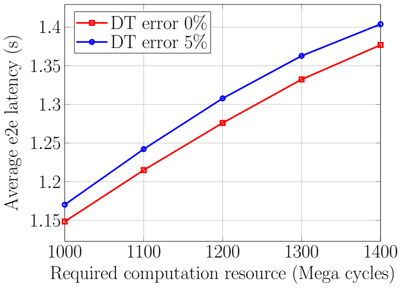

Fig. 3. Impacts of required computation resource and DT estimation error on the latency in satellite-enabled MEC.

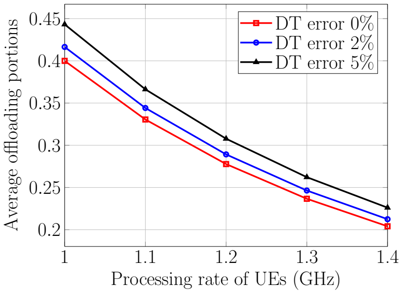

Fig. 4. Impacts of UE processing rate and DT estimation error on the offloading behaviour.

V. RESEARCH CHALLENGES AND FUTURE WORKS

Full DT models hold significant promise for future wireless systems. With the advancements in software engineering, AI, and robust information infrastructure, the full DT-enabled ISTNs in real-world applications will be widely deployed in various networked systems to support many domains such as autonomous vehicles, ubiquitous wireless services, tracking, and envisioned to be a part of 6G. However, due to the early stage of employment of DT models for ISTNs, some challenges require more research.

A. Two-way real-time communications

DT networks receive real-time updating data from physical elements in ISTNs and transmit control signals back. Therefore, two-way real-time communications in DT-ISTNs require reliability and low latency. As mentioned in Section II-A, P2V should be the connection between satellites and the servers, resulting in the requirement of two-way connections between satellites and the other network elements, e.g., IoT devices, UAVs, aeroplanes.

Different from cellular networks where user devices connect to BSs through two-way connections, the uplink transmission directly from IoT devices in ISTNs is a big challenge since the power of IoT devices is low and limited. Using UAVs or ground BSs as relay nodes is a promising solution to this problem. Additionally, due to the long-distance transmission and the quickly changing environment, the links between satellites and servers may suffer from connection loss, high attenuation, and high delay. In this case, AI algorithms can be used for predicting missing information before the real one is updated; however, they can require a huge amount of data for training. Thus, this can be an open topic for future work.

B. DT modelling challenge

For the replication of the physical world, it is important to quantify the impact of imperfect estimation of the physical parameters on the DT network performance as illustrated in Sec. IV. DT networks usually consist of DT nodes and V2V connections, described by mathematical equations and computational functions since DT networks operate completely in the digital/virtual world. With the inherent complexity of integrating satellite and terrestrial networks, a massive number of such equations and programming functions, which are originated from diverse multidisciplinary theories such as physics, mechanics, and wireless communications, are required to build DT networks for ISTNs. Clearly, to replicate a physical system more accurately, these functions are expected to be complex, while the computational model is required to be simple to allow fast responses and low energy consumption; this represents a challenge in designing the DT-ISTNs. A feasible approach for building a DT for ISTNs is by starting with one or more simpler models that tackle practical tasks of high priority such as resource allocation, predicting data overhead, and interference mitigation, following extending the remaining parts with the support of cloud computing, distributed computing, and ML.

Additionally, data exchange to adjust the digital networks as same as possible to the physical ones is also a challenge in DT-ISTNs. The key difference between the DT technology and other traditional simulators is real-time status updating. However, to provide global connections, ISTNs have to be built on many satellites, terrestrial BSs, gateways, devices, and many other elements. This leads to the requirement of exchanging a massive amount of data for updating information between DT networks and physical networks in DT-ISTNs. Therefore, to avoid data overhead, big data techniques and ML algorithms are required to analyse, process, integrate, and manage the massive volume of data.

C. Quantum technology in DT-ISTNs

Quantum technology is an emerging solution that can make resolutions for real-time computing and ultra-secure communications in wireless networks, including ISTNs, with two key properties: superposition and entanglement. Firstly, quantum computers may outweigh conventional ones in solving some optimisation problems in ISTNs. For example, in terms of evolutionary algorithms, conventional computers search for the optimal solution serially or even parallelly with a limited number of processors, while some qubits in quantum computers can represent all feasible solutions, resulting in finding the optimal solution much more quickly. Thus, real-time optimisation which is achieved by using quantum computing is applied to solve large-scale and highly complex optimisation problems in ISTNs. Secondly, the superposition of qubits can also be used to create unbreakable cryptographic keys to build a scenario of ISTNs with perfect security. In DT, the updating data of network nodes' status is collected to a simulation point; this poses the threat of private data leak. Quantum communication with ultra-high security can be used for data exchange between DT and physical nodes. Thirdly, thanks to the characteristic of entanglement, quantum teleportation, which transfers instantly the quantum state from one particle to another regardless of how far they are, can be used as a new resource for communications together with conventional ones such as power, frequency, and bandwidth. This is suitable for creating a new communication method not affected by high-attenuation and long-latency links in ISTNs. However, it requires much effort in research to employ and leverage the full potential of quantum technology in ISTNs due to the challenges of hardware, software, the integration between quantum and classical systems, and even cost.

VI. CONCLUSIONS

In this paper, we first presented the dominance and featured models for digital twin (DT)-enabled integrated satellite-terrestrial networks (ISTNs) to support the global coverage vision in 6G. Then, we introduced three potential technologies (i.e., real-time optimisation, AI, and URLLC-aided MES) that leverage the DT technology to improve the network performance in ISTNs. As a case study, we investigated a MEC system where the satellites act as remote edge servers; at the same time, DT is used for estimating the computing capacity and optimising the usage. Finally, we presented fundamental important challenges and future works to facilitate the full potential of DT-enabled ISTNs to be efficiently applicable in real systems such as two-way real-time communications, modelling, and quantum technology.

ACKNOWLEDGEMENT

The work of V. Sharma was supported by the UK Department for Science, Innovation and Technology under the Future Open Networks Research Challenge project TUDOR (Towards Ubiquitous 3D Open Resilient Network). The work of H. Shin was supported by the National Research Foundation of Korea Grant funded by the Korean Government (MSIT) under Grant 2022R1A4A3033401. The work of O. A. Dobre was supported in part by the Canada Research Chairs Program CRC-2022-00187. The work of T. Q. Duong was supported in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109.

REFERENCES

[1] M. Series, "Minimum requirements related to technical performance for IMT-2020 radio interface (s)," ITU, Geneva, Switzerland, Tech. Rep. 2410-0, Nov. 2017.

[2] P. Jonsson *et al.*, "Ericsson mobility report," *Ericsson*, Nov. 2022.

[3] X. Zhu and C. Jiang, "Integrated satellite-terrestrial networks toward 6G: Architectures, applications, and challenges," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 437–461, Jan. 2022.

[4] H. Saarnisaari *et al.*, "A 6G white paper on connectivity for remote areas," *arXiv:2004.14699*, Apr. 2020.

[5] T. Q. Duong *et al.*, "Machine learning-aided real-time optimized multibeam for 6G integrated satellite-terrestrial networks: Global coverage for mobile services," *IEEE Netw.*, vol. 37, no. 2, pp. 86–93, Mar. 2023.

[6] ——, “From digital twin to metaverse: The role of edge intelligence-based 6G ultra-reliable and low latency communications,” *IEEE Wireless Commun. Mag.*, vol. 30, no. 3, pp. 140–146, Jun. 2023.

[7] L. Zhao *et al.*, “INTERLINK: A digital twin-assisted storage strategy for satellite-terrestrial networks,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 5, pp. 3746–3759, Oct. 2022.

[8] X. Jiang *et al.*, “Research on satellite QoS routing algorithm based on digital twin,” in *Proc. ICWOC*, Chongqing, China, Jun. 2023, pp. 118–122.

[9] Z. Lai *et al.*, “Space digital twin for secure satellite Internet: Vulnerabilities, methodologies and future directions,” *IEEE Netw.*, Nov. 2023.

[10] S. Kassing *et al.*, “Exploring the ’internet from space’ with HYPATIA,” in *Proc. ACM Internet Meas. Conf.*, Oct. 2020, pp. 214–229.

[11] Y. Wu, K. Zhang, and Y. Zhang, “Digital twin networks: A survey,” *IEEE Internet Things J.*, vol. 8, no. 18, pp. 13 789–13 804, Sep. 2021.

[12] P. Zhang *et al.*, “AI-enabled space-air-ground integrated networks: Management and optimization,” *IEEE Netw.*, Apr. 2023.

[13] M. Bennis, M. Debbah, and H. V. Poor, “Ultra-reliable and low-latency wireless communication: Tail, risk, and scale,” *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.

[14] T. Q. Duong *et al.*, “Real-time optimal multibeam and power allocation in 5G satellite–terrestrial IoT networks,” in *Proc. IEEE GLOBECOM*, Rio de Janeiro, Brazil, Dec. 04-08, 2022, pp. 5619–5624.

[15] Q. Tang, Z. Fei, B. Li, and Z. Han, “Computation offloading in LEO satellite networks with hybrid cloud and edge computing,” *IEEE Internet of Things J.*, vol. 8, no. 11, pp. 9164–9176, Jun. 2021.

Tinh T. Bui (Student Member, IEEE) was with Queen’s University Belfast, UK. He is now working toward his PhD degree at Memorial University, Canada.

Long D. Nguyen (Member, IEEE) is currently with Dong Nai University as an Assistant Professor and Duy Tan University as an Adjunct Assistant Professor.

Berk Canberk (Senior Member, IEEE) is a professor at Edinburgh Napier University, UK and a professor at Istanbul Technical University, Turkey.

Vishal Sharma (Senior Member, IEEE) is currently working as a Senior Lecturer (Associate Professor) at Queen’s University Belfast (QUB), Belfast, U.K.

Octavia A. Dobre (Fellow, IEEE) is currently a Professor and the Canada Research Chair Tier 1 at Memorial University, St. John’s, NL, Canada.

Hyundong Shin (Fellow, IEEE) is now a Professor at the Department of Electronic Engineering and Department of Electronics and Information Convergence Engineering.

Trung Q. Duong (Fellow, IEEE) is a Canada Excellence Research Chair and professor at Memorial University, Canada. He is also an adjunct professor at Queen’s University Belfast, UK and a Research Chair of Royal Academy of Engineering.