Guest Editorial Introduction to the Special Section on Reloading Feature-Rich Information Networks

THE growing availability of multi-facetedrelational data gives rise to unprecedented opportunities for unveiling complex real-world behaviors and phenomena. This also supports the proliferation of complex network models where the expressive power of the graph-based relational structure is enhanced through exposing several types of features that are peculiar of the domain-specific environment (e.g., social media platforms, biological environment, geographical location). Examples of feature-rich networks include heterogeneous information networks, multi-layer networks, temporal networks, location-aware networks, and probabilistic networks.

The aim of the special section on Reloading Feature-rich Information Networks was to address challenging issues and emerging trends in feature-rich information networks that can be mined in several domains, including not only long studied contexts such as social media, but also less investigated or even new frontiers for network science, such as finance, engineering, geography, and many others. Although the use of feature-rich networks can intuitively be perceived as beneficial for most graph-based research tasks, their expressive power has not been fully valued yet in most domains. Therefore, there is an emergence for providing insights into how the study of complex network models can pave the way for solving domain-specific problems that might not be adequately addressed by existing graph models.

Thanks to the extensive efforts of the reviewers and the great support from the former Editor-in-Chief, Dr. Dapeng Oliver Wu, and currently Editor-in-Chief, Jianwei Huang, we were able to accept 8 contributed articles covering several timely and challenging research topics, ranging from opinion maximization and link prediction, to community detection, from mitigation and restoration in response to cascading node failures, to network anonymization, to new models and applications in finance network, and communication networks. A brief overview of each of the contributions is reported as follows.

In "Behavioral Information Diffusion for Opinion Maximization in Online Social Networks," Hudson and Khamfroush propose a behavioral independent cascade (BIC) diffusion model that accounts for the personalities and opinions of user nodes when computing propagation probabilities for diffusion. The "Big Five" model from the social sciences is adopted to model the personality of users in the network. Under the BIC model, the opinion maximization problem is studied and proved to be NP-hard with a non-submodular objective function. An adaptation of a linear-time iterative approach for efficient influence maximization is then proposed and experimentally evaluated for the proposed opinion maximization problem.

In "Scalable Community Detection for Complex Data Graphs via Hyperbolic Network Embedding and Graph Databases," Tsitseklis *et al.* propose an enhancement of a Girvan-Newman-like community detection approach to capitalize on favorable aspects of the hyperbolic network embedding space and graph databases in order to generate meaningful and accurate communities for arbitrarily large networks and associated attributes. The authors evaluate their approach in the analysis and exploration of RDF data as case in point for linked data from diverse domains.

Graph anonymization aims at reducing the ability of an attacker to identify the nodes of a graph by obfuscating its structural information. In this context, Minello *et al.* address in "k-Anonymity on Graphs using the Szemerédi Regularity Lemma" the k-anonymity problem, which is to make each node indistinguishable from at least other k-1 nodes. They propose an algorithm to enforce k-anonymity based on the Szemerédi regularity lemma, which roughly states that every sufficiently large and dense graph can be approximated by the union of random-like bipartite graphs called regular pairs. The proposed algorithm enables the formation of anonymous groups that are resilient to any type of structural attack while minimizing the structural information loss.

Cascading failures are low-probability high-impact events that affect the vulnerability of interdependent networks, producing substantial social and economic impacts. These phenomena have received attention especially in contexts of critical infrastructure systems, which provide necessary services such as those related to energy, water supply, transportation, communication. In "Failure Mitigation and Restoration in Interdependent Networks via Mixed-integer Optimization," Chen *et al.* propose a new optimization model for determining optimal mitigation and restoration strategies for coupled interdependent networks to preserve and/or restore the maximum flow through the entire networked system, subject to cascading node failures that may be caused by disruptions of a subset of seed nodes.

Link prediction is another challenging problem, especially when coping with scarce network data, or in cold start scenarios. In this context, Güven *et al.* show in "Applying ASP for

Digital Object Identifier 10.1109/TNSE.2021.3073824

Knowledge-Based Link Prediction with Explanation Generation in Feature Rich Networks" how to apply answer set programming (ASP) for formalizing link prediction in feature-rich networks by a hybrid approach. Domain knowledge, which in the ASP setting is implemented in the form of simple logical predicates and rules, can profitably be used to enhance structure-based link prediction leading to knowledge-based feature-rich networks.

In "Multilayer Feature-Rich Satellite Network Analysis: An Application-Oriented and Time-Evolving Approach," Gao and Fang study satellite communication networks based on a multilayer feature-rich model and in accordance with various system and application requirements. More specifically, application-oriented energy consumption of satellite payload is characterized and measured through linear and nonlinear factors, and a physical layer energy model is incorporated with travelling-wave tube amplifiers based energy consumption. A time-evolving analysis approach is developed by formulating multi-objective optimization problems corresponding to the linear and nonlinear models.

In "Advancing Receivable Financing via a Network-based Approach," Bordino *et al.* propose a novel, network-based approach to receivable financing, which is the process whereby cash is advanced to firms against receivables their customers have not yet paid. The key idea in the proposed approach is to enable customers of the same funder to autonomously pay each other as much as possible, thus providing benefits to both the funder (by reduced cash anticipation and exposure risk) and its customers (by smaller fees and lightweight service establishment). The authors contribute with a principled formulation of the network-based receivable-settlement strategy, and show how to achieve all algorithmic challenges posed by the design of this strategy.

The study in "Leveraging the Users Graph and Trustful Transactions for the Analysis of Bitcoin Price" proposed by Crowcroft *et al.* is concerned with the most popular cryptocurrency nowadays, namely Bitcoin, and its exchange rate. They provide a formalization of the Bitcoin user graph, which is a partially deanonymized version of the Bitcoin transaction graph obtained by applying the common input heuristics to the transaction graph extracted from the blockchain. The authors also introduce the novel concept of trustful transactions and their induced graphs to model the special behavior of 0-confirmations transactions. A temporal analysis of the evolution of selected properties of the Bitcoin user graph is conducted, and autoregressive distributed-lag linear regression is applied to assess whether and to what extent a change in the considered features is likely to influence the exchange rate up to a predefined time horizon.

In summary, the contributing articles to this special section have offered advances on feature-rich network modeling and mining, with both timely theoretical and application-oriented studies that help us to enhance our understanding of complex phenomena and problems at convergence of network science, data mining, optimization, and machine learning. We expect that this special section will trigger further development on feature-rich information networks and related topics.

ANDREA TAGARELLI, *Lead Guest Editor*
Department of Computer Engineering, Modeling, Electronics, and Systems Engineering
University of Calabria
87036 Rende (CS), Italy
e-mail: andrea.tagarelli@unical.it

SABRINA GAITO, *Guest Editor*
Department of Computer Science
University of Milan
20122 Milan, Italy
e-mail: gaito@di.unimi.it

ROBERTO INTERDONATO, *Guest Editor*
Cirad UMR TETIS
34398 Montpellier, France
e-mail: roberto.interdonato@cirad.fr

TSUYOSHI MURATA, *Guest Editor*
Department of Computer Science
Tokyo Institute of Technology
Tokyo 152-8550, Japan
e-mail: murata@c.titech.ac.jp

ALESSANDRA SALA, *Guest Editor*
Shutterstock
Dublin 2, Ireland
e-mail: asala@shutterstock.com

MY T. THAI, *Guest Editor*
University of Florida
Gainesville, FL 32611, USA
e-mail: mythai@cise.ufl.edu

Andrea Tagarelli is currently an Associate Professor of computer engineering with the Department of Computer Engineering, Modeling, Electronics, and Systems Engineering, University of Calabria, Rende, Italy, where he leads the effort on data science and machine learning for social network and media. His research interests include data mining, machine or deep learning, web and network science, and information retrieval.

Sabrina Gaito (Member, IEEE) is currently an Associate Professor with the Computer Science Department, University of Milan, Milan, Italy, where she teaches social media mining, network science, and artificial intelligence. Her research interests include network science and data mining applied to social media.

Roberto Interdonato is currently a Research Scientist with Cirad, UMR TETIS, Montpellier, France. His research interests include data mining, machine learning applied to complex network analysis, which include social media networks, trust networks, semantic networks, and bibliographic networks, and to remote sensing.

Tsuyoshi Murata is currently a Professor with the Department of Computer Science, Tokyo Institute of Technology, Tokyo, Japan. He is doing research on artificial intelligence, especially complex networks, machine learning, and data mining.

Alessandra Sala is currently the Director of AI and Data Science at Shutterstock, Dublin, Ireland. Formerly, she worked for Bell Labs in Ireland, as head of Analytics Research.

My T. Thai (Fellow, IEEE) is currently a UF Research Foundation Professor of computer and information sciences and engineering, and an Associate Director of the Nelms Institute for the Connected World, University of Florida, Gainesville, FL, USA. Her research interests include the optimization and algorithmic mathematics underpinnings of billion-scale data mining, machine learning, and cybersecurity.