Toward Digital Twin Oriented Modeling of Complex Networked Systems and Their Dynamics: A Comprehensive Survey

JIAQI WEN, BOGDAN GABRYS, (Senior Member, IEEE), AND KATARZYNA MUSIAL

Complex Adaptive Systems Laboratory, Data Science Institute, University of Technology Sydney, Sydney, NSW 2007, Australia

Corresponding author: Jiaqi Wen (14125367@student.uts.edu.au)

ABSTRACT This paper aims to provide a comprehensive critical overview on how entities and their interactions in Complex Networked Systems (CNS) are modelled across disciplines as they approach their ultimate goal of creating a Digital Twin (DT) that perfectly matches the reality. We propose four complexity dimensions for the network representation and five generations of models for the dynamics modelling to describe the increasing complexity level of the CNS that will be developed towards achieving DT (e.g. CNS dynamics modelled offline in the 1st generation v.s. CNS dynamics modelled simultaneously with a two-way real time feedback between reality and the CNS in the 5th generation). Based on that, we propose a new framework to conceptually compare diverse existing modelling paradigms from different perspectives and create unified assessment criteria to evaluate their respective capabilities of reaching such an ultimate goal. Using the proposed criteria, we also appraise how far the reviewed current state-of-the-art approaches are from the idealised DTs. Finally, we identify and propose potential directions and ways of building a DT-orientated CNS based on the convergence and integration of CNS and DT utilising a variety of cross-disciplinary techniques.

INDEX TERMS Complex network systems, digital twins, dynamic processes, network dynamics.

I. INTRODUCTION

A complex network can be seen as a universal concept used for representation and analyses of complex systems. Given the growing interest in real complex systems and a fast development of modelling techniques, the complex networked system (CNS) area has become a highly cross-disciplinary field that involves multiple modelling approaches with various research aims posed and achieved over the years.

There is a considerable literature about complex networks and various researchers have published several surveys reviewing and exploring the topic from different perspectives and application areas. Those include works on complex networks and their applications covering multiple application areas [1] or orientated towards specific topics such as networks of cryptocurrency transactions [2], vehicular networks [3], supply chain networks [4], internet of things [5] and networks of short written text [6].

When it comes to the models of complex networks in the context of their structure, there is also a body of work surveying a variety of network topologies [6], [7] and dynamics. The networks dynamics can be either considered as: (i) dynamic processes over networks, that involve surveys on spreading processes like epidemic processes, information spreading processes [8], [9], or (ii) dynamic networks with evolving structures and features [10], [11].

When it comes to modelling techniques that have been applied to build complex networked systems, there are also surveys that review certain types of modelling approaches for complex networks, including Graph Neural Network [10], game theory [12] or non-parametric Bayesian modeling [13].

Complex networked systems are modelled with the goal of accurate reflection of reality with the aim of simulation, prediction and/or control. Over the years, proposed CNSs models have become more and more accurate with more

The associate editor coordinating the review of this manuscript and approving it for publication was Yang Tang.

realistic networks topologies, characteristics and evolving dynamics modelled. Increasingly, they can capture features of real world scenarios and behave like their twins. Researchers have already focused on the studies of Digital Twining of real systems across disciplines, which involves a wide range of applications but so far complex networks area has been marginalised in this development space. Digital Twins are expected to provide an accurate reflection and extension of the reality fulfilling simultaneously multiple aims of modelling. In our review of the current state-of-the-art of the modelling of CNS, a DT is treated as an ultimate goal for the representation and modelling of the Complex Networked Systems, which hasn't been realized yet but is expected to be steadily approached with the ability to take more complex information into account, capture and represent dynamics of and on the networks as well as simulate and interact with the modelled reality in real time.

Therefore, the convergence of Digital Twinning and Complex Networked Systems emerges as an exciting research focus with a potential to address some of the outstanding modelling and representational challenges ultimately leading to the establishment of a DT-orientated CNS area as its main goal. And in this space, the modelling of CNS in the context of DT is to represent the observable information faithfully in a form of a network together with dynamics of and on this network, while enabling the fulfilment of model's aim and, in its ultimate realization, seamlessly intertwining and interacting with reality in real time.

Multiple modelling paradigms, with their ways of data analysis, representation and modelling, have been applied to answer various questions about CNSs. By modelling paradigm we understand a family of approaches, which are built around common principle (e.g. rule–based systems), that can be used to build a model of a system. There is a need to review and explore where, how and why complex networked systems are modelled across many different disciplines. However, most of the surveys only account for certain types of complex networked systems from a specific perspective or a single discipline (e.g. supply chain networks [4], network topology [6], [7], etc.). Therefore, to fill in this gap, we review conceptually and compare research on complex networks from a holistic point of view while trying to deal with the questions concerning diverse modelling paradigms and their distance to the idealised DTs. We devise a framework that enables to compare various modelling paradigms and evaluate their respective distances to DTs, which involves answering the following four fundamental questions: (1) What is the aim of the modelling?; (2) How to represent information about a system in a form of a network?; (3) How to model the dynamics in a networked system?; and (4) How do we approach the ultimate goal of building a CNS that models the reality at a Digital Twin level?

Hence, while answering the above questions, this critical survey aims to integrate and overview the current, relevant state-of-the-art from multiple disciplines and inform future research directions and foci for this new multidisciplinary area. Four complexity dimensions and five generations of models are proposed to describe the increasing complexity level of CNS that has been/will be developed towards a DT in terms of the network representation and the dynamics modelling. Another contribution of this survey is a unified assessment framework that is proposed to evaluate how far the reviewed CNSs are from the idealised DTs.

Our paper is organised as follows. Section II-B illustrates the prerequisites of setting and fulfilling model's aims, discusses different model's aims across different disciplines and summarises their research foci. This is followed by two sections concerned with complex networked systems with section III defining and reviewing different network representations and modelling paradigms and section IV focusing on modelling dynamics in networked systems. In Section V a new framework is proposed for evaluating models of complex network systems given the ultimate goal of achieving DTs faithfully representing the reality. The research gaps and future research directions are also identified and discussed in this section. Finally, the Conclusions are provided in the last section.

II. THE AIM OF MODELLING COMPLEX NETWORKED SYSTEMS

There is always a need to determine the aim of developing a new model at the start of the modelling process, where, under the constraints of observability, we model a Complex Networked System in a way to fulfill this aim. The choice of modelling paradigms, which can be understood as the approaches employed to model real state or dynamics of the system, is largely dependent on the model's aims determined based on the research questions. The aims of modelling Complex Networked Systems include solving a wide range of tasks such as link prediction, network detection, mimicking of real systems, etc. The CNS models together with their respective primary aims are also typically developed within different disciplinary focuses such as epidemiology, sociology, microeconomics, etc. We categorise the modelling paradigms based on their concrete model's aims, which helps us to understand what is expected from modelling complex network systems and serves as the basis for further discussions of detailed modelling paradigms. In this section, having in mind various model's aims that influence how we build such a model, we also set an ultimate goal of modelling complex networked systems, which is to create a Digital Twin (DT) of a real world system. Achieving such a goal would enable the fulfilment of multiple aims simultaneously through accurate reflection and mimicking of the reality.

A. PREREQUISITES OF MODELLING
Selecting the right modelling paradigm together with the availability of the needed input data are the prerequisites of setting and fulfilling model's aims. Given varying complexity of reality which can only be partially observed and modelled (e.g. instantaneously emergent patterns motivated by intractable dynamics), there are no concrete standards or

measures to ensure and prove the implementation of these prerequisites. Thus the model is always built on the strong assumptions of appropriate variable and model selection. The discrepancy between these assumptions and the truth may result in incapability of predicting dramatic changes due to partially observable input space [14] or overfitting problem that is typical for overly complex models [15]. Therefore, researchers start to evaluate and reduce their distance to these prerequisites by considering observability.

Observability, in structural way, is about the ability of reconstructing the state of a system from a limited set of measured variables in finite time [16], while from the perspective of dynamics, observability is related to a deeper understanding about how the selected variables interact and evolve to change the states of the system. The structural observability is dependent on the available data set while the dynamical observability is analysed based on a model of dynamics, which is inevitably wrong to some extent but may be useful either in reflecting the reality or diagnosing an input space that is partially observable. Linear time-invariant system proposed by [17] is one typical modelling paradigm widely used in the studies on observability of dynamical systems [18].

In complex networked systems (CNS), structural observability determines the number of nodes to be measured to render a network observable, while dynamical observability considers node dynamics and coupling when selecting the best driver (sensor) node to modulate (observe) the whole network activity [16]. The design and analysis of CNS for each model's aim involves the observability of topology [18], variables for coupling nodes [19] and node dynamics [20]. These structural observability and dynamical observability are closely related to variable and model selection, and serve as the prerequisites of representing and modelling reality under each model's aim in different application scenarios.

The measures of observability vary depending on the linearity of systems. Linear networks without symmetries have been well studied research objects with respect to the topic of observability, where the observability matrix based on the dynamic model for a linear (time-invariant) system proposed by [17] is widely used. Nonlinearity of dynamical networks have been recently considered in the studies on dynamical observability. For example, [21] quantify the observability and controllability of nonlinear networks with explicit symmetries that shows the connection between symmetries and nonlinear measures of observability and controllability. [22] propose a nonlinear graph-based theory for dynamical network observability from the Jacobian matrix of the governing equations of nonlinear systems.

B. CURRENTLY STUDIED CNS MODELLING AIMS

The model's aims can be categorised as: (i) specific and (ii) abstract goals. Specific model's aims, such as community discovery, link prediction, anomaly detection, synchronization and controllability of networked systems, focus on specific external tasks for observable research

TABLE 1. Different goals of building CNS models.

Model's aim	References
Community discovery	[11], [23], [24], [25], [26], [27], [28], [29]
Link prediction	[30], [31], [32], [33], [34], [35], [36], [37], [38]
Node classification	[39], [40], [41], [42]
Synchronization	[43], [44], [45], [46], [47], [48], [49], [50], [51]
Controllability	[11], [23], [52], [26], [27], [28], [29]
Anomaly detection	[53], [54], [55], [56], [57], [58], [59], [60], [61]
Topological analysis	[6], [62], [63], [46], [64], [65], [66], [51]
Mimics of reality	[67], [9], [68], [69], [70], [71], [51], [72], [60], [73], [74].

object (i.e. CNS in our case) with measurable model outputs. Abstract model's aims, like topological feature analysis and the mimics of real life systems, approach the inner rules of real dynamics, analyse observables and simulate unobservables for further research on specific model's aims. Examples of different CNS models' aims together with the relevant references are shown in Table 1:

Community discovery aims to decompose complex networks into meaningful sub-networks that better describe local phenomena [11]. The local phenomena refers to a set of entities that share some closely correlated sets of actions with the other entities of the community [11], [75]. This has been explored and discussed in a wide range of applications, including the detection of community structure hidden in real social networks [23]–[25], collaboration network analysis like detecting citation patterns [26], improving routing of telecommunication network [27], reconfiguration of the brain network [28] or political affiliation [29].

Link prediction aims to infer the behaviour of the network link formation process by predicting missed or future relationships based on observed links and the attributes of both nodes and relationships [31], [32]. Link prediction involves questions of dealing with missing links or link labels of networks and predicting links in changing networks, including social networks [30], [33], [34], food webs [35], networks in collaborative recommendation tasks [36], knowledge graphs [37] and biochemical networks of protein interaction [38] and metabolism [35].

Node classification aims to provide a labeling for unlabeled nodes in a network composed of partially labelled nodes and edges [39]. Node classification, as an important way to explore node features and links, has been widely studied in social networks [39], [40], citation networks [41] and co-author networks [42].

Synchronisation of complex networks implies that the states of two or more interacting nodes in a network with different initial conditions gradually approach each other and finally reach the same state [43]. The applications of synchronisation in complex networks range from the stability of power grids [44], [45], controllablity of neuronal networks [46]–[48], optimising timetables for transportation [49] to the synchronization patterns affected by network topology in chemical systems [50] and IoT systems [51].

Controllability of networks represents the ability of controlling the networks, which is independent of the way

that the outputs are formed, while its related concept of observability, depends only on the outputs but not on the inputs [76]. The studies on controllability are often combined with observability, which range from social networks [77], protein interaction networks [78], brain networks [79] to transportation networks [80].

Anomaly detection of networks is about finding objects, relationships, or points in time that are unlike the rest [53]. There are many studies on anomaly detection in various application scenarios, ranging from anomaly detection in social networks [54]–[56], public health [55], IP networks [57], wireless sensor networks [58] to intrusion of networks [59]. Similarly to the principles of anomaly detection, pattern recognition is also used for diagnostic analysis [60] and heterogeneous component detection in knowledge networks [61].

Topological feature analysis is a very popular model's aim in both real-life networks and artificial networks. This is a research topic that probes network topological features from real data for applications across disciplines based on network-based models, such as the probed topological features of text network for language organisation [6], [62], probed boundary features within the small world complex networks for image classification [63] and organisation of brain network [46]. Local network topological features like three-node motif [81], directed closure [82] and quadrangle structure [83] are researched through topological feature analysis of social networks. There are also discussions of artificial network topological effects on network based models, which involves studies of topological effects on P2P trading in financial market [64], artificial neural networks in computation tasks [65], epidemic spreading [66] and enhancing synchronization of IoT systems [51].

Mimics of reality helps to deal with questions about analyses of dynamics over networks or features of networks [67]. There are studies of spreading processes in artificial networks including epidemic dynamics [9], [68], opinion dynamics [69], [70] and meme diffusion [71]. There are also studies that simplify the real complex systems as data-driven networks to assist further analyses, which involves applications ranging from Digital Twins of IoT systems [51], [72], image information representation [60], mimics of transportation systems [73], design of supply chain networks [84], [85], [85]–[88] and the representation of human-object interactions via networks [74].

C. DIGITAL TWIN: AN ULTIMATE GOAL

Researchers focus on the studies of twining real systems across many disciplines and those efforts have already resulted in a development of a field on its own known as Digital Twins (DT). DT serves as an "almighty" paradigm of mimics across spatial and temporal scales. It has also grown to become an ultimate goal of modelling complex networked systems due to its reality-friendly nature, integration of model functions and the wide range of applications.

Digital Twin is a virtual extension of reality, which not only allows to compare current conditions with historical data to provide meaningful information to assist in decision-making, but also enables forecasting and feedback of eventualities that have never happened before [89]. Researchers have defined DTs from different perspectives across the application scenarios. In a fully digitalized product life cycle, DT is a comprehensive virtual product model with the features of real-time monitoring, simulation and forecasting [90]. For mechanical and cyber-physical systems, DT is a linked collection of digital artefacts that evolves with the real system along the whole life cycle and integrates currently available knowledge with the purpose of describing behaviour and deriving solutions for the real system [91].

DTs have three elementary components repeatedly emphasised: the digital (virtual part), the real physical product and the connection between them [92], while there are also other imperative components added with the accumulation of practice, including data, service, machine learning, and DT Performance evaluation [93]–[95]. DTs also resemble a series of models with integrated functions like simulation, optimization and data analytics [96] and features of real-time processing and continuous updates [97]. This makes DTs impossible to be replaced by any single tool and ideal modelling paradigms for health monitoring [98], planning of manufacturing [99], management of smart city [100], accurate healthcare [101] and anomaly detection [102] within a wide range of complex systems, including complex networked systems like IoT systems [103], [104] and blockchain-encapsulated systems [105], [106].

The above mentioned components, integrated functions and universal application of DTs differentiate them from any other simulation tool or modelling paradigms by emphasising the properties of real-time data acquisition of observations and feedback, and self-evolution through continuous machine learning analysis. They contribute to DTs' status as a powerful tool for the mimicking of a series of realities and an ultimate goal for modelling complex networked systems across disciplines.

Modelling complex networked systems using a Digital Twin paradigm has a potential to build a universal model that can be adapted to fulfil multiple, different models' aims discussed in this section. But before this can be attempted, we need to review and assess how the modelling of complex networks, their dynamics as well as dynamics on those networks is approached currently and this is the focus of the following sections.

III. HOW TO REPRESENT INFORMATION IN A FORM OF A NETWORK

This question involves two important issues in building complex networked systems, which are: (1) what types of networks are needed in modelling certain phenomena including the ways of building the required network topology with appropriate complexity using real data and simulations, and (2) how to obtain these networks from different

perspectives of obtaining and processing data to realise such complexity. To achieve a faithful representation of a network that preserves as much information as needed in the modelling of networked systems (see section III) for certain models' aims (see II-B), one needs to collect and process observable data and information about the network structures and associated dynamics with modelling paradigms that minimise the information loss in the process of a network generation.

A. HOW TO REPRESENT NETWORKED DATA

According to [107] network topology is a representation of the physical connections that exist among entities in a communication network. This definition can be and has been easily expanded across disciplines as network topology describes how the entities of any type relate to each other in any type of network. The establishment of network topology is a crucial step for modelling complex networked system, and its effect on the dynamics has been a popular research focus.

a: COMPLEXITY DIMENSIONS

Network topologies vary in complexity when they represent networked information that can vary in data availability and modelling necessity. Complexity of the topology results from different types of nodes, edges and their attributes. As shown in equation 1, the network at time t can be represented as G_t, with its complexity level dependent on the heterogeneous network components and their changes over time: the set of nodes $V_t = \{v_{1,t}, \cdots, v_{N,t}\}$, the set of edges $E_t = \{(v_{i,t}, v_{j,t})|v_{i,t}, v_{i,t} \in V, i \neq j\}$, the attributes of each node $A_t = \{a(v_{1,t}), \cdots, a(v_{N,t})\}$ and the attributes of each edge $W_t = \{w(v_{i,t}, v_{j,t})|v_{i,t}, v_{i,t} \in V, i \neq j\}$ [24],

$$G_t = (V_t, E_t, W_t, A_t) \tag{1}$$

In Figure 1, we propose to describe this complexity alongside and using the following four dimensions: (i) a structural dimension connected to the scale and diversity of the topology, (ii) a temporal dimension concerned with time-to-live of different components of the network, (iii) a spatial dimension connected with the space in which the topology can be embedded, and (iv) a dynamics dimension connected to the topology's exerted or encapsulated dynamics. The complexity in structural, temporal and spatial dimensions describes the necessary reality required to be represented via networks, while the dynamics complexity depends on the models selected to explore the complex reality. These modelling paradigms are shown in Figure 1. Given available networked data set that is observable from the perspective of the above mentioned four complexity dimensions, in topological feature analysis, only structural information about networks need to be represented and analysed based on a model that reveals the inner rules of network formation, while for prediction of link formation over time, temporal complexity is further incorporated and a complex model that performs well in such an external task is employed.

Structural complexity of the scale and diversity of nodes, edges and their attributes involves discussions of non-attributed and attributed networks in terms of their diversity, as well as small/big networks and sparse/dense networks in terms of their scale, where structural complexity increases with more detailed information needed to be represented.

Research studies usually start from the exploration of non-attributed networks that are built only with homogenous nodes and edges. Such networks represent simplified real world scenarios and are thoroughly studied in terms of their topological features, like the analysis of potential energy landscape with both a small-world and scale- free character [108] or the conduction of specific model's aim e.g. community discovery which only uses network topology to find partitions [109].

Attributed networks can better represent real-world networked interactions and information as they introduce the auxiliary information via node or edge attributes [110], [111]. The node attributes describe the features of nodes within interactions or relations, while the edge attributes capture information about how the adjacent nodes interact with others in the networks [110]. These attributes vary with application scenarios. Taking online social networks as an example, nodes represent users and are attributed with user profiles [112], while edges represent online relationships and have attributes like nature of relation, direction, intensity and durability [113]. Many researchers study the structural complexity of node-attributed networks having in mind such modelling aims as community discovery [24], [114], link prediction [115], anomaly detection [112], controllability [116], topological feature analysis and mimicking of reality [117]. However, few researches focus on the representation and modelling of generalised edge-attributed networks. Edge-attributes of such networks within most researches typically take the numerical or categorical forms [110], [118], where directions [119] and edge weights [120], [121] are often studied. Compared with the large number of studies on node-attributed networks with various model's aims, only a small number of studies focuses on community discovery [122] and anomaly detection [110] of edge-attributed networks.

The structural complexity also increases with the requirements of representing more information within large-scale networks and the difficulty of processing networks with sparse edge information. Structural complexity of large-scale networks with thousands and millions of nodes results from a complicated and higher-order inner structure [123], which are common in DTs like city IoT [117] and DT of manufacturing with big data [124]. Structural complexity of sparse networks, given fewer edges, lies in their restrictions of attributes processing [111], [125] and optimal modelling [126]. There is an even more complex case for large sparse networks where both complicated large-scale inner structure and problematic sparse edges are involved [126].
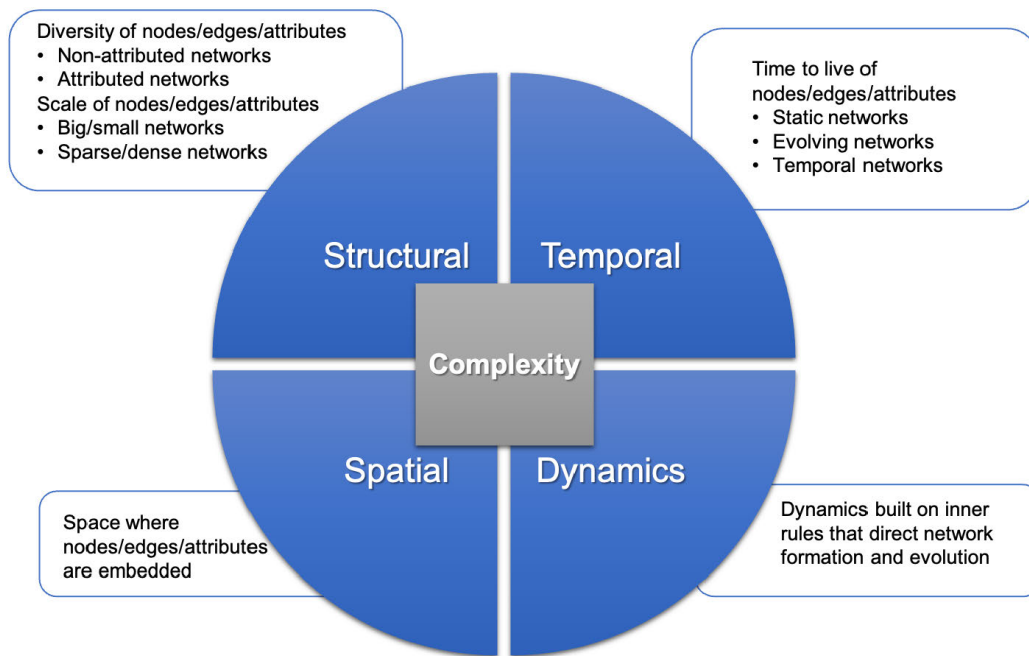
FIGURE 1. Four dimensions of complexity related to a network's topology.

Temporal complexity of networks increases when more temporal information about nodes, edges and their attributes can be captured and modelled meaning that less information is lost. Networks can be conceptually described as static, edge-weighted, evolving and temporal [10], [11] as they are transforming to be instantaneous and approaching the state of continuity without temporal aggregation.

The basic modelling of real-world phenomena and systems using CNS starts with a static network topology where nodes and edges are fixed and they are assumed to be "frozen" in time. Such an assumption greatly simplifies the modelling process but fails to capture the evolving features of real-world systems.

The attempts of accounting for temporal information in the network modelling process have gained attention as they usually improve the model performance. For example, social network analysis initially viewed networks as static rather than changing over time [127], [128]. As the field developed, social interactions started being represented using temporal networks to capture dynamics and instantaneous character of the contacts [121], [129]. In terms of biomolecular networks, studies on protein-protein networks first employed static data-driven networks to represent and analyse the protein-protein interaction [130] until the dynamic protein-protein networks have been found to benefit the study on the molecular systems of protein complexes [131]. The analyses of transportation networks also initially used static networks [132] and then shifted to dynamic management of transportation systems for better model performance [133]. However, the transition from the static networks under the most stringent assumptions to the temporal networks that

result in smaller loss of information can not be achieved overnight, while the scale of involving time dimension also differs from case to case.

Some of the researches focus on edge-weighted networks with temporal weights in the analysis of social relations [121] and the mobility in wireless networks [120]. Other researches focus on evolving networks where network topology changes slowly over time so that its instantaneous snapshot yields a well defined network [10]. Some evolving networks just represent more stable relations, rather than instantaneous interactions between the nodes, which can be captured with durable edges like citations [134] and friendships [113], [135], while some researches use evolving networks built with snapshots to aggregate the temporal information of interactions within a time window for a more stable representation in the analysis of instantaneous features [26], [27], [52], [135]. There are also studies on temporal networks that have non-trivial topology changes and can not be represented via instantaneous snapshots. They preserve all the temporal information and build networks in a more faithful way, such as instantaneous contacts of communication via e-mail, text messages, or phone calls with temporal edges of networks [136].

Spatial complexity involves discussion of spatial networks, which are proposed and defined as networks where nodes are located in a space equipped with a metric of the usual euclidean distance [137]. For example, spatial networks that represent urban street patterns can be built with a metric of distance, which is measured not just in topological terms (steps), but in properly spatial terms (meters, miles) [138]. Considerable applications of spatial
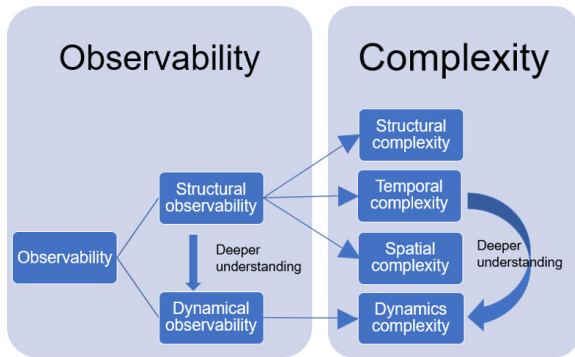
FIGURE 2. Observability and complexity with corresponding relations.

networks involve modelling of human activities that take place on a spatial matrix obtained from largely three types of transportation networks including matter (streets, roads, highways, railways, airport networks), energy (the power grid) and information (Internet, telephone networks) [139], [140].

Researchers have been characterising and understanding the structure and the evolution of networks under the impact of spatial constraints [137], where topology alone does not contain all the information to understand the dynamics of networks and the spatial information is needed. Some social-spatio networks consider the interplay of social networks and locations, such as social interactions under the impact of transportation [141], employment outcomes in a labour market driven by social contacts under an explicit geographic structure [142] and the social properties of Twitter users' networks with the spatial proximity of the networks [143]. Traffic under the constraints of transportation networks also involves studies on commuter flows constrained in large transportation networks [144] and planning of unobstructed paths in traffic-aware spatial networks [145].

Dynamics complexity is about the dynamics of networks that enables to develop a deeper understanding of temporal complexity based on structural and spatial complexity by investigating the rules of network evolution via simulation or modelling. Temporal and dynamics complexity, which are respectively derived from structural and dynamical observability, represent varying degrees of reality. The temporal complexity describes which components of a network change over time, while the dynamics complexity, as a deeper understanding of temporal complexity, considers why and how these components change. The relations between observability and complexity dimensions are shown in Figure 2.

For dynamics built on inner rules that direct network formation and evolution, the dynamics complexity increases as networks are generated with less human involvement based on more realistic rules ranging from statistical relations [146], [147] to realistic principles like homophily [148], [149]. This dynamics can either be exerted dynamics that controls the network changes to a desired state of CNS [150]–[152],

or encapsulated dynamics that motivates the network formation and evolution with higher degrees of automation [40], [148], [149]. For example, exerted dynamics used for social network intervention can control each step of network evolution by changing attributes of nodes that are identified based on different man-made strategies [151]. Encapsulated dynamics mainly focuses on the edge formation mechanisms, e.g. based on preferential attachment principles [40], [149] which are typical examples of highly-automated self-evolution without involvement of human decision.

The synergy effects of complexity dimensions describe the "1+1>2" effect of the combined complexities from different dimensions, which are more complex but closer to reality. Spatial temporal networks are typical examples for combined complexity of temporal and spatial dimension, which are proposed for a more faithful representation of reality with the influence of space on constraining the structure of temporal networks considered. Some researches employ temporal networks to capture and process temporal information under consecutive frames, while they construct spatial networks to extract certain static features. Such spatial-temporal networks have been widely used in computer vision like facial expression [153], video-based person reidentification [154] and identification of human-object interaction [74]. Some researches introduce temporal information to spatial networks. Taking recommendation task as an example, [129] build spatial temporal networks with temporal edge-weights by incorporating time dimension to user-location graphs and using sessions that capture the co-locations among two or more users during a time window.

It is clear that different complexity dimensions are intertwined. They influence and build on each other and this is one of the challenges that need to be considered when modelling of complex networked systems is attempted. Research has been conducted in each of the complexity dimensions but building overarching framework that would enable to flexibly adjust a level of complexity of each of the dimensions and simulate various what if scenarios is still an outstanding challenge. With the recent developments in the Digital Twin space, modelling CNS using DT paradigm is a promising way forward.

b: DATA-DRIVEN VS SIMULATION-BASED NETWORKS
Researchers have made a lot of effort to faithfully represent the information from real world systems by developing variety of modelling approaches. This involves data-driven networks, simulation-based networks and networks that are built by combining these two approaches. These networks are featured with varying degrees of complexity. Components of different network types are shown in Figure 3.

Data-driven networks built on rich real data sets may not necessarily capture all the information as they may be confined by relatively simple model's aims under strict assumptions, like the ignorance of temporal information for temporal networks analysed in a "static" time scale [127]. As researchers explore the real data sets with more complex
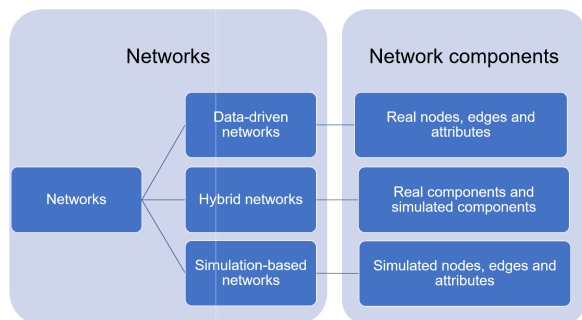
FIGURE 3. Different types of networks and their components.

modelling aims and advanced techniques, data-driven networks gradually approach reality as more complexity is allowed to be introduced to network topology with more relaxed assumptions. For example, the analysis of social networks started from statistical analysis of static networks and their topology [127] and then evolved into a more complex modelling aims such as community discovery in networks that are increasingly more and more complex, starting from static [155] to evolving [23], [156], attributed [125] and both evolving and attributed networks [11].

When real data is not available, simulated (synthetic) networks can be generated and used to analyse various phenomena [8]. These simulated networks enable to model network phenomena with different levels of complexity. The statistics-based simulations based on predetermined network statistics involve classical examples like the Barabasi-Albert model for the scale-free network [157], the Watts-Strogatz small-world model for the small-world network [158] or the Erdos-Renyi model for the random graph network [146], [147]. The simulation-based approaches increase their complexity and flexibility as more complex rules governing creation of structure and its dynamics are introduced. The capability of generating networks with distinctive features enabled the similutaion-based approaches to become a universal tool for topological feature analysis. They also help with assessing the impact of network topology on dynamic processes, involving researches like synchronization of IoT systems using networks ranging from scale-free to small world models [51].

Another type of simulation-based networks, the principle-based simulations, are built according to different connection principles like homophily [148], [149], triadic structure [149], geographic proximity [152], [159]. These networks have typically higher degrees of temporal and dynamics complexity than statistics-based network simulations as they self-evolve with highly autonomous and interpretable edge formation process and generate temporal networked information. However, few researches involve principle-based network simulations with various edge attributes [40] or network simulations embedded in spatial dimension [160]. They are more complex but closer to reality, which calls for future research on network simulations with the combination of different levels of

complexity from structural, temporal, spatial and dynamics space.

There are also networks built using combination of real data and simulations. They are featured with the mix of different levels of complexity either captured from real data or represented via simulation. Such networks consist of real attributed nodes and connections between the nodes are created using network simulation dynamics. For example, [151] build social networks by extracting necessary information about nodes' attributes from real database and simulate the edges using scale-free networks based on a network density from referenced literature. Some researches simulate missing edges for real networks, where the single imputation methods including null-tie imputation and reconstruction summarised by [161] are typical examples. The complexity of networks built using this hybrid approach increases as more data-driven features are captured by having the network simulation dynamics trained to fit observable real network components. For example, as an improvement of single imputation, researches propose a multiple imputation method that fits an exponential random graph model (ERGM) to the real data and simulates missing ties via inference [161].

B. HOW TO GENERATE NETWORKS USING DIFFERENT MODELLING PARADIGMS

Networks, no matter to what extent real data or simulation is used, are created to faithfully represent information about a given system in preparation for further analyses. Depending on the modelling goal, networks used for analysis will differ with respect to the four introduced above complexity dimensions and will have varying settings of nodes, edges and attributes. There are several modelling paradigms that are used to obtain desired networks and those include rule–based, agent–based and event–based approaches that focus on fundamental generative process from a local perspective of network formation, as well as basic graph–based, probabilistic graph–based and network–embedding based approaches that aim at condense network representation from the global perspective. These modelling paradigms, of local or global level, differ in ways of observing and processing networks but converge to a faithful representation of reality that aims at minimising the information loss between reality and the model. Observed local interactions lead to global emerging characteristic behaviours observed and analysed via graphs, while graphs lose less information as the four complexity dimensions are introduced via local level observations.

c: A LOCAL VIEW
Modelling paradigms of networks from a local view focus on mechanisms ruling the network formation. They take local level observations as a starting point and explore how they construct the characteristic CNS at a global level, corresponding to the discussion of observability considering the structural, temporal, spatial and dynamics complexity. The rule–based, agent–based and event–based

approaches at a local level are introduced below with more information observed or simulated for the four complexity dimensions.

Rule–based paradigm generates networks under explicit dependence laws from predetermined assumptions or rules detected from real world cases, which not only focuses on network topology, but also investigates dynamics of network structures via simplified rules of edge formation and evolution.

Rule–based paradigms are controlled by a limited number of parameters according to a rule-based mathematical function, where randomness is introduced via certain probability distributions involved in edge formation or variable changes, like the scale-free network generation via a scale-free power-law distribution [162], edge formation with a bias probability dependent on the similarity of node attributes [149], and the changes of node property according to a uniform distribution [40]. For rule-based paradigms, there is inevitable information loss resulting from the divergence of rules and reality, as the rules greatly simplify the partially observable and complex real world scenarios. Researchers have made much effort to bridge this gap, involving studies on rule-based dynamics transforming from simulated [146], [147], [157], [158] to trained to better fit real networks [149], as well as rules evolving from statistics-based [161], [163]–[166] to principle-based [149], [167]. They include temporal complexity by incorporating temporal changes of network topology into the rule–based network generation process, and sometimes, to better characterise network generation, introduce the impact of node attributes with increasing structural complexity.

Rule–based paradigm started from simulated dynamics of networks either based on statistical rules [146], [147], [157], [158] or principles [148], [149], [168] (as is mentioned in section III-A0.b.), where they can be tuned to approach reality via seeking optimal parameters for the rule-based mathematical functions to fit real data and make inference [149]. For example, the ERGM involved in multiple imputation of networks can fit real data and simulate missing ties via inference [161], where the involved dynamics represent data-driven features to some extent, but neither preserve rich information about node/edge attributes nor explore topological information other than typical interactive patterns of ERGM rules.

Statistics-based network generators are able to fit real networks via statistical inference based on an explicit likelihood function, like scale-free model [163], ERGM [161], [164] and geometric branching growth model [166]. As statistics-based dynamics on typical interactive patterns can hardly represent diverse real-world networks, principle-based network generators have been studied with their flexible and adaptable design of principles, which fit the statistics of real networks via likelihood-free inference such as approximated Bayesian computation [149]. For example, scale-free structure is empirically rare in social networks [163], while a principle-based simulator based on cumulative effects of

triadic closure and homophily is able to reveal social network dynamics [149].

Event–based paradigm refers to the network representations with two elementary components: nodes and their local pairwise interactions referred to as events [169] or nodes of events and their logical relationships [170]–[172], which involves event-based network representation and analysis [171]–[173] as well as stochastic point processes for events that can perform network inferential tasks [174].

Event–based paradigm started from a simple network representation that captures richer heterogeneous information of interactions related to events with increasing structural complexity. As a typical example, event-based social networks (EBSNs) are enabled to further capture and use the information of offline social interactions, in addition to the online social relationships included in conventional online social networks [173]. The enriched information observed about events assist further analysis and modelling tasks. For example, the consideration of both online and offline interactions for EBSNs provides adequate information for global level analysis and modelling via graph-based models [172] and improves the prediction power for event recommendation tasks [173].

Then it comes to stochastic point processes that are enabled to make inference about nodes [174] or edges (events) [175] via estimating dependencies between events or dependencies between events and latent space models with increasing dynamics complexity. Taking Hawkes process as an example, [174] employ a self-exciting point process on the edge to perform an identity-inference task, considering the effect of available observations of geographically distributed interactions (edges). [175] apply mutually exciting point process on the edge that includes the effect of node-specific latent vectors to assess the significance of previously unobserved connections for anomaly detection tasks.

Agent–based paradigm applied in the context of complex networked system refers to interactive structures that are typically composed based on three basic components: agents, interaction rules, and space (this could be geo-space or some other abstract space) [176], which enables more degrees of freedom in building networks from rich data or simulations as they can select more detailed information about interactions from microscopic perspective of agents.

Agent–based paradigms vary in degrees of information loss with different dimensions and degrees of complexity required for various research goals. The interactive structure can either be static with neighbour sets determined once and for all, or dynamic with evolution along time depending on model assumptions [177], while the interaction rules may either be simulation-based under certain constraints of space [64], [178], [179] or data-driven based on realistic scenarios [180]–[182].

The information about interactions of agents is initially represented in a fixed network topology, where decisions of agents are affected by network structures over which

they interact. Such networks serve as an environment and a constraint for agents' behaviours [177]. Most static networks involved in the researches on agent-based models are simulation-based, focusing on the network effect on interactions in agent-based systems like trading behaviours in double auction [64], [183] and tax compliance and evasion [179]. Recently, these agent-based networks approach reality by introducing real data with features of time and space. [180] proposed a data-driven agent-based model for forecasting emerging infectious diseases, where data-driven networks are built with spatial information and social contact of realistic synthetic population. [181] also proposed an agent-based computational model under a data-driven decision-making framework for supply chain networks given their complicated micro structures, macro emergencies and dynamic evolution.

Multi-agent systems are also important simulation tools for modelling evolving networked systems with multiple agents interacting under certain constraints. [52] proposed a multi-agent system that replays the evolution of a network and reproduces the rise and fall of communities with the strength of adapting to real-time, changing problems. [184] research on the constrained concensus and optimisation of multi agent networks, where multiple agents align their estimates with a particular value over a network with time-varying connectivity in different local constraint sets.

To better capture the structural patterns and instantaneous dynamics of networks, event-driven models (also named as activity-driven models) have been proposed with an activity potential, a time invariant function characterising the agents' interactions and encoding the instantaneous time description of the network dynamics [185]. These paradigms include the rich information observed or simulated from microscopic views of agents and model their connections activated by an event trigger, involving observations varying from binary interactions [185] to simplicial complexes [186], as well as applications ranging from event-based consensus [187] to contagion problems [186], [188].

d: A GLOBAL VIEW

Modelling paradigms of networks from a global view aim at representing high-dimensional and heterogeneous networked information in a way that can be easily analysed. More information can be observed at a global level as more information about interactions is captured via local methods, which involves discussion on how global performance can change through controlling and modifying those interactions, corresponding to the research aim about controllability. The basic graph–based, probabilistic graph–based and network embedding–based methods at a global level are introduced below with increasing complexity as they are able to preserve more observable information in network representation and modelling.

Basic graph–based paradigm is based on the graph theory and it can be seen as a set of selection principles for microscopic laws of behaviour in network science [189] which typically involves a simplified graph representation

and analysis of networks just concerned with nodes and their connections, e.g. [190], [191].

Graph theory began when, in 1735, Leonhard Euler presented the first mathematical demonstration based on geometry of position to solve the seven bridges of Köningsberg puzzle [192], [193]. Graph theory focuses on providing rigorous proofs for graph properties, such as graph enumeration, coloring, and covering [189], [194], while the evolution of random graphs motivated graph theory to generate a new branch of network science for a separate direction: quantifying the structure and dynamics of real-world complex systems [189].

Graph-based paradigm simply represents networks as basic graphs composed of nodes and connections, enabling readily available graph analysis but taking the cost of information loss to varying degrees especially in terms of structural complexity [190], [191], [195]. For example, from an accumulation of experimental data on biomolecules, graph-based models for cell biology build the graph only with cellular components (nodes) and their interactions (edges), which allows for network topology analyses using graph-theoretical concepts but lose information other than the graph-structure [190].

Graphs are one of the widely studied data structures in computer science and discrete mathematics [196], while the graph-based models are also widely applied in the modelling of CNS across disciplines, such as analyses of graphic characteristics for networks in cell biology [190], anomaly detection of computer networks using protocal graphs [195], graph representation of vulnerability relations for industry IoT network [191].

Probabilistic graph–based paradigm models networks with uncertainties on the relationships between nodes [197], which has two elementary components: a graph that defines the network structure and a set of local functions whose product is the joint probability of this compact representation [198]. The network representation with probabilistic graph–based paradigm is typically flexible in directed or undirected, static or dynamic dimensions, each corresponding to varying degrees of structural complexity and temporal complexity. The exploration of the local functions for dependence rules or cause-effect relationships enables the inference and learning of real networks in sophisticated models [199].

The probabilistic graphs can either be undirected with symmetric relations like conditional random fields [200] and Markov networks [201], [202], or directed with cause-effect relationships between the nodes, such as sigmoid belief networks [203], Bayesian networks [204] and hidden Markov model [205]. The usage of directed or undirected graphs depends on the application scenarios. For example, Markov networks, as a typical undirected graph, represent relational dependencies without the hindrance of acyclicity constraint and thus are well suited for discriminative training [201]. They have been widely used in argumentation tasks like finding labellings or probabilistic inference tasks by

deciding credulous and sceptical acceptance [206]. While Bayesian networks, as a directed acyclic graph with explicit cause-effect assumptions for interactions [204], can handle problems including missing data, prediction and data over-fitting [207]. Some researchers also employ Bayesian networks in DTs due to efficient linear computation [103], [208], [209], while nonparametric Bayesian networks are more flexible in capturing time varying features with computation efficiency [13], [210].

Probabilistic graph–based paradigm can either be static or dynamic based on whether the probabilistic graphs represent variables at a time point or across different times [198], such as the Bayesian networks [204] versus the dynamic Bayesian networks [208]. They are transforming from static to dynamic with the trend of modelling networks with more temporal complexity. For example, continuous time Bayesian networks [211] and continuous time Markov networks [212] are proposed to capture changeable variables given the evolving features of structured stochastic processes and the graphical structure of Markov networks and Bayesian networks.

Probabilistic graph–based paradigms are able to conduct inference that helps in answering different probabilistic queries based on the model and some evidence as well as learning that estimates the graph structure and parameters of probabilistic graph–based paradigms' local functions [198]. As exact inference is often intractable, researchers tend to use approximation algorithms to find distributions that are close to the correct posterior distribution [213], like Gibbs sampling [214] and belief propagation [215]. While maximum likelihood estimation methods and the expectation maximization (EM) algorithms [216] are respectively employed for learning problems with or without hidden variables. To deal with data uncertainty and to integrate imprecise data, fuzzy probabilistic regression models are employed to model the functional relationship between the entities [217], [218]. However, probabilistic graph–based paradigms have difficulty representing, inferring and learning high-dimensional, heterogeneous networks, which calls for combined application of network embedding methods.

Network embedding–based paradigm aims at network construction and network inference via embedding node information in the network into low-dimensional space [219], which is featured with a concatenation of an encoder and a decoder [10], [220], [221].

Network embedding starts from dimensionality reduction techniques that are also applicable in scenarios other than networks, which include stochastic multidimensional scaling [222], isometric mapping (ISOMAP) [223], principle component analysis (PCA) [224], linear discriminant analysis (LDA) [225], stochastic neighbor embedding (SNE) [226] and t-distributed stochastics neighbor embedding (t-SNE) [227]. There are also basic models that merely focus on network embedding tasks, which, as summarised by [228], are respectively built upon the skip-gram models [229] and matrix factorization models [230]. These models are able to encode network information into a low-dimensional space, but do not focus on decoding information to reconstruct networks or perform inferential tasks.

To preserve more information in the modelling process and perform inferences, deep learning (DL) techniques have been widely utilised to embed highly diverse, heterogeneous, high-dimensional network information into a low-dimensional latent space. DL-based network embedding, also referred to as a network representation learning, is able to make inferences and assist network analytic tasks including node classification, link prediction, clustering, recommendation, similarity search and visualization, involving unsupervised and semi-supervised learning methods [231]. As it is categorised by [232], DL based graph embedding methods are either random walk based like Skip-Gram based deep learning models [233]–[235], or without random walk but directly utilising DL methods on a whole graph or its proximity matrix via autoencoder [236], deep neural networks [237] and graph convolutional networks [238].

Graph Neural Networks (GNN) refer to a widely used DL-based embedding method that encode graph structures via a neural networks architecture, which is able to decrease information loss by aggregating features of neighbouring nodes together [10]. This motivates the combination of GNN with other models, which is typically composed of an encoder, a generative model and a decoder. For example, graph Bayesian networks [239] and graph Markov Neural Networks [240], [241] can be employed to infer graph parameters (statistical relational learning) based on encoder/decoder - graph convolutional neural networks = and generative model = the involved PGM. The GNN combined with ordinary differential equation (ODE) is able to infer the latent states of irregular observations (encoding), learn the state transition in latent space via generative model = ODE, and make predictions continuously (decoding) [242].

e: AN OVERALL VIEW

Three trends can be discovered if we look through the modelling paradigms of generating networks from either local or global view: (i) increasing complexity of network topology and dynamics, (ii) decreasing interpretability with the increase of complexity, (iii) models' aims transforming from abstract to specific. These trends are shown in Figure 4:

Complexity of networks increases when they are represented and generated more faithfully with less information loss.

Models at a local level focus on the network formation process. Rule–based paradigms are usually statistics-based and generate static networks that are just composed of fixed nodes and edges [162]. They recently turned to be principle-based while allowing for the introduction of node attributes with increasing structural complexity and edge addition over time with increasing temporal complexity [40], [149]. Given the above complexity dimensions, rule–based paradigms generate networks under the simplest rules of edge addition with lowest level of dynamics complexity. To incorporate
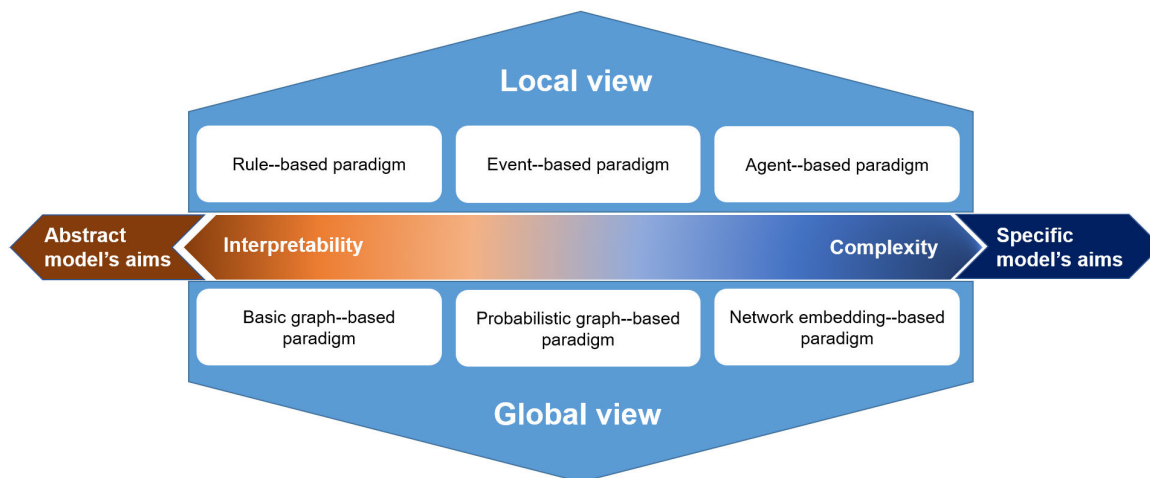
FIGURE 4. An overall view of modelling paradigms.

more information about attributes and temporal changes of network topology including edge addition and removal, event–based paradigms based on more complex rules of local pairwise interactions, such as the stochastic point processes, are employed with increasing dynamics complexity [174]. Agent–based paradigms introduce more complex interaction rules to account for more detailed information of agents (nodes), including their attributes and various actions that result in edge addition or removal over time [52].

Models at a global level aim at a condense representation of high-dimensional and heterogeneous networked information. Basic graph–based paradigms focus on a simplified graph representation of networks that are just about nodes and their connections [190]. Probabilistic graph–based paradigms further incorporate information of network components with increasing structural and temporal complexity via the introduction of node attributes, edge directions or addition of edges over time [32], [198]. These models, with increasing dynamics complexity, also enable the inference and learning of real networks via modelling the uncertainties on the relationships between nodes [197]. Network embedding–based paradigms can conduct network construction and network inference via embedding highly diverse, heterogeneous, high-dimensional node information into a low-dimensional space. This is characterised with the highest level of complexity in terms of network representation and dynamics.

Interpretability represents the ability to explain or to present in understandable terms to a human [243]. The interpretability of networks is about the understanding of network representation and the corresponding network dynamics. It decreases as the network complexity increases with more information represented and modelled. From a local view, more complex rules of network formation enable smaller information loss but may result in less interpretable global emerging characteristic behaviours observed in networks. For example, compared with rule–based paradigms, agent–based paradigms may generate less interpretable networks as

they incorporate the impact of various observables from microscopic perspective of agents [52]. From a global view, more complex models can embed more complex networked information but are characterised by less interpretable embedding process and inference results, such as the network embedding paradigms which are less interpretable in encoding or decoding but can represent highly diverse and heterogeneous networks.

Measurability represents the ability to measure a characteristic of a class of objects [39]. Measurability of networks generated for abstract model's aims is about measuring the similarity between the inner rules of a networked system and that of real networks, while the measurability of networks under specific model's aims focuses on the measurable output for external tasks. Networks of higher interpretability but lower complexity, like non-attributed static networks generated by rule–based or graph–based paradigms, can be easily analysed, measured and compared in terms of network components and dynamics. These networks are usually involved in researches with abstract model's aims including topological feature analysis [157], [162], [244], [245] and mimics of reality [164], [192]. Networks of high complexity but low interpretability like the attributed temporal networks are always generated by event–based, agent–based, probabilistic graph–based or network embedding–based paradigms. They incorporate rich information especially in structural and temporal dimensions while preserve enough dynamics complexity for the fulfilment of specific model's aims including link prediction, node classification, community discovery, anomaly detection, synchronization and controllability.

To conclude, modelling paradigms of network representation, either from a local or a global perspective, get more complex with compromised interpretability while transferring from the abstract models' aims to the specific model's aims. Such a trade-off between complexity and interpretability in the current CNSs does not indicate their

incompatibility but reveals a research direction where new CNSs models can be both complex and interpretable. And the division between the abstract and the specific model's aims will also be eliminated when CNSs are modelled as a Digital Twin and are enabled to undertake multiple external tasks thanks to faithful representation of real systems.

IV. HOW TO MODEL THE DYNAMICS IN NETWORKED SYSTEMS

In this section, we explore the answers to the question "how?" about the modelling of dynamics in complex networked systems. This involves a discussion about two key concepts of complex networked systems: (i) dynamics of networks and (ii) dynamics over networks. Dynamics of networks considers the network's generation process and network's changes over time (evolution), while the dynamics over networks refers to the dynamic processes that occur on networks over time. In addition to the process and network complexity, we also emphasise the complexity resulting from various interrelations between dynamic processes and network dynamics, where unilateral or mutual influence from one-way or two-way interactions are reviewed and discussed. Given complex networks (obtained via methods from section III) and observable information of dynamic processes, their dynamics and interrelations should be modelled in a way that models' aims (see section II-B) are fulfilled.

A. AN OVERVIEW OF DYNAMICS AND THEIR INTERRELATIONS

Taking the networks obtained via modelling paradigms introduced in section III as a starting point, the modelling of dynamics within the CNS, at the appropriate complexity level as per complexity dimensions introduced in section III-A, enables a deeper understanding of changeable network structures and processes over them with and contributes to achieving certain models' aims under the constraints of observability.

The complexity of modelling CNS dynamics comes from three primary elements: the modelling of networks, processes over networks, as well as the interrelation between the two. Their states can be respectively represented as G_t, P_t and I_t, with the dynamics of and over the networks modelled as $F_g(x_t)$ and $F_p(y_t)$. The parameter x_t and y_t of these two dynamics and their interrelation I_t, with the capability to change over time, result in the complexity of modelling CNS.

$$\begin{cases} G_t = F_g(x_t) \\ P_t = F_p(y_t) \end{cases}$$
$$s.t. \ I_t(x_t, y_t) = c$$
$$x_t, y_t \in R$$
$$c \in \{R, \text{NA}\} \tag{2}$$

To have an overview of various CNS resulting from those three elements, we introduce a 5-generation modelling framework (see Table 2 for brief descriptions and Figure 5

for details) to navigate a pathway through different levels of complexity of CNS.

As is shown in Figure 5a, in the first generation (generation 1), the research concentrates on static networks, where spreading processes are introduced without changing their parameters. The static networks with fixed nodes and edges just provide necessary spatial or structural information rather than one-way influence that might trigger the parameter changes (evolution) of dynamic processes. Some studies employ spatial networks that utilize a metric to embed spatial information [137], [139], where the dynamic processes are modelled under the constraints of space, like the human activity on transportation networks [139]. Majority of research on dynamic processes over complex networks can be classified as the first generation of models [8].

Assuming a virus *a* spreading in the society where all relevant information is observable, in the generation 1, we can simplify this scenario into an epidemic spreading process of virus *a* with fixed parameter (e.g. infection rate) on a static non-attributed networks that are just built with fixed nodes (e.g. people) and edges (e.g. social contact). This simplified CNS increases complexity in structural and spatial dimensions once we incorporate more information about node attributes (e.g. age, gender, location, etc.) and edge attributes (e.g. direction, weights, distance, etc.).

With the temporal complexity introduced in CNS in the second generation (generation 2), parameter changes of dynamic processes or network topology changes over time can be observed via snapshots in discrete time steps. In the generation 2a, CNS represents the interactions via evolving dynamic processes (captured in snapshots) and static networks (see Figure 5b), where the spreading process on static networks can evolve through parameter changes [246]. In the generation 2b, CNS represents the interactions directly via evolving networks (captured in network snapshots) and dynamic process with no parameter changes (see Figure 5c). In this context, researchers aim to transform the snapshots of CNS into latent states and model their transition process discretely, involving the networks that switch arbitrarily between different adjacency matrices according to stochastic mechanisms like positive linear switched systems [247] and Markov switching rules [248], [249].

Taking the above mentioned spreading of virus *a* as an example, in the generation 2a, a parameter (e.g. infection rate) of virus *a* can change just arbitrarily or according to an external factor outside of CNS (e.g. temperature, etc.), while in the generation 2b, edge addition (e.g. social contact over time) or edge removal (e.g. enabled by the implemented policies of social distancing) can naturally lead to observations of evolving networks.

The modelling framework steps into the third generation(generation 3) as the dynamic process starts to co-evolve with the evolving networks (see Figure 5d). The involved evolving dynamic processes are captured via evolving snapshots with discrete parameter changes, where the changes of parameters and the transition of the latent states of evolving
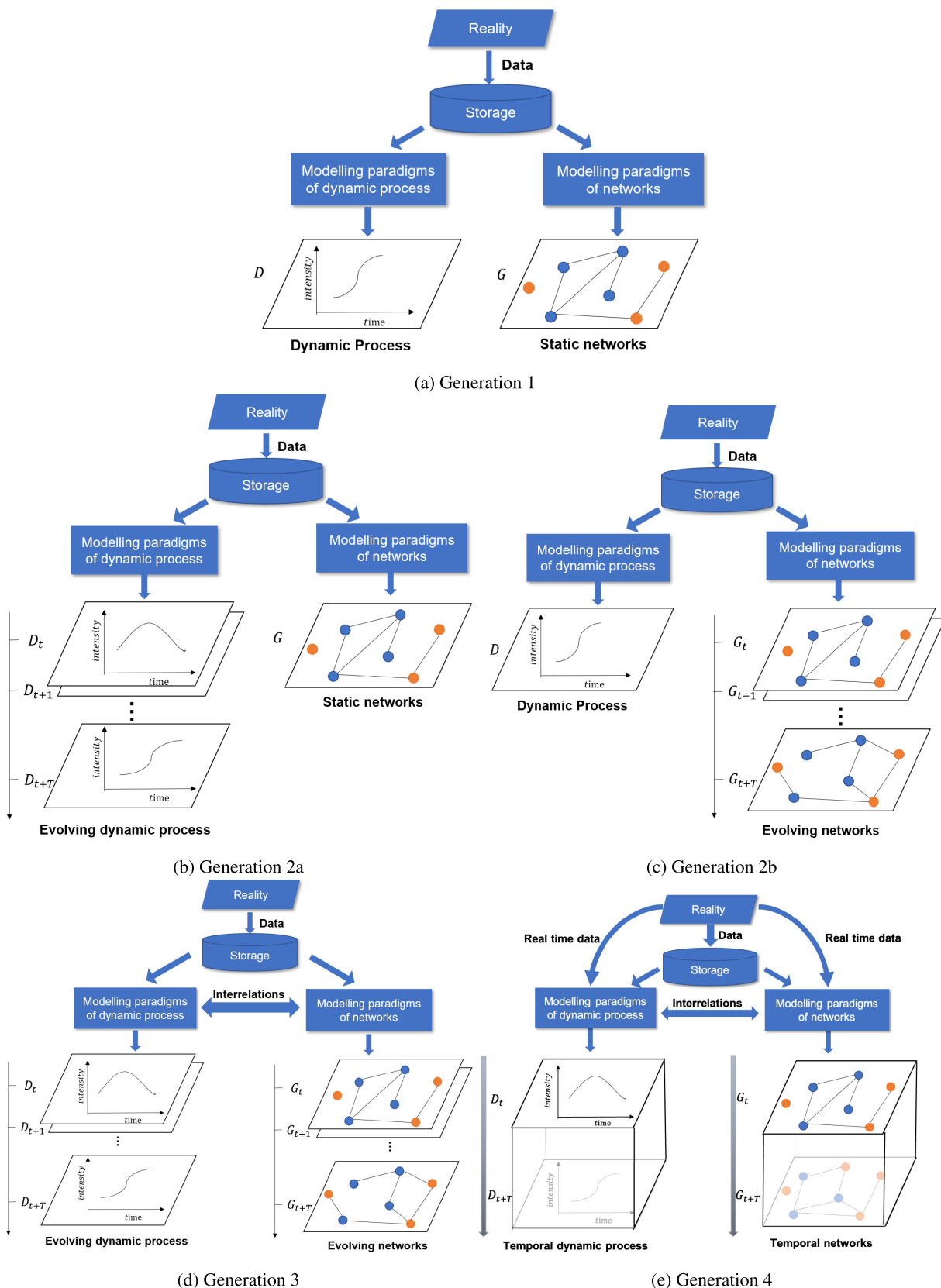
FIGURE 5. Generations of modelling framework.

networks can either be independent or interrelated. The interrelations, including unilateral and mutual influences, enable the modelling of co-evolution of dynamics and greatly increase the complexity of modelling CNS. For example, [250] model the co-evolution process of dynamic social networks and the opinion migration on networks via introducing mutual influence between these two dynamics, where opinion migrates based on social structure and social networks evolve considering the similarity of opinions. Generation 3 can easily collapse into the generation 2a when there is only one snapshot of dynamic processes and no interrelation is considered and vice versa for the generation 2b.

To have a better illustration of the interrelation in the generation 3, we also start with the simplest scenario where the virus *a* spreads on a static social network (a single snapshot of evolving networks). In the generation 3, we can allow the network topology to change over time in response to the spreading of virus *a* (e.g. people die of virus *a* and are removed from the social networks). These evolving networks are interrelated with dynamic processes and are characterised with increasing temporal complexity. The parameter (e.g. infection rate) of virus *a* can also change according to the network topology (e.g. number of social contacts, etc.) or node attributes (e.g. age, gender, etc.). In this way, the complexity of the CNS also increases with an introduction of the evolving dynamic process and interrelations. The CNS will become even more complex if we further incorporate a vaccination *b* in the modelling framework of the generation 3, where the spreading of the vaccination *b* can affect the node attribute (e.g. vaccinated or not) and directly change the parameter (e.g. infection rate) of the virus *a*.

The modelling framework then goes through a fundamental change in the fourth generation (generation 4) as the time gaps between the CNS snapshots are narrowed, in the limit, to zero and the co-evolving dynamics are represented and modelled continuously with the introduction of real time data. As is shown in Figure 5e, based on both historical data from a data storage and the real time data, the temporal networks capture and model all the instantaneous interactions, while the temporal dynamic processes are modelled with continuous parameter changes. The complexity of the CNS increases as more complex interrelations between continuous dynamics are introduced and modelled. There are studies on modelling continuous changes of networks via introducing ordinary differential equations (ODE) to the GNN methods [242], but none of them involve the most complex case of modelling continuous co-evolving dynamics of interrelated dynamic processes and dynamic networks. CNS in the generation 4 can approach the ultimate goal of DTs when the model output of CNS is additionally fed back to and can influence the reality in a real time manner as a reference for practice. We refer to this scenario as the fifth generation of modelling framework (generation 5), where a closed feedback loop of real time monitoring, simulation, forecasting and deriving solutions for

TABLE 2. Summary of CNS modelling generations.

Generation		Brief description
First generation	(generation 1)	Dynamic process on static networks
Second generation	(generation 2a)	Evolving dynamic process on static networks
	(generation 2b)	Dynamic process on evolving networks
Third generation	(generation 3)	Evolving dynamic processes on evolving networks with interrelations between them
Fourth generation	(generation 4)	Temporal dynamic processes on temporal networks with interrelations between them and the acquisition of real time information
Fifth generation	(generation 5)	Temporal dynamic processes on temporal networks with interrelations between them, as well as the real time two-way feedback between the reality and the CNSs, enabling an idealised state required by a DT

reality is formed and enables the CNS to approach DTs as an extension of reality.

To be more specific and continuing with our illustrative example, in the generation 4, the constant event streams about the spreading of the virus *a* and the temporal networks can be observed and modelled simultaneously. The instantaneous social contact and the virus *a* spreading can be captured instantly. CNS in this context can react to sudden changes observed in reality and evolve in real time. Any changes to the temperature, recorded all the time, can trigger continuous parameter (e.g. infection rate) changes of the spreading virus *a*. The vaccination *b* can also be introduced at any time and trigger network attribute (e.g. vaccinated or not) change right away. In generation 5, the simulation result of CNS about the spread of the virus *a* can be fed back to the policy makers and trigger, for instance, a launch of a "promote vaccination *b*" campaign, where the spread of vaccination *b* will be simultaneously monitored and modelled by CNS with a real time output about social networks fed back to the policy makers.

The above mentioned modelling framework shows with examples of real scenarios that CNSs can be represented and modelled with increasing complexity through generations and finally reach the goal of a DT in generation 5. To be more aware of the progress of studies on CNSs under this modelling framework, we need to review and discuss what kind of model's aims can be achieved by modelling the dynamics in networked systems, and how these dynamics are modelled in network dimension, process dimension and both of these dimensions.

B. MODELLING DYNAMICS OF NETWORKS
Dynamics of networks can result in networks with different structural characteristics and their changes over time. Their modelling fulfils models' aims via analysing and learning

these resulting patterns and changes of nodes, edges and attributes. There are overlapping areas between modelling dynamics of networks (in a way that models' aims can be fulfilled) and modelling paradigms of network generation (in a way that minimises information loss), as a good model that approaches real network dynamics can naturally accomplish these two tasks simultaneously. Studies on modelling dynamics of networks start from patterns of non-attributed static networks and then turn to dynamic networks that consist of nodes and edges that change over time. Those networks include temporal information about characteristics and behaviours of their components. Time dimension is introduced to break the assumption of static networks with fixed nodes and edges [10], [11], [53].

Modelling of the network dynamics, considering the patterns and changes of nodes, edges or attributes from a structural view, can be categorised as models that allow for: (i) no changes, (ii) topology changes, (iii) attribute changes, (iv) both attribute and topology changes and (v) structural pattern changes. They differ in whether network structures change and how they change over time, each to varying degrees coping with the models' aims including prediction and classification of network components, as well as the pattern discovery of network structures. As networks with only attribute changes are only involved in studies of dynamic processes on networks, we only discuss models of network dynamics for category (i), (ii),(iv) and (v) just from the perspective of a network dimension.

f: NO CHANGES

The modelling of network dynamics starts from static networks without any changes of structural components over time and focuses on the prediction of unobservable network components.

Networks built only with fixed nodes and edges are of the least complexity within this no change category, which involves missing link prediction fulfilled with approaches that are applicable for static networks. Some techniques take a link prediction of these networks as an unsupervised ranking problem based on a score for each non-observed link, which can be calculated either with a structural similarity index or probabilistic and statistical functions [32]. Structural similarity–based methods assume that nodes tend to form links with other similar nodes, which solely consider the local or global topological information of networks [31]. Probabilistic and statistical function–based models abstract the network structure and then predict probabilities of the missing links using the learned model [32], [251], where rule–based modelling paradigms in section III-B0.c and probabilistic graph–based paradigms mentioned in section III-B0.d can be employed to fit network topology for a probability score of each link. There are also techniques that take link prediction as a supervised binary classification problem about whether each pair of nodes are connected or not, where machine learning models like logistic regression and decision tree can be applied to account for

the effects of different topological similarity metrics [252], [253]. Another universally applied supervised approach is the network embedding modelling paradigm mentioned in section III-B0.d, where the low-dimensional latent space representation of nodes is learned and their connections can be inferred via dependencies of latent space.

Networks increase in structural complexity when attributes are gradually introduced for nodes and edges. Some approaches use node attributes to assist prediction of links [38], where above introduced basic approaches for link prediction have been altered to incorporate node features. For unsupervised ranking algorithms, as is summarised by [252], the similarity score between two attributed nodes can be calculated to assist a link prediction using methods including vertex feature aggregation [254], kernel feature conjunction [255], extended graph formulation [256] and generic SimRank method [257]. In addition, probabilistic and statistical function–based models can also incorporate attributes to model the probability of links between each pair of nodes [252]. Supervised models that take link prediction as a binary classification problem can also be improved based on the above mentioned information about attributed nodes. For example, [256] introduces node attributes via vertex feature aggregation to the machine learning algorithms like a decision tree or SVM in the link prediction tasks. There are also improved network embedding methods extensively reviewed by [219] for the link prediction task considering an effect of both network topology and node attribute. An example of such an approach is the deep attributed network embedding method designed by [258] using a deep neural network based on the topology proximity and attribute proximity.

Some studies additionally introduce attributes to the edges and transform the link prediction into a classification task for multi-relational networks. Once the transformation is completed, this problem can be approached by the above mentioned probabilistic and statistical function–based approaches and supervised learning approaches considering the similarity of nodes and dependencies of inner principles. For example, [259] uses relational Markov network to investigate the probability of link labels given the known node attributes. [260] performs link prediction in multi-relational networks using a non-negative matrix factorization algorithm based on relational similarity.

Other methods introduce attributes to the networks and focus on the node classification tasks, which employ models considering the effect including node attributes [258], [261], [262] and edge attributes such as weights [258], [258], [263]–[265]. Given the sparsity of graphs with fully labelled nodes and the time consuming manual labelling, most studies employ partially labelled graphs and train a classifier for the prediction of unlabelled nodes. Referring to the models for a node classification, very well summarised by [266], we further categorise these modelling approaches into three types: (i) unsupervised learning approaches including probabilistic and statistical relational learning [263], metric

modelling [264], spectral partitioning [265] and graph clustering [267]; (ii) supervised learning approaches that build classifiers like logistic regression model given features of labelled nodes [261]; and (iii) the semi-supervised learning approaches given the labels of few nodes, where network embedding methods like random walk–based [268] and GCN–based Network embedding [258], [269] are employed to account for both topological and attribute proximity effect [258].

g: TOPOLOGY CHANGES

The modelling of network dynamics becomes more complex as network components start to change over time. The changes of topology include the edge addition and removal, node addition and removal. As the network formation and evolution can be captured in a time series of static snapshots, modelling approaches of link prediction for static networks mentioned in section IV-B0.f can be extended and improved to learn one or several types of topology changes.

Studies on network topology changes start from networks built with fixed nodes and addition of edges, where the edges can be predicted and inferred via edge formation process extracted from networks over time using structural similarity–based models or probablistic and statistic function–based models discussed in section IV-B0.f. For example, [270] employ a Markov model based on both topological and semantic features similarity between two nodes to evaluate the probability of a link. [34] predict formation of new links based on a combined popularity and similarity measure, which incorporates both global and local topological information via the introduction of Newton's gravitational law.

Modelling network dynamics becomes more complex when it comes to networks built with fixed nodes and temporal edges that are added or removed over time. Link prediction tasks for these networks focus on the temporal topology information and its evolution, where modelling approaches for prediction of missing links in static networks have been improved to account for the effect of historical topology. For unsupervised similarity ranking–based models, the link prediction task is based on the predicted similarity scores calculated using past structural similarity scores via a time series forecasting model such as ARIMA [271]. To deal with the model capacity and computational efficiency problem for probabilistic and statistical function–based models, efficient learning algorithms can be introduced to account for influence of topology, such as the proposed neighbour influence clustering algorithm in a conditional temporal restricted Boltzmann Machine for a prediction of temporal edges [272]. As for supervised learning approaches, graph convolution network (GCN) is widely used to learn a node structure of a network snapshot for each time slide and LSTM is employed to performs a temporal feature learning for all the network snapshots [41], [273]. The discussed so far approaches only use network topology, but the heterogeneous prior information, such as node attributes, is suggested

to be used to further improve the accuracy [274] but is not commonly explored in the research community. [275] propose a nonparametric link prediction algorithm that can use both topology and node labels for the calculation of linkage probability with seasonality linkage patterns.

There is an even more complex case when the networks shrink or grow in size as they evolve with temporal nodes and edges. However, as the networks can be modified to include all observed nodes from all the snapshots of temporal networks, it is always assumed that there is a fixed set of nodes for all the networks at different time points [276], where the above mentioned modelling approaches for link prediction of networks built with fixed nodes and temporal edges can be used.

To conclude, the existing studies focus on modelling topology changes where nodes are fixed. The current methods consider one of the two scenarios: (i) edge addition and (ii) edge addition or removal. There is space for further research on modelling networks with temporal nodes and edges, where addition and removal of nodes, as well as the resulting change of network sizes should be considered.

h: ATTRIBUTE AND TOPOLOGY CHANGES

The modelling of network dynamics becomes more complex as we allow, on top of the topology changes, for network attributes to change over time. As current models of topology evolution are limited to networks built with a fixed set of nodes and edge changes over time, the models of networks where both attributes and topology change also have the same limitation. These models, given an addition or removal of edges, consider: (i) edge attributes changes, (ii) node attributes changes and (iii) both of these changes. Modelling approaches for the link prediction and node classification of static networks presented in section IV-B0.f and the network topology changes introduced in section IV-B0.g can be extended to learn both attribute changes and topology changes.

The topology change, linked to the edge addition or removal, can be accompanied with changes of edge attributes. There are studies on link prediction in temporal networks that have edge weights [41], [274], [276] and directions [276]. The typical supervised learning approach for topology changes mentioned in the section IV-B0.g, where GCN explores the local topology of each snapshot and LSTM characterises the evolving features of dynamic networks, can be improved by introducing the generative adversarial network (GAN) to tackle the sparsity and the wide-value-range problem of edge weights [41]. Network embedding methods based on matrix factorisation can also be improved to include the information about edge weight or direction into adjacency matrix for a prediction of temporal edges [274], [276]. Further research is needed for a variation of other edge attribute changes.

When node attributes change in networks built with fixed nodes and non-attributed edges that change over time, modelling approaches of node classification can be employed

to learn the changes of the node attributes by considering the evolution of attributes and topology. [40] use GCN to conduct node classification task on the social networks built with a fixed number of attributed nodes, changeable node labels and non-attributed edges that are added over time. GCN implemented in this research not only considers the local topology and its attributes, but also uses the similarity–based matrices to account for patterns of high-order neighborhood. Further research is needed for node classification in networks built with dynamic labelled nodes and temporal edges. There also comes the most complex case where node attributes, edge attributes and network topology all can change at the same time. Currently, there is no research in this space and further studies are required to model this very complex scenarios.

i: STRUCTURAL PATTERN CHANGES

The above mentioned structural components and their changes result in various structural patterns and corresponding changes. Structural patterns refer to the correlated combination of nodes, edges and attributes within a community or a network. Research in the space of structural patterns includes patterns and their dynamics discovery, analysis and prediction. The common models' aims here are e.g. a community discovery and an anomaly detection.

A community discovery starts from defining a community. This characterises the structural patterns of the sub-networks to be discovered via generally unsupervised way of modelling. A community in a complex network, as is defined by [11], [75] in a generic way, is a set of entities that share some closely correlated sets of actions with the other entities of the community. To fulfil the requirement of reflecting certain features of reality, the closeness within each community can be measured based on density, vertex similarity, actions of nodes or influence spread, all corresponding to different types of community discovery algorithms well summarised by [75].

Considering the increasing temporal and structural complexity of networks resulting from the topology variations and an introduction of attributes, the community discovery approaches have also been further developed and categorised from the perspective of process that is seen as the inner structures of algorithms. To deal with the community instability problem in dynamic networks, the temporal smoothing operations have been included in community discovery models to smooth-out the evolution of communities, which accordingly involve a new category based on varying extent of temporal-smoothness [11]. For attributed networks, built with attributed nodes and edges, a fusion procedure has been introduced to community discovery models to account for both effects of topology and node features, where another category based on when and how they use and fuse network structure and attributes is proposed [125]. To deal with the directed networks that are featured with edge directions and asymmetrical matrices, different community discovery models have been proposed and well summarised depending on the way directed edges are treated [277].

However, the perfect community discovery algorithm does not exist despite of all the above mentioned attempts, as each of them performs well on one specific declination of the general problem and can achieve different partitions even for the same networks [11]. Based on that, [278] further categorise the community detection algorithms according to the similarity of their results, which attempts to confirm the valid definitions of a community and help with the choice of algorithms for future research. There is still a need for further research on the community discovery of network variations with varying degrees of temporal complexity and structural complexity, given the superposed challenges of community definition, temporal smoothness as well as topology and attribute information incorporation.

Studies on anomaly detection focus on the rare occurrences of structural components, patterns as well as their changes, involving detection of anomalous nodes, edges, subgraphs, events and graphs [53], [56], [279]. They start from static networks built with fixed nodes and edges, where anomaly detection of nodes, edges and subgraphs can be realised via traditional non-deep learning approaches based on the network statistical features or using representation learning methods, as is summarised by [279], [280]. These methods have also been improved respectively when it comes to attributed networks with richer information about network structures [279], [280].

As temporal complexity is introduced to networks, two types of anomaly detection methods can be distinguished. More specifically, there is a two-stage approach that maps networks into a vector of real numbers and then employs an anomaly detector on it for a node, edge or subgraph anomalies, involving well categorised community, compression, decomposition, distance, and probabilistic model–based models [53]. There are also deep learning approaches respectively applied to the anomaly detection of nodes, edges, subgraphs and graphs [279]. There is also research gap that calls for further study on anomaly detection models incorporating both attribute and temporal information of networks.

C. MODELLING DYNAMIC PROCESSES

In this subsection, we focus on the dynamic processes that can take place over networks. These dynamic processes interact with the dynamics of networks and vice versa. Dynamic processes can result in three types of changes: (i) a network topology change; (ii) a network attribute change; and (iii) a parameter change of dynamic processes [281]. The network topology change and the attribute change can also lead to the parameter change of dynamic processes [282]. There are many interesting variations of the combination of dynamic processes and dynamic networks, which differ depending on the research objectives and application scenarios. Corresponding studies start from single parameterized dynamic processes on the static networks and then turn to multiple dynamic processes with dynamically changing parameters on the dynamic networks, which

is featured with increasing process complexity, network complexity, as well as the complexity resulting from various interrelations between dynamic processes and network dynamics

Modelling of the dynamic processes research can be categorised into the following groups: (i) a single dynamic process, (ii) independent multiple processes and (iii) interrelated multiple processes. They differ in a number of dynamic processes and how they interact, each to a varying degree mimicking the real world. Within each category, researchers either focus on parameterized dynamics or dynamically changing dynamics.

j: SINGLE DYNAMIC PROCESSES

Spreading dynamics on complex networked structures obtained via methods in section III, as a typical dynamic process, is covered by considerable literature on the spread of phenomena/medium ranging from virus [283]–[286], meme [71], opinion [69], [70], [287], decision making [288], nutrition [289], [290] and social contagion [291].

The modelling of spreading dynamics on networks starts from classic population models with the simplest analysis that considers the evolution of the state for all individuals rather than the state of each individual [9]. They include the stochastic population model that describes the evolution of the population state via a Markov process, as well as its approximation, deterministic population model with deterministic definitions of the population state [9], [292]. To model the states of each individual independently and allow for arbitrary interactions among them, as is summarised by [9], these population models are extended and improved to faithfully learn spreading processes on networks.

A spreading process on static networks is an example of a simplified representation of various real world scenarios where networks underlying the dynamic processes are simulated or represented under the most stringent assumption of nodes and edges that do not change. When a dynamic process evolves much faster than the network of interactions, static networks can serve as accurate proxies of slowly switching topologies. This real world situation can be approximated as dynamic processes over static networks [292].

The modelling of dynamic processes on static networks start from analysing the impact of pairwise interactions using the extended classic models, including stochastic network model and deterministic network model [9], [292], where networks with varying features and structures can be introduced to the studies. Stochastic network models assume the state transition for each node as a Markov jump process or its extension under relaxed Markovian assumptions [292]–[295]. Deterministic network models, as approximations of stochastic network models with deterministic definitions of node states, are also widely used [296], [297]. A threshold model is one example of a typical deterministic model for an epidemic process [291]. The modelling of

spreading dynamics becomes more complex and realistic as complex contagion is also considered with the incorporation of group interactions, which currently has been realised via simplicial complexes [298].

A spreading process on dynamic networks refer to one of the most common real-world application scenarios. The underlying networks of interactions are not static, but dynamically change while co-evolving with and influenced by the dynamic processes over the networks [299]. Models that capture their co-evolution at comparable time-scales have been well categorised into temporal-switching, activity-driven, and edge-Markovian networks [292]. For dynamic processes on dynamic networks, temporal switching networks model them as snapshots switching arbitrarily between a set of topologies according to stochastic mechanisms such as Markov switching rules [248], [249]. Activity-driven approaches focus on the networks' interactions generated according to a time invariant function characterising individual properties, which involves a series of extensions with the introduction of epidemic threshold due to its analytical tractability, as is summarised in detail by [300]. Edge-Markovian dynamic graphs can model stochastic evolution of dynamic networks, which also involves analytically tractable extensions with spreading dynamics [301], [302].

Spreading dynamics on dynamic networks can also be modelled with data-driven machine learning approaches. They focus on transforming spatial information and other temporal features involved in a spreading process to well handled temporal information, where deep learning–based predictive models like Recursive Neural Networks and Convolutional Neural Networks can be employed to predict spreading process, as is well summarised by [303]. Researchers also start to employ network embedding approaches, as is mentioned in section III-B0.d, to incorporate network information into the predictive systems, which involves typical examples of predicting epidemic spreading with graph neural networks [304] or using node regression based on transfer learning [305].

Parameter changes of processes are discussed in terms of the above mentioned single dynamics. The existing modelling approaches to a single dynamic process propagating over the network are generally parameterized without any change of parameters implemented. Only few studies focus on the evolution of spreading processes on networks. [246] incorporate mutation of pathogen strains and corresponding changes of epidemic transmission probability, which trigger the evolutionary adaptations of the spreading processes with dynamically changing epidemic threshold. For this example, transmissibility changes between limited number of fixed values and is controlled via mutation and transition probabilities.

k: INDEPENDENT MULTIPLE PROCESSES

Dynamic processes can result in three types of changes, including: (i) a network topology change; (ii) a network attribute change; and (iii) a parameter change of dynamic

processes [281]. As almost all the models of multiple processes in section IV-C0.j are based on probability of state transition, the parameters of multiple processes in section IV-C0.k and section IV-C0.l refer to the transmissibility, adoptability or probability of an entity being infected or activated or an entity adopting a given behaviour or state.

Independent multiple processes take place independently without direct influence on their parameters as they ignore the dependence of multi-spread in co-infected status or just exclude the concurrent infections of multi-spread. Since no research can be found on independent multiple processes where co-infected status is allowed, we mainly focus on independent multiple processes that exclude co-infected status and interact via changes of network topology, attributes or structural patterns.

Mutually exclusive processes refer to the multiple processes propagating over the network where the concurrent infection by more than one spread is not possible. The goal is to investigate circumstances under which the dominance of a single spread can emerge [306]–[308].

These dynamic processes interact on the same network while preserving their independence via assumptions including temporal separation [307], structural separation [282], cross immunity [306], [309], [310] and cross adoption [311], [312]. Under temporal separation, two pathogens can spread independently in separate time steps and interact via network topology changes, such as node removal as a result of death or immunity [307]. Studies on concurrent multiple processes generally adopt cross immunity or cross adoption assumption to deal with the concurrent infection that is not allowed for mutually exclusive competing processes. Under cross immunity, the states of an infected/recovered network vertex are changed to be immune to any other infections [306], [309], [310]. Under cross adoption, the infected network nodes can transit to be just infected by another spread with a specified probability [311]–[313]. There is also research on concurrent multiple processes that use structural separation, where network nodes are grouped and are only available for specific spreading dynamics, such as the simple contagion and the complex contagion that work out for predetermined vertex groups [282].

Under the above mentioned assumptions and basic settings, the models widely employed in studies of single spreading dynamics in section IV-C0.j can be extended and used to model independent, multiple processes. In addition, there are already studies using SI [282], SIS [309], SIR [310], [311], independent cascade model [308], percolation model [307], [310] and Reed–Frost model [306]. For example, [282] propose dynamic message-passing equations for two SI-type competing processes to incorporate the message-passing into the parameters. [309] use the extended model, SI_1I_2S, to model the propagation of two concurrent epidemic spreading processes.

Parameter changes of dynamic processes are discussed in terms of independent multiple processes. There are

scenarios (e.g. election information spread and entertainment information spread in the same social networks) concerning parameterized independent multiple processes where co-infected status is allowed under the ignorance of the dependence of multi-spread without changing their parameters, though no research can be found here.

Most of the mutually exclusive processes involved in the existing studies do not consider dynamically changing parameters of the process [306], [307], [309], [310]. A small number of studies introduce dynamics with parameter changes over time under the impact of node states, where transmissibility varies from nodes' groups (structural patterns) [311], [312], states of neighboring nodes [313], message-passing [282]. The parameters can either change between limited number of fixed values [311], [312] or change continuously according to the network attributes resulting from another spreading process [282].

l: INTERRELATED MULTIPLE PROCESSES

Interrelated multiple processes are characterised by direct unilateral or mutual influence of processes on their parameters. They can interact not only by changing network topology or attributes, but also via parameter changes.

Partially inclusive processes refer to the multiple processes that allow the concurrent infections of nodes while also incorporate the dependence between spreads themselves.

Interrelated multiple processes can have suppressing [314], [315] or supporting [282] relations, which are involved in only limited number of studies and so far all of them consider static networks [282], [314], [315]. These spreading dynamics change parameters with the transition of concurrent infection states. For example, epidemic spreading process, under the suppressing impact of awareness spreading, has different infection probabilities given nodes' different levels of awareness [314]–[316]. Similarly, collaborative multiple processes also have different probabilities under their supporting impact [282].

Under the above mentioned assumptions of concurrent infection, the models mentioned in section IV-C0.j and section IV-C0.k can also be extended and used, where some of them have already been used to model interrelated multiple processes, including SI [282], SIS [316], SIR [314] and $SIS − SIRS$ [315].

Parameter changes of processes here are discussed in the context of interrelated multiple processes. For almost all the existing research, at least one of the considered spreading dynamics has dynamically changing parameters in response to the impact of another spreading dynamics [314]–[316]. [282] further introduce collaborative multiple processes that all change parameters with the message passing of nodes. Similarly to the case of independent multiple processes, parameters of interrelated multiple processes can either change between limited number of fixed values [314]–[316] or change continuously according to the network attributes resulting from another spreading process [282].

D. COMBINATION OF THE NETWORK AND PROCESS DIMENSIONS

In this section, we focus on the superposition of the network dimension and the process dimension, as well as the increasing complexity of modelling CNS considering the interactions and interrelations between the network and a dynamic process.

m: SUPERPOSITION OF NETWORKS AND PROCESSES

Propagation process dynamics, either on static networks or dynamic networks, has been extensively studied using non-machine learning approaches [9], [292], where data-driven machine learning approaches have recently been a popular choice of incorporating more structurally and temporally complex network information [304], [305]. Dynamic networks involved in these studies only allow for topology changes [248], [249], [304], [305] which influences the result of spreading dynamics.

The majority of the independent multiple processes considered in the existing and current research take place on static networks [282], [306], [308]–[311] and only a small number of studies can be found for those on dynamic networks [307], [312]. As a typical example of competing epidemic processes, [307] allows removal of nodes and their edges over time as a representation of death or immunity and this results in the topology changes. However, interrelated multiple processes are involved in only a limited number of studies and so far all of them take place on static networks [282], [314], [315]. Modelling approaches on multiple processes generally employ non-machine learning approaches to model single dynamics on networks, where further research is needed for modelling multiple processes using both non- and data-driven machine learning approaches.

n: INTERACTIONS OF NETWORKS AND PROCESSES

Network dimension and dynamic process dimension can either be interrelated or independent based on whether one dimension can trigger the dynamics of another dimension to change. Parameter changes of dynamic processes as well as network changes of (i) topology, (ii) attributes and (iii) structural pattern, each indicates the changes of dynamics in the process dimension or the network dimension. Thus, the interrelations exist in two scenarios: (a) certain states of networks trigger the parameter changes of a dynamic process, (b) a dynamic process results in one of the above mentioned three types of the network changes. The interrelations can either be described as a one-way influence that is just about (a) or (b), or a mutual influence that refers to both (a) and (b).

Independent VS Interrelated relations between a network and a dynamic process are discussed considering their changes and the corresponding causes of change within the CNS.

An independent relation between the network and a dynamic process is common in terms of dynamic processes on networks and the research space is dominated by this approach [9], [249], [292]. In this case, a dynamic process only influences and causes changes of the node attributes connected with the process itself (e.g. whether a node, as a result of the process, has been infected or adopted new behaviour) rather than a change of the network structure or dynamics. In this scenario parameters of a dynamic process are not altered by the changes in the network.

Networks and dynamic processes with interrelated relations between them consider their mutually or unilaterally triggered changes. Multiple processes that interact with each other via changing network attributes are typical examples of a one-way influence of the type (a), where parameters of spreading dynamics can change with network attributes [282], [311]–[316]. For example, in a rumour-truth mixed spreading scenario, the truth spreading rate gets lower when nodes are attributed as rumour-believers [313]. There are also cases for a one-way influence of the type (b), where networks change topology in response to the spread. For example, disease spreading through the network can leave some nodes dead and get them removed [307]. Currently no research is found on interrelation about a mutual influence between the process dimension and the network dimension so the situation where a closed feedback loop between the process and the network is considered.

Parameter changes of dynamic processes are here discussed in terms of interactions and interrelations between networks and processes. All the available examples of processes on networks that are dynamically changing under the impact of networks are about multiple processes, where networks serve as a media for their interactions [282], [282], [311], [311], [312], [312], [313]. Further research is required in the space of dynamic processes that are dynamically changing under the impact of network changes. Network topology that changes in response to the dynamically changing dynamics is another interesting and not addressed research gap.

E. CONTROL MECHANISMS ON CNS

Control mechanisms of CNS aim to find the optimal strategy to attain its desired state, which involves controllability and synchronization of networks and a control of dynamic processes on networks.

o: NETWORK CONTROL

Controllability of networks is achieved using model–based or data–driven approaches [317]. Model–based approaches aim to find an optimal set of driver nodes for CNS under the assumption of a tractable model for network dynamics, where a linear time-invariant systems are often employed to approximate the nonlinear processes that drive the directed networks [18], [318], [319]. This approach identifies the driver nodes via maximum matching approach [320], which enables the calculation of the structural controllability [318] and exact (state) controllability [319]. In this context, many studies investigate the impact of topology variations on the controllability of directed networks, ranging from

degree distributions [318], connection types [18], topology switching [20] to even all possible network structures [320]. To measure the robustness of controllability, simulations of node removal and edge removal attacks are also conducted [320] and convolutional neural networks can be further utilised to improve computation efficiency [321]. Data–driven based approaches, on the contrary, learn controls from network data without knowing network dynamics [317]. Relevant studies generally focus on undirected, directed or weighted networks, where machine learning methods like reinforcement learning are used to find their optimal control parameters for desired network states [317], [322].

Synchronization of CNS, as another type of control towards the desired synchronized state, also involves intensive studies that are well summarised from the perspective of phase oscillator models, stability of synchronised state, and synchronisation in complete or spare networks [43], [323]. In addition, given the similar definitions of synchronization and consensus problem of multi-agent systems [324], they can also be studied from a unified view point by employing ideas about consensus problems across disciplinary areas to complex networks [325]–[327].

p: PROCESS CONTROL

Controllability of dynamic processes is achieved via changing networks or introducing another dynamic process.

Control via changes of networks refers to the control of spreading dynamics via changing network topology and attributes, involving non-deep learning–based, deep learning–based and manual strategy–based approaches. Taking control of epidemic spreading processes as an example, researchers seek an optimal set of control actions including topology changes like node removal and edge removal as well as attribute changes via an antidote allocation to minimise infections [9], [292], [322]. Non-deep learning–based approaches, as is well summarised in [292], mimick spreading dynamics, with the non-deep learning models reviewed in section IV-C0.j, and optimise the action sequence under corresponding model constraints by a mean-field approximation or geometric programming. Deep learning–based approaches use machine learning methods to seek the optimal action sequence over graphs based on the network information that is incorporated by the network embedding. For example, [328] control the epidemic processes over a temporal attributed network using reinforcement learning as a ranking module for actions of changing node attributes, where GNN is encapsulated to embed information about networks and epidemic process. Manual strategy–based approaches focus on the simulation and comparison of manual strategies that change network topology or attributes, like different social network-based distancing strategies proposed and compared to reduce infections of Covid-19 [152].

Control via interactions between processes refers to the control of spreading dynamics realised via introducing another spreading dynamics with their competitive,

suppressing or supporting impact. For example, the disease containment of a single epidemic spreading dynamics *A* can be controlled by introducing its competitive process *B*, which is realised by an optimal allocation of a limited number of *B* spreaders to minimise the spreading of *A* [282], [308]. The advertising campaign *A* can also be controlled via introducing its collaborative spreading process *B*, where the best joint advertising campaign can be designed via the optimal allocation of spreader *B* with the aim of maximising the number of susceptible nodes [282].

V. HOW DO WE APPROACH THE ULTIMATE GOAL?

In this section, we answer the question of how to approach the ultimate goal of modelling complex networked systems: building a Digital Twin (DT) for real world networked systems. This question is decomposed into three sub-questions: (1) What have we done so far to achieve the goal? (2) How far are we from building the DT for CNS? and (3) How can we move forward?

Existing research on modelling networked systems and their dynamics aims at representing the complex reality through networked structures that minimise information loss and ensure that the model's aims can be fulfilled. The complementary effect of the model's aims fulfilment and minimised information loss contributes to a good model of CNS and its convergence to a DT. To assess the networked systems models and narrow their gaps with DTs, we build an assessment framework from two perspectives: (i) the CNS model's aim fulfilment that diverges when a specific model's aims and an abstract model's aims each focuses on the external tasks and inner rules; and (ii) a DT faithful representation and modelling that helps to merge the requirement of both the specific model's aims about external tasks and the abstract model's aims about its inner rules.

A. WHAT HAVE WE DONE SO FAR?

Researchers have done a lot of work on modelling real world networked systems. A small number of studies have already attempted to develop Digital Twins of complex networked systems for specific application contexts, like IoT systems [103], [104] and blockchain-encapsulated systems [105], [106]. However, while many recent studies on modelling, simulation and control of complex networked systems started taking into account the necessary details to faithfully represent aspect of complex reality, none of them explicitly attempted to create a DT of CNS with all its implications which we will now discuss in more detail.

q: BOTTOM-UP VIEW: CNS-BASED ATTEMPTS

Researchers have been trying to build a Complex Networked System (CNS) that can faithfully represent and adapt to the real world situation. The networks, as the basic representations of CNS, are approaching reality with more faithful representation of real world information and infusion of evolving dynamics. These attempts partially enable CNS to meet the requirements of Digital Twins in terms of similarity

to reality by incorporating complex inner rules and fulfilling external tasks.

Networks are approaching reality as the structural, temporal, spatial and dynamics dimensions are gradually taken into account. For networks constructed in a data-driven way based on readily observable data set, the description of the the impact of time enables the modelling of data-driven networks to capture the evolving feature of real-world systems, while spatial information represented in the network structure serves as space where dynamics take place. Spatial temporal networks, where temporal networks are modelled under the constrain of space that influences the structure of the networks, are proposed to encompass both temporal and spatial information, which are the closest to the reality data-driven network structures considered until now. For simulation-based networks and hybrid networks, the rule-based simulations have been developed to approach reality as more complex inner rules of complex networked systems are simulated to incorporate the above mentioned complexity dimensions. Networks simulated from microscopic views of agents enable the representation of either real or simulated information in a flexible and faithful manner. From another perspective, the networks have also been widely employed in different scenarios across disciplines, where a wide range of research objects can be represented and analysed with network structures. As is detailed in section III-B, networks are employed to represent information in agent-based systems and graph structures, relations that are statistically or semantically extracted and constructed, complex systems infused in networks like IoTs and Blockchains.

Dynamics of and on CNSs with different possible interrelations between those two is another key aspect of building an accurate model of CNS. With the infusion of dynamics into the networks, the CNS is getting further on the road to a DT. Referring to section IV, there are built-in dynamics that trigger the evolution of networks, while there are also dynamics that take place on the networks under the influence of network structures. CNS composed of networks and dynamics are approaching reality as researchers try to understand and model the interrelations among network dynamics and dynamic processes. Until now, there is considerable literature on dynamic processes on networks that are either static or dynamic. Among different interrelations between dynamics and networks, we find that some studies consider the unilateral interrelation between dynamic processes and static networks. However, we can hardly find any literature on the interrelation of dynamic processes and network dynamics, which is actually the closest to the real world situation.

r: TOP-DOWN VIEW: DT-DIRECTED METHODS

The digital twining tasks vary for different cases and require to be adopted and adapted in the context of a CNS. For example, [89] model the DT of an urban-integrated hydroponic farm, where they decompose the modelling

process into three crucial elements: data creation that enables an extensive monitoring system for a virtual representation of the farm through data, data analysis that helps to identity key influencing variables, data modelling that enables forecast and feedback. [329] try to shape the actual state and a possible future of the Product Data Technologies from a Closed-Loop Product Lifecycle Management (C-L PLM) perspective, where they see an intelligent product as a product system which contains sensing, memory, data processing, reasoning and communication capabilities at four intelligence levels. [330] view the physical asset and its digital twin as two coupled dynamical systems that evolve over time through their respective state spaces, where the digital twin acquires and assimilates observational data from the asset (e.g., data from sensors or manual inspections) and uses this information to continually update its internal models so that they reflect the evolving physical system. These up-to-date internal models can then be used for analysis, prediction, optimisation and control of the physical system. Referring to the previous studies, we generally decompose the modelling of a DT as tasks including (a) data processing that includes data creation and data integration, (b) data analysis with the purpose of parameter selection, (c) data modelling that enables forecast of eventualities and feedback of real system under the impact of the decision that is made by reference to the forecast.

Data processing, as the fundamental task of modelling a DT, is composed of two parts: (i) data creation enabled by an extensive and robust monitoring system that tracks the observable information, (ii) data integration that features with the record, management and retrieval of information in a real time. In the data creation stage, taking the DT of hydroponic farm built by [89] as an example, they track changing environmental conditions and crop growth through unstructured manual records and a wireless sensor network that sends data in real time to a server. For data integration function, semantic modelling that includes the application of ontology has also been employed to equip the DT with context awareness through record of data, answering queries and information retrieval [91], [331]–[335]. To accommodate and deal with a system's context and uncertainty when selecting an interaction mode that reconfigures the system's model in response to the model output, a fuzzy logic rule based layer has often been employed in complex systems, which enables the interaction mode selection while considering the intermediate ranges between the modes [84], [103]. While there also emerges another popular data integration approach that employs blockchain technology. The blockchain can serve as the middleware of IoT with improved interoperability, privacy, security, reliability and scalability [336]–[338]. However, in some cases, it is hard to collect and process readily real-time data for a well-established DT in an efficient way and there are studies attempting to deal with problems of data integrity. For example, [339] propose a collaborative city digital twin based on federated learning, where multiple city DT can learn

a shared model while keeping all the training data locally. This is a promising solution to accumulate the insights from multiple data sources efficiently and avoid violating privacy rules.

Data analysis and variable selection task is of importance to the establishment of DTs, which is presented by [89] as something that includes: (a) the influence of the environment on physical asset, (b) the influence of operable controls on the environment, and (c) the influence of manual changes on the operational controls, where, within the limitations of the data, this exercise identifies the variables which are crucial to track and forecast. Data analysis that enables variable selection is of importance to the modelling of DT. However, most studies ignore this process and just predetermine the observable variables, while the fundamental question of how to identify the minimum number of observable variables has been understudied over the years and needs a systematic research and answer. As is detailed by [330], a well-designed digital twin should be comprised of models that provide a sufficiently complex digital state space, capturing variation in the physical asset that is relevant for diagnosis, prediction, and decision-making in the application of interest. On the other hand, the digital state space should be simple enough to enable tractable estimation of the digital state, even when only partially observable. Specifically, as a rare example of selecting variables, [89] identify key influencing variables on energy use and crop yield by analysing the relationships between the broad data collected based on temperature, visible radiation, and CO_2 levels.

Data modelling process for a DT is essentially a forecasting model that predicts and provides feedback on real world system to help control the DT, which includes two fundamental tasks, a forecast of extended future and a feedback of real-world system. Prediction and inference of reality that has happened before is the basic function of forecasting models, while DTs can further forecast the extended reality by predicting facts that have never happened before. For example, [340] propose a disaster city DT for enhancing disaster response and emergency management processes, where disasters that have never happened before are simulated and real world systems are extensively forecast to enable increased visibility into network dynamics of complex disaster management and humanitarian actions. The digital twining of complex networked systems are also featured with a decision-making feedback loop with dynamically updated asset specific computational models infused [330]. Especially in cases of solving multi-objective optimisation problems for complex systems using DT, which are common in analyses of entire product lifecycle in manufacturing, researchers have proposed DT frameworks aimed at multi-objective optimisation with effective feedback from different dynamics. [341] enhance DTs of autonomous manufacturing systems through reinforcement learning of continuous data fed back from DT, where residual errors between DT and its physical counterpart are compensated and an improved autonomous system can be established.

[342] propose a bi-level iterative coordination mechanism to achieve optimal design performance for AFMS, where an effective feedback of collected decision-support information from the intelligent multi-objective optimisation of the dynamic execution is presented. To deal with uncertain information in real life scenarios involving multi-objective optimization problems, a fuzzy programming method can be employed to obtain optimal solutions [343]

B. HOW FAR ARE THE CNSs, AS THEY ARE CURRENTLY MODELLED, FROM DTs?

In most cases, CNS models presented in the literature, while fulfilling the relatively simple model's aims under relatively stringent assumptions, have not been developed with the goal of becoming DTs of their modelled aspects of reality. In effect they only posses partial features of DTs. Therefore, in order to bring the two areas closer together, we propose a unified assessment approach by discussing and attempting to answer the following three questions: (1) What constitutes a good DT? (2) What is a good CNS model? (3) To what extent the current CNS models approach a DT? We try to answer the first two questions with measures that aim at assessing the performance of CNS and metrics that evaluate the quality of DT, as is shown in Fig.6, and try to answer question (3) in the context of a good CNS that performs well under certain model's aims and thus has the potential of becoming a DT.

s: WHAT CONSTITUTES A GOOD DT AND HOW TO ASSESS IT?

DTs are featured with integrated functions like simulation, optimization and data analytics [96]. DTs use real-time processing and updates characterised by: (1) real-time connection with the physical entity, (2) self-evolution that enables a DT to learn and adapt in real-time by providing feedback to both the physical asset and the DT, (3) continuous machine learning analysis (dependent on the frequency of the synchronisation), not just a one-time output forecasting, (4) availability of time-series (or time continuous) data for monitoring, (5) level of autonomy that defines if a DT could either make changes to the physical asset itself or if it relies on a human in control who could make changes to the DT, where the property of a DT to be autonomous, not autonomous, or partly autonomous is case-dependent and (6) synchronisation which could be partly continuous or partly event-based [95]. A good DT should meet the requirement of researchers using relatively simple models while preserving the trust in the data, model and their updates [344], [345]. Based on the above features, the assessment of a good DT in the context of CNS can be categorised into two parts: (i) efficiency of data processing and modelling and (ii) similarity with reality from the perspectives of multiple model's aims, self-evolving dynamics and model updates (see Fig.6).

The evaluation of efficiency includes data processing efficiency and modelling efficiency. Data processing efficiency involves data quality such as validity and reliability
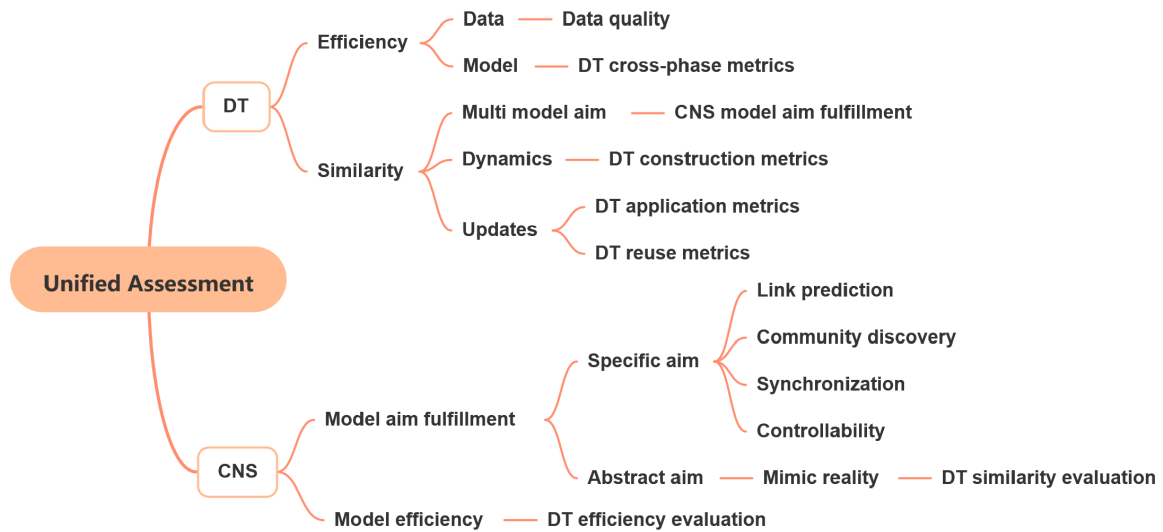
FIGURE 6. The plot of unified assessment criterion for CNS and its distance to a DT.

enabled by handling the imperfection of real-time data ranging from imprecision, uncertainty, incompleteness to ambiguity during the process of information retrieval and data integration [346]. Metadata ("data about the data") captures aspects of the measurement process that may affect the reliability and future usability of the data, which partially addresses trust in data gathered by sensors [344]. The belief function theory is also utilised to estimate the reliability of the information sources [346]. However, as the efficiency of data processing is centered on the observability of experimental physical asset confined by availability of time-series and the real-time connection with the physical entity, it is hard to create quantifiable measures for various evolving application scenarios. For modelling efficiency of a DT, cost, model maturity and model adaptability summarised as a DT cross-phase metrics can be utilised in the assessment [345], where a high-quality model may cost less in maintenance and reuse. A more mature model gives the expected outcomes and meets application requirements better as time and frequency of using the model increase and a highly-adaptable model recreates the status of the real system better. The cost and model maturity are quantifiable in each application scenario [347], [348], while the adaptability is hard to be quantified but can be enhanced via parameter sensitivity analysis [349] and continuos monitoring of the model's accuracy over time.

Similarity level between modelled dynamics and reality can be evaluated from the perspectives of multiple model's aims, dynamics and the model updates of parameters in response to the real time data integration and feedback from real systems. A good DT is characterised with multiple model's aim fulfilment and well-handled trade-off between model performance and model complexity. The validation and verification of a good DT is dependent on the model-aim directed evaluation of model output for external tasks and the faithful representation of the inner rules of real systems. In case of a CNS modelled using a DT approach, the evaluation involves the comprehensive application of CNS model-aim evaluation methods (see section V-B0.t). Similarity of the modelled network dynamics and dynamic processes with that of real world systems can be evaluated using DT construction metrics summarised by [345], including quantifiable credibility, fidelity, maturity and qualitative description of complexity and DT standardisation, while the similarity in the context of model updates with the evolving reality can also be evaluated using DT application metrics including failure rate and qualitative description of decoupling ability and parallelizabiiity as well as DT reuse metrics including degree of reconfigurability, reconstructibility and composability. Based on the forementioned methods, the capability of forecasting events that have never happened before and the synchronisation of nodes existing in the CNS can be assessed and enhanced. However, there is no unified quantifiable measures across application scenarios, where the assessment and the comparison of DT modelling methods can be further studied.

t: IS IT A GOOD CNS?

To answer this question, both measures and standards proposed for CNS and DT can be considered towards the goal of building a good DT with high data processing efficiency and similarity of dynamics with the reality. The assessment of CNS can be divided into two parts: (i) model's aim fulfilment assessment, and (ii) model efficiency assessment (see Fig.6).

There is considerable literature on the specific model's aims with measurable outputs such as link prediction, community discovery, synchronization, observability and controllability. The evaluation methods for community discovery can be summarised as the internal and external quality evaluation, where more detailed measures can be

found in [11]. The assessment of synchronization of networks focuses on the stability of identical states [43], where evaluation of consensus of multi-agent system can also be utilised as they have similar definitions and can be studied from a unified point [325]–[327]. The observability, with its dual: controllability, can be categorised as structural and dynamical, each representing observability of topology [18], and variables for coupling nodes and node dynamics [19], [20]. For its evaluation, observability matrix based on the dynamic model for a linear (time-invariant) system proposed by [17] has been widely used and extended, while the observability and controllability of nonlinear networks are also studied to investigate the effect of nonlinear dynamical interdependences among variables or the connection with symmetries of networks [21], [22]. Particularly, as an evolving model can be mapped to a link prediction algorithm, performance metrics for link prediction can also assist the quantitative comparison of the accuracies of different evolving models [31]. For example, link prediction can be utilised to validate dynamic social network simulators with graph convolutional neural networks (GCN) [40]. Link prediction performance can be evaluated using precision [31], Area Under the Precision–Recall (AUPR) curve [350], Receiver Operating Characteristic (ROC) curves, Area Under the ROC (AUC) [31], Geometric Mean of AUC and PRAUC (GMAUC) [351], Error Rate [273], SumD [272], Kendall's Tau Coefficient (KTC) [253] and Micro/Macro/Weighted Average Precision/Recall/F1 Score [352].

For abstract model's aims such as the mimics of reality using simulation-based networks, we can evaluate these models based on their similarity with the characteristics of reality that they are required to capture, where DT construction metrics [345], as well as observability and controllability measures for CNS can be utilised. In addition, when the simulation-based networks are encapsulated in complex networked systems with forementioned specific model's aims, the evaluation measures for specific model's aims can also reflect to what extent the network simulations reach the research goals, especially in case of link prediction where similarity of nodes, edges and their dynamics is the focus.

DT feature assessment that is applicable in CNS is mainly about the efficiency of simulation or modelling, as modelling efficiency is a good quality persued by both CNS and DT. DT application metrics and the DT cross-phase metrics utilised in the DT modelling efficiency are also applicable in the context of a CNS, even though it is not a DT. The data processing efficiency can be assessed and persued via observability and controllability measures for CNS, as well as the measures utilised to ensure the trust in the data in the evaluation of data processing efficiency for a DT,

u: TO WHAT EXTENT IS IT APPROACHING A DT?

The assessment of the distance of a CNS model to a DT is built on the prerequisite that the CNS is a good CNS, where
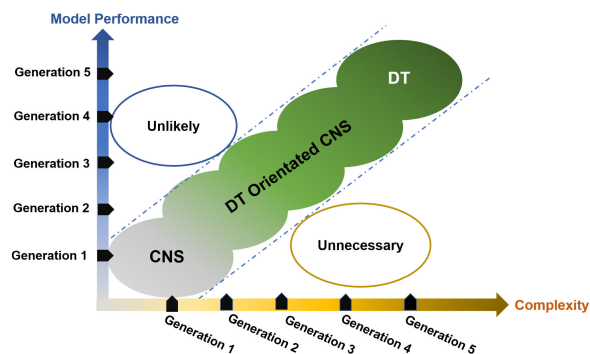

FIGURE 7. The plot of assessing the distance of a CNS to a DT.

a good CNS has partial DT features and has the potential to approach a DT.

A DT has an appropriate level of complexity that enables it, with good model performance in terms of faithful representation of real systems, to meet the models' aim. As we have discussed in the previous sections, a CNS can approach a DT with better model performance through appropriately increased complexity. The distance between a CNS and a DT can also be discussed from these two perspectives: (i) complexity and (ii) model performance (See Fig. 7).

There are no clear boundaries for the development path of a CNS towards a DT, either in terms of model performance or complexity. CNSs with an unnecessary level of complexity, under the lower bound of development path, can be identified when there exists a less complex CNS with equally good or even better model performance. The upper bound of model performance for a CNS naturally exists under the limit of modelling paradigms. When a CNS achieves better model performance through increasing complexity, while falling out of the "unlikely" scope and the "unnecessary" scope, it gets closer to a DT. The bounds of development path for DT-orientated CNS can also be updated with empirical findings in this space.

Complexity of a CNS is hard to measure using one, concrete measure, but we are able to rank the complexity of (i) network representation in each complexity dimension (see section III-A), and (ii) CNS modelling based on the 5-generation framework (see section IV). Based on that, a complexity metric can be identified for each generation of CNSs, including their two components: (i) process and (ii) network representation (see Fig. 8).

For generations of CNSs shown in Fig. 8, their components, the process dimension and the network dimension, are each represented with the coloured symbols of G_1, G_{2a}, G_{2b}, G_3, G_4 and G_5. The CNSs in each generation of models vary in dynamics complexity and temporal complexity, while for structural and spatial complexity, they can be built with any complexity level from those two dimensions. The temporal complexity of a CNS increases as its process, network representation, or both of them, start to change over time in a manner from static (frozen in the time scale), evolving
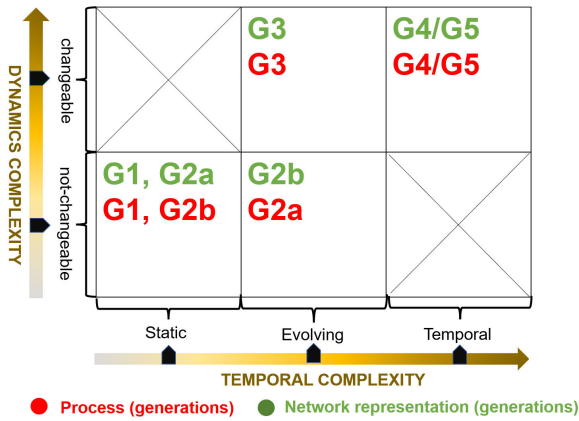
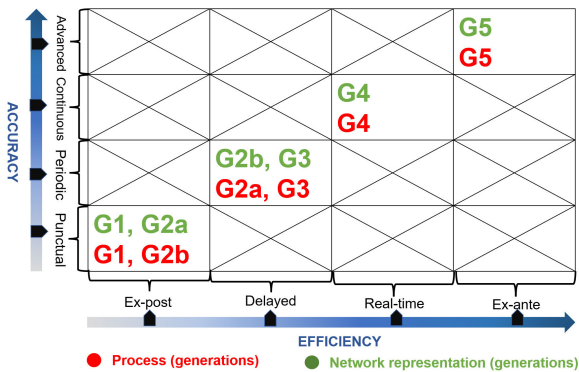**FIGURE 8.** The complexity metric of a CNS.

FIGURE 9. The model performance metric of a CNS.

(captured in time windows) to temporal (continuous). The dynamics complexity of a CNS also increases when the dynamics are changeable, with the process modelled based on the changing parameters or the networks evolving via the changes of inner rules. We can compare the complexity of one component of CNSs in one complexity dimension based on this complexity metric. For example, compared with the CNSs in G_{2a}, CNSs in G_{2b} are characterised with the temporal complexity that is higher in network representation but lower in process dimension.

Model performance of a CNS can be assessed based on the two requirements of building a good CNS: (i) the model's aim fulfilment, and (ii) the model efficiency (see section V-B0.t and Fig. 6). There are both quantitative measures and qualitative description for a CNS assessment, but how to combine them for a comprehensive assessment and to deal with the multi-objective optimisation problem for a good CNS, still requires further study. To gain a rough understanding of the levels of model performance for each generation of CNSs, a model performance metric is built based on (i) the accuracy from the perspective of model's aim fulfilment and (ii) the efficiency (see Fig. 9).

As is shown in Fig. 9, the efficiency is described as ex-post, delayed, real-time and ex-ante based on the CNSs' ways of

data processing and modelling. CNSs and their components in the ex-post group have the lowest ranking of efficiency due to the completely post-hoc modelling, like CNSs in G_1. CNSs in the delayed group are characterised with streams of snapshots feeding into the systems across time windows with a time lag behind the real systems. CNSs in G_4 fall in the real-time group as they conduct real-time data processing and modelling. CNSs in G_5, also termed as DTs, are classified in the ex-ante group as they are not only reactive to the observations of real systems in a real-time manner, but also proactive to the things that have never happened before (enabled by the closed feedback loop). The other perspective, i.e. the accuracy, represents a generalised conception of model performance considering how accurately the model' aims are fulfilled. It is classified as punctual, periodic, continuous and advanced. These groups each requires a faithful representation and modelling of the information at only one static time point, within a discrete period, captured continuously or simulated in advance. The required accuracy level increases with the upgraded assessment criterion and the paradigm shift from G_1 to G_5. For example, the evaluation metrics of community discovery, like the modularity and error rate, are widely used for static networks in G_1. They can be further supplemented with a relative reconstruction error rate to analyse the temporal evolution of dynamic networks and communities in G_2 [156].

Current CNS, in terms of data efficiency, tends to rely more on complex simulation-based networks to be able to capture more realistic features or employ data-driven networks introduced with observable temporal and spatial information. However, the data quality is case-dependent in each application scenario and confined by data sparsity, data security, as well as data processing and representation techniques. It is hard to find studies on CNS built and modelled with real time information because of its observability and the difficulties of building realistic real time data simulator. Though there are studies on CNS built with big data [123], it is still hard to achieve data efficiency at a "real-time" level. Though in some applications of CNS in a DT like IoT, where networked information can be gathered in a real time by sensors and integrated into a Knowledge Graph or a block chain, such a method is not applicable in all application scenarios given the requirement of equipment. Therefore, there exists a large gap when it comes to data efficiency between a good CNS in current studies and a good DT. On the other hand, for model efficiency, some CNS start to approach a DT with the development of modelling and computation techniques like parallel computation [353], edge computing [354], [355] and cloud computing [356], but it is hard to find such empirical research on CNS except for CNS encapsulated in a DT like IoT.

When it comes to the similarity of dynamism of CNS to that of the reality, there are studies on dynamics of the spatio-temporal networks where both time and space are considered to mimic the reality [74], [153], [154]. While there are some studies that model dynamics over dynamic networks

with interrelations between dynamics, including interrelations of dynamic processes over networks [8] and unilateral influence of dynamic process on static networks [357]. However, we can hardly find CNS with mutual influence between dynamics on the networks and dynamics of the networks or dynamic processes over spatio-temporal networks with interrelations between dynamics across temporal and spatial dimensions. Therefore, as the state-of-the-art CNS models have relatively simple aims and often are developed under some strict assumptions, it is hard to find a CNS that fulfills DT standards, unless we focus on DT systems with encapsulated networks like IoT [358] and blockchain [337], where the networked systems are built using DT-oriented and directed methods and the network features like topology and interaction of networked node dynamics are ignored.

In terms of model updates, most studies model evolving dynamics in an offline manner without model updates. There are very few attempts to enable model updates e.g. where complex systems are built on networked mobile devices utilising Federated Learning (FL) [359], [360]. FL selects random subsets of devices in an offline manner to collect the local model updates and share the updated global model with the devices. This method is also used to deploy distributed data processing and learning in wireless networks in a blockchain encapsulated in a DT [361], [362]. The divergence of model updates remain a future research gap. In addition, the assessment methods of CNS are also in need of updates, as they are required to be more dynamic and able to evolve with model changes.

C. HOW CAN WE GO FURTHER?

To answer the question of how to go further to achieve the ultimate goal of building complex networked systems model that faithfully reflects a real system, we build a framework of CNS-based DT, where the modelling can be decomposed into a series of tasks that require modelling methods from both DT and CNS spaces while setting the half-way point as DT-orientated CNS. We start with listing the research gaps that will guide us in setting goals and tasks for future research.

v: CURRENT RESEARCH GAPS

Based on the five generations modelling framework proposed in section IV and the conducted review of the state-of-the-art, the current research can realise modelling framework of generation 1 and generation 2 and a small number of approaches can reach generation 3. There is no research on generations 4 and 5, where further studies are required to model this very complex scenarios. To build CNSs under generation 4 or 5, and in this way achieving DT-orientated CNS, seven research gaps need to be tackled:

1) fulfilment of external tasks (model aim's) while faithfully mimicking the inner rules of the real system;
2) meaningful feature extraction and model selection that enables network representation to preserve as much information as needed for the model's aim fulfilment;

3) network simulation via models built with interpretable inner rules which are able at the same time to deal with structural observability and dynamical observability problems;
4) dynamics of networks that not only focuses on the topology change but also incorporates attribute changes;
5) modelling dynamically changing dynamic processes in a way that allows for continuous parameter changes under the impact of network changes;
6) real time data acquisition and processing for CNS modelling;
7) the establishment of a feedback loop that enables continuous updates of CNS and the changes of real systems referring to the CNS modelling.

w: SET THE HALF-WAY POINT: DT-ORIENTATED CNS

DT-orientated CNS emerges with the convergence of DT and CNS modelling approaches, where CNS models approach reality by introducing DT features to the modelling process while a DT incorporates networked information through blockchain, knowledge graph or IoT to assist data processing and modelling. Studies on CNS generally focus on a single model's aim and simple research objectives with predefined set of assumptions. Also, a vast majority of them use historical rather than streaming and continuous data. DTs can encompass various functions of tools like simulation, optimization and data analytics [96] via real-time processing and updates, with research objects ranging from a single product to the society and relaxed assumptions that allow for a CNS representation and modelling in structural, temporal, spatial and dynamics complexity dimensions. When CNS approaches DT more assumptions are being relaxed, more model's aims fulfilled and more complex features can be modelled via more efficient data processing and modelling.

There are still several challenges to overcome and trade-offs to be made on the way to DT-orientated CNS: (i) trade-off between model performance and multiple model's aims, (ii) trade-off between controllability and complexity, (iii) the trade-off between efficiency and accuracy. The existence of multiple model's aims poses a demanding challenge for modelling, which can only be fulfilled by compromising the model performance for a certain model's aim. As complexity of CNS increases with richer information about network components represented and more complex inner rules modelled, it becomes more difficult to control the CNS with limited number of features. To achieve higher accuracy, which is a measure of model performance for external tasks, also requires more complex CNS structures and dynamics, diminishing the efficiency of CNS representation and modelling.

A good DT-orientated CNS refers to the CNS that simulates or models necessary reality to achieve the predetermined model's aim with a DT-level efficiency, which may not need to be a ready DT but is required to deal with the before mentioned trade-offs. Unified assessment criteria made of

mathematical measures for particular model's aims of CNS and the DT evaluation metrics that reflect model efficiency can be used to assess DT-orientated CNS in a dynamic way, as an assistance of model selection and updates. Modelling paradigms of DT across disciplines can also be introduced to CNS to go further on the road to a DT. Therefore, the modelling tasks of DT-oriented CNS and the research gaps are mainly about the resolution of trade-offs between performance (output accuracy, input controllability, model efficiency) and complexity, dynamic assessment of evolving dynamics and the introduction of DT features and paradigms to CNS.

x: EMBARKING ON THE MODELLING TASKS

The modelling tasks of DT-orientated CNS should be based on the modelling tasks of CNS and DT that are detailed in section V-A and the assessment approach from a unified view point of CNS and DT in section V-B. The network representation involves tasks of data processing, data analysis and variable selection, while the training of a node and relationships dynamics and dynamics over networks can be summarised as data modelling process, where trust in data, model and updating procedure should be considered with the requirements of observability, similarity and synchronization. Therefore, the modelling tasks of DT-orientated CNS can be generally categorised and presented as:

1) Data processing that cope with imperfection of data;
2) Data analysis and feature selection that considers observability and controllability;
3) Network representation based on the selected variables and the similarity measures;
4) Modelling of real-time self-evolving dynamics;
5) Model updates enabled by reconfiguration and reconstruction;
6) Model evaluation over the entire process.

For the task (1) on data processing and data management, uncertainty analysis has been utilised to deal with data imperfection, while data fusion emerges as a prevalent way to capture reliable, valuable and accurate information. Also knowledge graphs as well as blockchain have been popular choices for data integration and information retrieval. However, it is challenging to ensure the efficiency of data processing and data management, especially given the ''real-time'' feature of a DT and the requirements of data quality, where further research is needed. There is also an issue of data sparsity which requires further research on simulation of CNS to deal with the unobservability and unavailability of data.

For the task (2), data analysis and variable selection that considers observability and controllability, has been studied in the context of CNS over the years. However, more effort is needed given the high demand for adaptability of CNS as they approach reality. Specifically, when it comes to simulation-based networks, how to choose the changeable variables that drive the evolution of networks while preserving the characteristics of real world situation is

an interesting research gap given the problem of data scarcity resulting from the rules of keeping data security.

For the task (3), network representation based on the selected variables and the similarity measures, is thoroughly studied research area, while CNS built with DT approaches with the emphasis on network properties has not been deeply studied. Spatio-temporal network, together with the interrelation and interconnection of dynamics within or over such networks are very interesting topics that can be further studied.

The task (4) on modelling of real-time self-evolving dynamics enabled by continuous machine learning is the core element of building a DT, where the interrelations between dynamics in CNS remain an unexplored area especially when it comes to the mutual influence of dynamics on and of networks. More specifically, the evolving dynamics on and of a network is the most relevant to the case-dependent autonomy of DT system, which can be autonomous, not autonomous, or partly autonomous, where the research on the interventions in networked systems can be further introduced with context-awareness and autonomy.

For the task (5), model updates enabled by the reconfiguration and reconstruction, is closely related to the task (4), where the construction of feedback loop is crucial for the continuous modelling process. There is research on a DT built from the perspective of state transition in a discrete way, where how to narrow the time gap between states and extend the state transition to the continuous modelling is a research gap.

For the task (6), model evaluation over the entire modelling process of DT-orientated CNS needs a unified assessment framework, where concrete measures that consider features of both DT and CNS should be explored. There is already literature on measures of network analysis and DT analysis, though the principled combination of these two or proposing new integrated quality measures remain an outstanding research gap.

The above tasks show the complexity of the research that is needed to be accomplished along the way of working towards DT-orientated Complex Networked Systems.

VI. CONCLUSION

This survey focuses on the modelling approaches of Complex Networked Systems that pave the path for its convergence to the ultimate goal: a Digital Twin of a CNS.

We review and discuss the CNS from three perspectives: (i) model's aims that have been studied for CNS (see section II), (ii) modelling paradigms that enable to represent a networked system in a way that preserves as much information as needed (see section III), (iii) modelling approaches for dynamics of networks and dynamics over networks that enable to meet model's aims (see section IV). Those themes are discussed through the lenses of four, proposed by us, complexity dimensions of complex systems: (i) structural, (ii) temporal, (iii) dynamics and (iv) spatial. A discussion that considers those complexity dimensions

enables to better understand current modelling challenges and quantify how far we are from achieving Digital Twin modelling capabilities when representing networked systems.

The model's aims for CNS distinguish when they focus either on the external task to be performed on the system or modelling of the inner rules of real systems, and this division and specialisation can be eliminated when a Digital Twin is considered since it is able to undertake multiple external tasks by faithfully covering and reflecting the complexities of real systems. Models of CNS proposed over the years have been found to approach real systems with increasing complexity in structural, temporal, spatial and dynamics dimensions. To generate and preserve this heterogeneous networked information, modelling paradigms of network representation get more complex with compromised interpretability. These models either focus on inner rules of network generation at a local level or aim at a compressed network representation at a global level, but all converge to the goal for a faithful representation of real systems.

Dynamics of networks, dynamic processes on networks as well as their interrelations are three elementary sources of complexity for dynamics in the CNS. To navigate a pathway through different levels of complexity of modelling CNS, we devise a modelling framework of CNS that considers all these three elements and consists of five generations reflecting the progress of work that has been done in this field. Each generation builds upon the previous one meaning that the next generation encompasses higher complexity levels than the previous one.

This modelling framework is agnostic to the model's aim so any of the discussed aims can be attempted using models built within each of the generations. Though, one needs to remember that models from different generations will enable to achieve selected aim to a different extent. The proposed framework also shows how Complex Networked Systems' models approach a Digital Twin with more complexity through generations, (i) generation 1: dynamic process on static networks, (ii) generation 2 with two variations: dynamic process on evolving networks and evolving dynamic process on static networks, (iii) generation 3: evolving dynamic processes on evolving networks with interrelations between them, (iv) generation 4: temporal dynamic processes on temporal networks with interrelations between them and the acquisition of real time information and finally (v) generation 5 that further introduces modelling framework of generation 4 with an information feedback from CNS's model to the real system. From generation 1 to generation 5, the real system can be represented more faithfully with richer information captured and finally a CNS-based DT can be created in generation 5. Current studies have made good progress under modelling framework of generation 1 and generation 2. Only small number of approaches reach generation 3. For generation 4 and 5, there is no research in this space and further studies are required to model this very complex scenarios to achieve better performance in the context of any of the presented model's aims.

To be more aware of how to approach a Digital Twin with CNS, we propose an assessment framework (see section V) that aims at quantifying the distance of CNS to DT from the perspective of CNS model's aim fulfilment and the perspective of a DT's faithful representation of reality. A half-way point referred to as a DT-orientated CNS is proposed to bridge the gap between the current approaches to modelling of CNS and the ultimate goal of a DT (generation 5 models) for future study.

The goal of the future research in the space of complex networked systems and network science more broadly is to develop a DT-orientated CNS that will enable to address research gaps presented in Section V-C. Integrating dynamic networks with dynamic processes and allowing for mutual influence between them (allowing at the same time for continuous adaptation of the system using streaming data as an input) will make it possible to create the DT-orientated CNSs. This will be a major breakthrough in the space of modelling CNS.

REFERENCES

[1] L. d. F. Costa, O. N. Oliveira Jr., G. Travieso, F. A. Rodrigues, P. R. V. Boas, L. Antiqueira, M. P. Viana, and L. E. C. Rocha, "Analyzing and modeling real-world phenomena with complex networks: A survey of applications," *Adv. Phys.*, vol. 60, no. 3, pp. 329–412, 2011.

[2] X. F. Liu, X.-J. Jiang, S.-H. Liu, and C. K. Tse, "Knowledge discovery in cryptocurrency transactions: A survey," *IEEE Access*, vol. 9, pp. 37229–37254, 2021.

[3] C. Celes, A. Boukerche, and A. A. F. Loureiro, "Mobility trace analysis for intelligent vehicular networks: Methods, models, and applications," *ACM Comput. Surveys*, vol. 54, no. 3, pp. 1–38, Apr. 2022.

[4] E. B. Tirkolaee, A. Goli, and G.-W. Weber, "Multi-objective aggregate production planning model considering overtime and outsourcing options under fuzzy seasonal demand," in *Proc. Int. Sci.-Technical Conf. Manuf.*, Cham, Switzerland: Springer, 2019, pp. 81–96.

[5] K. Batool and M. A. Niazi, "Modeling the Internet of Things: A hybrid modeling approach using complex networks and agent-based models," *Complex Adapt. Syst. Model.*, vol. 5, no. 1, pp. 1–19, Dec. 2017.

[6] D. R. Amancio, "Probing the topological properties of complex networks modeling short written texts," *PLoS ONE*, vol. 10, no. 2, Feb. 2015, Art. no. e0118394.

[7] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas, "Characterization of complex networks: A survey of measurements," *Adv. Phys.*, vol. 56, no. 1, pp. 167–242, 2007.

[8] P. Brodka, K. Musial, and J. Jankowski, "Interacting spreading processes in multilayer networks: A systematic review," *IEEE Access*, vol. 8, pp. 10316–10341, 2020.

[9] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *IEEE Control Syst.*, vol. 36, no. 1, pp. 26–46, Feb. 2016.

[10] J. Skarding, B. Gabrys, and K. Musial, "Foundations and modeling of dynamic networks using dynamic graph neural networks: A survey," *IEEE Access*, vol. 9, pp. 79143–79168, 2021.

[11] G. Rossetti and R. Cazabet, "Community discovery in dynamic networks: A survey," *ACM Comput. Surv.*, vol. 51, no. 2, pp. 1–37, Jun. 2018.

[12] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J.-P. Hubaux, "Game theory meets network security and privacy," *ACM Comput. Surv.*, vol. 45, no. 3, pp. 1–39, Jun. 2013.

[13] M. N. Schmidt and M. Morup, "Nonparametric Bayesian modeling of complex networks: An introduction," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 110–128, May 2013.

[14] F. Degenhardt, S. Seifert, and S. Szymczak, "Evaluation of variable selection methods for random forests and omics data sets," *Briefings Bioinf.*, vol. 20, no. 2, pp. 492–503, Mar. 2019.

[15] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: Model selection and overfitting," *Nature methods*, vol. 13, no. 9, pp. 703–705, 2016.

[16] L. A. Aguirre, L. L. Portes, and C. Letellier, "Structural, dynamical and symbolic observability: From dynamical systems to networks," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0206180.

[17] T. Kailath, *Linear Systems*, vol. 156. Englewood Cliffs, NJ, USA: Prentice-Hall, 1980.

[18] D. Leitold, Á. Vathy-Fogarassy, and J. Abonyi, "Controllability and observability in complex networks—The effect of connection types," *Sci. Rep.*, vol. 7, no. 1, pp. 1–9, Dec. 2017.

[19] I. Sendina-Nadal and C. Letellier, "Observability of dynamical networks from graphic and symbolic approaches," in *Proc. Int. Workshop Complex Netw.*, Cham, Switzerland: Springer, 2019, pp. 3–15.

[20] F. L. Iudice, F. Sorrentino, and F. Garofalo, "On node controllability and observability in complex dynamical networks," *IEEE Control Syst. Lett.*, vol. 3, no. 4, pp. 847–852, Oct. 2019.

[21] A. J. Whalen, S. N. Brennan, T. D. Sauer, and S. J. Schiff, "Observability and controllability of nonlinear networks: The role of symmetry," *Phys. Rev. X*, vol. 5, no. 1, Jan. 2015, Art. no. 011005.

[22] C. Letellier, I. Sendiña-Nadal, and L. A. Aguirre, "Nonlinear graph-based theory for dynamical network observability," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 98, no. 2, Aug. 2018, Art. no. 020303.

[23] H. Alvari, A. Hajibagheri, and G. Sukthankar, "Community detection in dynamic social networks: A game-theoretic approach," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2014, pp. 101–107.

[24] M. Qin, D. Jin, K. Lei, B. Gabrys, and K. Musial-Gabrys, "Adaptive community detection incorporating topology and content in social networks," *Knowl. Based Syst.*, vol. 161, pp. 342–356, Dec. 2018.

[25] D. Jin, B. Gabrys, and J. Dang, "Combined node and link partitions method for finding overlapping communities in complex networks," *Sci. Rep.*, vol. 5, no. 1, pp. 1–8, Aug. 2015.

[26] M. Rosvall and C. T. Bergstrom, "Mapping change in large networks," *PLoS ONE*, vol. 5, no. 1, Jan. 2010, Art. no. e8694.

[27] N. P. Nguyen, T. N. Dinh, Y. Xuan, and M. T. Thai, "Adaptive algorithms for detecting community structure in dynamic social networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2282–2290.

[28] U. Braun, "Dynamic reconfiguration of frontal brain networks during executive cognition in humans," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 37, pp. 11678–11683, 2015.

[29] K. T. Macon, P. J. Mucha, and M. A. Porter, "Community structure in the united nations general assembly," *Phys. A, Stat. Mech. Appl.*, vol. 391, nos. 1–2, pp. 343–361, 2012.

[30] F. Gao, K. Musial, and B. Gabrys, "A community bridge boosting social network link prediction model," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Jul. 2017, pp. 683–689.

[31] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Phys. A, Statist. Mech. Appl.*, vol. 390, no. 6, pp. 1150–1170, 2011.

[32] V. Martínez, F. Berzal, and J.-C. Cubero, "A survey of link prediction in complex networks," *ACM Comput. Surv.*, vol. 49, no. 4, pp. 1–33, Dec. 2017.

[33] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 58, no. 7, pp. 1019–1031, 2003.

[34] A. Wahid-Ul-Ashraf, M. Budka, and K. Musial, "How to predict social relationships—Physics-inspired approach to link prediction," *Phys. A, Stat. Mech. Appl.*, vol. 523, pp. 1110–1129, Jun. 2019.

[35] A. Clauset, C. Moore, and M. E. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, no. 7191, pp. 98–101, 2008.

[36] F. Fouss, A. Pirotte, and M. Saerens, "An experimental investigation of graph kernels on a collaborative recommendation task," in *Proc. 6th Int. Conf. Data Mining (ICDM)*, Dec. 2006, pp. 863–868.

[37] M. Nayyeri, G. M. Cil, S. Vahdati, F. Osborne, M. Rahman, S. Angioni, A. Salatino, D. R. Recupero, N. Vassilyeva, E. Motta, and J. Lehmann, "Trans4E: Link prediction on scholarly knowledge graphs," *Neurocomputing*, vol. 461, pp. 530–542, Oct. 2021.

[38] L. Lü, C.-H. Jin, and T. Zhou, "Similarity index based on local paths for link prediction of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 80, no. 4, Oct. 2009, Art. no. 046122.

[39] G. B. Rossi, "Measurability," *Measurement*, vol. 40, no. 6, pp. 545–562, 2007.

[40] A. Wahid-Ul-Ashraf, M. Budka, and K. Musial, "Simulation and augmentation of social networks for building deep learning models," 2019, *arXiv:1905.09087*.

[41] K. Lei, M. Qin, B. Bai, G. Zhang, and M. Yang, "GCN-GAN: A nonlinear temporal link prediction model for weighted dynamic networks," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2019, pp. 388–396.

[42] Y. Wang, W. Wang, Y. Liang, Y. Cai, J. Liu, and B. Hooi, "NodeAug: Semi-supervised node classification with data augmentation," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 207–217.

[43] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, "Synchronization in complex networks," *Phys. Rep.*, vol. 469, no. 3, pp. 93–153, Dec. 2008.

[44] F. Dörfler, and F. Bullo, "Synchronization and transient stability in power networks and nonuniform Kuramoto oscillators," *SIAM J. Control. Optim.*, vol. 50, no. 3, pp. 1616–1642 2012.

[45] F. Dörfler, M. Chertkov, and F. Bullo, "Synchronization in complex oscillator networks and smart grids," *Proc. Nat. Acad. Sci. USA*, vol. 110, pp. 2005–2010, Feb. 2013.

[46] E. Bullmore and O. Sporns, "The economy of brain network organization," *Nature Rev. Neurosci.*, vol. 13, pp. 336–349, Apr. 2012.

[47] Y. Tang, H. Gao, and J. Kurths, "Distributed robust synchronization of dynamical networks with stochastic coupling," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 5, pp. 1508–1519, May 2014.

[48] C. Hammond, H. Bergman, and P. Brown, "Pathological synchronization in Parkinson's disease: Networks, models and treatments," *Trends Neurosci.*, vol. 30, no. 7, pp. 357–364, Jul. 2007.

[49] R. C. W. Wong, T. W. Y. Yuen, K. W. Fung, and J. M. Y. Leung, "Optimizing timetable synchronization for rail mass transit," *Transp. Sci.*, vol. 42, no. 1, pp. 57–69, Feb. 2008.

[50] M. Wickramasinghe and I. Z. Kiss, "Spatially organized dynamical states in chemical oscillator networks: Synchronization, dynamical differentiation, and chimera patterns," *PLoS ONE*, vol. 8, no. 11, Nov. 2013, Art. no. e80586.

[51] I. Sohn, "Small-world and scale-free network models for IoT systems," *Mobile Inf. Syst.*, vol. 2017, pp. 1–9, Jan. 2017.

[52] R. Cazabet and F. Amblard, "Simulate to detect: A multi-agent system for community detection," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.*, Aug. 2011, pp. 402–408.

[53] S. Ranshous, S. Shen, D. Koutra, S. Harenberg, C. Faloutsos, and N. F. Samatova, "Anomaly detection in dynamic networks: A survey," *Wiley Interdiscipl. Rev., Comput. Stat.*, vol. 7, no. 3, pp. 223–247, 2015.

[54] P. V. Bindu, P. S. Thilagam, and D. Ahuja, "Discovering suspicious behavior in multilayer social networks," *Comput. Hum. Behav.*, vol. 73, pp. 568–582, Aug. 2017.

[55] W. Yu, W. Cheng, C. C. Aggarwal, K. Zhang, H. Chen, and W. Wang, "NetWalk: A flexible deep embedding approach for anomaly detection in dynamic networks," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 2672–2681.

[56] L. Kendrick, K. Musial, and B. Gabrys, "Change point detection in social networks—Critical review with experiments," *Comput. Sci. Rev.*, vol. 29, pp. 1–13, Aug. 2018.

[57] M. Thottan and C. Ji, "Anomaly detection in IP networks," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2191–2204, Aug. 2003.

[58] S. Rajasegarar, C. Leckie, and M. Palaniswami, "Anomaly detection in wireless sensor networks," *IEEE Wireless Commun.*, vol. 15, no. 4, pp. 34–40, Aug. 2008.

[59] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 303–336, 1st Quart., 2014.

[60] Cancer Genome Atlas Research Network, "Comprehensive genomic characterization defines human glioblastoma genes and core pathways," *Nature*, vol. 455, no. 7216, p. 1061, 2008.

[61] B. Jia, C. Dong, Z. Chen, K.-C. Chang, N. Sullivan, and G. Chen, "Pattern discovery and anomaly detection via knowledge graph," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 2392–2399.

[62] R. Ferrer i Cancho, R. V. Solé, and R. Köhler, "Patterns in syntactic dependency networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 5, May 2004, Art. no. 051915.

[63] A. R. Backes, D. Casanova, and O. M. Bruno, "A complex network-based approach for boundary shape analysis," *Pattern Recognit.*, vol. 42, no. 1, pp. 54–67, 2009.

[64] J. Zhang, P. McBurney, and K. Musial, "Convergence of trading strategies in continuous double auction markets with boundedly-rational networked traders," *Rev. Quant. Finance Accounting*, vol. 50, no. 1, pp. 301–352, Jan. 2018.

[65] S. Kaviani and I. Sohn, "Influence of random topology in artificial neural networks: A survey," *ICT Exp.*, vol. 6, no. 2, pp. 145–150, Jun. 2020.

[66] A. Ganesh, L. Massoulie, and D. Towsley, "The effect of network topology on the spread of epidemics," in *Proc. 24th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, vol. 2, Mar. 2005, pp. 1455–1466.

[67] D. Król, D. Fay, and B. Gabryś, *Propagation Phenomena in Real World Networks*. Cham, Switzerland: Springer, 2015.

[68] M. J. Keeling and K. T. D. Eames, "Networks and epidemic models," *J. Roy. Soc. Interface*, vol. 2, no. 4, pp. 295–307, Sep. 2005.

[69] F. Amblard and G. Deffuant, "The role of network topology on extremism propagation with the relative agreement opinion dynamics," *Phys. A, Stat. Mech. Appl.*, vol. 343, pp. 725–738, Nov. 2004.

[70] Y. Cho, J. Hwang, and D. Lee, "Identification of effective opinion leaders in the diffusion of technological innovation: A social network approach," *Technolog. Forecasting Social Change*, vol. 79, no. 1, pp. 97–106, Jan. 2012.

[71] X. Ye, L. Dang, J. Lee, M.-H. Tsou, and Z. Chen, "Open source social network simulator focusing on spatial MEME diffusion," in *Human Dynamics Research in Smart and Connected Communities*. Cham, Switzerland: Springer, 2018, pp. 203–222.

[72] L. Atzori, A. Iera, G. Morabito, and M. Nitti, "The social Internet of Things (SIoT)—When social networks meet the Internet of Things: Concept, architecture and network characterization," *Comput. Netw.*, vol. 56, no. 16, pp. 3594–3608, Nov. 2012.

[73] H. Soh, S. Lim, T. Zhang, X. Fu, G. K. K. Lee, T. G. G. Hung, P. Di, S. Prakasam, and L. Wong, "Weighted complex network analysis of travel routes on the Singapore public transportation system," *Phys. A, Stat. Mech. Appl.*, vol. 389, no. 24, pp. 5852–5863, Dec. 2010.

[74] R. Morais, V. Le, S. Venkatesh, and T. Tran, "Learning asynchronous and sparse human-object interaction in videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 16041–16050.

[75] M. Coscia, F. Giannotti, and D. Pedreschi, "A classification for community discovery methods in complex networks," *Stat. Anal. Data Mining, ASA Data Sci. J.*, vol. 4, no. 5, pp. 512–546, 2011.

[76] R. E. Kalman, "Mathematical description of linear dynamical systems," *J. Soc. Ind. Appl. Math. Ser. A Control*, vol. 1, no. 2, pp. 152–192, 1962.

[77] M. Cremonini and F. Casamassima, "Controllability of social networks and the strategic use of random information," *Comput. Social Netw.*, vol. 4, no. 1, pp. 1–22, Dec. 2017.

[78] S. Wuchty, "Controllability in protein interaction networks," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 19, pp. 7156–7160, May 2014.

[79] T. M. Karrer, J. Z. Kim, J. Stiso, A. E. Kahn, F. Pasqualetti, U. Habel, and D. S. Bassett, "A practical guide to methodological considerations in the controllability of structural brain networks," *J. Neural Eng.*, vol. 17, no. 2, Apr. 2020, Art. no. 026031.

[80] M. Rinaldi, "Controllability of transportation networks," *Transp. Res. B, Methodol.*, vol. 118, pp. 381–406, Dec. 2018.

[81] K. Juszczyszyn, P. Kazienko, K. Musial, and B. Gabrys, "Temporal changes in connection patterns of an email-based social network," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.*, Dec. 2008, pp. 9–12.

[82] M. Jia, B. Gabrys, and K. Musial, "Directed closure coefficient and its patterns," *PLoS ONE*, vol. 16, no. 6, Jun. 2021, Art. no. e0253822.

[83] M. Jia, B. Gabrys, and K. Musial, "Measuring quadrangle formation in complex networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 2, pp. 538–551, Mar. 2022.

[84] R. Lotfi, Y. Z. Mehrjerdi, M. S. Pishvaee, A. Sadeghieh, and G.-W. Weber, "A robust optimization model for sustainable and resilient closed-loop supply chain network design considering conditional value at risk," *Numer. Algebra, Control Optim.*, vol. 11, no. 2, p. 221, 2021.

[85] E. B. Tirkolaee and N. S. Aydin, "Integrated design of sustainable supply chain and transportation network using a fuzzy bi-level decision support system for perishable products," *Expert Syst. Appl.*, vol. 195, Jun. 2022, Art. no. 116628.

[86] R. Lotfi, B. Kargar, A. Gharehbaghi, and G.-W. Weber, "Viable medical waste chain network design by considering risk and robustness," *Environ. Sci. Pollut. Res.*, pp. 1–16, Oct. 2021, doi: 10.1007/s11356-021-16727-9.

[87] A. Goli, E. B. Tirkolaee, and G.-W. Weber, "A perishable product sustainable supply chain network design problem with lead time and customer satisfaction using a hybrid whale-genetic algorithm," in *Logistics Operations and Management for Recycling and Reuse*. Cham, Switzerland: Springer, 2020, pp. 99–124.

[88] E. B. Tirkolaee, A. Goli, P. Ghasemi, and F. Goodarzian, "Designing a sustainable closed-loop supply chain network of face masks during the COVID-19 pandemic: Pareto-based algorithms," *J. Cleaner Prod.*, vol. 333, Jan. 2022, Art. no. 130056.

[89] M. Jans-Singh, K. Leeming, R. Choudhary, and M. Girolami, "Digital twin of an urban-integrated hydroponic farm," *Data-Centric Eng.*, vol. 1, pp. 1–26, Dec. 2020.

[90] S. Haag and R. Anderl, "Digital twin—Proof of concept," *Manuf. Lett.*, vol. 15, pp. 64–66, Jan. 2018.

[91] S. Boschert, C. Heinrich, and R. Rosen, "Next generation digital twin," in *Proc. TMCE*. Las Palmas de Gran Canaria, Spain, 2018, pp. 209–218.

[92] M. Grieves. *Digital Twin: Manufacturing Excellence Through Virtual Factory Replication*. Accessed: Aug. 5, 2019. [Online]. Available: www.apriso.com/library/Whitepaper_Dr_Grieves_DigitalTwin_ManufacturingExcellence.php

[93] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *Int. J. Adv. Manuf. Technol.*, vol. 94, pp. 3563–3576, Mar. 2018.

[94] F. Tao, F. Sui, A. Liu, Q. Qi, M. Zhang, B. Song, Z. Guo, S. C.-Y. Lu, and A. Y. C. Nee, "Digital twin-driven product design framework," *Int. J. Prod. Res.*, vol. 57, no. 12, pp. 3935–3953, 2018.

[95] A. Sharma, E. Kosasih, J. Zhang, A. Brintrup, and A. Calinescu, "Digital twins: State of the art theory and practice, challenges, and open research questions," 2020, *arXiv:2011.02833*.

[96] D. Ivanov, A. Dolgui, A. Das, and B. Sokolov, "Digital supply chain twins: Managing the ripple effect, resilience, and disruption risks by data-driven optimization, simulation, and visibility," in *Handbook of Ripple Effects in the Supply Chain*. Cham, Switzerland: Springer, 2019, pp. 309–332.

[97] A. Madni, C. Madni, and S. Lucero, "Leveraging digital twin technology in model-based systems engineering," *Systems*, vol. 7, no. 1, p. 7, Jan. 2019.

[98] F. Tao, M. Zhang, Y. Liu, and A. Y. C. Nee, "Digital twin driven prognostics and health management for complex equipment," *CIRP Ann.*, vol. 67, no. 1, pp. 169–172, May 2018.

[99] T. H.-J. Uhlemann, C. Lehmann, and R. Steinhilper, "The digital twin: Realizing the cyber-physical production system for industry 4.0," *Proc. CIRP*, vol. 61, pp. 335–340, Jan. 2017.

[100] J. Du, Q. Zhu, Y. Shi, Q. Wang, Y. Lin, and D. Zhao, "Cognition digital twins for personalized information systems of smart cities: Proof of concept," *J. Manage. Eng.*, vol. 36, no. 2, Mar. 2020, Art. no. 04019052.

[101] Y. Liu, L. Zhang, Y. Yang, L. Zhou, L. Ren, F. Wang, R. Liu, Z. Pang, and M. J. Deen, "A novel cloud-based framework for the elderly healthcare services using digital twin," *IEEE Access*, vol. 7, pp. 49088–49101, 2019.

[102] A. Castellani, S. Schmitt, and S. Squartini, "Real-world anomaly detection by using digital twin systems and weakly supervised learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4733–4742, Jul. 2021.

[103] K. M. Alam and A. El Saddik, "C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems," *IEEE Access*, vol. 5, pp. 2050–2062, 2017.

[104] D. Singh, G. Tripathi, and A. J. Jara, "A survey of Internet-of-Things: Future vision, architecture, challenges and services," in *Proc. IEEE World Forum Internet Things (WF-IoT)*, Mar. 2014, pp. 287–292.

[105] Y. He, H. Li, X. Cheng, Y. Liu, C. Yang, and L. Sun, "A blockchain based truthful incentive mechanism for distributed P2P applications," *IEEE Access*, vol. 6, pp. 27324–27335, 2018.

[106] W. Hao, J. Zeng, X. Dai, J. Xiao, Q.-S. Hua, H. Chen, K.-C. Li, and H. Jin, "Towards a trust-enhanced blockchain P2P topology for enabling fast and reliable broadcast," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 2, pp. 904–917, Jun. 2020.

[107] Y. Breitbart, M. Garofalakis, C. Martin, R. Rastogi, S. Seshadri, and A. Silberschatz, "Topology discovery in heterogeneous IP networks," in *Proc. IEEE INFOCOM Conf. Comput. Commun. 19th Annu. Joint Conf. IEEE Comput. Commun. Societies*, Mar. 2000, pp. 265–274.

[108] J. P. K. Doye, "Network topology of a potential energy landscape: A static scale-free network," *Phys. Rev. Lett.*, vol. 88, no. 23, May 2002, Art. no. 238701.

[109] J. R. Tyler, D. M. Wilkinson, and B. A. Huberman, "Email as spectroscopy: Automated discovery of community structure within organizations," in *Communities and Technologies*. Cham, Switzerland: Springer, 2003, pp. 81–96.

[110] N. Shah, A. Beutel, B. Hooi, L. Akoglu, S. Gunnemann, D. Makhija, M. Kumar, and C. Faloutsos, "EdgeCentric: Anomaly detection in edge-attributed networks," in *Proc. IEEE 16th Int. Conf. Data Mining Workshops (ICDMW)*, Dec. 2016, pp. 327–334.

[111] C. Hou, S. He, and K. Tang, "RoSANE: Robust and scalable attributed network embedding for sparse networks," *Neurocomputing*, vol. 409, pp. 231–243, Oct. 2020.

[112] T. Huang, Y. Pei, V. Menkovski, and M. Pechenizkiy, "Hop-count based self-supervised anomaly detection on attributed networks," 2021, *arXiv:2104.07917*.

[113] K. Musiał and P. Kazienko, "Social networks on the internet," *World Wide Web*, vol. 16, no. 1, pp. 31–72, 2013.

[114] D. Jin, Z. Liu, D. He, B. Gabrys, and K. Musial, "Robust detection of communities with multi-semantics in large attributed networks," in *Proc. Int. Conf. Knowl. Sci., Eng. Manage.*, Cham, Switzerland: Springer, 2018, pp. 362–376.

[115] J. Li, K. Cheng, L. Wu, and H. Liu, "Streaming link prediction on dynamic attributed networks," in *Proc. 11th ACM Int. Conf. Web Search Data Mining*, Feb. 2018, pp. 369–377.

[116] J. Li, L. Dueñas-Osorio, C. Chen, and C. Shi, "Connectivity reliability and topological controllability of infrastructure networks: A comparative assessment," *Rel. Eng. Syst. Saf.*, vol. 156, pp. 24–33, Dec. 2016.

[117] H. Xia, L. Li, X. Cheng, C. Liu, and T. Qiu, "A dynamic virus propagation model based on social attributes in city IoT," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8036–8048, Sep. 2020.

[118] X. Liu, W. Song, K. Musial, X. Zhao, W. Zuo, and B. Yang, "Semi-supervised stochastic blockmodel for structure analysis of signed networks," *Knowl.-Based Syst.*, vol. 195, May 2020, Art. no. 105714.

[119] S. Liu, Z. Ji, and H. Ma, "Jordan form-based algebraic conditions for controllability of multiagent systems under directed graphs," *Complexity*, vol. 2020, pp. 1–18, Feb. 2020.

[120] T. V. T. Duong and D. Q. Tran, "An effective approach for mobility prediction in wireless network based on temporal weighted mobility rule," *Int. J. Comput. Sci. Telecommun.*, vol. 3, no. 2, pp. 29–36, 2012.

[121] A. G. Nair, S. L. Brunton, and K. Taira, "Networked-oscillator-based modeling and control of unsteady wake flows," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 97, no. 6, Jun. 2018, Art. no. 063107.

[122] Q. Kang, Y. Kang, and H. Kong, "Edge-attributed community search for large graphs," in *Proc. 2nd Int. Conf. Big Data Res. (ICBDR)*, 2018, pp. 114–118.

[123] H. D. Bedru, S. Yu, X. Xiao, D. Zhang, L. Wan, H. Guo, and F. Xia, "Big networks: A survey," *Comput. Sci. Rev.*, vol. 37, Aug. 2020, Art. no. 100247.

[124] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *Int. J. Adv. Manuf. Technol.*, vol. 94, nos. 9–12, pp. 3563–3576, 2018.

[125] P. Chunaev, "Community detection in node-attributed social networks: A survey," *Comput. Sci. Rev.*, vol. 37, Aug. 2020, Art. no. 100286.

[126] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, "Spectral redemption in clustering sparse networks," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 52, pp. 20935–20940, Dec. 2013.

[127] P. V. Marsden and N. Lin, *Social Structure and Network Analysis*. Newbury Park, CA, USA: Sage, 1982.

[128] J. E. Perry-Smith and C. E. Shalley, "The social side of creativity: A static and dynamic social network perspective," *Acad. Manage. Rev.*, vol. 28, no. 1, pp. 89–106, Jan. 2003.

[129] P. Kefalas, P. Symeonidis, and Y. Manolopoulos, "Recommendations based on a heterogeneous spatio-temporal social network," *World Wide Web*, vol. 21, no. 2, pp. 345–371, 2018.

[130] X. Li, M. Wu, C.-K. Kwoh, and S.-K. Ng, "Computational approaches for detecting protein complexes from protein interaction networks: A survey," *BMC Genomics*, vol. 11, no. 1, pp. 1–19, 2010.

[131] B. Chen, W. Fan, J. Liu, and F.-X. Wu, "Identifying protein complexes and functional modules-from static PPI networks to dynamic PPI networks," *Briefings Bioinf.*, vol. 15, no. 2, pp. 177–194, Mar. 2014.

[132] C. E. Mandl, "Evaluation and optimization of urban public transportation networks," *Eur. J. Oper. Res.*, vol. 5, no. 6, pp. 396–404, 1980.

[133] Y. H. Wu and H. J. Miller, "Computational tools for measuring space-time accessibility within dynamic flow transportation networks," *J. Transp. Statist.*, vol. 4, no. 2, pp. 1–14, 2001.

[134] M. W. Bruner, K. Erickson, B. Wilson, and J. Côté, "An appraisal of athlete development models through citation network analysis," *Psychol. Sport Exercise*, vol. 11, no. 2, pp. 133–139, Mar. 2010.

[135] F. Gao and K. Musial-Gabrys, "Hybrid structure-based link prediction model," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2016, pp. 1221–1228.

[136] P. Holme and J. Saramäki, "Temporal networks," *Phys. Rep.*, vol. 519, no. 3, pp. 97–125, 2012.

[137] M. Barthélemy, "Spatial networks," *Phys. Rep.*, vol. 499, no. 1, pp. 1–101, 2011.

[138] P. Crucitti, V. Latora, and S. Porta, "Centrality measures in spatial networks of urban streets," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 73, no. 3, Mar. 2006, Art. no. 036125.

[139] J. Buhl, J. Gautrais, N. Reeves, R. V. Solé, S. Valverde, P. Kuntz, and G. Theraulaz, "Topological patterns in street networks of self-organized urban settlements," *Eur. Phys. J. B Condens. Matter Complex Syst.*, vol. 49, no. 4, pp. 513–522, Feb. 2006.

[140] A. Cardillo, S. Scellato, V. Latora, and S. Porta, "Structural properties of planar graphs of urban street patterns," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 73, no. 6, Jun. 2006, Art. no. 066107.

[141] F. Goetzke, R. Gerike, A. Páez, and E. Dugundji, "Social interactions in transportation: Analyzing groups and spatial networks," *Transportation*, vol. 42, no. 5, pp. 723–731, Sep. 2015.

[142] G. Topa, "Social and spatial networks in labour markets," *Oxford Rev. Econ. Policy*, vol. 35, no. 4, pp. 722–745, Dec. 2019.

[143] M. Stephens and A. Poorthuis, "Follow thy neighbor: Connecting the social and the spatial networks on Twitter," *Comput., Environ. Urban Syst.*, vol. 53, pp. 87–95, Sep. 2015.

[144] Y. Ren, M. Ercsey-Ravasz, P. Wang, M. C. González, and Z. Toroczkai, "Predicting commuter flows in spatial networks using a radiation model based on temporal ranges," *Nature Commun.*, vol. 5, no. 1, pp. 1–9, Dec. 2014.

[145] S. Shang, J. Liu, K. Zheng, H. Lu, T. B. Pedersen, and J.-R. Wen, "Planning unobstructed paths in traffic-aware spatial networks," *GeoInformatica*, vol. 19, no. 4, pp. 723–746, 2015.

[146] R. Solomonoff and A. Rapoport, "Connectivity of random nets," *Bull. Math. Biophys.*, vol. 13, no. 2, pp. 107–117, Jun. 1951.

[147] E. N. Gilbert, "Random graphs," *Ann. Math. Statist.*, vol. 30, no. 4, pp. 1141–1144, 1959.

[148] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, no. 1, pp. 415–444, 2001.

[149] A. Asikainen, G. Iñiguez, J. Ureña-Carrión, K. Kaski, and M. Kivelä, "Cumulative effects of triadic closure and homophily in social networks," *Sci. Adv.*, vol. 6, no. 19, May 2020, Art. no. eaax7310.

[150] D. Verhoeven, K. Musial, S. Palmer, S. Taylor, S. Abidi, V. Zemaityte, and L. Simpson, "Controlling for openness in the male-dominated collaborative networks of the global film industry," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0234460.

[151] L. Shi, L. Zhang, and Y. Lu, "Evaluating social network-based weight loss interventions in Chinese population: An agent-based simulation," *PLoS ONE*, vol. 15, no. 8, Aug. 2020, Art. no. e0236716.

[152] P. Block, M. Hoffman, I. J. Raabe, J. B. Dowd, C. Rahal, R. Kashyap, and M. C. Mills, "Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world," *Nature Hum. Behav.*, vol. 4, no. 6, pp. 588–596, Jun. 2020.

[153] K. Zhang, Y. Huang, Y. Du, and L. Wang, "Facial expression recognition based on deep evolutional spatial-temporal networks," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4193–4203, Sep. 2017.

[154] S. Xu, Y. Cheng, K. Gu, Y. Yang, S. Chang, and P. Zhou, "Jointly attentive spatial-temporal pooling networks for video-based person re-identification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4733–4742.

[155] D. Cai, Z. Shao, X. He, X. Yan, and J. Han, "Mining hidden community in heterogeneous social networks," in *Proc. 3rd Int. Workshop Link Discovery*, 2005, pp. 58–65.

[156] W. Wang, P. Jiao, D. He, D. Jin, L. Pan, and B. Gabrys, "Autonomous overlapping community detection in temporal networks: A dynamic Bayesian nonnegative matrix factorization approach," *Knowl.-Based Syst.*, vol. 110, pp. 121–134, Oct. 2016.

[157] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.

[158] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, Jun. 1998.

[159] J. Zhang, J. Wen, and J. Chen, "Modeling market fluctuations under investor sentiment with a Hawkes-contact process," *Eur. J. Finance*, pp. 1–16, Jul. 2021, doi: 10.1080/1351847X.2021.1957699.

[160] J.-S. Kim, H. Kavak, C. O. Rouly, H. Jin, A. Crooks, D. Pfoser, C. Wenk, and A. Züfle, "Location-based social simulation for prescriptive analytics of disease spread," *SIGSPATIAL Special*, vol. 12, no. 1, pp. 53–61, Jun. 2020.

[161] R. W. Krause, M. Huisman, C. Steglich, and T. Snijders, "Missing data in cross-sectional networks—An extensive comparison of missing data treatment methods," *Social Netw.*, vol. 62, pp. 99–112, Jul. 2020.

[162] S. Fortunato, A. Flammini, and F. Menczer, "Scale-free network growth by ranking," *Phys. Rev. Lett.*, vol. 96, no. 21, May 2006, Art. no. 218701.

[163] A. D. Broido and A. Clauset, "Scale-free networks are rare," *Nature Commun.*, vol. 10, no. 1, pp. 1–10, Dec. 2019.

[164] D. R. Hunter, S. M. Goodreau, and M. S. Handcock, "Goodness of fit of social network models," *J. Amer. Stat. Assoc.*, vol. 103, no. 481, pp. 248–258, Mar. 2008.

[165] K. Musial, M. Budka, and K. Juszczyszyn, "Creation and growth of online social network," *World Wide Web*, vol. 16, no. 4, pp. 421–447, Jul. 2013.

[166] A. Godoy-Lorite and N. S. Jones, "Inference and influence of network structure using snapshot social behavior without network data," *Sci. Adv.*, vol. 7, no. 23, Jun. 2021, Art. no. eabb8762.

[167] J. Zhang, J. Wang, and J. Shao, "Finite-range contact process on the market return intervals distributions," *Adv. Complex Syst.*, vol. 13, no. 5, pp. 643–657, Oct. 2010.

[168] J. Zhang and J. Wang, "Modeling and simulation of the market fluctuations by the finite range contact systems," *Simul. Model. Pract. Theory*, vol. 18, no. 6, pp. 910–925, Jun. 2010.

[169] V. Arora and M. Ventresca, "Action-based modeling of complex networks," *Sci. Rep.*, vol. 7, no. 1, pp. 1–10, Dec. 2017.

[170] B. K. Choi and D. Kang, *Modeling and Simulation of Discrete Event Systems*. Hoboken, NJ, USA: Wiley, 2013.

[171] A. Q. Macedo, L. B. Marinho, and R. L. T. Santos, "Context-aware event recommendation in event-based social networks," in *Proc. 9th ACM Conf. Recommender Syst.*, Sep. 2015, pp. 123–130.

[172] T.-A.-N. Pham, X. Li, G. Cong, and Z. Zhang, "A general graph-based model for recommendation in event-based social networks," in *Proc. IEEE 31st Int. Conf. Data Eng.*, Apr. 2015, pp. 567–578.

[173] X. Liu, Q. He, Y. Tian, W.-C. Lee, J. McPherson, and J. Han, "Event-based social networks: Linking the online and offline social worlds," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2012, pp. 1032–1040.

[174] E. W. Fox, M. B. Short, F. P. Schoenberg, K. D. Coronges, and A. L. Bertozzi, "Modeling e-mail networks and inferring leadership using self-exciting point processes," *J. Amer. Stat. Assoc.*, vol. 111, no. 514, pp. 564–584, 2016.

[175] F. S. Passino and N. A. Heard, "Mutually exciting point process graphs for modelling dynamic networks," 2021, *arXiv:2102.06527*.

[176] J. M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling*, vol. 13. Princeton, NJ, USA: Princeton Univ. Press, 2006.

[177] L. Bargigli and G. Tedeschi, "Interaction in agent-based economics: A survey on the network approach," *Phys. A, Stat. Mech. Appl.*, vol. 399, pp. 1–15, Apr. 2014.

[178] W. Wang, J. Zhang, S. Zhao, and Y. Zhang, "Simulation of asset pricing in information networks," *Phys. A, Stat. Mech. Appl.*, vol. 513, pp. 620–634, Jan. 2019.

[179] A. L. Andrei, K. Comer, and M. Koehler, "An agent-based model of network effects on tax compliance and evasion," *J. Econ. Psychol.*, vol. 40, pp. 119–133, Feb. 2014.

[180] S. Venkatramanan, B. Lewis, J. Chen, D. Higdon, A. Vullikanti, and M. Marathe, "Using data-driven agent-based models for forecasting emerging infectious diseases," *Epidemics*, vol. 22, pp. 43–49, Mar. 2018.

[181] Q. Long, "Data-driven decision making for supply chain networks with agent-based computational experiment," *Knowl.-Based Syst.*, vol. 141, pp. 55–66, Feb. 2018.

[182] W. Wang, S. Zhao, and J. Zhang, "Multi-asset pricing modeling using holding-based networks in energy markets," *Finance Res. Lett.*, vol. 46, May 2022, Art. no. 102483.

[183] J. Zhang, "Influence of individual rationality on continuous double auction markets with networked traders," *Phys. A, Stat. Mech. Appl.*, vol. 495, pp. 353–392, Apr. 2018.

[184] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.

[185] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, "Activity driven modeling of time varying networks," *Sci. Rep.*, vol. 2, no. 1, pp. 1–7, Dec. 2012.

[186] G. Petri and A. Barrat, "Simplicial activity driven model," *Phys. Rev. Lett.*, vol. 121, no. 22, Nov. 2018, Art. no. 228301.

[187] X. Meng and T. Chen, "Event based agreement protocols for multi-agent networks," *Automatica*, vol. 49, no. 7, pp. 2125–2132, 2013.

[188] M. Ogura, V. M. Preciado, and N. Masuda, "Optimal containment of epidemics over temporal activity-driven networks," *SIAM J. Appl. Math.*, vol. 79, no. 3, pp. 986–1006, 2019.

[189] G. Iñiguez, F. Battiston, and M. Karsai, "Bridging the gap between graphs and networks," *Commun. Phys.*, vol. 3, no. 1, pp. 1–5, Dec. 2020.

[190] T. Aittokallio, "Graph-based methods for analysing networks in cell biology," *Briefings Bioinf.*, vol. 7, no. 3, pp. 243–255, May 2006.

[191] G. George and S. M. Thampi, "A graph-based security framework for securing industrial IoT networks from vulnerability exploitations," *IEEE Access*, vol. 6, pp. 43586–43601, 2018.

[192] L. Euler, "Solutio problematis ad geometriam situs pertinentis," in *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, vol. 8, 1736, pp. 128–140.

[193] G. Alexanderson, "About the cover: Euler and Königsberg's bridges: A historical view," *Bull. Amer. Math. Soc.*, vol. 43, no. 4, pp. 567–573, 2006.

[194] J. A. Bondy and U. S. R. Murty, *Graph Theory With Applications*, vol. 290. London, U.K.: Macmillan, 1976.

[195] H. Djidjev, G. Sandine, C. Storlie, and S. V. Wiel, "Graph based statistical analysis of network traffic," in *Proc. 9th Workshop Mining Learn. With Graphs*, 2011, pp. 1–8.

[196] T. Washio and H. Motoda, "State of the art of graph-based data mining," *ACM SIGKDD Explor. Newslett.*, vol. 5, no. 1, pp. 59–68, Jul. 2003.

[197] N. Hassanlou, M. Shoaran, and A. Thomo, "Probabilistic graph summarization," in *Proc. Int. Conf. Web-Age Inf. Manage.*, Cham, Switzerland: Springer, 2013, pp. 545–556.

[198] J. A. Gámez and A. Salmerón, "Probabilistic graphical models," *Int. J. Intell. Syst.*, vol. 18, no. 2, pp. 149–151, Feb. 2003.

[199] Y. Anzai, *Pattern Recognition and Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2012.

[200] J. Lafferty, A. McCallum, and F. C. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proc. 18th Int. Conf. Mach. Learn. (ICML)*, 2001, pp. 282–289.

[201] B. Taskar, P. Abbeel, M.-F. Wong, and D. Koller, "Relational Markov networks," in *Introduction to Statistical Relational Learning*. Cambridge, MA, USA: MIT Press, 2007, pp. 175–200.

[202] D. Jin, B. Zhang, Y. Song, D. He, Z. Feng, S. Chen, W. Li, and K. Musial, "ModMRF: A modularity-based Markov random field method for community detection," *Neurocomputing*, vol. 405, pp. 218–228, Sep. 2020.

[203] I. Sutskever and G. E. Hinton, "Deep, narrow sigmoid belief networks are universal approximators," *Neural Comput.*, vol. 20, no. 11, pp. 2629–2636, Nov. 2008.

[204] S. H. Chen and C. A. Pollino, "Good practice in Bayesian network modelling," *Environ. Model. Softw.*, vol. 37, pp. 134–145, Nov. 2012.

[205] M. J. Beal, Z. Ghahramani, and C. E. Rasmussen, "The infinite hidden Markov model," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 1, 2002, pp. 577–584.

[206] N. Potyka, "Abstract argumentation with Markov networks," in *Proc. ECAI*. Amsterdam, The Netherlands: IOS Press, 2020, pp. 865–872.

[207] D. Heckerman, "A tutorial on learning with Bayesian networks," in *Innovations in Bayesian Networks*. Cham, Switzerland: Springer, 2008, pp. 33–82.

[208] C. Li, S. Mahadevan, Y. Ling, S. Choze, and L. Wang, "Dynamic Bayesian network for aircraft wing health monitoring digital twin," *AIAA J.*, vol. 55, no. 3, pp. 930–941, 2017.

[209] J. Yu, Y. Song, D. Tang, and J. Dai, "A digital twin approach based on nonparametric Bayesian network for complex system health monitoring," *J. Manuf. Syst.*, vol. 58, pp. 293–304, Jan. 2021.

[210] P. Orbanz and D. M. Roy, "Bayesian models of graphs, arrays and other exchangeable random structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 2, pp. 437–461, Feb. 2015.

[211] H. Boudali and J. Bechta Dugan, "A continuous-time Bayesian network reliability modeling, and analysis framework," *IEEE Trans. Rel.*, vol. 55, no. 1, pp. 86–97, Mar. 2006.

[212] D. F. Anderson and T. G. Kurtz, "Continuous time Markov chain models for chemical reaction networks," in *Design and Analysis of Biomolecular Circuits*. Cham, Switzerland: Springer, 2011, pp. 3–42.

[213] B. J. Frey and N. Jojic, "A comparison of algorithms for inference and learning in probabilistic graphical models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1392–1416, Sep. 2005.

[214] M. Plummer, "JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling," in *Proc. 3rd Int. Workshop Distrib. Stat. Comput.*, Vienna, Austria, vol. 124, 2003, pp. 1–10.

[215] E. B. Sudderth, A. T. Ihler, M. Isard, W. T. Freeman, and A. S. Willsky, "Nonparametric belief propagation," *Commun. ACM*, vol. 53, no. 10, pp. 95–103, 2010.

[216] J. V. Graça, K. Ganchev, and B. Taskar, "Expectation maximization and posterior constraints," in *Advances in Neural Information Processing System*, vol. 20. San Francisco, CA, USA: Curran Associates, 2007.

[217] E. Kropat, A. Özmen, G.-W. Weber, S. Meyer-Nieberg, and O. Defterli, "Fuzzy prediction strategies for gene-environment networks—Fuzzy regression analysis for two-modal regulatory systems," *RAIRO Oper. Res.*, vol. 50, no. 2, pp. 413–435, Apr. 2016.

[218] E. Kropat, G.-W. Weber, and S. Belen, "Dynamical gene-environment networks under ellipsoidal uncertainty: Set-theoretic regression analysis based on ellipsoidal OR," in *Dynamics, Games and Science I*. Cham, Switzerland: Springer, 2011, pp. 545–571.

[219] P. Cui, X. Wang, J. Pei, and W. Zhu, "A survey on network embedding," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 5, pp. 833–852, May 2018.

[220] M. Budka, K. Juszczyszyn, K. Musial, and A. Musial, "Molecular model of dynamic social network based on e-mail communication," *Social Netw. Anal. Mining*, vol. 3, no. 3, pp. 543–563, Sep. 2013.

[221] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: Methods and applications," 2017, *arXiv:1709.05584*.

[222] K. Rajawat and S. Kumar, "Stochastic multidimensional scaling," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 3, no. 2, pp. 360–375, Jun. 2017.

[223] B. Li, Y. He, F. Guo, and L. Zuo, "A novel localization algorithm based on Isomap and partial least squares for wireless sensor networks," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 2, pp. 304–314, Feb. 2013.

[224] J. Bollen, H. Van de Sompel, A. Hagberg, and R. Chute, "A principal component analysis of 39 scientific impact measures," *PLoS ONE*, vol. 4, no. 6, Jun. 2009, Art. no. e6022.

[225] P. P. Paul, M. L. Gavrilova, and R. Alhajj, "Decision fusion for multimodal biometrics using social network analysis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 11, pp. 1522–1533, Nov. 2014.

[226] Z. Yang, J. Peltonen, and S. Kaski, "Optimization equivalence of divergences improves neighbor embedding," in *Proc. Int. Conf. Mach. Learn.*, 2014, pp. 460–468.

[227] F. Olivon, N. Elie, G. Grelier, F. Roussi, M. Litaudon, and D. Touboul, "MetGem software for the generation of molecular networks based on the t-SNE algorithm," *Anal. Chem.*, vol. 90, no. 23, pp. 13900–13908, Dec. 2018.

[228] Y. Wang, Y. Yao, H. Tong, F. Xu, and J. Lu, "A brief review of network embedding," *Big Data Mining Anal.*, vol. 2, no. 1, pp. 35–47, Mar. 2019.

[229] D. Guthrie, B. Allison, W. Liu, L. Guthrie, and Y. Wilks, "A closer look at skip-gram modelling," in *Proc. LREC*, vol. 6. Princeton, Nj, USA: Citeseer, 2006, pp. 1222–1225.

[230] A. P. Singh and G. J. Gordon, "A unified view of matrix factorization models," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2008, pp. 358–373.

[231] D. Zhang, J. Yin, X. Zhu, and C. Zhang, "Network representation learning: A survey," *IEEE Trans. Big Data*, vol. 6, no. 1, pp. 3–28, Mar. 2020.

[232] H. Cai, V. W. Zheng, and K. C. Chang, "A comprehensive survey of graph embedding: Problems, techniques, and applications," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 9, pp. 1616–1637, Feb. 2018.

[233] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 701–710.

[234] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 855–864.

[235] P. Yanardag and S. V. N. Vishwanathan, "Deep graph kernels," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2015, pp. 1365–1374.

[236] Y. Gong, Y. Niu, W. Zhang, and X. Li, "A network embedding-based multiple information integration method for the MiRNA-disease association prediction," *BMC Bioinf.*, vol. 20, no. 1, pp. 1–13, Dec. 2019.

[237] S. Cao, W. Lu, and Q. Xu, "Deep neural networks for learning graph representations," in *Proc. AAAI Conf. Artif. Intell.*, vol. 30, no. 1, 2016, pp. 1145–1152.

[238] H. Chen, H. Yin, X. Sun, T. Chen, B. Gabrys, and K. Musial, "Multi-level graph convolutional networks for cross-platform anchor link prediction," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 1503–1511.

[239] J. Sun, W. Guo, D. Zhang, Y. Zhang, F. Regol, Y. Hu, H. Guo, R. Tang, H. Yuan, X. He, and M. Coates, "A framework for recommending accurate and diverse items using Bayesian graph convolutional neural networks," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 2030–2039.

[240] M. Qu, Y. Bengio, and J. Tang, "GMNN: Graph Markov neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 5241–5250.

[241] D. Bacciu, F. Errica, and A. Micheli, "Contextual graph Markov model: A deep and generative approach to graph processing," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 294–303.

[242] Z. Huang, Y. Sun, and W. Wang, "Coupled graph ODE for learning interacting system dynamics," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*, Aug. 2021, pp. 1–11.

[243] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," 2017, *arXiv:1702.08608*.

[244] K. Musial, B. Gabrys, and M. Buczko, "What kind of network are you?: Using local and global characteristics in network categorisation tasks," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2013, pp. 1366–1373.

[245] J. Zhang, Y. Xu, and D. Houser, "Vulnerability of scale-free cryptocurrency networks to double-spending attacks," *Eur. J. Finance*, vol. 27, no. 12, pp. 1235–1249, 2021.

[246] R. Eletreby, Y. Zhuang, K. M. Carley, O. Yağan, and H. V. Poor, "The effects of evolutionary adaptations on spreading processes in complex networks," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 11, pp. 5664–5670, Mar. 2020.

[247] M. Ait Rami, V. S. Bokharaie, O. Mason, and F. R. Wirth, "Stability criteria for SIS epidemiological models under switching policies," 2013, *arXiv:1306.0135*.

[248] M. R. Sanatkar, W. N. White, B. Natarajan, C. M. Scoglio, and K. A. Garrett, "Epidemic threshold of an SIS model in dynamic switching networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 46, no. 3, pp. 345–355, Mar. 2016.

[249] H. Zhao and J. H. Park, "Dynamic output feedback consensus of continuous-time networked multiagent systems," *Complexity*, vol. 20, no. 5, pp. 35–42, May 2015.

[250] Y. Gu, Y. Sun, and J. Gao, "The co-evolution model for social network evolving and opinion migration," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2017, pp. 175–184.

[251] Y. Wang, D. Jin, K. Musial, and J. Dang, "Community detection in social networks considering topic correlations," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, 2019, pp. 321–328.

[252] M. Al Hasan and M. J. Zaki, "A survey of link prediction in social networks," in *Social Network Data Analytics*. Cham, Switzerland: Springer, 2011, pp. 243–275.

[253] Z. Bu, Y. Y. Wang, H.-J. Li, J. Jiang, Z. Wu, and J. Cao, "Link prediction in temporal networks: Integrating survival analysis and game theory," *Inform. Sci.*, vol. 498, pp. 41–61, Sep. 2019.

[254] A. L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek, "Evolution of the social network of scientific collaborations," *Phys. A, Stat. Mech. Appl.*, vol. 311, pp. 590–614, Aug. 2001.

[255] J. Basilico and T. Hofmann, "Unifying collaborative and content-based filtering," in *Proc. 21st Int. Conf. Mach. Learn. (ICML)*, 2004, p. 9.

[256] M. A. Hasan, V. Chaoj, S. Salem, and M. Zaki, "Link prediction using supervised learning," in *Proc. SDM Workshop Link Anal. Counterterrorism Secur.*, vol. 30, no. 9, 2006, pp. 798–805.

[257] G. Jeh and J. Widom, "SimRank: A measure of structural-context similarity," in *Proc. 8th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2002, pp. 538–543.

[258] R. Hong, R. He, L. Wu, Y. Ge, and X. Wu, "Deep attributed network embedding by preserving structure and attribute information," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 3, pp. 1434–1445, Mar. 2021.

[259] B. Taskar, M.-F. Wong, P. Abbeel, and D. Koller, "Link prediction in relational data," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, 2003, pp. 659–666.

[260] C. Dai, L. Chen, B. Li, and Y. Li, "Link prediction in multi-relational networks based on relational similarity," *Inf. Sci.*, vols. 394–395, pp. 198–216, Jul. 2017.

[261] H. Gao and H. Huang, "Deep attributed network embedding," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 3364–3370.

[262] X. Liu, B. Yang, W. Song, K. Musial, W. Zuo, H. Chen, and H. Yin, "A block-based generative model for attributed network embedding," *World Wide Web*, vol. 24, no. 5, pp. 1439–1464, Sep. 2021.

[263] B. Taskar, E. Segal, and D. Koller, "Probabilistic classification and clustering in relational data," in *Proc. Int. Joint Conf. Artif. Intell.*, vol. 17, no. 1, 2001, pp. 870–878.

[264] J. Kleinberg and E. Tardos, "Approximation algorithms for classification problems with pairwise relationships: Metric labeling and Markov random fields," *J. ACM*, vol. 49, no. 5, pp. 616–639, Sep. 2002.

[265] F. McSherry, "Spectral partitioning of random graphs," in *Proc. 42nd IEEE Symp. Found. Comput. Sci.*, Oct. 2001, pp. 529–537.

[266] S. Bhagat, G. Cormode, and S. Muthukrishnan, "Node classification in social networks," in *Social Network Data Analytics*. Boston, MA, USA: Springer, 2011, pp. 115–148.

[267] Y. Zhou, H. Cheng, and J. X. Yu, "Graph clustering based on structural/attribute similarities," *Proc. VLDB Endowment*, vol. 2, no. 1, pp. 718–729, Aug. 2009.

[268] X. Zhu, Z. Ghahramani, and J. D. Lafferty, "Semi-supervised learning using Gaussian fields and harmonic functions," in *Proc. 20th Int. Conf. Mach. Learn. (ICML)*, 2003, pp. 912–919.

[269] S. Abu-El-Haija, A. Kapoor, B. Perozzi, and J. Lee, "N-GCN: Multi-scale graph convolution for semi-supervised node classification," in *Proc. Uncertainty Artif. Intell.*, 2020, pp. 841–851.

[270] C. Wang, V. Satuluri, and S. Parthasarathy, "Local probabilistic models for link prediction," in *Proc. 7th IEEE Int. Conf. Data Mining (ICDM)*, Oct. 2007, pp. 322–331.

[271] I. Güneş, S. Gündüz-Ögüdücü, and Z. Çataltepe, "Link prediction using time series of neighborhood-based node similarity scores," *Data Mining Knowl. Discovery*, vol. 30, no. 1, pp. 147–180, Jan. 2016.

[272] X. Li, N. Du, H. Li, K. Li, J. Gao, and A. Zhang, "A deep learning approach to link prediction in dynamic networks," in *Proc. SIAM Int. Conf. Data Mining*, Apr. 2014, pp. 289–297.

[273] J. Chen, X. Wang, and X. Xu, "GC-LSTM: Graph convolution embedded LSTM for dynamic network link prediction," *Appl. Intell.*, vol. 52, pp. 7513–7528, Sep. 2021.

[274] X. Ma, P. Sun, and Y. Wang, "Graph regularized nonnegative matrix factorization for temporal link prediction in dynamic networks," *Phys. A, Stat. Mech. Appl.*, vol. 496, pp. 121–136, Apr. 2018.

[275] P. Sarkar, D. Chakrabarti, and M. Jordan, "Nonparametric link prediction in large scale dynamic networks," *Electron. J. Statist.*, vol. 8, no. 2, pp. 2022–2065, Jan. 2014.

[276] H. Chen and J. Li, "Exploiting structural and temporal evolution in dynamic link prediction," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2018, pp. 427–436.

[277] F. D. Malliaros and M. Vazirgiannis, "Clustering and community detection in directed networks: A survey," *Phys. Rep.*, vol. 533, no. 4, pp. 95–142, 2013.

[278] M. Coscia, "Discovering communities of community discovery," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2019, pp. 1–8.

[279] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q. Z. Sheng, H. Xiong, and L. Akoglu, "A comprehensive survey on graph anomaly detection with deep learning," *IEEE Trans. Knowl. Data Eng.*, early access, Oct. 8, 2021, doi: 10.1109/TKDE.2021.3118815.

[280] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: A survey," *Data Mining Knowl. Discovery*, vol. 29, no. 3, pp. 626–688, 2015.

[281] S. Funk, M. Salathé, and V. A. A. Jansen, "Modelling the influence of human behaviour on the spread of infectious diseases: A review," *J. Roy. Soc. Interface*, vol. 7, no. 50, pp. 1247–1256, May 2010.

[282] H. Sun, D. Saad, and A. Y. Lokhov, "Competition, collaboration, and optimization in multiple interacting spreading processes," *Phys. Rev. X*, vol. 11, no. 1, Mar. 2021, Art. no. 011048.

[283] I. Ameen, D. Baleanu, and H. M. Ali, "An efficient algorithm for solving the fractional optimal control of SIRV epidemic model with a combination of vaccination and treatment," *Chaos, Solitons Fractals*, vol. 137, Aug. 2020, Art. no. 109892.

[284] F. Darabi Sahneh and C. Scoglio, "Competitive epidemic spreading over arbitrary multilayer networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdisc. Top.*, vol. 89, no. 6, Jun. 2014, Art. no. 062817.

[285] I. Cooper, A. Mondal, and C. G. Antonopoulos, "A SIR model assumption for the spread of COVID-19 in different communities," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110057.

[286] S. Syafruddin and M. Noorani, "A sir model for spread of dengue fever disease (simulation for south Sulawesi, Indonesia and Selangor, Malaysia)," *World J. Model. Simul.*, vol. 9, no. 2, pp. 96–105, 2013.

[287] P. P. Klamser, M. Wiedermann, J. F. Donges, and R. V. Donner, "Zealotry effects on opinion dynamics in the adaptive voter model," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdisc. Top.*, vol. 96, no. 5, Nov. 2017, Art. no. 052315.

[288] T. Xu, H. Zhu, H. Zhong, G. Liu, H. Xiong, and E. Chen, "Exploiting the dynamic mutual influence for predicting social event participation," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 6, pp. 1122–1135, Jun. 2019.

[289] C. Laur, J. Bell, R. Valaitis, S. Ray, and H. Keller, "The sustain and spread framework: Strategies for sustaining and spreading nutrition care improvements in acute care based on thematic analysis from the more-2-eat study," *BMC Health Services Res.*, vol. 18, no. 1, pp. 1–11, Dec. 2018.

[290] A. E. Kendig, E. T. Borer, E. N. Boak, T. C. Picard, and E. W. Seabloom, "Host nutrition mediates interactions between plant viruses, altering transmission and predicted disease spread," *Ecology*, vol. 101, no. 11, Nov. 2020, Art. no. e03155.

[291] P. D. Karampourniotis, S. Sreenivasan, B. K. Szymanski, and G. Korniss, "The impact of heterogeneous thresholds on social contagion with multiple initiators," *PLoS ONE*, vol. 10, no. 11, Nov. 2015, Art. no. e0143020.

[292] L. Zino and M. Cao, "Analysis, prediction, and control of epidemics: A survey from scalar to dynamic network models," 2021, *arXiv:2103.00181*.

[293] I. Z. Kiss, G. Röst, and Z. Vizi, "Generalization of pairwise models to non-Markovian epidemics on networks," *Phys. Rev. Lett.*, vol. 115, no. 7, Aug. 2015, Art. no. 078701.

[294] Q. Liu and P. Van Mieghem, "Burst of virus infection and a possibly largest epidemic threshold of non-Markovian susceptible-infected-susceptible processes on networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdisc. Top.*, vol. 97, no. 2, Feb. 2018, Art. no. 022309.

[295] A. Melo, C. L. Beck, J. I. Pena, and P. E. Pare, "Knowledge transfer from universities to regions as a network spreading process," in *Proc. IEEE Int. Syst. Eng. Symp. (ISSE)*, Oct. 2018, pp. 1–8.

[296] P. Van Mieghem, "The N-intertwined SIS epidemic network model," *Computing*, vol. 93, nos. 2–4, pp. 147–169, Dec. 2011.

[297] H. J. Ahn and B. Hassibi, "Global dynamics of epidemic spread over complex networks," in *Proc. 52nd IEEE Conf. Decis. Control*, Dec. 2013, pp. 4579–4585.

[298] I. Iacopini, G. Petri, A. Barrat, and V. Latora, "Simplicial models of social contagion," *Nature Commun.*, vol. 10, no. 1, pp. 1–9, Dec. 2019.

[299] T. Gross and B. Blasius, "Adaptive coevolutionary networks: A review," *J. Roy. Soc. Interface*, vol. 5, no. 20, pp. 259–271, Mar. 2008.

[300] J. Leitch, K. A. Alexander, and S. Sengupta, "Toward epidemic thresholds on temporal networks: A review and open questions," *Appl. Netw. Sci.*, vol. 4, no. 1, pp. 1–21, Dec. 2019.

[301] I. Z. Kiss, L. Berthouze, T. J. Taylor, and P. L. Simon, "Modelling approaches for simple dynamic networks and applications to disease transmission models," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 468, no. 2141, pp. 1332–1355, May 2012.

[302] M. Taylor, T. J. Taylor, and I. Z. Kiss, "Epidemic threshold and control in a dynamic network," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdisc. Top.*, vol. 85, no. 1, Jan. 2012, Art. no. 016103.

[303] F. Baldo, L. Dall'Olio, M. Ceccarelli, R. Scheda, M. Lombardi, A. Borghesi, S. Diciotti, and M. Milano, "Deep learning for virus-spreading forecasting: A brief survey," 2021, *arXiv:2103.02346*.

[304] S. Mežnar, N. Lavrač, and B. Škrlj, "Prediction of the effects of epidemic spreading with graph neural networks," in *Proc. Int. Conf. Complex Netw. Their Appl.*, Cham, Switzerland: Springer, 2020, pp. 420–431.

[305] S. Mežnar, N. Lavrač, and B. Škrlj, "Transfer learning for node regression applied to spreading prediction," 2021, *arXiv:2104.00088*.

[306] B. Karrer and M. E. J. Newman, "Competing epidemics on complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdisc. Top.*, vol. 84, no. 3, Sep. 2011, Art. no. 036106.

[307] M. E. J. Newman, "Threshold effects for two pathogens spreading on a network," *Phys. Rev. Lett.*, vol. 95, no. 10, Sep. 2005, Art. no. 108701.

[308] C. Budak, D. Agrawal, and A. El Abbadi, "Limiting the spread of misinformation in social networks," in *Proc. 20th Int. Conf. World Wide Web (WWW)*, 2011, pp. 665–674.

[309] B. A. Prakash, A. Beutel, R. Rosenfeld, and C. Faloutsos, "Winner takes all: Competing viruses or ideas on fair-play networks," in *Proc. 21st Int. Conf. World Wide Web*, Apr. 2012, pp. 1037–1046.

[310] B. Min and M. S. Miguel, "Competing contagion processes: Complex contagion triggered by simple contagion," *Sci. Rep.*, vol. 8, no. 1, pp. 1–8, Dec. 2018.

[311] G. Fu, F. Chen, J. Liu, and J. Han, "Analysis of competitive information diffusion in a group-based population over social networks," *Phys. A, Stat. Mech. Appl.*, vol. 525, pp. 409–419, Jul. 2019.

[312] I. Koprulu, Y. Kim, and N. B. Shroff, "Battle of opinions over evolving social networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 2, pp. 532–545, Apr. 2019.

[313] C. Pan, L.-X. Yang, X. Yang, Y. Wu, and Y. Y. Tang, "An effective rumor-containing strategy," *Phys. A, Stat. Mech. Appl.*, vol. 500, pp. 80–91, Jun. 2018.

[314] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen, "The spread of awareness and its impact on epidemic outbreaks," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 16, pp. 6872–6877, 2009.

[315] S. Pei and H. A. Makse, "Spreading dynamics in complex networks," *J. Stat. Mech., Theory Exp.*, vol. 2013, no. 12, Dec. 2013, Art. no. P12002.

[316] X.-X. Zhan, C. Liu, G. Zhou, Z.-K. Zhang, G.-Q. Sun, J. J. Zhu, and Z. Jin, "Coupling dynamics of epidemic spreading and information diffusion on complex networks," *Appl. Math. Comput.*, vol. 332, no. 9, pp. 437–448, 2018.

[317] G. Baggio, D. S. Bassett, and F. Pasqualetti, "Data-driven control of complex networks," *Nature Commun.*, vol. 12, no. 1, pp. 1–13, Dec. 2021.

[318] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, pp. 167–173, May 2011.

[319] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang, and Y.-C. Lai, "Exact controllability of complex networks," *Nature Commun.*, vol. 4, no. 1, pp. 1–9, Dec. 2013.

[320] Y. Lou, L. Wang, K.-F. Tsang, and G. Chen, "Towards optimal robustness of network controllability: An empirical necessary condition," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 9, pp. 3163–3174, Sep. 2020.

[321] Y. Lou, Y. He, L. Wang, and G. Chen, "Predicting network controllability robustness: A convolutional neural network approach," *IEEE Trans. Cybern.*, vol. 52, no. 5, pp. 4052–4063, May 2022.

[322] B. Kiumarsi, K. G. Vamvoudakis, H. Modares, and F. L. Lewis, "Optimal and autonomous control using reinforcement learning: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2042–2062, Jun. 2018.

[323] F. Dörfler and F. Bullo, "Synchronization in complex networks of phase oscillators: A survey," *Automatica*, vol. 50, no. 6, pp. 1539–1564, Jun. 2014.

[324] Y. Tang, F. Qian, H. J. Gao, and J. Kurths, "Synchronization in complex networks and its application—A survey of recent advances and challenges," *Annu. Rev. Control*, vol. 38, no. 2, pp. 184–198, 2014.

[325] D. Fiore, G. Russo, and M. di Bernardo, "Exploiting nodes symmetries to control synchronization and consensus patterns in multiagent systems," *IEEE Control Syst. Lett.*, vol. 1, no. 2, pp. 364–369, Oct. 2017.

[326] J. Chen, B. Chen, and Z. Zeng, "Synchronization and consensus in networks of linear fractional-order multi-agent systems via sampled-data control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 8, pp. 2955–2964, Aug. 2020.

[327] W. Lu and T. Chen, "QUAD-condition, synchronization, consensus of multiagents, and anti-synchronization of complex networks," *IEEE Trans. Cybern.*, vol. 51, no. 6, pp. 3384–3388, Jun. 2021.

[328] E. Meirom, H. Maron, S. Mannor, and G. Chechik, "Controlling graph dynamics with reinforcement learning and graph neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 7565–7577.

[329] D. Kiritsis, "Closed-loop PLM for intelligent products in the era of the Internet of Things," *Comput.-Aided Des.*, vol. 43, pp. 479–501, May 2011.

[330] M. G. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, "A probabilistic graphical model foundation for enabling predictive digital twins at scale," *Nature Comput. Sci.*, vol. 1, no. 5, pp. 337–347, May 2021.

[331] E. Kharlamov, F. Martin-Recuerda, B. Perry, D. Cameron, R. Fjellheim, and A. Waaler, "Towards semantically enhanced digital twins," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 4189–4193.

[332] A. Banerjee, R. Dalal, S. Mittal, and K. P. Joshi, "Generating digital twin models using knowledge graphs for industrial production lines," in *Proc. Workshop Ind. Knowl. Graphs (ACM Web Sci. Conf.)*, 2017, pp. 1–6.

[333] T. Hubauer, S. Lamparter, P. Haase, and D. M. Herzig, "Use cases of the industrial knowledge graph at Siemens," in *Proc. Int. Semantic Web Conf. (P&D/Industry/BlueSky)*, 2018, pp. 1–2.

[334] J. Lu, X. Zheng, A. Gharaei, K. Kalaboukas, and D. Kiritsis, "Cognitive twins for supporting decision-makings of Internet of Things systems," in *Proc. 5th Int. Conf. Ind. 4.0 Model Adv. Manuf.*, Cham, Switzerland: Springer, 2020, pp. 105–115.

[335] C. Fan, C. Zhang, A. Yahja, and A. Mostafavi, "Disaster city digital twin: A vision for integrating artificial and human intelligence for disaster management," *Int. J. Inf. Manage.*, vol. 56, Feb. 2021, Art. no. 102049.

[336] C. Mandolla, A. M. Petruzzelli, G. Percoco, and A. Urbinati, "Building a digital twin for additive manufacturing through the exploitation of blockchain: A case analysis of the aircraft industry," *Comput. Ind.*, vol. 109, pp. 134–152, 2019.

[337] H.-N. Dai, Z. Zheng, and Y. Zhang, "Blockchain for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8076–8094, Oct. 2019.

[338] A. E. Azzaoui, T. W. Kim, V. Loia, and J. H. Park, "Blockchain-based secure digital twin framework for smart healthy city," in *Advanced Multimedia and Ubiquitous Engineering*. Singapore: Springer, 2021, pp. 107–113.

[339] J. Pang, Y. Huang, Z. Xie, J. Li, and Z. Cai, "Collaborative city digital twin for the COVID-19 pandemic: A federated learning solution," *Tsinghua Sci. Technol.*, vol. 26, no. 5, pp. 759–771, Oct. 2021.

[340] C. Chen, J. He, N. Bliss, and H. Tong, "Towards optimal connectivity on multi-layered networks," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 10, pp. 2332–2346, Oct. 2017.

[341] C. Cronrath, A. R. Aderiani, and B. Lennartson, "Enhancing digital twins through reinforcement learning," in *Proc. IEEE 15th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2019, pp. 293–298.

[342] H. Zhang, Q. Liu, X. Chen, D. Zhang, and J. Leng, "A digital twin-based approach for designing and multi-objective optimization of hollow glass production line," *IEEE Access*, vol. 5, pp. 26901–26911, 2017.

[343] S. Rahmanzadeh, M. S. Pishvaee, and K. Govindan, "Emergence of open supply chain management: The role of open innovation in the future smart industry using digital twin network," *Ann. Oper. Res.*, pp. 1–29, Jan. 2022, doi: 10.1007/s10479-021-04254-2.

[344] L. Wright and S. Davidson, "How to tell the difference between a model and a digital twin," *Adv. Model. Simul. Eng. Sci.*, vol. 7, no. 1, pp. 1–13, Dec. 2020.

[345] L. Zhang, L. Zhou, and B. K. P. Horn, "Building a right digital twin with model engineering," *J. Manuf. Syst.*, vol. 59, pp. 151–164, Apr. 2021.

[346] H. B. Sta, "Quality and the efficiency of data in 'smart-cities'," *Future Gener. Comput. Syst.*, vol. 74, pp. 409–416, Sep. 2017.

[347] J. K. Crawford, *Project Management Maturity Model*. Boca Raton, FL, USA: Auerbach Publications, 2006.

[348] P. Zhang, C. Huo, and M. Kezunovic, "A novel measure of omponent importance considering cost for all-digital protection systems," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Jun. 2007, pp. 1–6.

[349] J. Wang, L. Ye, R. X. Gao, C. Li, and L. Zhang, "Digital twin for rotating machinery fault diagnosis in smart manufacturing," *Int. J. Prod. Res.*, vol. 57, no. 12, pp. 3920–3934, 2019.

[350] Y. Dong, J. Zhang, J. Tang, N. V. Chawla, and B. Wang, "CoupledLP: Link prediction in coupled networks," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2015, pp. 199–208.

[351] R. R. Junuthula, K. S. Xu, and V. K. Devabhaktuni, "Evaluating link prediction accuracy in dynamic networks with added and removed edges," in *Proc. IEEE Int. Conferences Big Data Cloud Comput. (BDCloud), Social Comput. Netw. (SocialCom), Sustain. Comput. Commun. (SustainCom) (BDCloud-SocialCom-SustainCom)*, Oct. 2016, pp. 377–384.

[352] R. Patel and Y. Guo, "Graph based link prediction between human phenotypes and genes," 2021, *arXiv:2105.11989*.

[353] G. Cybenko, "Parallel computing for machine learning in social network analysis," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshops (IPDPSW)*, May 2017, pp. 1464–1471.

[354] R. Yang, F. R. Yu, P. Si, Z. Yang, and Y. Zhang, "Integrated blockchain and edge computing systems: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1508–1532, 2nd Quart., 2019.

[355] J. Sun, L. Yin, M. Zou, Y. Zhang, T. Zhang, and J. Zhou, "Makespan-minimization workflow scheduling for complex networks with social groups in edge computing," *J. Syst. Archit.*, vol. 108, Sep. 2020, Art. no. 101799.

[356] M. S. Ahuja, R. Kaur, and D. Kumar, "Trend towards the use of complex networks in cloud computing environment," *Int. J. Hybrid Inf. Technol.*, vol. 8, no. 3, pp. 297–306, Mar. 2015.

[357] Y. Zhang, Y. Guo, Z. Zhang, M. Chen, S. Wang, and J. Zhang, "Automated discovery of interactions and dynamics for large networked dynamical systems," 2021, *arXiv:2101.00179*.

[358] Z. Yang, L. Shangguan, W. Gu, Z. Zhou, C. Wu, and Y. Liu, "Sherlock: Micro-environment sensing for smartphones," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3295–3305, Dec. 2014.

[359] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.

[360] M. S. H. Abad, E. Ozfatura, D. GUndUz, and O. Ercetin, "Hierarchical federated learning ACROSS heterogeneous cellular networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 8866–8870.

[361] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4177–4186, Jun. 2020.

[362] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 5098–5107, Jul. 2021.

BOGDAN GABRYS (Senior Member, IEEE) received the M.Sc. degree in electronics and telecommunication from Silesian Technical University, Gliwice, Poland, in 1994, and the Ph.D. degree in computer science from Nottingham Trent University, Nottingham, U.K., in 1998.

Over the last 25 years, he has been working at various universities and research and development departments of commercial institutions. He is currently a Professor of data science, University of Technology Sydney, Sydney, NSW, Australia. He has authored over 200 research papers, chaired conferences, workshops, and special sessions. His research interests include data science, complex adaptive systems, computational intelligence, machine learning, predictive analytics, and their diverse applications. He is a member of the IEEE Computational Intelligence Society and a fellow of the Higher Education Academy (HEA), U.K. He has been on program committees of a large number of international conferences with the data science, computational intelligence, machine learning, and data mining themes. He is frequently invited to give keynote and plenary talks at international conferences and lectures at internationally leading research centers and commercial research labs.

KATARZYNA MUSIAL received the M.Sc. degree in computer science from the Wrocław University of Science and Technology (WrUST), Poland, the M.Sc. degree in software engineering from the Blekinge Institute of Technology, Sweden, in 2006, and the Ph.D. degree from WrUST, in November 2009.

In November 2009, she was appointed as a Senior Visiting Research Fellow with Bournemouth University (BU), where she has been a Lecturer in informatics, since 2010. In November 2011, she joined Kings College London, as a Lecturer in computer science. In September 2015, she returned to BU as a Principal Academic in computing, where she was a member of the Data Science Institute. In September 2017, she joined as an Associate Professor of network science with the School of Software, University of Technology Sydney, where she is currently a member of the Advanced Analytics Institute. Her research interests include complex networked systems, analysis of their dynamics and its evolution, adaptive and predictive modelings of their structure and characteristics, and adaptation mechanisms that exist within such systems.

• • •

JIAQI WEN received the B.S. degree in applied statistics from the Central University of Finance and Economics (CUFE), Beijing, China, in 2019. She is currently pursuing the Ph.D. degree with the University of Technology Sydney.

She worked as a Research Assistant with the School of Economics and Management, Beihang University, Beijing. Her research interest includes modelling complex networked systems in the digital twin space.