Electronic Research Archive

http://www.aimspress.com/journal/era

ERA, 30(10): 3786–3810.
DOI: 10.3934/era.2022193
Received: 29 May 2022
Revised: 21 July 2022
Accepted: 01 August 2022
Published: 23 August 2022

Perspective

A survey on temporal network dynamics with incomplete data

Xing Wu, Shuai Mao, Luolin Xiong and Yang Tang*

The Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China

* **Correspondence:** Email: yangtang@ecust.edu.cn.

Abstract: With the development of complex network theory, many phenomena on complex networks, such as infectious disease transmission, information spreading and transportation management, can be explained by temporal network dynamics, to reveal the evolution of the real world. Due to the failure of equipment for collecting data, human subjectivity, and false decisions made by machines when the high accuracy is required, data from temporal networks is usually incomplete, which makes the samples unrepresentative and the model analysis more challenging. This survey concentrates on the pre-processing strategies of incomplete data and overviews two categories of methods on data imputation and prediction, respectively. According to whether each layer in temporal networks has the coupling process, this survey overviews the dynamic modeling approaches in terms of both a single process and coupling processes on complex temporal networks. Moreover, for complex temporal networks with incomplete data, this survey summarizes various characteristic analysis methods, which concentrate on critical nodes identification, network reconstruction, network recoverity, and criticality. Finally, some future directions are discussed for temporal networks dynamics with incomplete data.

Keywords: temporal network; incomplete data; data processing; coupling process; network characteristics

1. Introduction

Complex networks widely exist in the human society, such as the social network [1], the railway network [2], the infectious disease transmission network [3], etc. Exploring the dynamics of these networks can help understand the intrinsic properties of various pratical complex systems fundamentally. In order to describe various complex systems in the real world, different models have been proposed, including the random network, the small-world network, and the scale-free network [4–7]. Temporal networks are an extension of traditional networks in the time dimension, which are a more accurate description for complex systems [8–11]. The diagram of temporal networks is shown in Figure 1. The combination of nodes and edges in temporal networks can explain some pratical complex systems' evolution mechanism in the dimension of time, which may prevent and control large-scale outbreaks of diseases or rumors among the population [12, 13]. Therefore, the study on temporal network dynamics is theoretically significant.

Due to the instability and lack of device for data collection, human subjectivity, and random noise, the data obtained from the real complex system is usually incomplete [2]. For instance, in the remote sensing applications, if insufficient sensors are placed in some areas, it may result in the incompleteness of data [14]. Incomplete data problems are often circumvented by data imputation with specific values or incomplete data deletion [15]. Classical data imputation schemes include the missing data filling with zero, unconditional or conditional averages [16]. Since data deletion often causes the loss of important information, the survey concentrates on the data imputation and generation. With the development of machine learning, supervised and unsupervised classification algorithms are widely utilized to the classification of incomplete data [17, 18]. The accurately classified data plays an important role in filling incomplete data.

In complex systems, including adaptive networks, temporal networks, and high-order networks, processing incomplete data is a challenge in studying complex systems [2]. Due to the different characteristics of the networks, there are differences in the methods of processing incomplete data in these networks. Adaptive networks consider fusion of incomplete data for analysis, which focusing on coevolution among nodes. In adaptive networks, the main methods of processing incomplete data are adaptive fusion graph networks [19], Euclidean embedding [20], matrix completion methods [21] and so on. Traditional networks only consider the relationship between two nodes. Incomplete data in realistic conditions is influenced by interactions of several factors. High-order networks can explain for missing data in real-world environments. In high-order networks, the main methods of dealing with incomplete data are graph learning attention neural networks [22], spatio-temporal imputation [23], and high-order dynamic Bayesian networks(BNs) [24]. In temporal networks, the main methods of dealing with incomplete data are multiple imputation (MI), data generation, bayesian inference, causal learning, Recurrent Neural Network (RNN) with Ordinary Differential Equations (ODEs) (or additional input features) and Long Short-Term Memory (LSTM) networks. Methods of incomplete data processing in temporal networks are summarized in Table 1.

The majority of existing literature concentrates on the analysis of temporal networks with complete data [25–27]. To better describe complex temporal networks, it is necessary to summarize analysis methods for temporal network dynamics with incomplete data. The main contributions of this survey are as follows. Firstly, this survey concentrates on the recent developments in dealing with the incomplete data for complex temporal networks and divides the existing methods and results into two categories, incomplete data imputation and prediction. Secondly, this survey overviews two categories of dynamical process modeling methods on single-layer and multilayer temporal networks: a single process on complex networks, and several coupling processes on different layers. Thirdly, this survey summarizes various analysis methods for the network characterization with incomplete data, which are utilized to analyze the local and global characteristics of complex temporal networks, respectively.

This survey is organized as follow. Section II introduces two categories of methods for data imputation and prediction. Section III concentrates on the temporal network modeling for both a single process on complex networks and the coupling process on different layers respectively. Section IV overviews some characteristics of temporal networks with incomplete data and the corresponding analysis methods. Section V summarizes some future directions such as high-order interaction, open environment, heterogeneity and few-shot learning. The overall structure of this survey is shown in Figure 2.

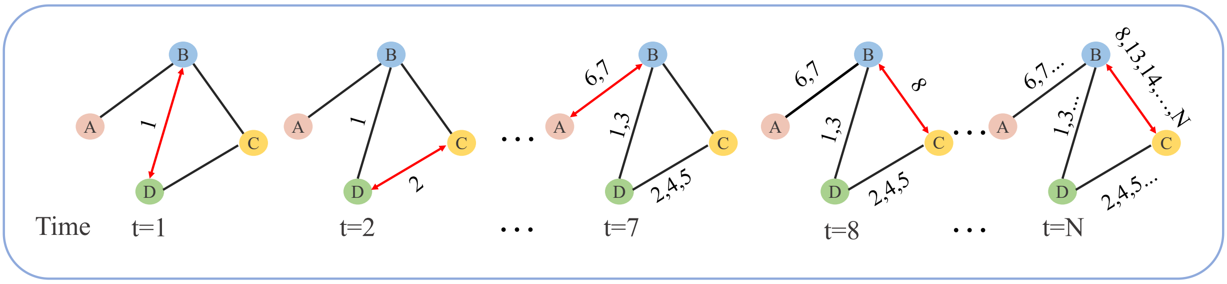

Figure 1. Diagram of temporal networks [8].

Table 1. Methods of incomplete data processing in temporal networks.

Year	Reference	Method
2019	Garcia *et al.* [15]	MI
2017	Dzunic *et al.* [28]	Bayesian inference
2015	Westreich *et al.* [29]	Causal learning
2021	Zhu *et al.* [30]	Data generation
2021	Zhou *et al.* [31]	RNN with ODEs
2018	Rahman *et al.* [32]	RNN with additional input features
2021	Yang *et al.* [33]	LSTM

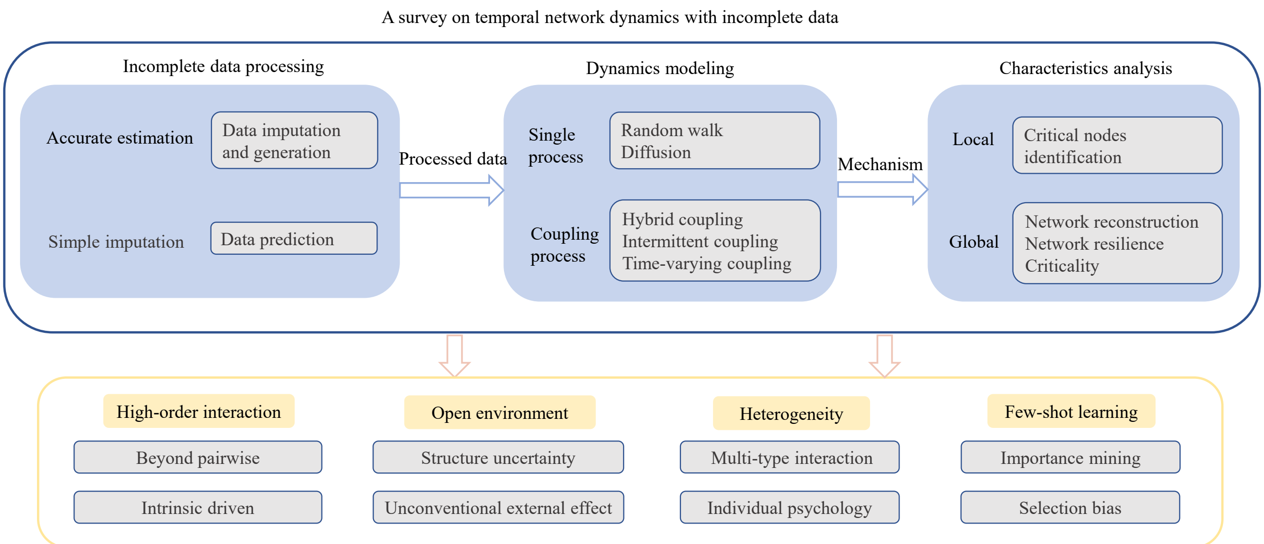

Figure 2. The overall structure of the survey.

2. Incomplete data processing in temporal networks

The incomplete data in temporal networks is mainly classified into three types: missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR) [15]. To eliminate the adverse effects of incomplete data, there are two kinds of processing methods. The first is incomplete data imputation and generation, which input the processed complete data into the model [16]. The other one is incomplete data prediction, which develops a prediction model for missing data with minimal pre-processing cost [34]. The difference between these two strategies is that the former requires data pre-processing to get accurate estimated data, which is more computationally intensive but highly accurate [35]. The latter does not require the complete imputation of initial data before inputting, and relies on the prediction model with minimal pre-processing cost. The accuracy is closely involved with the performance of the prediction model. The approaches of incomplete data processing are summarized in Table 2.

Table 2. Summary of incomplete data processing.

Year	Reference	Type of incomplete data	Processing Type	Approach	Application	Interpretability
2019	Garcia *et al*. [15]	MCAR	MI	An incremental learning method	Industry system	
2019	Venugopalan *et al*. [16]	MCAR	MI	An alternating least squares PCA method	Medical System	
2016	Yu *et al*. [36]	MAR	MI	Time series matrix through method	Climatology	
2019	Liu *et al*. [14]	MAR	MI	K-means method	Industry system	
2020	Liu *et al*. [37]	MAR	MI	Consensus clustering matrix with regularization	Industry system	
2020	Liu *et al*. [1]	MAR	MI	Combinatorial analysis	Sociology	✓
2020	Mancuso *et al*. [38]	MAR	MI	Dynamic property analysis	Industry system	✓
2022	Gunn *et al*. [39]	MAR	MI	Lasso linear regression	Psychology	✓
2017	Dzunic *et al*. [28]	NMAR	Bayesian inference	A switching Bayesian model	Industry system	
2020	Zhao *et al*. [40]	NMAR	Bayesian inference	A semi-supervised sparse bayesian regression model	Industry system	
2019	Benjumeda *et al*. [41]	NMAR	Bayesian inference	A Bayesian model	Industry system	
2017	Hasan *et al*. [6]	MAR	Bayesian inference	a semi-Markov algorithm for continuous time	Sociology	
2017	Hasan *et al*. [42]	NMAR	Bayesian inference	A crowdsourcing Bayesian model	Sociology	
2015	Westreich *et al*. [29]	MAR	Causal learning	MI and the parametric g-formula	Industry system	
2020	Ray *et al*. [43]	MAR	Causal learning	Semiparametric Bayesian causal method	Industry system	✓
2020	Nguyen *et al*. [44]	MAR	Causal learning	Inverse Bayes formula	Sociology	✓
2019	Tikka *et al*. [3]	MAR	Causal learning	Bivariate missing data analysis	Medical System	✓
2021	Athey *et al*. [45]	MAR	Causal learning	Matrix causal structure estimators	Sociology	✓
2020	Richens *et al*. [46]	MAR	Causal learning	Counterfactual diagnostic method	Medical System	✓
2021	Zhu *et al*. [30]	MAR	Data generation	A conditional GAN	Chemical industry	
2021	Hu *et al*. [25]	MAR	Data generation	A GAN based on trinetworks form	Pipeline Networks	
2021	Gao *et al*. [12]	MAR	Data generation	An attention GAN	Medical System	
2020	Xu *et al*. [2]	MAR	Data generation	A GAN based on graph embedding	Transportation networks	
2022	Xu *et al*. [26]	MAR	Data generation	A deep hashing model with GANs	Transportation networks	
2021	Zhou *et al*. [31]	NMAR	Prediction	Discrete optimization method	Transportation networks	
2019	Raissi *et al*. [34]	NMAR	Prediction	A deep RNN model with ODEs	Medical System	
2018	Rahman *et al*. [32]	NMAR	Prediction	Dynamic Layered-RNN method	Power grid	
2018	Tian *et al*. [17]	NMAR	Prediction	Multi-scale temporal smoothing analysis	Transportation networks	
2021	Yang *et al*. [33]	NMAR	Prediction	Graph Laplacian spatial regularizer	Transportation networks	

2.1. Incomplete data imputation and generation

Data imputation is the most general method to refine incomplete data [36]. It can be divided into the imputation with statistics and machine learning (ML). The statistical variables, such as mean, mode or median value, are generally utilized for filling incomplete data. But these methods ignore the important information about feature, structure types and so on. With the rapid development of computation power, ML is now predominantly used in many scenarios, including MI [14], data generation [30], bayesian inference [6], causal learning [43] and so on.

2.1.1. Multiple imputation

MI concentrates on the own time series evolution of missing values and the neighboring ones because different time series are often correlated with each other. In particular, MI considers the influence between temporal indicators [39]. The diagram of MI is shown in Figure 3. By learning the global characteristics of time series matrix through the matrix decomposition, the temporal characteristic matrix is approximated by a low rank matrix [36]. This algorithm has low computational complexity, and can handle larger scale data under MAR. A temporal regularized matrix factorization framework is proposed to support data-driven temporal learning and forecasting in [36]. This framework utilizes the scalable matrix factorization methods that are eminently suited for incomplete high-dimensional temporal data. Considering the incompleteness of the time matrix under MAR, a new algorithm is presented to combine the inputs with clustering in [14], which makes the incomplete matrices complement each other. Based on the combination scheme in [14], the prior knowledge of incomplete data under MAR is utilized to improve the algorithm [37]. The consensus clustering matrix with regularization is presented to improve the filling sub-classification performance. Considering the time-series prediction in an online context with incomplete data under MCAR, a modified evolving granular fuzzy-rule-based model is proposed in [15]. This model provides the parameter estimation in a nonlinear and time-varying way.

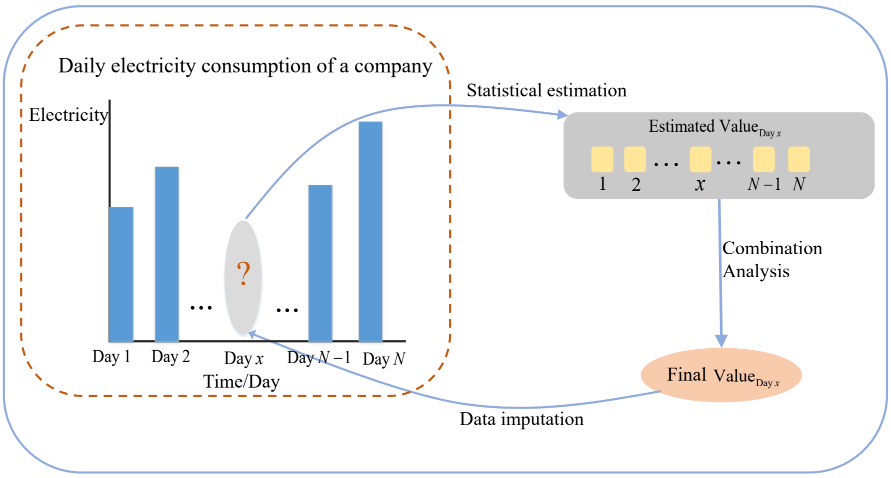

Figure 3. Diagram of multiple imputation.

In many scenarios, filling incomplete data requires giving specific rules, and makes the final result understandable and trustworthy to the public. So data imputation needs to improve interpretability. To meet this situation, it is difficult for the matrix decomposition to give an intuitive explanation. Therefore, it is necessary to combine multi-source tensors and their relationships by considering the external properties of the analyzed data with interpretable correlations [1]. Based on the neighboring users' electricity, it is considered to correlate the electricity consumption data of different users and complement the missing data under MAR [1]. Considering the dynamic characteristic in [1], a termed Sample-Lasso method is proposed to interlink dynamic characteristics of each new target sample [38]. This way of interlinking the expression of unmeasured genes demonstrates the biological interpretability. To further consider the interpretability of regularization methods like the Lasso, the hybrid estimation of fitting a Lasso for variable selection is utilized to fill the data under MAR by combining several different variables in [39]. It reduces the variance of the estimates and increases the interpretability of the model.

2.1.2. Bayesian inference

Bayesian inference is a statistical learning inference method, and it updates the probability of hypotheses to obtain more evidence and information [42]. Bayesian inference is a critical technique which is widely utilized in scientific research, engineering and other fields. Bayesian inference considers the uncertainty of data imputation and the correlation between data, which can combine the possibilities of all incomplete data to make optimal decisions. A switching Bayesian model is developed for dependency analysis, which relies on a state-space approach accounting for the noisy measurement processes and incomplete data under NMAR [28]. Considering the uncertainty of the Bayesian model, a semi-supervised sparse Bayesian regression model is proposed to deal with incomplete output under NMAR, through variational reasoning technology in [40]. The structural expectation maximization (EM) algorithm is introduced to enhance the uncertainty of the model in BNs from incomplete data sets. With sufficient data and a defined maximized objective function, the desired model can be obtained by training with the relevant parameters. Considering the high computation cost, with the improved EM algorithm and a semi-supervised Bayesian regression model, the polynomial cost in variables is utilized in the BNs for the lower inference complexity in [41]. Based on the incomplete data randomness under NMAR, an EM algorithm and a semi-Markov algorithm for continuous time are introduced to infer the data type information from incomplete trajectory data, which can predict a person's next action [6]. It settles the activity construction with incomplete information on online social media. Bayesian inference also has many extensions to accommodate more complex cases. To enhance the performance of Bayesian inference in precision, the concept of crowdsourcing is integrated into the Bayesian inference process, which reduces the cost of acquiring training samples [42].

2.1.3. Causal learning

Causality represents a type of correlation, and causal inference is used to learn about cause and effect relationships [29, 47]. The diagram of causal learning is shown in Figure 4. Two graphs in the diagram $X_1 \to X_2 \to X_3$ and $X_1 \leftarrow X_2 \leftarrow X_3$ imply the same conditional independence (X_1 and X_3 are independent given X_2). P represents the probability distribution. $d(t)$ represents the action at time t. A causal model represents a set of distributions, one for each possible intervention (indicated with a tool icon in Figure 4). In this sense, the causal relationship inference implies gaining the significant knowledge [48]. This knowledge is not limited to independently and identically distributed data. Some methods such as a general search-based approach learn the causal structure from the missing data and then estimate the causality of the data to complete missing data [3]. A causal inference method to fill the missing value under MAR is proposed in [29], which is similar to multiple imputation and the parametric g-formula. The result suggests that epidemiologists can benefit from thinking about the potential outcome's causal inference. Based on the uncertainty of incomplete data under MAR, a semiparametric Bayesian causal approach is proposed to determine the relationship in the incomplete data, to obatin the complete data [43]. Considering the common cause failure rate based on causal inference with missing data under MAR, the inverse Bayes formula combined with causal learning is developed to handle this problem in filling missing values [44]. This method also improves the universality of processing incomplete data.

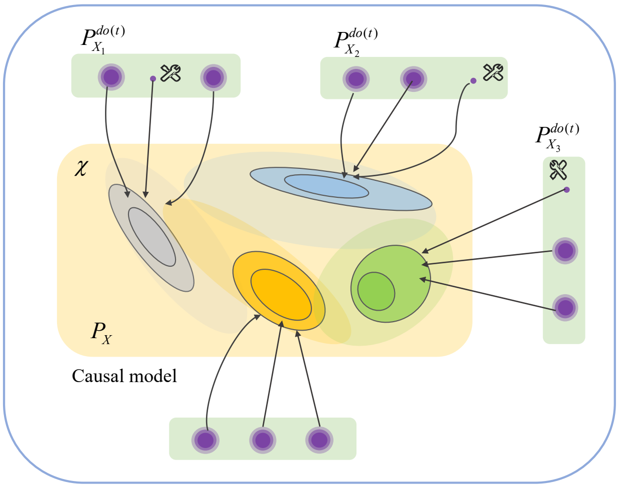

Figure 4. The framework of causal learning [48].

Counterfactual causality can be continuously verified by hypothesizing and constructing causal models to obtain more complete data. The causal relationship between the incomplete data can be closer to the real complete data through the above method [46]. Considering the counterfactual based on the causal relationships, a systematic analysis of the bivariate missing data problem in counterfactuals is proposed in [3]. For counterfactuals of the treatment unit/cycle combination, a class of matrix completion estimators that uses the observed elements of the matrix of control is utilized to deal with missing elements of the control outcome matrix under MAR [45]. To improve the interpretability and accuracy in the medical diagnosis, the diagnosis is reformulated as a counterfactual inference task, and a causal definition of diagnosis that is closer to the decision making of clinicians is presented in [46]. This approach derives counterfactual diagnostic algorithms for incomplete data under MAR.

2.1.4. Data generation

With the development of Generative Adversarial Networks (GANs) in recent years, many approaches based on GANs are utilized to fill incomplete data in time series [49]. The distribution features of time series are captured by generative models. The diagram of GANs is shown in Figure 5. It can regenerate the time series data and fill the original incomplete data [26]. In order to solve the problem that only few samples are available, a virtual sample generation method based on conditional GANs is proposed in [30]. It combines a local anomaly factor with the k-means algorithm to generate two output samples that match the overall output trend. This achieves the purpose of filling the incomplete data under MAR. Considering the multiple types of generated data, the GANs based on the ternary network form are developed to handle leak detection problems with incomplete sensor data under MAR [25].

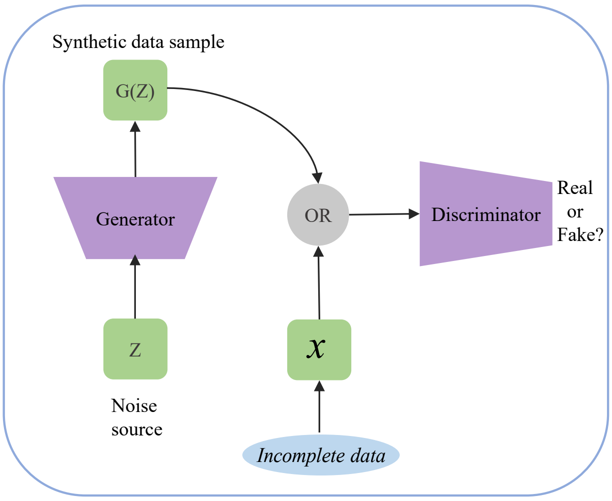

Figure 5. Diagram of GANs [49].

Based on the need for online learning, a deep learning (DL) framework is proposed in [2], which implements the representation of road networks in embedding graphs. In this case, GANs are used to generate the real-time traffic state information, so road traffic conditions are estimated based on information from adjacent links. Considering the heterogeneity of incomplete data under MAR, a DL framework, which integrates a task-induced pyramid-attention GAN with a pathwise transfer dense convolution network, is proposed in [12]. There is a better improvement in multi-type data processing. For the robustness of data processing under MAR, the loss in the reconstruction and the generation is utilized mainly to generating continuous features of the incomplete data in deep hash model [26]. It improves the accuracy of the generated data.

2.2. *Incomplete data prediction*

Rather than producing the accurate estimate of incomplete data, the primary purpose of the prediction-centric strategy is to construct a prediction layer for incomplete data, which data is processed with simple imputation [32]. This improves the ability to handle incomplete data under NMAR. Both Recurrent Neural Network with Ordinary Differential Equations (or additional input features) and Long Short-Term Memory networks, are usually utilized as the prediction layer.

2.2.1. RNN with ODEs

Neural ODEs use neural networks to parameterize the derivatives of hidden states [31]. It can adapt the evaluation strategy according to each input without partitioning or ordering of data dimensions. This method can be formulated commonly as

$$\mathbf{h}_{t+1} = \mathbf{h}_t + f(\mathbf{h}_t, \theta_t) \tag{2.1}$$

where $\mathbf{h}$ is the value of the hidden layer, t represents the serial number of the hidden layer, $t \in \{0 \ldots T\}$ and $\mathbf{h}_t \in \mathbb{R}^D$. These iterative updates can be seen as an Euler discretization of acontinuous transformation [50].

In the limit, we parameterize the continuous dynamics of hidden units using an ODE specified by a neural network as

$$\frac{d\mathbf{h}(t)}{dt} = f(\mathbf{h}(t), t, \theta) \tag{2.2}$$

where $\mathbf{h}(0)$ represents the input layer, $\mathbf{h}(T)$ represents the output layer, θ represents dynamics parameters. Starting from the input layer $\mathbf{h}(0)$, we can define the output layer $\mathbf{h}(T)$ to be the solution to this ODE initial value problem at some time T.

RNN with ODEs can naturally process arbitrary time intervals and common incomplete data under NMAR [50]. To balance the prediction accuracy and computational efficiency, the discrete optimization method is introduced in [31]. Meanwhile, the hidden states of RNN with ODEs are utilized to form a new urban flow prediction framework to improve the handling of missing values. Based on the robustness of the prediction model, combining discrete optimization methods with ODEs, the incomplete data and the signal-to-noise ratio in a high-dimensional state are interrelated in [51]. In order to account for the uncertainty of incomplete data under NMAR, the probabilistic latent ODE dynamics model parameterized by deep Bayesian neural networks is proposed in [52]. Neural networks, which combine the recursive bayesian filtering with the known state dynamic differential equations, are proven to well approximate the state dynamics corresponding to missing ODEs in [34]. The validity of this model is verified in a recognized model of human retinal blood circulation.

2.2.2. RNN with additional input features

By adding inputs with additional features, which can provide direct data information about the time delta and the linear relationship between missing data, RNN can further effectively handle incomplete data under NMAR [35]. Considering the specific operating environment and the robustness under device failures, a deep RNN model is adopted to develop a data calculation scheme about electricity consumption data, and adds the scenario feature information on the input data [32]. This scheme can obtain more complete data by merging scene information in the power grid but the types of data that can be predicted are limited. The information between the missing values and the observed values can also be provided by increasing the interval between observed values. A dynamic Layered-Recurrent Neural Network (L-RNN) approach is proposed to recover missing data under NMAR from the Internet of Medical Things [53], which can increase the dynamic observed interval in each layer to obtain more information. Its advantage over Deep RNN is that it divides the data into complete and incomplete data, so just a dynamic L-RNN trained by complete data can predict other incomplete data under NMAR. Considering the generality of the model, the approach which integrates improved incomplete data into two general regression neural networks is proposed in [54]. Meanwhile, an extension term with a binary variable indicating missingness is added to the neural input structure in the task of missing data management systems. This model can improve the overall prediction accuracy.

2.2.3. LSTM

LSTM is a class of RNNs, which is famous for the special network structure. LSTM adds three control units such as the input gate, the output gate and the forget gate. As the data inputs this network, the units in LSTM network will judge whether the information conforms to the rules, which can solve the long time series problem in neural networks [33]. This approach is utilized to confirm whether incomplete data under NMAR is retained when processing multiple data types. Considering the prediction accuracy, the incomplete data under NMAR is subdivided by a LSTM network prediction model into multiple types [55]. Then the refined data is utilized for retraining. Based on the multiple scales of temporal data, a multi-scale temporal smoothing method is introduced to settle the data consistency at different time scales [17]. For the correlation between incomplete data under NMAR, the graph Laplacian model is chosen as a spatial regularizer in the LSTM network [33]. It improves the prediction performance by exploiting the spatial correlation among the network sensors. Considering the robustness of the model, a stacked bidirectional unidirectional LSTM network structure is proposed in [18]. It can fill missing values under NMAR and assist in the traffic prediction by designing an estimation unit, which improves the model accuracy without adding additional computation burden.

3. Temporal network dynamics modeling

Incomplete data can lead to the final results that are far from reality. Therefore, it is necessary to increase the feasibility of temporal network dynamics models with incomplete data based on data processing [56]. In the following, this section will summarize the dynamics modeling on monolayer and multilayer networks according to whether there exists the coupling behavior, respectively. The diagram of interactive process dynamics in temporal networks is shown in Figure 6. The approaches of dynamics modeling and characteristics analysis in temporal networks are summarized in Table 3.

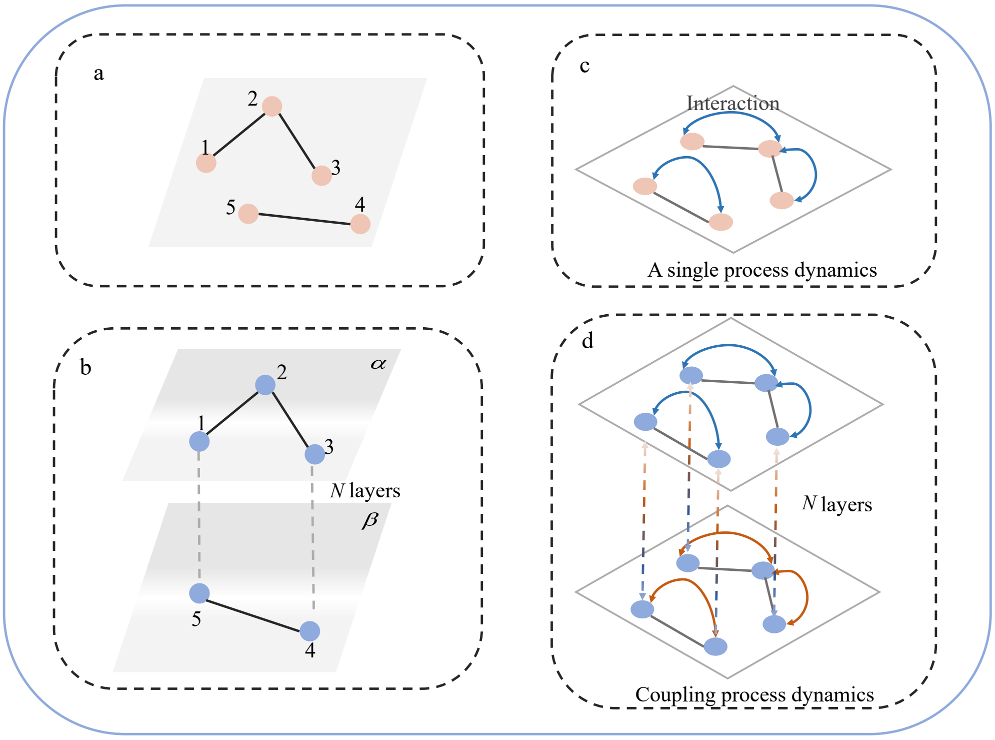

Figure 6. Diagram of interactive process dynamics in temporal networks. (a) A single-layer network, (b) Multilayer networks, (c) A single process dynamics in networks, (d) Coupling processes dynamics in different networks.

3.1. A single process in networks

3.1.1. Single-layer networks

The structure of social relationships between individuals, interactions between proteins, and many other cases can be represented by networks. The majority of research concentrates on networks made up of a single entity, where an entity is only interconnected by another entity [57]. Such networks are now known as single-layer networks [58]. When there is one process in single networks, the intrinsic mechanism modeling of temporal networks is considered. Based on the group characteristics in the Susceptible-Infectious-Recovered (SIR) model, the behavior characteristics of group communication are introduced in [59]. This model explains the information about the potential disease spreading in the population, and how the disease dynamics are influenced by consciousness. Considering the interpretability of the model, the parameters in the spatio-temporal volterra model are estimated in wireless sensing systems [57]. This model can explain the correlation of data processing tasks between wireless sensor network nodes at different time periods when analyzing wireless sensor networks with temporal networks. Considering the bursty activities on the above theory, a temporal network model based on the bursty node activation is proposed to explain the dynamic processes of various systems in [58].

Data-driven modeling is an effective solution when the amount of data is larger and the mechanism of the network is not clear. Based on the robustness of the model, a data-driven model with a new fundamental graph is developed to discover the critical density of a given road segment in the highway traffic networks [60]. It determines scenario-oriented event definitions in a completely unsupervised manner, independent of the noise level in the input data. To further consider stochastic nonlinearity in temporal networks, a data-driven model based on the mixed integer linear programming is proposed to enhance the real-time detection capability in [61]. The time-series data of the transmission system is utilized as input, which can quickly identify uncertain fault event scenarios and resolve missing sensor transmission data.

Table 3. Summary of dynamics modeling and characteristics analysis in temporal networks.

Year	Reference	The network type	Coupling process	Incomplete data processing type	Data-driven modeling	Approach	Characteristics
2020	Gupta *et al*. [57]	Single-layer		MI		Communication-efficient method	Critical nodes identification
2022	Zhao *et al*. [59]	Single-layer				Opinion leader theory	Criticality
2020	Hiraoka *et al*. [58]	Single-layer				Actual matching method	Criticality
2020	Jiang *et al*. [61]	Single-layer		Causal learning	✓	Mixed integer linear programming	Criticality
2018	Alesiani *et al*. [60]	Single-layer		Data generation	✓	AIP method	Criticality
2017	Mei *et al*. [62]	Multilayer				Random walk	Criticality
2018	Manfredi *et al*. [63]	Multilayer		MI		Random Walk	Criticality
2019	Indu *et al*. [13]	Multilayer		Data generation		Forest-fire method	Critical nodes identification
2020	Liu *et al*. [64]	Multilayer		MI		Variable propagation rate and perception mechanism	Critical nodes identification
2020	Liang *et al*. [65]	Multilayer		Data generation	✓	Cross-media semantic correlation learning	Critical nodes identification
2015	Domenicou *et al*. [56]	Multilayer		Data generation		Infomap search method	Critical nodes identification
2020	Hu *et al*. [4]	Multilayer	✓	MI		Direct error method	Network reconstruction
2019	Jia *et al*. [5]	Multilayer	✓			Nonsmooth analysis	Criticality
2021	Sun *et al*. [66]	Multilayer	✓			Derived method	Network resilience
2018	Zhan *et al*. [67]	Multilayer	✓	MI		Nonlinear analysis	Criticality
2020	Müller *et al*. [68]	Multilayer	✓	MI		Nonlinear analysis	Criticality
2022	Gao *et al*. [27]	Multilayer	✓		✓	Two-phase autonomous inference method	Network reconstruction
2017	Hasan *et al*. [6]	Multilayer	✓	Bayesian inference	✓	An expectation-maximization algorithm	Network reconstruction
2021	Yang *et al*. [33]	Multilayer	✓	LSTM	✓	Alternating algorithm and GRMF method	Criticality

3.1.2. Multilayer networks

A typical model of a single dynamic process on multilayer networks is the random walk. This model mainly considers the relative speeds and probabilities of steps within layers versus steps between layers which can affect the qualitative nature of the dynamics [56]. Considering the relevance of random walk, a theoretical modeling framework based on the compressed sensing and regularization is proposed in [62]. It achieves more accurate estimation of structural information between layers. Considering the further robustness of models in temporal networks, it is modeled from the perspective of the nodes which have a limited capacity of storing and processing the agents on multilayer networks [63]. This model improves the predictive ability of long time-series data in the random walk model.

Another typical model of a single dynamic process on multilayer networks is the information spreading on social media and multimodal transport systems. To more accurately identify the spreading characteristics of rumors, a model based on laws of the nature like forest fires is proposed in [13], which focuses on identifying best features of simulated rumor propagation through online social networks. This method connects the spread of rumors in social networks with the spread of wildfires in forests, and identifies the main characteristics of rumor spreading. Considering the dynamic variability of rumors, the variable propagation rate and perception mechanism are integrated with a rumor spreading model, which can change the attitude of ignorant spreaders [64]. It can also explain the internal motivation of the rumor but there is limited capability for computing more complex social networks. With the increase of the number of network layers, the integration of deep neural networks can break the computational limitations of the above models. Based on the framework of deep neural networks, a timetable-based modeling approach is developed to incorporate critical temporal factors into the travel time analysis [69]. However, these models cannot utilize similarities and correlations in temporal networks. Combining deep neural networks with hash networks, the deep network with multimodal data features is utilized to an unified optimization architecture [65]. This model better preserves the similarity and correlation of incomplete data in temporal networks.

3.2. Coupling processes in multilayer networks

Dynamics processes on different layers of multilayer networks can interact with each other, and studying different interaction processes on multilayer networks is a hot research topic in network science [6,70]. The dynamics of multilayer networks can be analyzed by coupling mechanism modeling. The diffusion process in multilayer networks is studied in [71]. Some physical phenomena is emphasized, which is related to the diffusion process generated by multilayer structures from the perspective of coupling mechanism. Considering the homogeneity of the coupling process, the network-based models are campared with homogeneous hybrid models in [72]. The differences in the predictions of homogeneous mixture models are pointed out and the dynamics of disease behavior processes on complex networks in statistical physics are also compared. Considering the intermittency of the coupling process, a direct error method based on the intermittently coupled temporal network model is introduced in [4]. It reveals the mechanism of intermittent coupling dynamics on multilayer networks. Considering the time variant in temporal networks, the Lyapunov function and the Laplace transform technique are introduced to model and analyze the coupling dynamics of complex multilayer networks in [5].

If the detailed and accurate process mechanisms are known, or a wealth of experience and knowledge is available, both mechanistic and data-driven modeling could work well. The increasing complexity of temporal networks makes these prerequisites no longer easy to meet. So it is inevitable to consider data-driven modeling. In order to accurately describe the dynamics of two interacted unidirectional diffusion processes, a data-driven model with the low complexity is developed to predict diffusion on a competitive scenario on various networks such as Lesmis, Football and Power [66]. This model can be used to predict the spread of coupled infectious diseases. Considering the nonlinearity in temporal networks, the large graphical limit of the nonlinear model is related to the stochastic prevalence model in [67]. It explains the coupling phenomenon of disease transmission in the population and network information spreading. In order to improve the online analysis of temporal network dynamics models, the coupling process of local neuronal synapses and the nonlinear dynamics modeling are integrated in [68]. To further consider the robustness of online analysis models, a two-stage approach is proposed to develop a data-driven model with the robust inference, which can study the dynamics of the early spread of influencing a global aviation network [27].

4. Characteristics analysis in temporal networks

The characteristics of complex networks represent the intrinsic mechanism of complex systems [4]. This section analyzes the characteristics mainly from these aspects: critical nodes identification, network reconstruction, network recoverity, and criticality.

4.1. Critical nodes identification

Critical nodes are generally nodes in the core position of temporal networks and have a significant influence on the structure and function of the whole network [73]. The critical node identification can obtain a priori knowledge about the importance of entities and thus predict the development of events, such like predicting critical components to prevent catastrophic failures in the grid [32]. To improve the local recognition, an algorithm to refine BNs structure is introduced to modify the local edges, which connect the nodes affected by misclassification in [74]. Considering the recognition of local features, a method based on a set of user-specific features is proposed to obtain more accurate critical nodes in [75]. To identify the current and future critical influencers on Twitter, it uses graph data mining to extract influence factors of the user-specific features. Considering the robustness of recognition, a new approach to identify critical communicators on social networks is proposed, which extracts and integrates differential features from various complex network connectivity structures [76]. Considering the global characteristics of nodes in temporal networks, the integration of global network structure features is utilized in [77], which can effectively identify critical nodes in virus spreading. Considering the heterogeneity of global features, an identification method of saliency analysis based on global features and the spectral is developed to identify critical nodes in an electricity meter network with incomplete data sets [78].

4.2. Network reconstruction

In the study of social, economic and biological systems, the need to solve the data scarcity of network structure has led to the birth of network reconstruction [79]. Combining incomplete activity information and diffusion times, a semi-Markov probabilistic modeling approach is proposed to reconstruct the user's activity trajectories on the location sequences in [6]. This approach is meaningful to predict an individual's next activity, duration and location with incomplete trajectory data. Based on the above theory, adding global information can make the prediction more accurate. On the basis of the zero model assumption, a data-driven model based on the stochastic diffusion process is proposed in [79], which reconstructs temporal networks based on the node diffusion time. Considering the integration of a priori information in the global features, deep convolutional neural networks are utilized to learn the correlation and prior information in the original data, which can reconstruct the network more efficiently [80]. To explain the robustness of the model, the full sampling conditions and differential phase contrast within the deep convolutional neural network are introduced to reconstruct the multilayer network [81].

4.3. Network resilience

The structure and function of a complex system may be failed and dysfunctional under the disturbance of external factors such as external attacks [82, 83]. In contrast, the structure and performance of the system can be gradually restored under self-resilient factors or human intervention [84]. In the perspective of local intervention, a data-driven model under incomplete data is developed to predict the recovery of dynamic temporal networks in fault states [85]. Based on the dynamic coupling process, a local node intervention scheme is proposed to recover the network resilience in [86]. Considering the model development day robustness on the basis of coupling process, a two-step recovery scheme is developed in [87]. It reconstructs failed temporal networks until the point where it can be recovered, and then dynamically intervenes to reignite the network's lost functionality.

In the perspective of global intervention, a probabilistic matrix decomposition of the data from the sensors is presented in [88], which adds the extended inputs to the neural structure. This approach is utilized to recover the Internet of Things from a fault state. To improve the computation efficiency, a method called the principal component analysis with the spatio-temporal tensor is proposed in [89], which utilizes tensor decomposition and low-dimensional representation. It can recover traffic networks from corrupted and incomplete data. This will significantly improve the safety and real-time performance in the traffic management. Considering the multi-scale analysis in network space, a general network recovery framework for big data is presented in [90], which can augment a generalized argument for inferring omnidirectional, multilayer, and multispace networks from any high dimension of data. This framework can retrieve dynamic information to infer various meaningful networks from static data.

4.4. Criticality

In complex networks, once the external disturbance exceeds a certain limit, the system will change from one equilibrium state to another equilibrium state, which is called the criticality of the system [91]. When a complex system reaches a critical value, we may have no time left for risk prediction and no valuable warning to trigger a sudden and dramatic change, which can cause the irreversible damage. Therefore, it is critical to analyze the criticality of complex systems. It is considered feasible to explain criticality in terms of the mechanism in complex temporal networks [91]. Based on the mean-field limit theory, the critical threshold for the scale-free diffusion behavior is given in [7]. It explores the spread of disease on a directed scale-free network and the effect of node thresholds on disease spread. On the basis of the above, through analyzing the classical maki-Thompson model of rumor spreading, the researchers discuss the critical point of the avalanche evolution and efficiency of information spreading in social networks [91]. Considering the heterogeneity of the spreading structures, the messaging model is introduced in the Maki-Thompson model, which can reproduce the observed diffusion characteristics and reveal how data incompleteness affects the criticality of news diffusion [92]. Considering the issue of controllability in the rumor spreading, two types of controlled and uncontrolled propagation behaviors are introduced in the rumor spreading model [93], which can better identify the spreading criticality in temporal networks. Based on the sensitivity analysis of models, two different social structures in temporal networks are proposed to predict the critical point [94]. It reveals that an infectious disease model can be constructed based on different assumptions.

Because of the DL's powerful ability to classify features of time series with large-scale data, DL is widely used in the prediction for critical points. Based on the DL framework, the growth dynamics of cell populations in microbial ecology is invoked to develop a basic growth model of modal popularity in online social networks, which can predict the criticality of the meme's development [95]. The contribution is to incorporate the general model of human interest dynamics into the basic model. In a complementary approach based on DL, the dynamic information of nodes is introduced to predict the critical point of disease spreading in [96]. Considering the multi-scale of incomplete data, the features from information in the general form and the scale behavior of all critical points in the vicinity of the dynamic system are extracted in [97]. The features are introduced into the DL algorithm for critical prediction. A robust machine learning framework is proposed to identify epidemic thresholds for susceptibility-infection-susceptibility dynamics in complex networks [98]. Reservoir Computing is a simplified cyclic neural network architecture, which has shown the excellent performance in predicting the dynamics and criticality of complex systems [99].

5. Discussion

In this survey, the processing of incomplete data for temporal networks is mainly introduced from the two perspectives of data imputation and prediction. Data imputation requires data pre-processing to get accurate estimated data, which is more computationally intensive but highly accurate [35]. Data prediction does not require the complete imputation of initial data before inputting, and relies on the prediction model with minimal pre-processing cost [32]. But the size of dataset and computational complexity are not considered. The dynamic modeling of temporal networks mainly summarizes the dynamics modeling on monolayer and multilayer networks according to whether there exists the coupling behavior, respectively. The single-layer networks are mainly modeling in mixed integer linear programming [61], AIP [60], opinion leader theory [59] and other methods. The methods of random walk [62, 63] and spreading [66, 68] are utilized in multilayer networks. However, the driving mechanisms of coupling behavior, the distinction of coupling dynamics in heterogeneous and homogeneous networks, coupling dynamics in high-order networks are not considered. This survey analyzes the characteristics mainly from these aspects: critical nodes identification [74], network reconstruction [80], network recoverity [87], and criticality [7]. The analysis methods from local to global are summarized and discussed in this survey, but the characteristic analysis in temporal networks is limited in such networks, which has none process of material, energy and information exchange between networks and the external environment.

6. Perspectives

Based on the theory of propagation dynamics for complex temporal networks, we summarize the data processing, dynamics modeling and characteristics analysis of temporal network dynamics with incomplete data, and look forward several research directions in the future.

6.1. High-order interaction

Low order systems are those in which only self-interaction or pairwise interactions take place (like edges in a graph), while higher-order systems display interactions in groups of more than two elements [100]. The high-order correlations of elements in systems can better explain some phenomena, such as the realization of brain functions, social communication, and ecological evolution [49]. So interactions in high-order systems cannot be simply described in terms of point-edges. There are crucial differences between pairwise modeling and higher-order interactions. The dynamics modeling beyond pairwise interactions in temporal networks receive less attention [27, 67, 68]. The higher-order structure of these complex temporal systems can be utilized to improve our modeling performance and predict their dynamic behavior precisely. There is a clear evidence indicating that neither interactions in social media nor in biological systems can explain why they emerge [5, 58, 66]. This modeling beyond pairwise interactions may be possible to account for the above phenomena. In addition, we still lack a general understanding that how beyond pairwise interactions affect dynamics systems. Through the intrinsic driven mechanisms, we may obtain a more accurate model to describe the temporal network dynamics with incomplete data.

6.2. Open environment

Temporal networks in the open environment mean that there is a process of material, energy and information exchange between networks and the external environment [101]. With the inputs and outputs changing, the network structure becomes more uncertain. The uncertainty of the network structure is mainly reflected in the nodes increasing (new sensors access), the nodes decreasing (sensors failure), and compensation of network links. In addition, unconventional external effects also require to be considered in open environments. Unconventional external effects, such like the bursty weather and the implementation of new policies, can cause the temporal networks to become uncertain. Considering the influence of strong external inputs, such as high noise and policy orientation, on the coupling behavior between nodes, and internal dominant factors, such as the dynamic change characteristics of internal nodes, the controllability analysis in temporal networks with incomplete data will be a major trend.

6.3. Heterogeneity

Considering the study of multilayer coupling dynamics, a standard SIR model is generally utilized on each layer [27, 71]. This approach may draw the connectivity pattern of each layer from the same standard random-graph model, which concentrates on multilayer homogeneous temporal networks [67, 68]. Due to the heterogeneity in the individual psychology and interaction patterns, it is significant to incorporate the heterogeneity into the temporal network dynamics. The heterogeneity of temporal network dynamics is reflected in the simultaneous information spreading both online and offline, which are multi-type interaction processes in temporal networks. The offline and online patterns of the information spreading are completely different network structures. In addition, the way to deal with the decision process in traditional social physics is to describe the belief dynamic evolution in temporal networks. These methods can reproduce some features of group decision-making, but cannot explain the individual psychology associated with these processes, especially in the heterogeneous group. The incorporation of individual psychology in dynamics modeling is beneficial to analyze the heterogeneity of temporal networks.

6.4. Few-shot learning

Due to the difficulty of collecting data in certain environments, the number of collected data samples is usually few and incomplete, which cannot meet the needs of current models. So few-shot learning is more suitable. Few-shot learning faces two main problems in temporal networks with incomplete data. One is that the important information of collected data obtained in many scenarios is incomplete. The problem caused by incomplete data in few-shot learning can be more serious. So the missing part of data in few-shot learning becomes particularly crucial. The significance identification of missing information is a challenge. By measuring the correlation between data and practical issues, it can selectively complete the dataset, while deleting unimportant data and filling important data [33, 65]. This can be conductive to deal with incomplete data more quickly and accurately. The other problem is data selection bias in few samples, where the data is selected in a way that differs from the target group. Due to data selection bias is usually unintentionally induced, such spurious correlations may be difficult to identify in advance. Stability learning and its variations are good choices for data selection bias, based on the latest stability learning ideas.

Acknowledgments

This work is supported by National Key Research and Development Program of China (2021YFB1714300), National Natural Science Foundation of China (Key Program: 62136003), Program of Shanghai Academic Research Leader (20XD1401300).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Z. Liu, Y. Yang, W. Huang, Z. Tang, N. Li, F. Wu, How do your neighbors disclose your information: social-aware time series imputation, in *The World Wide Web Conference*, (2019), 1164–1174. https://doi.org/10.1145/3308558.3313714
2. D. Xu, C. Wei, P. Peng, Q. Xuan, H. Guo, Ge-gan: a novel deep learning framework for road traffic state estimation, *Transp. Res. Part C: Emerging Technol.*, **117** (2020), 102635. https://doi.org/10.1016/j.trc.2020.102635
3. S. Tikka, A. Hyttinen, J. Karvanen, Causal effect identification from multiple incomplete data sources: a general search-based approach, *J. Stat. Software*, **99** (2021), 1–40. https://doi.org/10.18637/jss.v099.i05
4. C. Hu, H. He, H. Jiang, Synchronization of complex-valued dynamic networks with intermittently adaptive coupling: a direct error method, *Automatica*, **112** (2020), 108675. https://doi.org/10.1016/j.automatica.2019.108675

5. Y. Jia, H. Wu, Global synchronization in finite time for fractional-order coupling complex dynamical networks with discontinuous dynamic nodes, *Neurocomputing*, **358** (2019), 20–32. https://doi.org/10.1016/j.neucom.2019.05.036

6. S. Hasan, S. V. Ukkusuri, Reconstructing activity location sequences from incomplete check-in data: a semi-markov continuous-time bayesian network model, *IEEE Trans. Intell. Transp. Syst.*, **19** (2017), 687–698. https://doi.org/10.1109/TITS.2017.2700481

7. D. J. Korchinski, J. G. Orlandi, S. W. Son, J. Davidsen, Criticality in spreading processes without timescale separation and the critical brain hypothesis, *Phys. Rev. X*, **11** (2021), 021059. https://doi.org/10.1103/PhysRevX.11.021059

8. P. Holme, J. Saramäki, Temporal networks, *Phys. Rep.*, **519** (2012), 97–125. https://doi.org/10.1016/j.physrep.2012.03.001

9. Y. Sun, S. Leng, Y. Lai, C. Grebogi, W. Lin, Closed-loop control of complex networks: a trade-off between time and energy, *Phys. Rev. Lett.*, **119** (2017), 198301. https://doi.org/10.1103/PhysRevLett.119.198301

10. Y. Zhang, X. Li, A. V. Vasilakos, Spectral analysis of epidemic thresholds of temporal networks, *IEEE Trans. Cybern.*, **50** (2020), 1965–1977. https://doi.org/10.1109/TCYB.2017.2743003

11. J. Hou, H. Ma, D. He, J. Sun, Q. Nie, W. Lin, Harvesting random embedding for high-frequency change-point detection in temporal complex systems, *Natl. Sci. Rev.*, **9** (2022), 228. https://doi.org/10.1093/nsr/nwab228

12. X. Gao, F. Shi, D. Shen, M. Liu, Task-induced pyramid and attention gan for multimodal brain image imputation and classification in alzheimers disease, *IEEE J. Biomed. Health. Inf.*, **26** (2022), 36–43. https://doi.org/10.1109/JBHI.2021.3097721

13. V. Indu, S. M. Thampi, A nature-inspired approach based on forest fire model for modeling rumor propagation in social networks, *J. Network Comput. Appl.*, **125** (2019), 28–41. https://doi.org/10.1016/j.jnca.2018.10.003

14. X. Liu, X. Zhu, M. Li, L. Wang, E. Zhu, T. Liu, et al., Multiple kernel k-means with incomplete kernels, *IEEE Trans. Pattern Anal. Mach. Intell.*, **42** (2020), 1191–1204. https://doi.org/10.1109/TPAMI.2019.2892416

15. C. Garcia, D. Leite, I. Škrjanc, Incremental missing-data imputation for evolving fuzzy granular prediction, *IEEE Trans. Fuzzy Syst.*, **28** (2020), 2348–2362. https://doi.org/10.1109/TFUZZ.2019.2935688

16. J. Venugopalan, N. Chanani, K. Maher, M. D. Wang, Novel data imputation for multiple types of missing data in intensive care units, *IEEE J. Biomed. Health. Inf.*, **23** (2019), 1243–1250. https://doi.org/10.1109/JBHI.2018.2883606

17. Y. Tian, K. Zhang, J. Li, X. Lin, B. Yang, Lstm-based traffic flow prediction with missing data, *Neurocomputing*, **318** (2018), 297–305. https://doi.org/10.1016/j.neucom.2018.08.067

18. Z. Cui, R. Ke, Z. Pu, Y. Wang, Stacked bidirectional and unidirectional lstm recurrent neural network for forecasting network-wide traffic state with missing values, *Transp. Res. Part C: Emerging Technol.*, **118** (2020), 102674. https://doi.org/10.1016/j.trc.2020.102674

19. Z. Peng, H. Liu, Y. Jia, J. Hou, Adaptive attribute and structure subspace clustering network, *IEEE Trans. Image Process.*, **31** (2022), 3430–3439. https://doi.org/10.1109/TIP.2022.3171421

20. K. Michalak, Low-dimensional euclidean embedding for visualization of search spaces in combinatorial optimization, *IEEE Trans. Evol. Comput.*, **23** (2019), 232–246. https://doi.org/10.1109/TEVC.2018.2846636

21. R. Tripathi, K. Rajawat, Adaptive network latency prediction from noisy measurements, *IEEE Trans. Netw. Serv. Manage.*, **18** (2021), 807–821. https://doi.org/10.1109/TNSM.2021.3051736

22. X. Zou, K. Li, C. Chen, Multilevel attention based u-shape graph neural network for point clouds learning, *IEEE Trans. Ind. Inf.*, **18** (2020), 448–456. https://doi.org/10.1109/TII.2020.3046627

23. S. Wang, G. Mao, Missing data estimation for traffic volume by searching an optimum closed cut in urban networks, *IEEE Trans. Intell. Transp. Syst.*, **20** (2019), 75–86. https://doi.org/10.1109/TITS.2018.2801808

24. L. Zhou, J. Zheng, Z. Ge, Z. Song, S. Shan, Multimode process monitoring based on switching autoregressive dynamic latent variable model, *IEEE Trans. Ind. Electron.*, **65** (2018), 8184–8194. https://doi.org/10.1109/TIE.2018.2803727

25. X. Hu, H. Zhang, D. Ma, R. Wang, A tngan-based leak detection method for pipeline network considering incomplete sensor data, *IEEE Trans. Instrum. Meas.*, **70** (2021), 1–10. https://doi.org/10.1109/TIM.2020.3045843

26. L. Xu, X. Zeng, W. Li, L. Bai, Idhashgan: deep hashing with generative adversarial nets for incomplete data retrieval, *IEEE Trans. Multimedia*, **24** (2022), 534–545. https://doi.org/10.1109/TMM.2021.3054503

27. T. Gao, G. Yan, Autonomous inference of complex network dynamics from incomplete and noisy data, *Nat. Comput. Sci.*, **2** (2022), 160–168. https://doi.org/10.1038/s43588-022-00217-0

28. Z. Dzunic, J. G. Chen, H. Mobahi, O. Büyüköztürk, J. W. Fisher III, A bayesian state-space approach for damage detection and classification, *Dyn. Civ. Struct.*, **2** (2015), 171–183. https://doi.org/10.1007/978-3-319-15248-6_18

29. D. Westreich, J. K. Edwards, S. R. Cole, R. W. Platt, S. L. Mumford, E. F. Schisterman, Imputation approaches for potential outcomes in causal inference, *Int. J. Epidemiol.*, **44** (2015), 1731–1737. https://doi.org/10.1093/ije/dyv135

30. Q. Zhu, K. Hou, Z. Chen, Z. Gao, Y. Xu, Y. He, Novel virtual sample generation using conditional gan for developing soft sensor with small data, *Eng. Appl. Artif. Intell.*, **106** (2021), 104497. https://doi.org/10.1016/j.engappai.2021.104497

31. F. Zhou, L. Li, K. Zhang, G. Trajcevski, Urban flow prediction with spatial–temporal neural odes, *Transp. Res. Part C: Emerging Technol.*, **124** (2021), 102912. https://doi.org/10.1016/j.trc.2020.102912

32. A. Rahman, V. Srikumar, A. D. Smith, Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks, *Appl. Energy*, **212** (2018), 372–385. https://doi.org/10.1016/j.apenergy.2017.12.051

33. J. Yang, Z. Peng, L. Lin, Real-time spatiotemporal prediction and imputation of traffic status based on lstm and graph laplacian regularized matrix factorization, *Transp. Res. Part C: Emerging Technol.*, **129** (2021), 103228. https://doi.org/10.1016/j.trc.2021.103228

34. M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.*, **378** (2019), 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

35. P. B. Weerakody, K. W. Wong, G. Wang, W. Ela, A review of irregular time series data handling with gated recurrent neural networks, *Neurocomputing*, **441** (2021), 161–178. https://doi.org/10.1016/j.neucom.2021.02.046

36. H. F. Yu, N. Rao, I. S. Dhillon, Temporal regularized matrix factorization for high-dimensional time series prediction, in *30th Conference on Neural Information Processing System*, 2016. Available from: https://proceedings.neurips.cc/paper/2016/file/85422afb467e9456013a2a51d4dff702-Paper.pdf.

37. X. Liu, M. Li, C. Tang, J. Xia, J. Xiong, L. Liu, et al., Efficient and effective regularized incomplete multi-view clustering, *IEEE Trans. Pattern Anal. Mach. Intell.*, **43** (2020), 2634–2646. https://doi.org/10.1109/TPAMI.2020.2974828

38. C. A. Mancuso, J. L. Canfield, D. Singla, A. Krishnan, A flexible, interpretable, and accurate approach for imputing the expression of unmeasured genes, *Nucleic Acids Res.*, **48** (2020), e125. https://doi.org/10.1093/nar/gkaa881

39. H. J. Gunn, P. Hayati Rezvan, M. I. Fernández, W. S. Comulada, How to apply variable selection machine learning algorithms with multiply imputed data: a missing discussion, *Psychol. Methods*, 2022. https://doi.org/10.1037/met0000478

40. J. Zhao, L. Chen, W. Pedrycz, W. Wang, A novel semi-supervised sparse bayesian regression based on variational inference for industrial datasets with incomplete outputs, *IEEE Trans. Syst. Man Cybern.: Syst.*, **50** (2020), 4773–4786. https://doi.org/10.1109/TSMC.2018.2864752

41. M. Benjumeda, S. Luengo-Sanchez, P. Larrañaga, C. Bielza, Tractable learning of bayesian networks from partially observed data, *Pattern Recognit.*, **91** (2019), 190–199. https://doi.org/10.1016/j.patcog.2019.02.025

42. C. Ye, H. Wang, W. Lu, J. Li, Effective bayesian-network-based missing value imputation enhanced by crowdsourcing, *Knowledge-Based Syst.*, **190** (2020), 105199. https://doi.org/10.1016/j.knosys.2019.105199

43. K. Ray, A. van der Vaart, Semiparametric bayesian causal inference, *Ann.Stat.*, **48** (2020), 2999–3020. https://doi.org/10.1214/19-AOS1919

44. H. Nguyen, E. Gouno, Bayesian inference for common cause failure rate based on causal inference with missing data, *Reliab. Eng. Syst. Saf.*, **197** (2020), 106789. https://doi.org/10.1016/j.ress.2019.106789

45. S. Athey, M. Bayati, N. Doudchenko, G. Imbens, K. Khosravi, Matrix completion methods for causal panel data models, *J. Am. Stat. Assoc.*, **116** (2021), 1716–1730. https://doi.org/10.1080/01621459.2021.1891924

46. J. G. Richens, C. M. Lee, S. Johri, Improving the accuracy of medical diagnosis with causal machine learning, *Nat. Commun.*, **11** (2020), 3923. https://doi.org/10.1038/s41467-020-17419-7

47. S. Leng, H. Ma, J. Kurths, Y. Lai, W. Lin, K. Aihara, et al., Partial cross mapping eliminates indirect causal influences, *Nat. Commun.*, **11** (2020), 2632. https://doi.org/10.1038/s41467-020-16238-0

48. B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, et al., Toward causal representation learning, *Proc. IEEE*, **109** (2021), 612–634. https://doi.org/10.1109/JPROC.2021.3058954

49. C. Zhang, J. Wang, G. G. Yen, C. Zhao, Q. Sun, Y. Tang, et al., When autonomous systems meet accuracy and transferability through ai: a survey, *Patterns*, **1** (2020), 100050. https://doi.org/10.1016/j.patter.2020.100050

50. R. T. Chen, Y. Rubanova, J. Bettencourt, D. K. Duvenaud, Neural ordinary differential equations, in *32nd Conference on Neural Information Processing Systems*, 2018. Available from: https://proceedings.neurips.cc/paper/2018/file/69386f6bb1dfed68692a24c8686939b9-Paper.pdf.

51. C. Wang, Y. C. Eldar, Y. Lu, Subspace estimation from incomplete observations: a high-dimensional analysis, *IEEE J. Sel. Top. Signal Process.*, **12** (2018), 1240–1252. https://doi.org/10.1109/JSTSP.2018.2877405

52. C. Yildiz, M. Heinonen, H. Lahdesmaki, Ode(2)vae: deep generative second order odes with bayesian neural networks, in *33rd Conference on Neural Information Processing System*, 2019. Available from: https://proceedings.neurips.cc/paper/2019/file/99a401435dcb65c4008d3ad22c8cdad0-Paper.pdf.

53. H. Turabieh, A. A. Salem, N. Abu-El-Rub, Dynamic l-rnn recovery of missing data in iomt applications, *Future Gener. Comput. Syst.*, **89** (2018), 575–583. https://doi.org/10.1016/j.future.2018.07.006

54. I. Izonin, R. Tkachenko, V. Verhun, K. Zub, An approach towards missing data management using improved grnn-sgtm ensemble method, *Eng. Sci. Technol. Int. J.*, **24** (2021), 749–759. https://doi.org/10.1016/j.jestch.2020.10.005

55. J. Zhang, P. Yin, Multivariate time series missing data imputation using recurrent denoising autoencoder, in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, (2019), 760–764. https://doi.org/10.1109/BIBM47256.2019.8982996

56. M. De Domenico, A. Lancichinetti, A. Arenas, M. Rosvall, Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems, *Phys. Rev. X*, **5** (2015), 011027. https://doi.org/10.1103/PhysRevX.5.011027

57. S. Gupta, A. K. Sahoo, U. K. Sahoo, Wireless sensor network-based distributed approach to identify spatio-temporal volterra model for industrial distributed parameter systems, *IEEE Trans. Ind. Inf.*, **16** (2020), 7671–7681. https://doi.org/10.1109/TII.2020.3004159

58. T. Hiraoka, N. Masuda, A. Li, H. H. Jo, Modeling temporal networks with bursty activity patterns of nodes and links, *Phys. Rev. Res.*, **2** (2020), 023073. https://doi.org/10.1103/PhysRevResearch.2.023073

59. J. Zhao, H. He, X. Zhao, J. Lin, Modeling and simulation of microblog-based public health emergency-associated public opinion communication, *Inf. Process. Manage.*, **59** (2022), 102846. https://doi.org/10.1016/j.ipm.2021.102846

60. F. Alesiani, L. Moreira-Matias, M. Faizrahnemoon, On learning from inaccurate and incomplete traffic flow data, *IEEE Trans. Intell. Transp. Syst.*, **19** (2018), 3698–3708. https://doi.org/10.1109/TITS.2018.2857622

61. Y. Jiang, A. K. Srivastava, Data-driven event diagnosis in transmission systems with incomplete and conflicting alarms given sensor malfunctions, *IEEE Trans. Power Delivery*, **35** (2020), 214–225. https://doi.org/10.1109/TPWRD.2019.2947671

62. G. Mei, X. Wu, Y. Wang, M. Hu, J. A. Lu, G. Chen, Compressive-sensing-based structure identification for multilayer networks, *IEEE Trans. Cybern.*, **48** (2018), 754–764. https://doi.org/10.1109/TCYB.2017.2655511

63. S. Manfredi, E. Di Tucci, V. Latora, Mobility and congestion in dynamical multilayer networks with finite storage capacity, *Phys. Rev. Lett.*, **120** (2018), 068301. https://doi.org/10.1103/PhysRevLett.120.068301

64. L. Qiu, S. Liu, C-siw rumor propagation model with variable propagation rate and perception mechanism in social networks, *Discrete Dyn. Nat. Soc.*, **2020** (2020), 5712968. https://doi.org/10.1155/2020/5712968

65. M. Liang, J. Du, C. Yang, Z. Xue, H. Li, F. Kou, et al., Cross-media semantic correlation learning based on deep hash network and semantic expansion for social network cross-media search, *IEEE Trans. Neural Networks Learn. Syst.*, **31** (2020), 3634–3648. https://doi.org/10.1109/TNNLS.2019.2945567

66. H. Sun, D. Saad, A. Y. Lokhov, Competition, collaboration, and optimization in multiple interacting spreading processes, *Phys. Rev. X*, **11** (2021), 011048. https://doi.org/10.1103/PhysRevX.11.011048

67. X. Zhan, C. Liu, G. Zhou, Z. Zhang, G. Sun, J. J. Zhu, et al., Coupling dynamics of epidemic spreading and information diffusion on complex networks, *Appl. Math. Comput.*, **332** (2018), 437–448. https://doi.org/10.1016/j.amc.2018.03.050

68. E. J. Müller, B. R. Munn, J. M. Shine, Diffuse neural coupling mediates complex network dynamics through the formation of quasi-critical brain states, *Nat. Commun.*, 2020. https://doi.org/10.1101/2020.06.09.141416

69. Y. Wang, Q. Lu, X. Cao, X. Zhou, V. Latora, L. C. Tong, et al., Travel time analysis in the chinese coupled aviation and high-speed rail network, *Chaos, Solitons Fractals*, **139** (2020), 109973. https://doi.org/10.1016/j.chaos.2020.109973

70. W. Lin, H. Ma, Synchronization between adaptively coupled systems with discrete and distributed time-delays, *IEEE Trans. Autom. Control*, **55** (2010), 819–830. https://doi.org/10.1109/TAC.2010.2041993

71. M. De Domenico, C. Granell, M. A. Porter, A. Arenas, The physics of spreading processes in multilayer networks, *Nat. Phys.*, **12** (2016), 901–906. https://doi.org/10.1038/nphys3865

72. Z. Wang, M. A. Andrews, Z. X. Wu, L. Wang, C. T. Bauch, Coupled disease–behavior dynamics on complex networks: a review, *Phys. Life Rev.*, **15** (2015), 1–29. https://doi.org/10.1016/j.plrev.2015.07.006

73. K. Chen, W. He, Q. Han, M. Xue, Y. Tang, Leader selection in networks under switching topologies with antagonistic interactions, *Automatica*, **142** (2022), 110334. https://doi.org/10.1016/j.automatica.2022.110334

74. R. Sardinha, A. Paes, G. Zaverucha, Revising the structure of bayesian network classifiers in the presence of missing data, *Inf. Sci.*, **439** (2018), 108–124. https://doi.org/10.1016/j.ins.2018.02.011

75. Z. Z. Alp, Ş. G. Öğüdücü, Influence factorization for identifying authorities in twitter, *Knowledge-Based Syst.*, **163** (2019), 944–954. https://doi.org/10.1016/j.knosys.2018.10.020

76. A. Namtirtha, A. Dutta, B. Dutta, Weighted kshell degree neighborhood: a new method for identifying the influential spreaders from a variety of complex network connectivity structures, *Expert Syst. Appl.*, **139** (2020), 112859. https://doi.org/10.1016/j.eswa.2019.112859

77. A. Namtirtha, A. Dutta, B. Dutta, A. Sundararajan, Y. Simmhan, Best influential spreaders identification using network global structural properties, *Sci. Rep.*, **11** (2021), 2254. https://doi.org/10.1038/s41598-021-81614-9

78. M. Xu, R. Li, F. Li, Phase identification with incomplete data, *IEEE Trans. Smart Grid*, **9** (2018), 2777–2785. https://doi.org/10.1109/TSG.2016.2619264

79. X. Li, X. Li, Reconstruction of stochastic temporal networks through diffusive arrival times, *Nat. Commun.*, **8** (2017), 15729. https://doi.org/10.1038/ncomms15729

80. W. Cheng, Y. Wang, H. Li, Y. Duan, Learned full-sampling reconstruction from incomplete data, *IEEE Trans. Comput. Imaging*, **6** (2020), 945–957. https://doi.org/10.1109/TCI.2020.2996751

81. J. Fu, J. Dong, F. Zhao, A deep learning reconstruction framework for differential phase-contrast computed tomography with incomplete data, *IEEE Trans. Image Process.*, **29** (2020), 2190–2202. https://doi.org/10.1109/TIP.2019.2947790

82. Y. Tang, C. Zhao, J. Wang, C. Zhang, Q. Sun, W. Zheng, et al., An overview of perception and decision-making in autonomous systems in the era of learning: a survey, preprint, arXiv:2001.02319.

83. Y. Tang, X. Wu, P. Shi, F. Qian, Input-to-state stability for nonlinear systems with stochastic impulses, *Automatica*, **113** (2020), 108766. https://doi.org/10.1016/j.automatica.2019.108766

84. Y. Tang, D. Zhang, P. Shi, W. Zhang, F. Qian, Event-based formation control for nonlinear multiagent systems under dos attacks, *IEEE Trans. Autom. Control*, **66** (2021), 452–459. https://doi.org/10.1109/TAC.2020.2979936

85. T. Yabe, P. S. C. Rao, S. V. Ukkusuri, S. L. Cutter, Toward data-driven, dynamical complex systems approaches to disaster resilience, *PNAS*, **119** (2022), e2111997119. https://doi.org/10.1073/pnas.2111997119

86. M. M. Danziger, A. L. Barabási, Recovery coupling in multilayer networks, *Nat. Commun.*, **13** (2022), 955. https://doi.org/10.1038/s41467-022-28379-5

87. H. Sanhedrai, J. Gao, A. Bashan, M. Schwartz, S. Havlin, B. Barzel, Reviving a failed network through microscopic interventions, *Nat. Phys.*, **18** (2022), 338–349. https://doi.org/10.1038/s41567-021-01474-y

88. I. Izonin, R. Tkachenko, N. Kryvinska, K. Zub, O. Mishchuk, T. Lisovych, Recovery of incomplete iot sensed data using high-performance extended-input neural-like structure, *Procedia Comput. Sci.*, **160** (2019), 521–526. https://doi.org/10.1016/j.procs.2019.11.054

89. X. Feng, H. Zhang, C. Wang, H. Zheng, Traffic data recovery from corrupted and incomplete observations via spatial-temporal trpca, *IEEE Trans. Intell. Transp. Syst.*, **2022** (2022), 1–14. https://doi.org/10.1109/TITS.2022.3151925

90. R. Wu, L. Jiang, Recovering dynamic networks in big static datasets, *Phys. Rep.*, **912** (2021), 1–57. https://doi.org/10.1016/j.physrep.2021.01.003

91. J. T. Davis, N. Perra, Q. Zhang, Y. Moreno, A. Vespignani, Phase transitions in information spreading on structured populations, *Nat. Phys.*, **16** (2020), 590–596. https://doi.org/10.1038/s41567-020-0810-3

92. X. Wang, Y. Lan, J. Xiao, Anomalous structure and dynamics in news diffusion among heterogeneous individuals, *Nat. Hum. Behav.*, **3** (2019), 709–718. https://doi.org/10.1038/s41562-019-0605-7

93. D. Guilbeault, D. Centola, Topological measures for identifying and predicting the spread of complex contagions, *Nat. Commun.*, **12** (2021), 4430. https://doi.org/10.1038/s41467-021-24704-6

94. S. Contreras, J. Dehning, M. Loidolt, J. Zierenberg, F. P. Spitzner, J. H. Urrea-Quintero, et al., The challenges of containing sars-cov-2 via test-trace-and-isolate, *Nat. Commun.*, **12** (2021), 378. https://doi.org/10.1038/s41467-020-20699-8

95. L. Z. Wang, Z. D. Zhao, J. Jiang, B. H. Guo, X. Wang, Z. G. Huang, et al., A model for meme popularity growth in social networking systems based on biological principle and human interest dynamics, *Chaos*, **29** (2019), 023136. https://doi.org/10.1063/1.5085009

96. C. Murphy, E. Laurence, A. Allard, Deep learning of contagion dynamics on complex networks, *Nat. Commun.*, **12** (2021), 4720. https://doi.org/10.1038/s41467-021-24732-2

97. T. M. Bury, R. Sujith, I. Pavithran, M. Scheffer, T. M. Lenton, M. Anand, et al., Deep learning for early warning signals of tipping points, *PNAS*, **118** (2021), e2106140118. https://doi.org/10.1073/pnas.2106140118

98. Q. Ni, J. Kang, M. Tang, Y. Liu, Y. Zou, Learning epidemic threshold in complex networks by convolutional neural network, *Chaos*, **29** (2019), 113106. https://doi.org/10.1063/1.5121401

99. J. Pathak, B. Hunt, M. Girvan, Z. Lu, E. Ott, Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach, *Phys. Rev. Lett.*, **120** (2018), 024102. https://doi.org/10.1103/PhysRevLett.120.024102

100. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, et al., Networks beyond pairwise interactions: structure and dynamics, *Phys. Rep.*, **874** (2020), 1–92. https://doi.org/10.1016/j.physrep.2020.05.004

101. J. Wang, Y. Hong, J. Wang, J. Xu, Y. Tang, Q. Han, et al., Cooperative and competitive multi-agent systems: from optimization to games, *IEEE/CAA J. Autom. Sin.*, **9** (2022), 763–783. https://doi.org/10.1109/JAS.2022.105506

AIMS Press