Graph Neural Networks for Anomaly Detection in Industrial Internet of Things

Yulei Wu, *Senior Member, IEEE*, Hong-Ning Dai, *Senior Member, IEEE*, and Haina Tang

***Abstract*—The Industrial Internet of Things (IIoT) plays an important role in digital transformation of traditional industries towards Industry 4.0. By connecting sensors, instruments and other industry devices to the Internet, IIoT facilitates the data collection, data analysis, and automated control, thereby improving the productivity and efficiency of the business as well as the resulting economic benefits. Due to the complex IIoT infrastructure, anomaly detection becomes an important tool to ensure the success of IIoT. Due to the nature of IIoT, graph-level anomaly detection has been a promising means to detect and predict anomalies in many different domains such as transportation, energy and factory, as well as for dynamically evolving networks. This paper provides a useful investigation on graph neural networks (GNN) for anomaly detection in IIoT-enabled smart transportation, smart energy and smart factory. In addition to the GNN-empowered anomaly detection solutions on point, contextual, and collective types of anomalies, useful datasets, challenges and open issues for each type of anomalies in the three identified industry sectors (i.e., smart transportation, smart energy and smart factory) are also provided and discussed, which will be useful for future research in this area. To demonstrate the use of GNN in concrete scenarios, we show three case studies in smart transportation, smart energy, and smart factory, respectively.**

***Index Terms*—Industrial internet of things, Graph neural networks, Anomaly detection, Industry 4.0**

I. Introduction

Many traditional industries that were isolated from the public access, have started their digital transformation under the umbrella of Industry 4.0 [1]–[3]. Such industries include energy, health, manufacturing, water, just to name a few. The Industrial Internet of Things (IIoT) is an emerging paradigm that facilitates the process of industry digital transformation [4], [5]. It allows the networked interconnection of sensors, instruments, and other Internet of Things (IoT) devices to enable data collection, data analytics and automated control in an industry environment [6]–[9]. IIoT can in turn improve the productivity and efficiency of the business in the industry, as well as the resulting economic benefits. A recent report of "Bringing Smart Water Networks Into Focus" mentioned that up to 12.5 billion USD in annual savings can be achieved from a combination of actions ranging from *improved leakage and pressure management* to *streamlined water quality monitoring*[1]. These actions can attribute to the development of smart water systems that can be implemented by IIoT.

Anomaly detection is of paramount importance in IIoT [10]–[15]. On the one hand, traditional isolated industry systems are now exposed to the public access due to the introduction of IIoT [16]–[19]. The sensors, instruments, and other IoT devices that were originally designed with little security mechanisms can be easily compromised by malicious users like attackers [4], [16], [20]. Efficient detection of such anomalies is crucial to the security of IIoT and the success of related businesses. On the other hand, one of the benefits of digital transformation of many industries is to use the collected data to detect abnormal situations in a timely manner or even in advance of the actual happening of anomalies [21], [22]. The absence of appropriate anomaly detection may result in significant economic loss. For example, British Airways IT systems failure resulted in 58 million GBP in lost business and follow-up compensation claims[2]. Across England and Wales, nearly 3 billion litres of water is lost to leaks every day, resulting in considerable economic loss of water companies[3].

IIoT, depending on the type of devices and working conditions, may generate and collect a wide variety and large volume of data that can be used for anomaly detection [10], [23]. These data may include value, image, text, audio and video, and each device may generate and/or collect a combination of different types of data [24]–[26]. Artificial intelligence (AI), especially machine learning and deep learning, has been widely adopted for anomaly detection in a wide range of anomaly detection tasks [27]–[35]. Given the nature of IIoT, where devices are interconnected and the interconnection evolves as shown in Fig. 1, graph-level anomaly detection has been a promising means to detect anomalies in many different domains such as transportation and energy as well as for dynamically evolving networks.

Graph neural networks (GNN) [36]–[39] have recently been fast developed to model complex patterns in graph-structured data, and is a promising graph-level paradigm to carry out anomaly detection. In general, the state of a node is influenced

This work was partially supported by the Engineering and Physical Sciences Research Council (EPSRC) under Grant No. EP/R030863/1, the Macao Science and Technology Development Fund under Macao Funding Scheme for Key R & D Projects under Grant No. 0025/2019/AKP, the Open Fund of Zhejiang Lab under Grant No. 2019KE0AB03, and the National Natural Science Foundation of China (NSFC) under Grant No. 52071312. (Corresponding author: Yulei Wu)

Y. Wu is with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, EX4 4QF, U.K. e-mail: y.l.wu@exeter.ac.uk

H.-N. Dai is with Faculty of Information Technology, Macau University of Science and Technology, Macau. email: hndai@ieee.org

H. Tang is with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, 100049, China. e-mail: hntang@ucas.ac.cn

Manuscript received xxx; revised xxx.

[1] https://www.actu-environnement.com/media/pdf/dossiers/831-sensus-smart-water.pdf

[2] https://www.computerweekly.com/news/252468002/BA-IT-systems-failure-results-in-cancelled-flights-and-delays-at-London-airports

[3] https://www.bbc.co.uk/news/business-53274914

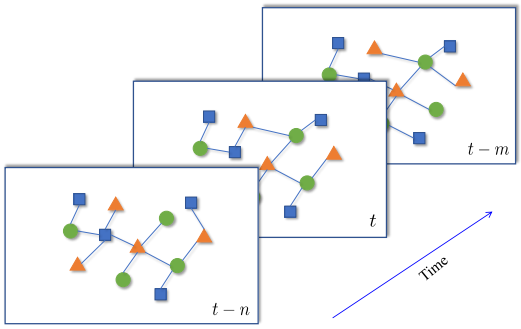

Fig. 1. The evolving interconnection of graph data in IIoT

by the states of its neighbors. Specifically, each node in GNN essentially aggregates the features of its neighboring nodes and works out its own feature representation. According to different variants of GNN, their neighboring nodes that are considered in the process of feature aggregation are different. Graph convolutional networks (GCN) consider the one-step neighbors of a node when calculating its aggregated features [40], [41]. Graph attention networks (GAN) adopt an attention function so that neighboring nodes can be assigned with different weights when being aggregated [42]. In order to maximize the effects of GNN on anomaly detection in different IIoT application areas, the unique characteristics of the IIoT applications need to be considered in the GNN model.

There are few existing studies investigating GNN-empowered anomaly detection solutions, and most of them are designed for a general scenario without an explicit consideration of specific industrial applications, such as [43]. The main contributions of this paper can be summarized as follows:

- We provide a deeper understanding of the three types of anomalies, i.e., point, contextual, and collective anomalies, in the context of specific IIoT applications.
- We explicitly bring the suitable GNN-based solutions to the context of IIoT applications, in terms of smart transportation, smart energy, and smart factory. The discussion and analysis can be treated as the basis for the further investigation and research in this area.
- We provide the useful public datasets for each type of anomalies in each of the above IIoT applications, thereby facilitating the future research.
- The research challenges and open issues of GNN-empowered anomaly detection for IIoT-enabled smart transportation, smart energy, and smart factory have been identified. This can be hopefully useful for the future research in advancing anomaly detection in IIoT.
- To demonstrate the use of GNN in concrete scenarios, we show three case studies in IIoT-enabled smart transportation, smart energy, and smart factory, respectively.

The remaining of this paper is organized as follows. Section II illustrates IIoT and the potential types of anomalies in IIoT. Section III summarizes some of the important GNN models. Sections IV – VI provide the useful public datasets, the GNN-based anomaly detection solutions, and research challenges and open issues for IIoT-enabled smart transportation, smart energy and smart factory, respectively. Section VII shows case studies of using GNN in the above three industry areas. Finally, Section VIII concludes this paper.

II. Industrial Internet of Things

In this section, IIoT will be first introduced, followed by the explanation of three types of anomalies in IIoT.

A. Industrial Internet of Things

IIoT is a similar term to IoT but has different focus. Different from IoT that focuses on consumer devices, IIoT has its focus on industrial purpose with more sophisticated and critical devices that are usually used in high-stakes industries such as water, health, energy and defence [44]–[46]. Essentially, IIoT provides an infrastructure that connects the industrial sensors, instruments and other industry devices through wired and/or wireless networks (e.g., industrial Ethernet, WiFi, 5G [47], and the future networks like 6G [48]), and enables data generation and collection through these connected devices [49]. More importantly, IIoT uses the generated and collected data to carry out effective data analytics (e.g., using AI techniques) and make useful decisions that can automate business operations and increase industry profits [50]. Among these, data analytics is of paramount importance [51]. The CBI Ministry of Foreign Affairs indicated in its recent report that the new 5G technology will allow more than 350,000 devices to be connected per square kilometre - 500 times more than comparable existing technologies[4]. This in turn significantly challenges the techniques of data analytics for IIoT.

IIoT has empowered multiple industrial applications, as shown in Fig. 2. Take smart energy as an example. Smart energy systems consist of energy generation, distribution, and energy consumption. Various energy sources include fuel power plants, nuclear power stations, hydro, wind, and solar energies. Energy has typically been transmitted in electricity via power lines from energy sources to substations, which then transmit and distribute electricity to various electricity customers, including commercial, industrial, and residential customers. It is challenging to assure the reliability of the complex energy systems. The proliferation of IIoT nodes and wireless/wired communication technologies brings opportunities to address this challenge, since IoT nodes, sensors, and smart meters can sense ambience and send the data to the control center. However, the sensory data can be tampered or wiretapped during the transmission. In addition, faults/errors that occur at the energy systems need to be detected and reported to the control center, which then takes immediate actions. Similarly, faults and anomalies should be detected at IIoT-enabled smart transportation and IIoT-enabled smart factory/manufacturing, as shown in Fig. 2.

B. Types of anomalies in IIoT

Efficient anomaly detection is a crucial factor to ensure the success of IIoT. In this section, the possible types of anomalies in IIoT will be presented [52].

[4]https://www.cbi.eu/market-information/outsourcing-itobpo/industrial-internet-things/market-potential

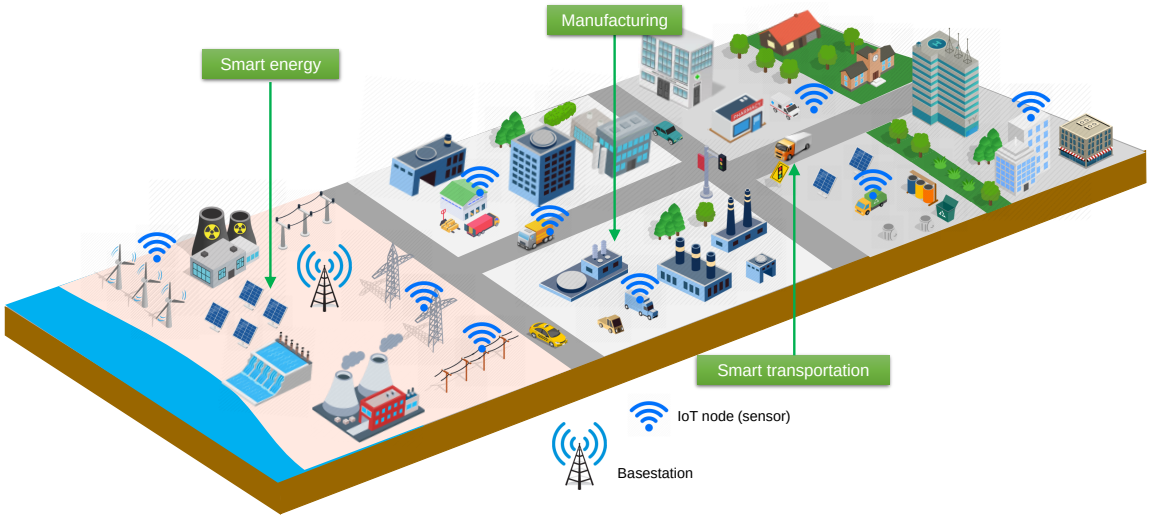

Fig. 2. Industrial Internet of Things with its applications

1) Point: Point anomalies often refer to an irregularity that happens randomly and may have no particular reason. For instance, an egress port of a router/switch in an industrial network with an instantaneous high volume of traffic seems a point anomaly since it significantly deviates from the normal volume of traffic of this port.

2) Contextual: A contextual anomaly (a.k.a conditional anomaly) represents an abnormal behavior happening within some specific context. This type of anomalies can be identified by considering both contextual and behavioural features. Time and space are usually used as the contextual features. The behavioral features may be a pattern of network traffic, e.g., the network traffic of an industry office at weekdays is usually much higher than that at weekends.

3) Collective: A collection of individual data points showing anomalies can be treated as collective anomalies. In this type of anomalies, each individual data point in isolation appears as normal data instances while observed in a group exhibit unusual behaviors. For example, the past five days at 1 am the network traffic of an industry production line is slightly higher than normal; this seems to be a potential case for collective anomalies. In addition, collective anomalies may also happen when the program of IIoT devices are patched/upgraded, but the controller that manages the behaviour of these devices is not upgraded.

III. Graph Neural Networks

GNN is a more generalized CNN. CNN can only handle the data with regular (Euclidean) structures such as 2-dimensional images and 1-dimensional text data, while GNN can process non-Euclidean data such as social media networks, 3-dimensional images, telecom networks, and the data in many industry settings [53], [54]. GNN propagates the node states in an iterative manner until reaching equilibrium, using a neural network. It outputs the state representation for each node. Similar to the basic graph theory, one of the important questions in GNN is to identify which parts of the data are nodes and which parts are edges. Then, the graph needs to be translated into the features for neural networks. Essentially, each node in GNN aggregates the features of its neighboring nodes and works out its own feature representation. In recent years, different variants of GNNs are being developed. Readers can refer to some survey papers that are dedicated for GNN for more details [53], [55]–[58]. As the main contribution of our paper focuses on the application of GNN on the anomaly detection in different industry sectors, the specific details of GNN models will not be included. In this section, we will briefly review some important variants of GNN that would be useful for facilitating the understanding of the rest of this paper.

A. Graph Convolutional Networks

GCN is considered as one of the basic variants of GNN, and thus it shares some key features of GNN such as working with non-Euclidean data. The way of how *convolution* works in GCN is the same as that in CNN, where input neurons are multiplied by a set of weights (kernels). There are basically two types of GCN: *spectral GCN* and *spatial GCN*. The spectral GCN can be treated as a message passing along the nodes within the graph. The convolutional operation calculates the eigen-decomposition of the graph Laplacian that helps to understand the graph structure [40]. It considers both node features and nodes connectivity as input features so that the model can learn the features of neighboring nodes.

As GCN does not consider the ordering of node neighbors, it cannot handle the graph data with such features, e.g., some IIoT datasets like smart factory data have geometric interpretation of the graph that shows an order according to their spatial positions. To address this issue, the spatial GCN was developed [59]. It uses the spatial features of nodes to aggregate information from the neighboring nodes. Graph SAmple and aggreGatE (GraphSAGE) [60] worked out the feature representation of a node by aggregating the features over a fixed-size neighboring nodes in an inductive manner. It works well on large-scale inductive benchmarks. In addition to Spatial GCN, Spatial-Temporal GCN (STGCN) can

TABLE I
TYPES OF GNNS AND THEIR APPLICATIONS IN ANOMALY DETECTION

Types of GNNs	Key features	Application domains	Types of anomalies
Spectral GCN	The basic variant of GNN; work with non-Euclidean data; consider both node features and node connectivity; learn the features of neighboring nodes	Applicable to all IIoT applications including smart transportation, smart energy, and smart factory	Point, contextual and collective anomalies
Spatial GCN	In addition to the features of spectral GCN, consider the ordering of node neighbors		
STGCN	Can capture correlations of spatial and temporal features		
GAN	Each neighboring nodes can contribute differently (different weights) when calculating the aggregated features of the central node		
GGNN/GGSNN	Can produce a sequence of outputs		
GRNN	Can analyze dynamic graphs		
Jump Knowledge Networks	Can leverage different neighborhood ranges (different number of neighboring nodes) when calculating the aggregated features of a central node		
Self-enhanced GNN	Considering and improving the quality of input data		

characterize correlations of both spatial and temporal features, thereby being used in smart transportation, smart energy and smart factory.

B. Graph Attention Networks

GCN works for the situation where a node has the same weights to all its neighboring nodes, i.e., each neighboring node contributes equally to the calculation of the feature representation of the central node. However, there are certain cases, where assigning different importance to each neighbor's contribution is needed when calculating the aggregated features of the central node. Such cases include smart grid, IoT and some biological networks. GAN [42] was then developed to address this limitation of GCN. In this work, the model can be applicable to inductive learning problems where the model can generalize to unseen graphs.

C. Gated Graph (Sequence) Neural Networks

GCN and GAN focus on models that produce a single output such as classification. However, many real-world problems require outputting a sequence, such as network planning [61] and virtual network embedding [62]. To address this issue, gate mechanisms like gated recurrent unit (GRU) [63] and long short-term memory (LSTM) [64] are used in the propagation step to improve the long-term propagation of information in the graph. Li et al. [37] developed gated graph neural networks (GGNN) where they used GRU and unrolled the recurrence for a fixed number of steps T, and used backpropagation through time to compute gradients. Gated graph sequence neural networks (GGSNN) used several GGNNs operating in sequence to produce an output sequence. Besides GGNN, Graph Recurrent Neural Networks (GRNN) [65] was also developed to analyze dynamic graphs.

D. Jump Knowledge Networks

Many GNN models like GCN, GAN and GraphSAGE use the fixed number of neighboring nodes to aggregate the representation of a central node, either in one-hop neighborhood or multiple hops. This may not work well in certain circumstances depending on the graph structure. Xu et al. [66] explored the jumping knowledge networks to develop a representation learning on graphs, where each node can flexibly leverage different neighborhood ranges (i.e., different number of neighboring nodes) to enable better structure-aware representation.

E. Self-Enhanced GNN

Most of GNN models focus on developing effective models without considering the quality of the input data. To this end, Yang et al. [67] proposed a self-enhanced GNN model to improve the quality of the input data using the outputs of existing GNN models for enhanced performance on semi-supervised classification problems, hence being named *self-enhanced*.

In summary, Table I summarizes major types of GNNs and their IIoT-applications in anomaly detection. It is worth mentioning that GNNs can be used for the three types anomalies, namely point, contextual, and collective anomalies.

IV. IIOT-ENABLED SMART TRANSPORTATION

We are experiencing the urbanization and the proliferation of various vehicles. It has been a hot topic to implement highly-efficient intelligent transportation systems (ITS) and develop autonomous vehicles (AVs), consequently realizing smart transportation [68]–[72]. Meanwhile, diverse traffic sensors, Global Navigation Satellite System (GNSS), radar, loop detectors, light detection and ranging (LIDAR), and Infrared (IR) cameras deployed at ITS and AVS also generate/collect massive diverse types of traffic data, which can be used to analyze the traffic status and identify possible faults in transportation systems. As shown in Fig. 3, the traffic anomalies include abnormal traffic congestion, traffic accident, damaged traffic infrastructure (e.g., a road).

TABLE II
A SUMMARY OF THE GNN-BASED ANOMALY DETECTION SOLUTIONS IN IIOT-ENABLED SMART TRANSPORTATION.

Type of anomalies	Public datasets/software	GNN-based solutions
Point anomaly	UCSD datase, U-turn dataset	[73] [74]
Contextual anomaly	NYC taxi dataset, Uber movement dataset, Chicago taxi dataset	[75] [76] [77]
Collective anomaly	Dataset based on NYC taxi dataset and NYC Bike dataset, UMN daset	[78] [79] [74] [80] [81]

Fig. 3. Anomalies in IIoT-enabled smart transportation

Traffic anomaly detection plays a critical role in fostering the urban ITS [82]. In particular, traffic anomalies are often indicators of traffic accidents, traffic congestion, traffic violations, and damaged traffic infrastructure [83]. The detection of traffic anomalies can help to rescue casualties, restore traffic, and make an urgent repair on damaged roads [84]. Table II summarizes the GNN-based solutions for smart transportation. We next present a detailed investigation on traffic anomaly detection based on GNNs in smart transportation.

A. Point

Point anomalies correspond to abnormal events which deviate from normal distribution. For example, a broken-down vehicle in a highway can be regarded as a point anomaly.

1) Public datasets: There are some recent studies on collecting point-anomaly datasets and conducting analysis on these datasets. In particular, the work [85] presents a framework to detection anomalies in crowds and also releases a crowd dataset collected from University of California San Diego (UCSD) campus[5]. The UCSD dataset consists of 98 videos, each of which has 200 frames. Besides the abnormal behaviours of crowds, there is another study [86] conducting an anomaly detection on abnormal vehicle-moving patterns on top of the U-turn dataset[6]. The work [87] also summarizes other datasets for anomaly detection based on visual analysis.

2) GNN solutions to anomaly detection: There are some recent studies in addressing anomaly detection in traffic and crowds. In particular, the work [73] presents a GCN-based method for anomaly detection based on noisy labels. Specifically, GCN was used to remove noises from the labelled data. The training process of this model is composed of cleaning and classification. The cleaning phase mainly aims to clean the noisy labels while the classification is to retrain the classifier based on the corrected labels. Experimental results on UCSD dataset [85] as well as other datasets demonstrate the effectiveness of the proposed method. Meanwhile, Yu et al. [74] propose a Deep STGCN to predict traffic accidents. The proposed STGCN consists of a spatial GCN layer and a combined GCN-CNN spatial-temporal layer to extract the correlated spatial features and the spatial-temporal features, respectively. In addition, an embedding layer is also adopted to give a semantic representation of external features. They also conducted experiments on traffic datasets consisting of vehicle traffic, weather conditions, and accident reports. The experimental results further confirm the effectiveness of the proposed STGCN.

B. Contextual

Contextual anomalies depend on the available context attributes [88] in traffic data. For example, the event that a car is driving in a high speed in a highway can be regarded as a normal event while it can be regarded as an anomaly in a dense road.

1) Public datasets: There are a number of public traffic datasets including taxi datasets and ride-sharing services datasets. For example, the New York City (NYC) taxi dataset [89] contains records of time and locations of pick-up and drop-off of a taxi in NYC. The Uber Movement dataset[7] includes data samples, each of which contains the travelling time, a source, and a destination of Uber ride-sharing services across multiple cities in USA. In addition, Chicago taxi dataset[8] offers a similar dataset for taxi services in Chicago; this dataset consists of taxi trip records from Jan. 1 to April 30, 2019 in the City of Chicago. However, most of these datasets do not contain anomalies. Thus, many recent studies like [75] manually inject some anomalies into the datasets such as Uber Movement dataset, NYC taxi dataset, and Chicago taxi dataset after following a similar approach to [90]. Moreover, some studies such as [76] integrate events (e.g., accidents) into traffic datasets by crawling Microblog data[9].

[5]http://www.svcl.ucsd.edu/projects/anomaly/dataset.html
[6]https://sites.google.com/view/ybenezeth/cvpr2009

[7]https://movement.uber.com/
[8]https://data.cityofchicago.org/Transportation/TaxiTrips/wrvz-psew
[9]These datasets are available at https://github.com/zzyy0929/AAAI2020-RiskOracle/.

2) GNN solutions to anomaly detection: It is a challenging task to conduct anomaly detection in multi-dimensional traffic data. There are recent advances in addressing this issue. The work [75] presents a Context-augmented Graph Autoencoder (namely Con-GAE) for anomaly detection in city traffic. This framework exploits graph embedding as well as context embedding so as to extract spatial features from traffic data. Extensive experiments on several representative datasets, such as Uber Movement dataset, NYC taxi dataset, and Chicago taxi dataset demonstrate the superior performance of the proposed method over other baseline methods. As a type of traffic anomalies, accidents have also received extensive attention. The work [76] presents a Differential Time-varying Graph neural network (DTGN) to analyze traffic data and achieve a minute-level accident forecasting. The DTGN is essentially an extension from GCN by incorporating time-varying overall affinity and differential GCN. Experimental results validate the effectiveness of the proposed framework. Moreover, in [77], Zhou et al. propose a framework to predict traffic accidents after considering both spatial-temporal features of traffic data and context factors.

C. Collective

Collective anomaly refers to a situation that a collection of data samples is anomalous to normal values while each individual sample may be within a normal range. For example, the event that a fleet of vehicles is moving slowly in a high way can be regarded as a collective anomaly.

1) Public datasets: There are several public traffic datasets available for collective anomalies. In [78], the authors obtain a new anomaly dataset based on NYC taxi dataset [89] and NYC Bike dataset[10] after inserting a number of anomalies based on anomaly reports of NYC. In addition, the work [91] presents a study on detecting collective anomalies in crowds based on the University of Minnesota (UMN) dataset, which contains videos of escaping scenarios[11]. Moreover, the work [79] also conducts anomaly detection on a dataset obtained from social networks (e.g., traffic accident reports at Twitter), remote sensing dataset and vehicle accidents.

2) GNN solutions to anomaly detection: There are several studies toward addressing anomaly detection in traffic data. In particular, the work [78] presents a spatiotemporal multi-modal fusion model (ST-MFM) to extract features from multiple crowd-flow datasets and predict anomalies. In the proposed ST-MFM, a GCN was adopted to extract spatial features. It is worth mentioning that the authors also construct a new anomaly dataset based on anomaly reports, bicycle traffic, and taxi traffic. Extensive experiments also validate the effectiveness of the proposed model. Moreover, the work [79] presents a multi-modal graph neural network to forecast traffic risks. In this framework, GANs were leveraged to further improve the forecasting accuracy. Extensive experiments on a real-world dataset constructed from traffic accidents, social networks, and remote sensing imagery also demonstrate the superior performance of the proposed model.

Besides traffic accidents, extreme weather events also affect the traffic. In [80], Wang et al. investigate the transportation resilience under extreme weather events based GNNs. In particular, they propose a graph convolutional recurrent neural network (GCRNN) to predict the traffic patterns under extreme weather events. Moreover, the authors also conduct experiments on DiDi Chuxing, i.e., an on-demand riding service in China. The experimental results also demonstrate the effectiveness of the proposed method.

In addition to vehicular transportation, the railway delay analysis based on STGCN was also conducted in [81]. In particular, the authors in this paper adopt STGCN to predict cascading delays in the British railway, where cascading delays are also regarded as anomalies since they are often deviated from normal operations of railways. Although the authors do not provide the dataset for their experiments, they explicitly describe the data collection approach to crawling the railway data from National Rail Enquiries Data Feeds[12]. Experimental results also demonstrate the superior performance than other statistical methods.

D. Challenges and Open Issues

Although GNNs have shown their strengths in anomaly detection in smart transportation, there are still some challenges that need to be well addressed before the formal adoption of GNNs in smart transportation. We next present several representative open issues as follows.

- *Data heterogeneity*. In urban transportation systems, there are multiple types of traffic data, such as videos (from surveillance cameras), traffic speeds, crowd flows, travelling time, etc. Moreover, external factors such as road structures and weather conditions can also influence traffic flows [92]. It is challenging to train GNNs to learn from heterogeneous traffic data.
- *Imbalance of traffic anomaly data*. Most of public traffic datasets only contain normal traffic data samples while few of them provide traffic anomalies. Many studies either manually inject traffic anomalies [75] or insert traffic anomalies through external sources from accident reports or social networks [78]. Compared with normal traffic data samples, traffic anomalies only occupy a small portion of the entire dataset. The imbalanced dataset often leads to poor performance [93]. It is a future direction to address the imbalanced traffic-anomaly data.
- *Dynamics of traffic data*. Both the road structure and transportation infrastructure are suffering from dynamics due to multiple factors, such as removal, relocation, and adding of road segments, intersections, and links. As a result, the well-trained GNN models based on massive historical traffic data cannot well handle these new scenarios. Meanwhile, it also takes a long time to re-train the entire GNN model. Thus, it is expected to address this emerging issue in the future, especially for anomaly detection in smart transportation.

[10] https://github.com/toddwschneider/nyc-citibike-data

[11] The UMN dataset is available through https://www.crcv.ucf.edu/projects/Abnormal_Crowd/

[12] https://www.nationalrail.co.uk/46391.aspx

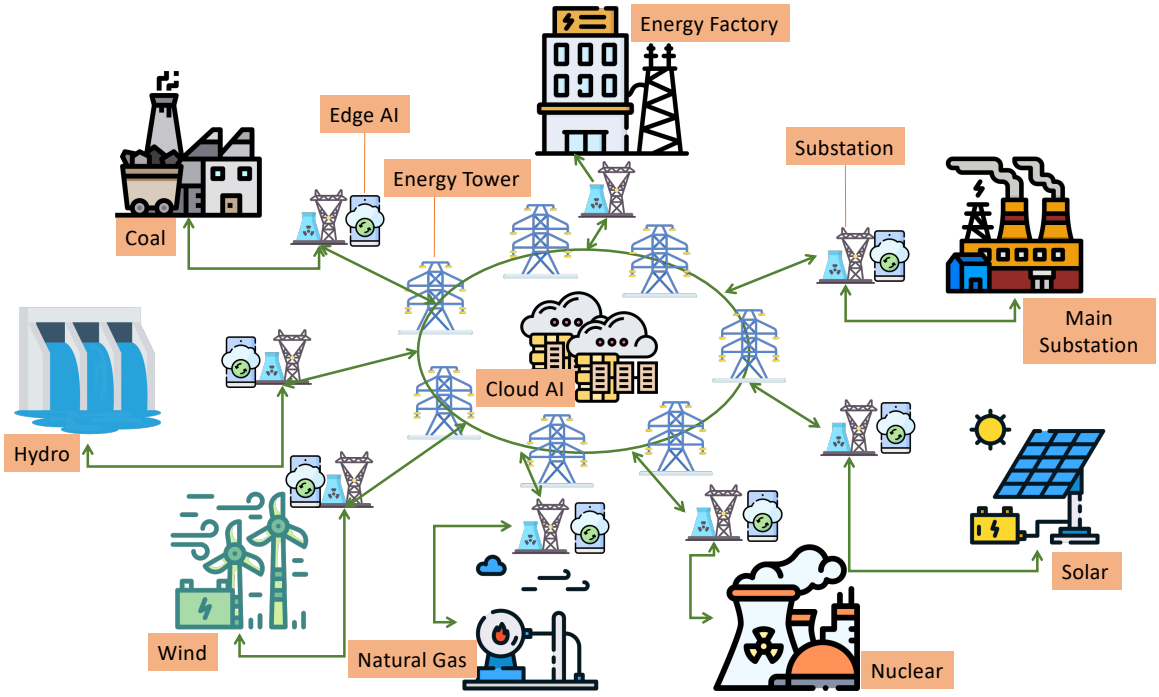

Fig. 4. IIoT for an energy industry

V. IIoT-enabled Smart Energy

The ever growing global population drives the huge demand on energy. However, the current energy supplies heavily depend on fossil fuels, consequently bringing huge Carbon dioxide (CO_2) emissions to the atmosphere. The recent renewable energies (REs) such as wind power, photovoltaic (PV) energy, and hydroelectric power can provide alternative energy sources to reduce the dependence on fossil fuels [94]. The new REs sources also promote the revolutionary upgrading of smart energy systems, which consist of distributed energy generation components, energy transmission, and energy consumption networks. However, both decentralization and complexity of the system also pose challenges of maintaining the reliability, safety, and security of smart energy systems.

The recent advances of IIoT bring opportunities to address the challenges of smart energy systems. Firstly, the wide adoption of IIoT can establish the advanced metering infrastructure (AMI) for energy systems. AMI equipped with sensors, smart meters, and controllers can measure the energy data from energy generation to energy consumption so as to provide better services [95]. Secondly, REs sources often suffer from unstable energy supply due to the fluctuated weather condition (e.g., wind, and drought). Thus, the auxiliary energy storage devices are often accompanied by REs sources to balance the supply while IoT devices deployed at both energy storage devices and REs sources can measure the voltage level to balance the entire system [96].

Fig. 4 shows an IIoT-enabled smart energy industry. The energy comes from a wide range of sources, including coal, hydro, wind, natural gas, nuclear, solar, and so on, forming a complex infrastructure for the energy industry. The data generated and/or collected by IIoT devices at each source site can be first processed by edge AI (i.e., edge computing and AI processing models) [97] and the results can be uploaded to the cloud AI (i.e., cloud computing and AI models) for further processing [98]. Alternatively, local IIoT data at each site can be partitioned and uploaded to both edge AI and cloud AI for processing, according to the application needs and the computing capacity at the edge. Data analysis in such an energy industry can benefit many aspects of business operations, e.g., balancing the energy requirement between energy sources, adjusting energy demands of a source site in the event of an anomaly happening at another source site, and facilitating the strategic policy changes such as reducing the nuclear energy demands. However, if the collected/generated data is tampered, the resulting decision from the data analysis may harm the business operation. In addition, if certain faults would happen in this complex energy infrastructure, efficient detection or prediction would save significant loss and even human lives. Since the integration of IIoT with smart energy systems becomes an inevitable trend, it is a necessity to assure the reliability and security of IIoT. Anomaly detection is a prerequisite to assure reliable and secure IIoT systems, as faults, errors, abnormal behaviors, and malicious activities can be identified to offer early warnings [99]. Table III summarizes the GNN-based solutions for anomaly detection in smart energy. We next enumerate the solutions for detecting anomalies in energy systems as follows.

A. Point

Faults can happen at every stage from power generation to power consumption. The point anomaly in smart energy system is often the event that diverges from normal measurements.

TABLE III
A SUMMARY OF THE GNN-BASED ANOMALY DETECTION SOLUTIONS IN IIOT-ENABLED SMART ENERGY.

Type of anomalies	Public datasets/software	GNN-based solutions
Point anomaly	Dissolved-gas dataset of power transformers	[100] [101]
Contextual anomaly	SGCC electricity theft dataset, US Energy Information Administration (EIA) dataset, energy consumption dataset of Ireland	[102] [103] [104] [105]
Collective anomaly	IEEE 123 bus, IEEE 39 bus, IEEE 14 bus, IEEE 118 bus, the 68-bus 16-machine 5-area system, distribution grid of LBNL	[106] [107] [108] [109]

1) Public datasets: There are several public datasets for point anomalies. In particular, the insulation at electricity transformers is a necessity to guarantee the safety of electricity systems. As one of widely-used insulation methods, the oil-paper insulation often suffers from aging and dissolving. The discharge faults or thermal faults at the transformers can fasten the dissolving process of oil-paper insulators. Since the dissolved process of oil-paper insulators also emits chemical gas like C_2H_2, C_2H_4, and C_2H_6, the analysis of dissolved gas can be used to detect insulation faults at transformers. Ref. [110] presents a dissolved-gas dataset offered by SGCC. This dataset contains a number of point anomalies, which deviate from normal values.

2) GNN solutions to anomaly detection: There are several studies toward solving anomaly detection. In particular, as shown in [111], there are strong correlations between the emitted gases and the fault types of transformers. However, it is challenging for traditional machine learning models to characterize the complex nonlinear relationship between the types of dissolved gases and the types of transformer faults. The recent study [100] proposes using GCNs for analyzing the nonlinear relationship mapping from the types of dissolved gases to the types of transformer faults. Extensive experiments on the above dissolved-gas dataset demonstrate a much higher diagnostic accuracy than other conventional machine learning methods.

Besides the fault-detection of transformers, GNN can also be used for fault classification in PV arrays as in [101]. As one of the main renewable energy sources, PV panels have been widely deployed in harsh environments causing faults of PV arrays. However, it is challenging to achieve automatic detection of the faults of PV arrays. Specifically, a spatial GCN is adopted in [101] for detecting PV faults with a limited number of labelled samples.

B. Contextual

Smart energy systems often consist of sensors, smart meters, and Phasor Measurement Units (PMUs) [112], which are interconnected through Power Line Communications (PLCs) and wireless communications. These IIoT devices can collect and generate massive smart energy data, which can be used to analyze faults, errors, and abnormal customer behaviours. Contextual anomalies in smart energy refer to data samples, which are anomalous to the remaining data samples in a certain context. For example, an extremely high temperature of a power line in winter can be regarded as an anomaly but be regarded as a normal value in summer. For another example, some malicious activities on energy consumption may be regarded as normal while the given reference electricity consumption can help to detect the abnormal energy consumption behaviours [113].

1) Public datasets: There are some public datasets available for contextual anomalies. In particular, the work [114] presents a study on investigating electricity thefts, whose electricity consumption is essentially anomalous to other normal electricity customers. The abnormal behaviours of electricity thefts include no obvious periodicity of electricity consumption in contrast to normal customers. The electricity-theft dataset[13] was adopted in this study that was released by State Grid Corporation of China (SGCC), contains the electricity consumption data of 42,372 electricity customers within more than two years (from Jan. 1, 2014 to Oct. 31, 2016). Meanwhile, the work [102] presents a study on power outages in New York City influenced by weather conditions. In particular, this study collects the Energy Information Administration (EIA) energy disturbance events (containing outages) occurred from January 2011 to December 2013 among 26,304 data samples[14] as well as weather measurement data according to historical observations[15] at the same period and weather stations at the same region. Moreover, the work [113] presents a labeled energy-consumption dataset from 500 energy customers in Ireland[16] though the labelled dataset is not officially released (available upon request to the authors).

2) GNN solutions to anomaly detection: There are several attempts in applying GNNs in contextual anomaly detection. In particular, the work [102] leverages GCN [103] to learn from weather measurement data to predict the power outages at a given region. The power-outage problem is converted into a contextual-anomaly detection problem, which is modeled by a graph, in which weather stations are nodes and an edge represents the correlation of the measurements between the two weather stations. The authors also manually label the power outage event at each weather station (i.e., occurred or not). Three variants of GCNs, clustering, selection GCN, and aggregation GCN that are considered in this paper, demonstrate the superior performance than conventional neural networks. Moreover, the work [104] proposes a GCN-based method for context anomaly detection (fault detection). In particular, a structural analysis was first used to convert pre-diagnose

[13] https://github.com/henryRDlab/ElectricityTheftDetection/

[14] From January 2011 to December 2013 https://www.eia.gov/electricity/monthly/

[15] http://wcai.wharton.upenn.edu/earth-networks-data-portal/

[16] https://www.ucd.ie/issda/data/commissionforenergyregulationcer/

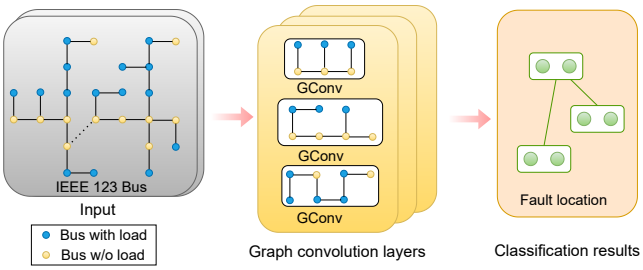

Fig. 5. GNN for fault location (reproduced from [106]).

results into graphs. Both graphs and measured datasets are fed into the GCN for the fault detection. Experiments were conducted on open-circuit fault datasets to evaluate the performance of the proposed method. Furthermore, a STGCN-based approach was proposed in [105] to investigate the short-term voltage stability with consideration faults (i.e., context anomalies). Realistic experiments on Guangdong Power Grid were conducted to evaluate the proposed method.

C. Collective

Collective anomalies are a collection of data samples, which are anomalous to the entire dataset. Regarding smart energy systems, collective anomalies can happen at energy generation, energy transmission, and energy consumption.

1) Public datasets: There are several public datasets available for collective-anomaly detection in smart energy systems. The IEEE 123 bus is a typical benchmark system for node test feeders in power systems [115]. In addition to the IEEE 123 bus system, IEEE 39 bus system [116]–[118], the IEEE 14 bus system [119], the IEEE 118 bus system, and the 68-bus 16-machine 5-area system [120] are alternative bus systems for power grid though there are few studies on applying GNNs for anomaly detection on them. Most of the above studies need to conduct simulations to obtain measured data. In addition to the above power grid systems, the work [121] also presents a measurement study on micro-PMU sensors deployed in smart grid. In particular, the dataset adopted for this study is collected from a distribution grid of the Lawrence Berkeley National Laboratory (LBNL)[17].

2) GNN solutions to anomaly detection: There are several recent advances in exploring GNNs for collective anomaly detection. The work [106] adopts a GCN to localize faults in power distribution networks. This fault-location problem is essentially collective-anomaly detection, in which multiple faults can be regarded as collective anomalies. Conventional CNN methods cannot be applied to solve this problem because the spatial-fault data in power distribution networks is no longer falling into Euclidean domain. Therefore, this paper applies a GCN framework to address this problem. In particular, the load (i.e., voltage and current) of each node (i.e., a bus) in the IEEE 123 bus system is measured. The measured data is then converted to a weighted undirected graph, which is then fed into GCN for further training, as shown in Fig. 5. Experimental results demonstrate the effectiveness of the GCN model in localizing faults. The GCN model adopted in this paper is essentially obtained from [107]. Another similar work [108] investigates the adoption of GCNs for power flow calculation based on IEEE Case 69 data. Although this study mainly focuses on deriving distribution characteristics of power flows, the main methodology can be further used for anomaly detection in the future. Moreover, the work [109] integrates GCN with long short-term memory (LSTM) network to construct the recurrent graph convolutional network (RGCN). Experiments on both IEEE 39 Bus and IEEE 300 Bus system verify the effectiveness of the proposed RGCN model for collective anomaly detection so as to ensure the stability of power grids.

[17]LBNL open power data: https://powerdata.lbl.gov/

D. Challenges and Open Issues

Despite the advances in applying GNNs in IIoT-enabled energy systems, there are a number of challenges and open issues to be solved in the future study. We enumerate several major research problems as follows.

- *Highly-reliable GNNs for smart energy systems*. Smart energy systems have a critical requirement on the reliability of power systems. However, most of existing GNN models cannot reach a quite high accuracy to ensure the high reliability for the entire system though they can assist incumbent systems to identify anomalies and faults.
- *Explainable GNNs for smart energy systems*. Like other deep neural networks, the explainability of GNNs is still not well explored, consequently limiting the wide adoption of GNNs in industrial systems, especially for smart energy systems, which have critical requirements on the explainability of models. It is worth investigating explainability of GNNs in the future [122].
- *Integration of multiple GNNs for smart energy systems*. There are diverse types of energy data in smart energy systems. For example, historical electricity consumption data often have the temporal correlation across data samples while fault-location data in smart grid have the spatial correlation. To process and analyze the diverse types of energy data (having both temporal and spatial features), the integration of multiple GNNs, such as GCNs with GRNN is a necessity in the future.
- *Deployment of GNNs in smart energy systems*. It takes extensive computational power to train complex GNN models. Since smart meters and IIoT nodes have limited computing capability, they may not be suitable for training GNN models. Thus, it is a necessity to train GNN models at remote cloud servers while the trained models can be downloaded to local edge computing nodes or IIoT nodes for consequent anomaly-detection tasks.

VI. IIoT-enabled Smart Factory

In recent years, with the fast development and deployment of 5G and IoT/IIoT [132]–[134], traditional factory and manufacturing environment is carrying out its digital transformation. While such a transformation brings benefits for economics, the security of the industry systems is being challenged [135]–[137]. This is largely due to the weakness and the lack of security considerations of traditional industry

TABLE IV
A SUMMARY OF THE GNN-BASED ANOMALY DETECTION SOLUTIONS IN IIOT-ENABLED SMART FACTORY.

Type of anomalies	Public datasets/software	GNN-based solutions
Point anomaly	Secure water treatment (SWaT), water distribution system (WADI), and critical infrastructure security showdown (CISS) datasets	[123], [124], [125]
Contextual anomaly	SWaT, WADI and BATADAL datasets, as well as the Xcos software and epanetCPA toolbox	[124], [125], [126], [127], [128]
Collective anomaly	LITNET-2020, M2M Using OPC UA, WUSTL-IIoT-2018 and KDD 1999 datasets, as well as the Xcos software and the epanetCPA tool	[129], [130], [131]

Fig. 6. Anomaly detection in smart factory

systems. Anomaly detection is an important tool to ensure an effective identification of anomalous system behaviours in smart factories [138], [139]. Fig. 6 shows several typical anomalies such as an overheated lathe, defected products, flaws with the package, and other faults. Follow-up maintenance actions can be performed in time to ensure the healthy of the operation and production systems. In this section, the three types of anomalies, i.e., point, contextual, and collective anomalies, in the context of IIoT-enabled smart factory and manufacturing will be illustrated and how GNN can enable an effective anomaly detection in this context will be elaborated. The challenges and open issues will be provided and discussed for guiding the future research in this field. A summary of this section is provided in Table IV.

A. Point

Point anomalies are the ones that are observed anomalous with respect to the rest of the data in the factory/manufacturing system without any prior indication. Such anomalies may be manifested in a single variable of a factory component, e.g., a meter reading. For example, in a time-series temperature data for a manufacturing machine, a point anomaly may refer to an anomalous reading returning to its previous normal state within a very short period. They may also be observed in multiple variables of a component where all the related variables are out of bounds at the same time, e.g., temperature and CPU usage, as the behaviors of some variables are interrelated in nature. In addition, point anomalies may also be the ones that are immediately observed without taking into account the temporal behavior. Such anomalies may not be detected in real time but would need to be detected effectively so as to avoid the "butterfly effect" to slow down or malfunction the whole manufacturing system.

1) Public datasets: Most of the available public datasets can be used for the detection of point anomalies although they are mainly used for detecting other types of anomalies. Singapore's Centre for Research in Cyber Security published a set of datasets for the research of anomaly detection in secure water treatment and water distribution systems[18]. They are maintaining four testbeds: electric power intelligent control (EPIC), internet of things automatic security testbed (IoTAS), secure water treatment (SWaT), and water distribution (WADI). SWaT is the dataset collected from continuous operations in the testbed in 11 days, where 7 days were under normal operations and 4 days were under attack scenarios. There were 41 attacks launched and the associated abnormal behaviors were labeled. WADI is a similar dataset collected from a different testbed, with 16 days of continuous operations (2 days were under attack scenarios with 15 types of attacks, the rest days were with normal operations). The data were collected from 123 sensors and actuators. There is another dataset named Critical Infrastructure Security Showdown (CISS) that is aimed at detecting cyber attacks launched in real-time on SWaT.

2) GNN solutions to anomaly detection: Point anomalies in smart factory and manufacturing environment can be detected using many techniques, such as bound/limit checking, rule-based, clustering and classification [140]. Many machine learning and deep learning techniques have been used to extract anomaly detection rules and then detect point anomalies, including traditional convolutional neural network (CNN) and LSTM. Due to the incapability of capturing the non-Euclidean data in many real-world manufacturing scenarios, GNN-based solutions have been proposed.

Due to the popularity of one class support vector machine in detecting outliers, Wang et al. [123] generalized it to graph data and proposed one class graph neural network (OCGNN) that is a one class classification framework for detecting anomalies in graph data. OCGNN can achieve the well-known one class objective using the powerful representation ability of GNN.

[18] https://itrust.sutd.edu.sg/itrust-labs_datasets/dataset_info/

For the concern of multiple variables in a time series data, Deng and Hooi [124] proposed a structure learning approach in combination of GNN and used attention weights to provide explainability for the detected anomalies. The proposed solution can handle high-dimensional time series data. Especially, it is able to capture the complex inter-variable relationships, and detects the deviation from the normal relationships.

In addition, Zhao et al. [125] also addressed the anomaly detection on multivariate time-series. They proposed a self-supervised approach where each univariate time-series was treated as an individual feature, and two graph-attention layers were used in parallel to learn the dependencies of multivariate time-series in both feature and temporal dimensions.

B. Contextual

The smart factory and manufacturing system usually consists of a considerable number of sensors, instruments, and other IoT devices interconnected to achieve many production purposes. Certain components may also connect with other systems in the factory, across factories or even a third-party system. With such a distributed environment, accurate anomaly detection shall consider contextual information. Such information may include *temporal context*, *spatial context*, and *external context* [141]. The temporal context is the one that is usually shown in the time-series data where the temporal correlation exists between observations. The spatial context usually refers to the position information of the devices where the spatial correlation exists when the devices are working together to carry out one task in the production environment. The external context explains how the externally connected systems affect the internal monitoring system. For example, temperature control of a production system might be related to the weather condition outside the factory. The weather condition could be measured by additional sensors from outside of the building.

1) Public datasets: Both SWaT and WADI datasets mentioned in Section VI-A1 contain the contextual data as well. Supervisory control and data acquisition (SCADA) is an important control and data collection mechanism in the industrial system. BATADAL data is the data on hourly historical SCADA operations[19]. It contains three datasets, where the first one contains the data for one-year-long normal operations, and the second and the third ones contain the data with 14 attacks. The Xcos software[20] developed by SciLab is an open source software that can design hybrid dynamical system models with the functionalities for modeling of mechanical systems (e.g., automotive and aeronautics), hydraulic circuits (e.g., dam and pipe modeling), control systems, etc. This software can be used to general industrial operational contextual data [142]. Meanwhile, epanetCPA[21] is an open-source MATLAB toolbox for modelling the hydraulic response of water distribution systems to cyber-physical attacks. It can be used to generate contextual dataset for anomaly detection.

[19] http://www.batadal.net/data.html
[20] https://www.scilab.org/software/xcos
[21] https://github.com/rtaormina/epanetCPA

2) GNN solutions to anomaly detection: Some of the methods mentioned in Section VI-A2 for anomaly detection on multivariate time-series data can also be adopted here for detecting contextual anomalies, e.g., [124], [125]. To achieve this purpose, the modeling purpose shall shift to finding the relationship among multiple variables.

To comprehensively consider all possible structural, context, and temporal features in anomaly detection, Zheng et al. [126] proposed an end-to-end anomalous edge detection framework, called AddGraph, based on an extended temporal GCN with the attention mechanism. The proposed model can capture both short-term and long-term patterns in dynamic graphs.

Statistical features of the data, such as network traffic are very important for detecting contextual anomalies, but they were usually carried out manually using expert knowledge. To avoid manual extraction of statistical features, Xiao et al. [127] developed an approach with two graphs: first-order graph and second-order graph. The former learns the latent features from a single entity such as a host or a variable, and the latter learns the latent features from a global point of view. The automatically extracted features can be used to train machine learning models for classifying network anomalies.

A graph based method was proposed in [128] to learn dependencies between variables in time-series data. Nodes in the graph represent individual observations or sequences of observations, where the weighting of the link between nodes represents the degree of dependencies on other nodes. Low weighting shows that node is flagged as anomalous.

C. Collective

Different from the contextual anomalies which focus on a data instance, collective anomalies essentially is a collection of related data instances where they are anomalous as a group with respect to the entire dataset. Note that the individual data instance in a collective anomaly may not show abnormal behaviors but their occurrence as a group exhibits the anomalous bahavior. For example, in an industry setting, the memory usage of a server may be normal individually compared with the historical records. But if the pattern of the memory consumption, as time goes, meets certain criteria, it could show a memory leak as result of a collective anomaly. In addition, on the way towards Industry 4.0, many industry devices/sensors will adopt machine learning models to make autonomous decisions. Some potential ethical issues may surface, resulting from a collection of sensors collectively behaving some anomalous actions e.g. making bias decisions. This can also be treated as collective anomalies.

1) Public datasets: The Xcos software and the epanetCPA tool mentioned in Section VI-B1 can be used to generate collective datasets for anomaly detection. In addition, Denial of Service/Distributed Denial of Service (DoS/DDoS) attack is a good example of collective anomalies. LITNET-2020 is an annotated network benchmark dataset obtained from the real-world academic network. It contains 85 network flow features of the dataset and 12 attack types. “M2M Using OPC UA” [143] is a dataset generated by injecting various attacks on a OPC UA based Cyber-Physical Production Systems testbed.

The attacks include DoS, eavesdropping or Man-in-the-middle attacks, and impersonation or spoofing attacks[22]. "WUSTL-IIoT-2018" is a dataset used for the SCADA cybersecurity research, where the attacks include port scanner, address scan, device identification, device identification (aggressive mode), and exploit[23]. It is also a useful dataset for carrying out the research of the detection on collective anomalies. In addition, the KDD 1999 dataset[24] has also been widely used to validate the effectiveness of collective anomaly detection algorithms.

2) GNN solutions to anomaly detection: Jiang et al. [129] devised a GCN-based anomaly detection model that can capture the entities' properties and structural information between them into graphs. With the proposed model, both abnormal behaviors of individuals and the associated anomalous groups can be detected.

Botnets are a major source for DoS/DDoS attacks which can result in collective anomalies. Zhou et al. [130] developed a GNN-based approach to detect the hierarchical structure of centralized botnets and the fast-mixing structure for decentralized botnets. The outcome will then be used for learning policies for automatic botnet detection.

Protogerou et al. [131] developed a multi-agent system to exploit the collaborative and cooperative nature of intelligent agents for anomaly detection. Each agent will be implemented using a GNN that can learn the representation of physical networks. This distributed detection approach can carry out the efficient monitoring of the entire network infrastructure and can be treated as a potential candidate solution for detecting collective anomalies.

D. Challenges and Open Issues

Although GNN has been used for enhancing the performance of anomaly detection in smart factory and manufacturing systems, there are still several challenges and open issues that need to investigate in future studies.

- *The effectiveness of GNN modeling.* Some data has a clear indication of node and link representation in GNN, while some may be not that straightforward. For example, the node in GNN can directly model the node in an industrial device/sensor, and the edge is the relationship between devices. In addition, the node in GNN could model a variable in a dataset, and the edge is the relationship between variables. It is still an important issue that which part of the data is the node in GNN and which part is the edge.
- *A combination of point, collective and contextual anomalies.* It is more challenging for anomaly detection if a combination of types of anomalies shows, such as detecting collective contextual anomalies. Dou, Yang and Poor [144] proposed a framework for discovering this type of anomalies in multiple time-series based on a combination of several techniques including deep learning, time-series modeling, and graph analysis. In addition, a recent study used graph autoencoders for group anomaly detection, where graph representation learning is achieved to detect collective anomalies by exploiting their graph structures[25]. The research in this horizon is still in its infancy especially in the field of anomaly detection in the smart factory and manufacturing.
- *Working with other machine learning techniques.* GNN is a model to effectively learn the node representation. This representation can then be used to work with other machine learning and deep learning techniques for carrying out anomaly detection. Factory and manufacturing environment is complex. It is important to carefully consider the output of GNN and the input of other machine learning models for more effective and accurate anomaly detection.

VII. Case Studies

In this section, we present several representative case studies to illustrate how GNN-based models work for anomaly detection.

A. STGNN-based model for detecting collective anomalies in a public transportation system

The collective anomalies in public transportation systems are often the root cause of traffic jams. It is a necessity to detect traffic anomalies in public transportation systems though it is often a challenging task since traffic anomalies are affected by multiple factors, such as accidents, gathering, criminals, and public events. However, existing deep learning models can either capture spatial features (e.g., road structures) or temporal features (e.g., number of vehicles across a road per hour), but not both. Thus, GNN-based models have potential to address this challenge since they have strong capability to learn from spatial-temporal features together.

We will next demonstrate that an STGNN-based model can be used for detecting collective anomalies in a public transportation system. Fig. 7 depicts the proposed STGNN for traffic anomaly detection. The proposed method works in the following three steps.

In Step 1, we need to obtain historical traffic-flow data from the public transportation system. The traffic-flow data includes station ID, station location, and historical crowd flows at a certain station and a certain time. The crowd flows can be essentially obtained by check-in/out records (i.e., recorded by NFC-based tokens or cards) of the public transportation system. In particular, we denote the incoming crowd flows at a station T_i by C_{in,T_i} and the outgoing crowd flows at T_i by C_{out,T_i}. In addition, we can obtain traffic anomalies from historical records of accidents, gathering events, criminal reports, and other public events. The anomaly data needs to be associated with the crow-flow data according the spatial-temporal correlations. In Step 2, we need to convert traffic-flow data into graphs (e.g., incidence graphs), which can be further processed by our STGNN. Meanwhile, historical anomalies need to integrated with the traffic flows. In Step 3, we then construct the STGNN, which consists of multiple layers. The

[22] https://ieee-dataport.org/open-access/m2m-using-opc-ua
[23] https://www.cse.wustl.edu/~jain/iiot/index.html
[24] http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
[25] https://grlearning.github.io/papers/85.pdf

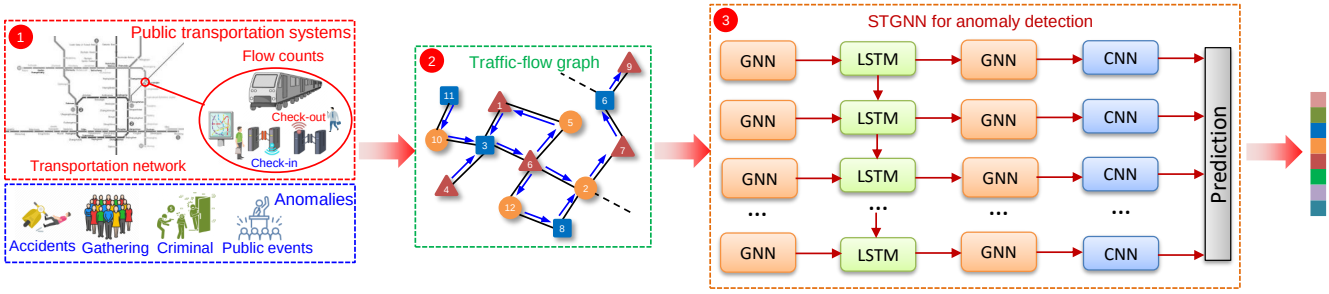

Fig. 7. Spatial-temporal graph neural network for traffic anomaly detection in public transportation systems

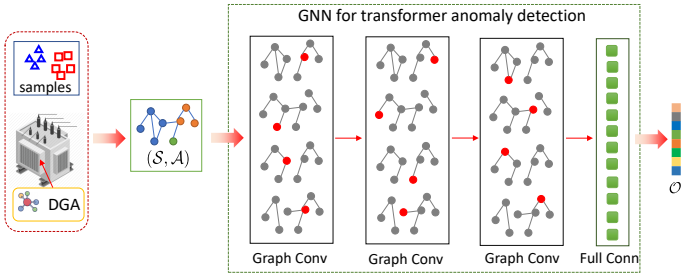

Fig. 8. GCN for power transformer anomaly detection

first GNN layer is used to capture the spatial features while the second LSTM layer is used to extract temporal features from the historical traffic flows. We next adopt another GNN layer to further explore the spatial-temporal dependencies obtained by the first two layers. We finally predict (i.e., classify the anomalies from normal traffic flows) after passing the spatial-temporal features by a CNN layer.

It is worth mentioning that this general framework can be further extended by replacing neural modules by other variants to further improve the performance. For example, we can replace the LSTM module by a GRU. Meanwhile, the attention mechanism can be also leveraged to improve the learning effectiveness.

B. GCN for detecting point anomalies in power transformers

Since power transformers play an important role between power transmission and power distribution, the reliability of power transforms is crucial for the safety assurance of electricity systems. The faulty transformers emit huge amount of dissolved gases such as C_2H_6, C_2H_2, and C_2H_4, which are indicators for transformer faults. The dissolved gases can be regarded as point anomalies for analysing faults of power transformers.

We therefore design a GCN-based model for detecting point anomalies in power transformers. Fig. 8 depicts the working flow of this model for anomaly detection of power transformers. Firstly, we can obtain historical DGA data from public datasets such as [110]. After applying the Siamese network and k-nearest neighbor (kNN) approaches, we can generate the well-formed input datasets, which consist of $(\mathcal{S},\mathcal{A})$, where $\mathcal{S}$ represents the feature matrix of dissolve gases and $\mathcal{A}$ denotes the adjacency matrix (i.e., characterizing the similarity between historical samples and current samples).

We then feed the input datasets into a GCN, as shown in Fig. 8. The benefits of using the GCN instead of conventional CNNs or other machine learning methods lie in the strong learning capability of GCNs in characterizing the nonlinear correlations between the types of dissolved gases and the types of transformer faults. Our GCN consists of three layers of graph convolution layers (each of which is denoted by Graph Conv). In particular, we have

$$\mathsf{Conv}(i) = g\big(\mathsf{Conv}(i-1),\mathcal{A}\big), \tag{1}$$

where $g(\cdot,\cdot)$ denotes a non-linear function, $i \in \{1,\cdots,N\}$, and N denotes the number of graph convolution layers. When $i=1$, $\mathsf{Conv}(1) = g(\mathcal{S},\mathcal{A})$.

Meanwhile, we also add a dropout-based hidden layer and a rectified linear activation unit (ReLU) between two adjacent graph convolution layers so as to avoid overfitting and improve learning effect. Finally, we apply a softmax function to finalize the classification results $\mathcal{O}$ after passing through a fully-connected layer (i.e., Full Conn).

We will explore the usage of GCN and other GNN models in anomaly detection in other smart energy systems, such as REs since GNN models have the strengths in capturing spatial correlations whereas there may existing spatial correlations amount multiple REs.

C. GCN-based model for detecting collective anomalies in the smart factory

Collective anomalies in a smart factory are usually not easy to detect. That is because, first, there are massive industrial internet of things (IIoT) devices, and second, individual devices seem to function as normal but the behaviour of many IIoT devices as a whole is abnormal. In this section, we will elaborate a concrete example of how GNN can be used to detect collective anomalies in the smart factory.

In a smart factory as shown in Fig. 9, PLC is an industrial digital computer that is designed for the control of manufacturing processes. Each PLC is associated with several IIoT devices, and a PLC is often networked to other PLC. IIoT devices usually need to be patched due to e.g. security reasons. Collective anomalies may happen when the program of several IIoT devices like environmental control devices and robot control devices is changed/patched but the corresponding programmable logic controller (PLC) is not upgraded.

Many traditional anomaly detection methods only consider the features of IIoT devices or PLC. However, in order to

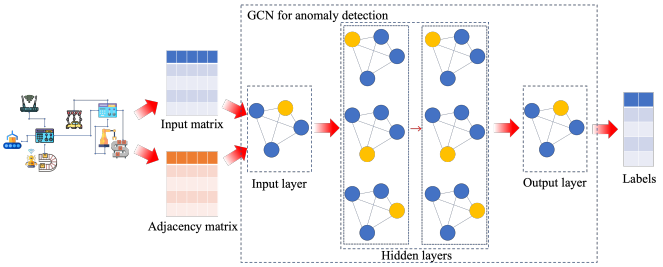

Fig. 9. GCN for anomaly detection in smart factory

detect collective anomalies, the relationships between IIoT devices or PLC is also important. GNN is a good candidate to capture such relationships, and thus it has the potential to effectively detect collective anomalies.

In this case study, we resort to a typical example of GNN, GCN for the detection of collective anomalies. First, we transform the industrial communication network and the behaviours of IIoT devices and PLC into a graph. The nodes of the graph denote the IIoT devices and PLC as well as other possible factory devices, and the edges between nodes are the structural information of these devices.

Then, we can design a GCN based anomaly detection model. It can be an n-layers GCN model with a softmax layer to train a binary classification model for "normal" and "abnormal". A softmax activation function, shown in Eq. (2) can be used to calculate the classification of each node.

$$\sigma(\vec{z_i}) = \frac{e^{z_i}}{\sum_{i=1}^{2} e^{z_k}}, \tag{2}$$

where σ denotes the softmax, $\vec{z_i}$ represents the input vector, e^{z_i} is the standard exponential function for the input vector, and e^{z_k} is the standard exponential function for the output vector. The input of the GCN model is the graph containing the IIoT devices and PLC as the nodes with features, and the connections between nodes as edges of the graph. In practice, the inputs are two matrices, one is the node feature matrix and the other is the node adjacency matrix. Batch gradient descent can be used to train the weights of the GCN model. The GCN model can output the binary classification of normal and abnormal for each node.

VIII. Conclusion

This paper provided a useful investigation for GNN-empowered anomaly detection solutions for IIoT-enabled smart transportation, smart energy, and smart factory. In particular, a deeper understanding of three types of anomalies, i.e., point, contextual, and collective anomalies, in the context of above IIoT applications was provided. In addition, the useful public datasets were provided for each type of anomalies in the corresponding IIoT applications. Further, important research challenges and open issues of GNN-based anomaly detection solutions for the three investigated IIoT applications were provided and discussed. Finally, we show three case studies of the use of GNN in addressing anomaly detection problems in IIoT-enabled smart transportation, smart energy, and smart factory. We hope that this paper provides useful guidance for the future research in this area.

References

[1] Y. Wu, H. N. Dai, and H. Wang, "Convergence of Blockchain and Edge Computing for Secure and Scalable IIoT Critical Infrastructures in Industry 4.0," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2300–2317, 2021.

[2] D. Wang, "An enterprise data pathway to industry 4.0," *IEEE Engineering Management Review*, vol. 46, no. 3, pp. 46–48, 2018.

[3] N. Zhang, Y. Li, Y. Wu, and Q. Zhang, "Guest editorial: Ai empowered communication and computing systems for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 4914–4916, 2021.

[4] Y. Wu, Z. Wang, Y. Ma, and V. C. Leung, "Deep reinforcement learning for blockchain in industrial IoT: A survey," *Computer Networks*, vol. 191, p. 108004, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621001213

[5] G. Culot, F. Fattori, M. Podrecca, and M. Sartor, "Addressing industry 4.0 cybersecurity challenges," *IEEE Engineering Management Review*, vol. 47, no. 3, pp. 79–86, 2019.

[6] R. D. Raut, A. Gotmare, B. E. Narkhede, U. H. Govindarajan, and S. U. Bokade, "Enabling technologies for industry 4.0 manufacturing and supply chain: Concepts, current status, and adoption challenges," *IEEE Engineering Management Review*, vol. 48, no. 2, pp. 83–102, 2020.

[7] P. Patel, M. I. Ali, and A. Sheth, "From raw data to smart manufacturing: Ai and semantic web of things for industry 4.0," *IEEE Intelligent Systems*, vol. 33, no. 4, pp. 79–86, 2018.

[8] W. Liang, W. Huang, J. Long, K. Zhang, K.-C. Li, and D. Zhang, "Deep reinforcement learning for resource protection and real-time detection in iot environment," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6392–6401, 2020.

[9] Z. Cai and Z. He, "Trading private range counting over big iot data," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, 2019, pp. 144–153.

[10] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong, J. Kang, and M. S. Hossain, "Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach," *IEEE Internet of Things Journal*, pp. 1–1, 2020.

[11] Y. Liu, N. Kumar, Z. Xiong, W. Y. B. Lim, J. Kang, and D. Niyato, "Communication-efficient federated learning for anomaly detection in industrial internet of things," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[12] B. Genge, P. Haller, and C. Enăchescu, "Anomaly detection in aging industrial internet of things," *IEEE Access*, vol. 7, pp. 74 217–74 230, 2019.

[13] P. Sun, E. Yuepeng, T. Li, Y. Wu, J. Ge, J. You, and B. Wu, "Context-aware learning for anomaly detection with imbalanced log data," in *2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2020, pp. 449–456.

[14] Y. Guo, Y. Wu, Y. Zhu, B. Yang, and C. Han, "Anomaly detection using distributed log data: A lightweight federated learning approach," in *2021 The International Joint Conference on Neural Networks (IJCNN)*, 2021.

[15] H. Cao, H. Tang, Y. Wu, F. Wang, and Y. Xu, "On accurate computation of trajectory similarity via single image super-resolution," in *2021 The International Joint Conference on Neural Networks (IJCNN)*, 2021.

[16] Y. Wu, Y. Ma, H.-N. Dai, and H. Wang, "Deep learning for privacy preservation in autonomous moving platforms enhanced 5G heterogeneous networks," *Computer Networks*, vol. 185, p. 107743, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S138912862031327X

[17] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion," *Information Fusion*, vol. 51, pp. 129–144, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253518304731

[18] X. Jing, Z. Yan, X. Jiang, and W. Pedrycz, "Network traffic fusion and analysis against ddos flooding attacks with a novel reversible sketch," *Information Fusion*, vol. 51, pp. 100–113, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253518305815

[19] Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, "Deep learning based inference of private information using embedded sensors in smart devices," *IEEE Network*, vol. 32, no. 4, pp. 8–14, 2018.

[20] H. Wang, S. Ma, C. Guo, Y. Wu, H.-N. Dai, and D. Wu, "Blockchain-based power energy trading management," *ACM Transactions on Internet Technology*, vol. 21, no. 2, 2021. [Online]. Available: https://doi.org/10.1145/3409771

[21] Y. Zuo, Y. Wu, G. Min, C. Huang, and K. Pei, "An intelligent anomaly detection scheme for micro-services architectures with temporal and spatial data analysis," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 548–561, 2020.

[22] C. Huang, G. Min, Y. Wu, Y. Ying, K. Pei, and Z. Xiang, "Time series anomaly detection for trustworthy services in cloud computing systems," *IEEE Transactions on Big Data*, pp. 1–1, 2017.

[23] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial iots," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.

[24] M. Aazam, S. Zeadally, and K. A. Harras, "Deploying fog computing in industrial internet of things and industry 4.0," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4674–4682, 2018.

[25] Z. Zhao, Y. Shi, B. Diao, and B. Wu, "Optimal data caching and forwarding in industrial iot with diverse connectivity," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2288–2296, 2019.

[26] X. Jing, H. Han, Z. Yan, and W. Pedrycz, "Supersketch: A multi-dimensional reversible data structure for super host identification," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2021.

[27] L. Ruff, J. R. Kauffmann, R. A. Vandermeulen, G. Montavon, W. Samek, M. Kloft, T. G. Dietterich, and K. R. Müller, "A unifying review of deep and shallow anomaly detection," *Proceedings of the IEEE*, pp. 1–40, 2021.

[28] X. Zhang, X. Ma, N. Huyan, J. Gu, X. Tang, and L. Jiao, "Spectral-difference low-rank representation learning for hyperspectral anomaly detection," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14, 2021.

[29] J. Lu, S. Jin, J. Liang, and C. Zhang, "Robust few-shot learning for user-provided data," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2020.

[30] D. Zhu, Y. Sun, H. Du, N. Cao, T. Baker, and G. Srivastava, "Huna: A method of hierarchical unsupervised network alignment for iot," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3201–3210, 2021.

[31] C. Huang, Y. Wu, Y. Zuo, K. Pei, and G. Min, "Towards experienced anomaly detector through reinforcement learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, Apr. 2018. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/12130

[32] C. Huang, Y. Wu, G. Min, and Y. Ying, "Kernelized convex hull approximation and its applications in data description tasks," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–8.

[33] Y. Wu, F. Hu, G. Min, and A. Zomaya, *Big Data and Computational Intelligence in Networking*. Boca Raton: CRC Press, 2017.

[34] C. Zhang, M. Dong, and K. Ota, "Accelerate deep learning in iot: Human-interaction co-inference networking system for edge," in *2020 13th International Conference on Human System Interaction (HSI)*, 2020, pp. 1–6.

[35] ——, "Enabling computational intelligence for green internet of things: Data-driven adaptation in lpwa networking," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 32–43, 2020.

[36] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 2, 2005, pp. 729–734 vol. 2.

[37] Y. Li, D. Tarlow, M. Brockschmidt, and R. S. Zemel, "Gated graph sequence neural networks," in *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2016. [Online]. Available: http://arxiv.org/abs/1511.05493

[38] H. Wang, Y. Wu, G. Min, and W. Miao, "A graph neural network-based digital twin for network slicing management," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2020.

[39] K. Li, G. Lu, G. Luo, and Z. Cai, *Seed-Free Graph De-Anonymiztiation with Adversarial Learning*. New York, NY, USA: Association for Computing Machinery, 2020, p. 745–754. [Online]. Available: https://doi.org/10.1145/3340531.3411970

[40] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. [Online]. Available: https://openreview.net/forum?id=SJU4ayYgl

[41] J. M.-T. Wu, Z. Li, G. Srivastava, M.-H. Tasi, and J. C.-W. Lin, "A graph-based convolutional neural network stock price prediction with leading indicators," *Software: Practice and Experience*, vol. 51, no. 3, pp. 628–644, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2915

[42] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018. [Online]. Available: https://openreview.net/forum?id=rJXMpikCZ

[43] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: a survey," *Data Mining and Knowledge Discovery*, vol. 29, no. 3, pp. 626–688, 2015. [Online]. Available: https://doi.org/10.1007/s10618-014-0365-y

[44] D. Miller, "Blockchain and the internet of things in the industrial sector," *IT Professional*, vol. 20, no. 3, pp. 15–18, 2018.

[45] P. Lade, R. Ghosh, and S. Srinivasan, "Manufacturing analytics and industrial internet of things," *IEEE Intelligent Systems*, vol. 32, no. 3, pp. 74–79, 2017.

[46] Y. Wu, H. Huang, C. Wang, and Y. Pan, *5G-Enabled Internet of Things*. Boca Raton: CRC Press, 2019.

[47] X. Cheng, Y. Wu, G. Min, A. Y. Zomaya, and X. Fang, "Safeguard network slicing in 5g: A learning augmented optimization approach," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 7, pp. 1600–1613, 2020.

[48] Y. Wu, S. Singh, T. Taleb, A. Roy, H. Dhillon, M. Kanagarathinam, and A. De, *6G Mobile Wireless Networks*. Springer, 2021.

[49] Y. Wu, "Cloud-edge orchestration for the internet-of-things: Architecture and ai-powered data processing," *IEEE Internet of Things Journal*, pp. 1–1, 2020.

[50] ——, "Robust learning enabled intelligence for the internet-of-things: A survey from the perspectives of noisy data and adversarial examples," *IEEE Internet of Things Journal*, pp. 1–1, 2020.

[51] Y. Wu, F. Hu, G. Min, and A. Zomaya, *Big Data and Computational Intelligence in Networking*. CRC Press, 2017.

[52] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *CoRR*, vol. abs/1901.03407, 2019. [Online]. Available: http://arxiv.org/abs/1901.03407

[53] J. Zhou, G. Cui, Z. Zhang, C. Yang, Z. Liu, and M. Sun, "Graph neural networks: A review of methods and applications," *CoRR*, vol. abs/1812.08434, 2018. [Online]. Available: http://arxiv.org/abs/1812.08434

[54] Z. Guo, K. Yu, Y. Li, G. Srivastava, and J. C.-W. Lin, "Deep learning-embedded social internet of things for ambiguity-aware social recommendations," *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2021.

[55] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2021.

[56] S. Georgousis, M. P. Kenning, and X. Xie, "Graph deep learning: State of the art and challenges," *IEEE Access*, vol. 9, pp. 22 106–22 140, 2021.

[57] Z. Zhang, P. Cui, and W. Zhu, "Deep learning on graphs: A survey," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–1, 2020.

[58] N. Vesselinova, R. Steinert, D. F. Perez-Ramirez, and M. Boman, "Learning combinatorial optimization on graphs: A survey with applications to networking," *IEEE Access*, vol. 8, pp. 120 388–120 416, 2020.

[59] T. Danel, P. Spurek, J. Tabor, M. Śmieja, Ł. Struski, A. Słowik, and Ł. Maziarka, "Spatial graph convolutional networks," in *Neural Information Processing*, H. Yang, K. Pasupa, A. C.-S. Leung, J. T. Kwok, J. H. Chan, and I. King, Eds. Cham: Springer International Publishing, 2020, pp. 668–675.

[60] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17. Red Hook, NY, USA: Curran Associates Inc., 2017, p. 1025–1035.

[61] Y. Zuo, Y. Wu, G. Min, and L. Cui, "Learning-based network path planning for traffic engineering," *Future Generation Computer Systems*, vol. 92, pp. 59–67, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X18313244

[62] Z. Yan, J. Ge, Y. Wu, L. Li, and T. Li, "Automatic virtual network embedding: A deep reinforcement learning approach with graph convolu-

tional networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1040–1057, 2020.

[63] K. Cho, B. van Merrienboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, A. Moschitti, B. Pang, and W. Daelemans, Eds. ACL, 2014, pp. 1724–1734. [Online]. Available: https://doi.org/10.3115/v1/d14-1179

[64] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[65] A. Hasanzadeh, E. Hajiramezanali, K. Narayanan, N. Duffield, M. Zhou, and X. Qian, "Variational graph recurrent neural networks," *Advances in neural information processing systems*, vol. 32, 2019.

[66] K. Xu, C. Li, Y. Tian, T. Sonobe, K. Kawarabayashi, and S. Jegelka, "Representation learning on graphs with jumping knowledge networks," in *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, ser. Proceedings of Machine Learning Research, J. G. Dy and A. Krause, Eds., vol. 80. PMLR, 2018, pp. 5449–5458. [Online]. Available: http://proceedings.mlr.press/v80/xu18c.html

[67] H. Yang, X. Yan, X. Dai, and J. Cheng, "Self-enhanced GNN: improving graph neural networks using model outputs," *CoRR*, vol. abs/2002.07518, 2020. [Online]. Available: https://arxiv.org/abs/2002.07518

[68] H. Li, G. Zhao, L. Qin, H. Aizeke, X. Zhao, and Y. Yang, "A survey of safety warnings under connected vehicle environments," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–17, 2020.

[69] K. Ren, Q. Wang, C. Wang, Z. Qin, and X. Lin, "The security of autonomous driving: Threats, defenses, and future directions," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 357–372, 2020.

[70] B. Yin, Y. Wu, T. Hu, J. Dong, and Z. Jiang, "An efficient collaboration and incentive mechanism for internet of vehicles (iov) with secured information exchange based on blockchains," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1582–1593, 2020.

[71] L. Cheng, J. Liu, G. Xu, Z. Zhang, H. Wang, H.-N. Dai, Y. Wu, and W. Wang, "Sctsc: A semicentralized traffic signal control mode with attribute-based blockchain in iovs," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1373–1385, 2019.

[72] L. Cui, D. Su, Y. Zhou, L. Zhang, Y. Wu, and S. Chen, "Edge learning for surveillance video uploading sharing in public transport systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2274–2285, 2021.

[73] J.-X. Zhong, N. Li, W. Kong, S. Liu, T. H. Li, and G. Li, "Graph convolutional label noise cleaner: Train a plug-and-play action classifier for anomaly detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[74] L. Yu, B. Du, X. Hu, L. Sun, L. Han, and W. Lv, "Deep spatio-temporal graph convolutional network for traffic accident prediction," *Neurocomputing*, vol. 423, pp. 135–147, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092523122031451X

[75] Y. Hu, A. Qu, and D. Work, "Graph convolutional networks for traffic anomaly," *arXiv preprint arXiv:2012.13637*, 2020. [Online]. Available: https://arxiv.org/abs/2012.13637

[76] Z. Zhou, Y. Wang, X. Xie, L. Chen, and H. Liu, "RiskOracle: A Minute-Level Citywide Traffic Accident Forecasting Framework," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020, pp. 1258–1265.

[77] Z. Zhou, Y. Wang, X. Xie, L. Chen, and C. Zhu, "Foresee urban sparse traffic accidents: A spatiotemporal multi-granularity perspective," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–1, 2020.

[78] R. Liu, S. Zhao, B. Cheng, H. Yang, H. Tang, and F. Yang, "ST-MFM: A Spatiotemporal Multi-modal Fusion Model for Urban Anomalies Prediction," in *Proceedings of the Twenty-fourth European Conference on Artificial Intelligence*, 2020.

[79] Y. Zhang, X. Dong, L. Shang, D. Zhang, and D. Wang, "A multi-modal graph neural network approach to traffic risk forecasting in smart urban sensing," in *2020 17th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2020, pp. 1–9.

[80] H.-W. Wang, Z.-R. Peng, D. Wang, Y. Meng, T. Wu, W. Sun, and Q.-C. Lu, "Evaluation and prediction of transportation resilience under extreme weather events: A diffusion graph convolutional approach," *Transportation Research Part C: Emerging Technologies*, vol. 115, p. 102619, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X19305868

[81] J. S. W. Heglund, P. Taleongpong, S. Hu, and H. T. Tran, "Railway delay prediction with spatial-temporal graph convolutional networks," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 2020, pp. 1–6.

[82] M. Zhang, T. Li, Y. Yu, Y. Li, P. Hui, and Y. Zheng, "Urban anomaly analytics: Description, detection and prediction," *IEEE Transactions on Big Data*, pp. 1–1, 2020.

[83] L. Tišljarić, S. Fernandes, T. Carić, and J. Gama, "Spatiotemporal traffic anomaly detection on urban road network using tensor decomposition method," in *Discovery Science*. Cham: Springer International Publishing, 2020, pp. 674–688.

[84] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," *arXiv preprint arXiv:2101.11174*, 2021. [Online]. Available: https://arxiv.org/abs/2101.11174

[85] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 18–32, 2014.

[86] Y. Benezeth, P. . Jodoin, V. Saligrama, and C. Rosenberger, "Abnormal events detection based on spatio-temporal co-occurences," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 2458–2465.

[87] K. K. Santhosh, D. P. Dogra, and P. P. Roy, "Anomaly detection in road traffic using visual surveillance: A survey," *ACM Comput. Surv.*, vol. 53, no. 6, Dec. 2020. [Online]. Available: https://doi.org/10.1145/3417989

[88] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, Jul. 2009. [Online]. Available: https://doi.org/10.1145/1541880.1541882

[89] "New york city (nyc) taxi dataset," 2021. [Online]. Available: https://www1.nyc.gov/site/tlc/index.page

[90] 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 *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 2672–2681.

[91] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 935–942.

[92] Z. Zheng, Y. Yang, J. Liu, H. Dai, and Y. Zhang, "Deep and embedded learning approach for traffic flow prediction in urban informatics," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3927–3939, 2019.

[93] J. Fang, D. Yan, J. Qiao, J. Xue, and H. Yu, "DADA: Driver Attention Prediction in Driving Accident Scenarios," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021.

[94] S. S. Ali and B. J. Choi, "State-of-the-art artificial intelligence techniques for distributed smart grids: A review," *Electronics*, vol. 9, no. 6, 2020.

[95] P. Gope, "PMAKE: Privacy-aware multi-factor authenticated key establishment scheme for Advance Metering Infrastructure in smart grid," *Computer Communications*, vol. 152, pp. 338–344, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366419313210

[96] L. Al-Ghussain, R. Samu, O. Taylan, and M. Fahrioglu, "Sizing renewable energy systems with energy storage systems in microgrids for maximum cost-efficient utilization of renewable energy resources," *Sustainable Cities and Society*, vol. 55, p. 102059, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670720300469

[97] J. Zhang, Y. Wu, G. Min, F. Hao, and L. Cui, "Balancing energy consumption and reputation gain of uav scheduling in edge computing," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1204–1217, 2020.

[98] D. Liu, Z. Yan, W. Ding, and M. Atiquzzaman, "A survey on secure data analytics in edge computing," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4946–4967, 2019.

[99] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power systems," *arXiv preprint arXiv:2101.10025*, 2021. [Online]. Available: https://arxiv.org/abs/2101.10025

[100] W. Liao, D. Yang, Y. Wang, and X. Ren, "Fault diagnosis of power transformers using graph convolutional network," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2, pp. 241–249, 2021.

[101] J. Fan, S. Rao, G. Muniraju, C. Tepedelenlioglu, and A. Spanias, "Fault classification in photovoltaic arrays using graph signal processing," in *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, vol. 1, 2020, pp. 315–319.

[102] D. Owerko, F. Gama, and A. Ribeiro, "Predicting power outages using graph neural networks," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018, pp. 743–747.

[103] F. Gama, A. G. Marques, G. Leus, and A. Ribeiro, "Convolutional Neural Network Architectures for Signals Supported on Graphs," *IEEE Transactions on Signal Processing*, vol. 67, no. 4, pp. 1034–1049, 2019.

[104] Z. Chen, J. Xu, T. Peng, and C. Yang, "Graph convolutional network-based method for fault diagnosis using a hybrid of measurement and prior knowledge," *IEEE Transactions on Cybernetics*, pp. 1–13, 2021.

[105] Y. Luo, C. Lu, L. Zhu, and J. Song, "Data-driven short-term voltage stability assessment based on spatial-temporal graph convolutional network," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 106753, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061520342988

[106] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault location in power distribution systems via deep graph convolutional networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 119–131, 2020.

[107] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Neural Information Processing Systems*, ser. NIPS'16. Red Hook, NY, USA: Curran Associates Inc., 2016, p. 3844–3852.

[108] D. Wang, K. Zheng, Q. Chen, G. Luo, and X. Zhang, "Probabilistic power flow solution with graph convolutional network," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, 2020, pp. 650–654.

[109] J. Huang, L. Guan, Y. Su, H. Yao, M. Guo, and Z. Zhong, "Recurrent graph convolutional network-based multi-task transient stability assessment framework in power system," *IEEE Access*, vol. 8, pp. 93 283–93 296, 2020.

[110] E. Li, "Dissolved gas data in transformer oil—fault diagnosis of power transformers with membership degree," 2019. [Online]. Available: https://dx.doi.org/10.21227/h8g0-8z59

[111] Y. Benmahamed, M. Teguar, and A. Boubakeur, "Application of SVM and KNN to Duval Pentagon 1 for transformer oil diagnosis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 6, pp. 3443–3451, 2017.

[112] G. Dileep, "A survey on smart grid technologies and applications," *Renewable Energy*, vol. 146, pp. 2589–2625, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148119312790

[113] A. Amara Korba, N. Tamani, Y. Ghamri-Doudane, and N. E. I. karabadji, "Anomaly-based framework for detecting power overloading cyberattacks in smart grid AMI," *Computers & Security*, vol. 96, p. 101896, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404820301693

[114] Z. Zheng, Y. Yang, X. Niu, H. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1606–1615, 2018.

[115] G. Dudgeon, "IEEE 123 Node Test Feeder in Simscape Power Systems," 2021. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/66599-ieee-123-node-test-feeder-in-simscape-power-systems

[116] M. Cupelli, C. Doig Cardet, and A. Monti, "Voltage stability indices comparison on the IEEE-39 bus system using RTDS," in *2012 IEEE International Conference on Power System Technology (POWERCON)*, 2012, pp. 1–6.

[117] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2, pp. 209–220, 2021.

[118] H. Karimipour and H. Leung, "Relaxation-based anomaly detection in cyber-physical systems using ensemble kalman filter," *IET Cyber-Physical Systems: Theory & Applications*, vol. 5, pp. 49–58(9), March 2020.

[119] H. Mohammadi Rouzbahani, H. Karimipour, A. Rahimnejad, A. Dehghantanha, and G. Srivastava, *Anomaly Detection in Cyber-Physical Systems Using Machine Learning*. Cham: Springer International Publishing, 2020, pp. 219–235.

[120] A. K. Singh, B. C. Pal *et al.*, "Report on the 68-bus, 16-machine, 5-area system," *IEEE PES Task Force on Benchmark Systems for Stability Controls. Ver*, vol. 3, 2013.

[121] A. Barua, D. Muthirayan, P. P. Khargonekar, and M. A. Al Faruque, "Hierarchical Temporal Memory Based Machine Learning for Real-Time, Unsupervised Anomaly Detection in Smart Grid: WiP Abstract," in *2020 ACM/IEEE 11th International Conference on Cyber-Physical Systems (ICCPS)*, 2020, pp. 188–189.

[122] H. Yuan, H. Yu, S. Gui, and S. Ji, "Explainability in graph neural networks: A taxonomic survey," *arXiv preprint arXiv:2012.15445*, 2020. [Online]. Available: https://arxiv.org/abs/2012.15445

[123] X. Wang, B. Jin, Y. Du, P. Cui, and Y. Yang, "One-class graph neural networks for anomaly detection in attributed networks," 2020.

[124] A. Deng and B. Hooi, "Graph neural network-based anomaly detection in multivariate time series," in *AAAI-2021*, 2021.

[125] H. Zhao, Y. Wang, J. Duan, C. Huang, D. Cao, Y. Tong, B. Xu, J. Bai, J. Tong, and Q. Zhang, "Multivariate time-series anomaly detection via graph attention network," in *20th IEEE International Conference on Data Mining, ICDM 2020, Sorrento, Italy, November 17-20, 2020*, C. Plant, H. Wang, A. Cuzzocrea, C. Zaniolo, and X. Wu, Eds. IEEE, 2020, pp. 841–850. [Online]. Available: https://doi.org/10.1109/ICDM50108.2020.00093

[126] L. Zheng, Z. Li, J. Li, Z. Li, and J. Gao, "Addgraph: Anomaly detection in dynamic graph using attention-based temporal gcn," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 7 2019, pp. 4419–4425. [Online]. Available: https://doi.org/10.24963/ijcai.2019/614

[127] Q. Xiao, J. Liu, Q. Wang, Z. Jiang, X. Wang, and Y. Yao, "Towards network anomaly detection using graph embedding," in *Computational Science – ICCS 2020*, V. V. Krzhizhanovskaya, G. Závodszky, M. H. Lees, J. J. Dongarra, P. M. A. Sloot, S. Brissos, and J. Teixeira, Eds. Cham: Springer International Publishing, 2020, pp. 156–169.

[128] H. Cheng, P.-N. Tan, C. Potter, and S. Klooster, *Detection and Characterization of Anomalies in Multivariate Time Series*. SIAM, 2009, pp. 413–424. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611972795.36

[129] J. Jiang, J. Chen, T. Gu, K. R. Choo, C. Liu, M. Yu, W. Huang, and P. Mohapatra, "Anomaly detection with graph convolutional networks for insider threat and fraud detection," in *MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM)*, 2019, pp. 109–114.

[130] J. Zhou, Z. Xu, A. M. Rush, and M. Yu, "Automating botnet detection with graph neural networks," *CoRR*, vol. abs/2003.06344, 2020. [Online]. Available: https://arxiv.org/abs/2003.06344

[131] A. Protogerou, S. Papadopoulos, A. Drosou, D. Tzovaras, and I. Refanidis, "A graph neural network method for distributed anomaly detection in iot," *Evolving Systems*, vol. 12, no. 1, pp. 19–36, 2021. [Online]. Available: https://doi.org/10.1007/s12530-020-09347-0

[132] X. Cheng, Y. Wu, G. Min, and A. Y. Zomaya, "Network function virtualization in dynamic networks: A stochastic perspective," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 10, pp. 2218–2232, 2018.

[133] W. Miao, G. Min, Y. Wu, H. Huang, Z. Zhao, H. Wang, and C. Luo, "Stochastic performance analysis of network function virtualization in future internet," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 3, pp. 613–626, 2019.

[134] B. Yin, H. Yin, Y. Wu, and Z. Jiang, "Fdc: A secure federated deep learning mechanism for data collaborations in the internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6348–6359, 2020.

[135] Y. Wu, H. N. Dai, H. Wang, and K. K. R. Choo, "Blockchain-based privacy preservation for 5g-enabled drone communications," *IEEE Network*, vol. 35, no. 1, pp. 50–56, 2021.

[136] H. Huang, Ed., *Blockchains for Network Security: Principles, technologies and applications*, ser. Computing. Institution of Engineering and Technology, 2020. [Online]. Available: https://digital-library.theiet.org/content/books/pc/pbpc029e

[137] Y. Ma, Y. Wu, and J. Ge, *Accountability and Privacy in Network Security*. Springer, 2020.

[138] F. Lopez, M. Saez, Y. Shao, E. C. Balta, J. Moyne, Z. M. Mao, K. Barton, and D. Tilbury, "Categorization of anomalies in smart manufacturing systems to support the selection of detection mechanisms," *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 1885–1892, 2017.

[139] V. Sharma, G. Choudhary, Y. Ko, and I. You, "Behavior and vulnerability assessment of drones-enabled industrial internet of things (iiot)," *IEEE Access*, vol. 6, pp. 43 368–43 383, 2018.

[140] J. Zhao, X. Jing, Z. Yan, and W. Pedrycz, "Network traffic classification for data fusion: A survey," *Information Fusion*, vol. 72, pp. 22–47, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S156625352100018X

[141] A. A. Cook, G. Mısırlı, and Z. Fan, "Anomaly detection for iot time-series data: A survey," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6481–6494, 2020.
[142] K. Zope, K. Singh, S. Nistala, A. Basak, P. Rathore, and V. Runkana, "Anomaly detection and diagnosis in manufacturing systems: A comparative study of statistical, machine learning and deep learning techniques," in *Annual Conference of the PHM Society*, vol. 11, no. 1, 09 2019.
[143] R. Pinto, "M2M using OPC UA ," 2020. [Online]. Available: https://dx.doi.org/10.21227/ychv-6c68
[144] S. Dou, K. Yang, and H. V. Poor, "PC2A: Predicting Collective Contextual Anomalies via LSTM With Deep Generative Model," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9645–9655, 2019.

Yulei Wu [Senior Member] is a Senior Lecturer with the Department of Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom. He received the B.Sc. degree (First Class Honours) in Computer Science and the Ph.D. degree in Computing and Mathematics from the University of Bradford, United Kingdom, in 2006 and 2010, respectively. His expertise is on intelligent networking, and his main research interests include computer networks, networked systems, software defined networks and systems, network management, and network security and privacy. He is an Associate Editor of IEEE Transactions on Network and Service Management, IEEE Transactions on Network Science and Engineering, and IEEE Access, as well as an Area Editor of Computer Networks (Elsevier). He is a Senior Member of the ACM and a Fellow of the HEA (Higher Education Academy).

Hong-Ning Dai [Senior Member] is currently with Faculty of Information Technology at Macau University of Science and Technology as an associate professor. He obtained the Ph.D. degree in Computer Science and Engineering from Department of Computer Science and Engineering at the Chinese University of Hong Kong. His current research interests include Internet of Things, big data, and blockchain technology. He has served as associate editors/editors for IEEE Transactions on Industrial Informatics, IEEE Systems Journal, IEEE Access, Ad Hoc Networks, and Connection Science. He is also a senior member of Association for Computing Machinery (ACM).

Haina Tang is an associate professor at the School of Artificial Intelligence, University of Chinese Academy of Sciences. She received the Ph.D. degree in Computer Software and Theory from the School of Information Science and Technology at Sun Yat-sen University. Her current research interests include network measurement, spatio-temporal data mining, social network analysis and graph mining.