

COMPLEX
NETWORKS

COMPLEX NETWORKS 2022

THE 11TH INTERNATIONAL CONFERENCE
ON COMPLEX NETWORKS
AND THEIR APPLICATIONS

08 - 10 November, 2022
Palermo, Italy

BOOK OF ABSTRACTS

COMPLEX NETWORKS 2022

The 11th International Conference on Complex Networks & Their Applications

08 - 10 November, 2022 Palermo, Italy

Published by the International Conference on Complex Networks & Their Applications

Editors

Hocine Cherifi

University of Burgundy, France

Rosario Nunzio Mantegna

Università degli Studi di Palermo, Italy

Luis Mateus Rocha

Indiana University, USA

Chantal Cherifi

University of Lyon, France

Salvatore Micciché

Università degli Studi di Palermo, Italy

COMPLEX NETWORKS 2022

e-mail: hocine.cherifi@u-bourgogne.fr

Copyright Notice COMPLEX NETWORKS 2022 and the Authors

This publication contributes to the Open Access movement by offering free access to its articles and permitting any users to read, download, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software. The copyright is shared by authors and the 10th International Conference on Complex Networks & Their Applications (COMPLEX NETWORKS 2021) to control over the integrity of their work and the right to be properly acknowledged and cited.

To view a copy of this license, visit <http://www.creativecommons.org/licenses/by/4.0/>

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use. While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained her.

ISBN: 978-2-9557050-6-3

Preface

The “Università degli Studi di Palermo” in Italy hosts the eleventh edition of the “International Conference on Complex Networks & their Applications” (COMPLEX NETWORKS 2022) from November 8 to November 10, 2022. Every year, COMPLEX NETWORKS brings together researchers from various scientific backgrounds to review the field’s current state and formulate new directions. The diversity of the attendees’ scientific interests (Finance, Medicine and Neuroscience, Biology and Earth Sciences, Sociology and Politics, Computer Science and Physics, etc.) is a unique opportunity for cross-fertilization between fundamental issues and innovative applications.

The quality of the contributors is undoubtedly an essential element for a successful edition. The success also goes to the keynote speakers. These leaders and visionaries in their fields present fascinating plenary lectures with big-picture ideas and unique perspectives to help attendees deepen their understanding of scientific challenges. We are delighted to bring together this great line-up of speakers.

- Luís A. NUNES AMARAL (Northwestern University, USA)
- Manuel CEBRIAN (Max Planck Institute for Human Development, Germany)
- Shlomo HAVLIN (Bar-Ilan University, Israel)
- Giulia IORI (City, University of London, UK)
- Melanie MITCHELL (Santa Fe Institute, USA)
- Ricard SOLÉ (Universitat Pompeu Fabra, Spain)

Our thanks also go to the speakers of the traditional tutorial sessions for delivering insightful talks on November 7, 2022.

- Michele COSCIA (IT University of Copenhagen, Denmark)
- Adriana IAMNITCHI (Maastricht University, Netherlands)

The success also relies in the deep involvement of many individuals, institutions, and sponsors.

We sincerely gratify the advisory board members for inspiring the essence of the conference:

Jon Crowcroft (University of Cambridge), Raissa D’Souza (University of California, Davis, USA), Eugene Stanley (Boston University, USA), and Ben Y. Zhao (University of Chicago, USA)

We record our thanks to our fellow members of the organizing committee:

The lightning sessions chairs:

Alessandro Rizzo (Politecnico di Torino, Italy), Giancarlo Francesco Ruffo (Università degli Studi di Torino, Italy), and Huijuan Wang (TU Delft, Netherlands)

III

The poster sessions chairs:

Manuel Marques Pita (Universidade Lusófona, Portugal), Michele Tumminello (Università degli Studi di Palermo, Italy), Laura Ricci (Università degli Studi di Pisa, Italy)

The tutorial chairs:

Luca Maria Aiello (Nokia-Bell Labs, UK) and Leto Peel (Maastricht university, Netherland),

The special issue chair:

Sabrina Gaito (University of Milan, Italy),

The publicity chairs:

Christophe Cruz (University of Burgundy, France), Xiangjie Kong (Zhejiang University of Technology, China)

The sponsor chairs:

Claudio Fazio (Università degli Studi di Palermo, Italy) and Roberto Interdonato (CIRAD - UMR TETIS, Montpellier, France)

The Web chairs:

Zakariya Ghalmane (LINEACT, CESI, France) and Stephany Rajeh (University of Burgundy, France),

Our profound thanks go to Matteo Zignani (University of Milan, Italy), publication chair, for the tremendous work in managing the submission system and the proceedings publication process.

We would also like to record our appreciation for the work of the local committee chair, Federico Musciotto (Università degli Studi di Palermo, Italy), and all the local committee members, Yuri Antonacci (Università degli Studi di Palermo, Italy), Giosuè Lo Bosco (Università degli Studi di Palermo, Italy), Luca Faes (Università degli Studi di Palermo, Italy), Giacomo Fiumara (Università degli Studi di Messina, Italy), Vincenzo Giuseppe Genova (Università degli Studi di Palermo, Italy), Matteo Milazzo (University of Catania, Italy), Pasquale de Meo (Università degli Studi di Messina, Italy), Alessandro Pluchino (Università degli Studi di Catania, Italy), Andrea Rapisarda (Università degli Studi di Catania, Italy), for their work in managing the sessions. They intensely participated to the success of this edition.

We would like to express our gratitude to our partner journals: *Advances in Complex Systems*, *Applied Network science*, *Complex Systems*, *Entropy*, *Plos One* and *Social Network Analysis and Mining*.

We are thankful to all those who have contributed to the success of this meeting. Sincere thanks to the authors for their creativity.

Finally, we would like to express our most sincere thanks to the program committee members for their considerable efforts in producing high-quality reviews in a minimal time.

These volumes make the most advanced contribution of the international community to the research issues surrounding the fascinating world of complex networks. Their breath, quality, and novelty demonstrate the profound impact of complex networks in understanding our world. We hope you enjoy the papers as much as we enjoyed organizing the conference and putting this collection of articles together.

Hocine Cherifi Rosario N. Mantegna Luis Mateus Rocha Chantal Cherifi Salvatore Micciche'

Hocine Cherifi
Rosario N. Mantegna
Luis Mateus Rocha
Chantal Cherifi
Salvatore Micciché

Table of Contents

Tutorials

Node Vector Distances: Methods and Applications	2
<i>Michele COSCIA</i>	
Modelling Information Diffusion in Social Media: Data-Driven Observations	3
<i>Adriana IAMNITCHI</i>	

Invited Speakers

An opinionated evaluation of where we are after a quarter century of complex networks	5
<i>Luís A. NUNES AMARAL</i>	
Networked and Crowdsourced Response to Time-critical Threats	6
<i>Manuel CEBRIAN</i>	
Network Science and Applications	7
<i>Shlomo HAVLIN</i>	
Information diffusion in trading networks	9
<i>Giulia IORI</i>	
Why AI is Harder Than We Think	11
<i>Melanie MITCHELL</i>	
Evolutionary transitions: universality, complexity and predictability	12
<i>Ricard SOLÉ</i>	
Boundary effects in diffusion of new products	14
<i>Gadi Fibich, Tomer Levin and Kenneth Gillingham</i>	
Early warning signals of multistage state transitions on complex networks	17
<i>Neil Maclaren, Prosenjit Kundu and Naoki Masuda</i>	
Socially-enhanced discovery processes	20
<i>Gabriele Di Bona, Enrico Ubaldi, Iacopo Iacopini, Bernardo Monechi, Vito Latora and Vittorio Loreto</i>	
The Biased-Voter model: How persuasive a small group can be?	23
<i>Christos Charalambous, Agnieszka Czaplicka, Raul Toral and Maxi San Miguel</i>	

Gender and the influence of the department network in topic selection on early career faculty	26
<i>Lluís Danus, Robert H. Davis, Marta Sales-Pardo and Roger Guimera</i>	
Network Topologies of Corporate Organization Charts and Their Correlation with Corporate Performance	29
<i>Hiroki Sayama and Junichi Yamanoi</i>	
Random Hypergraph Models Preserving Degree Correlation and Local Clustering	32
<i>Kazuki Nakajima, Kazuyuki Shudo and Naoki Masuda</i>	
Decomposing networks of cross-correlations with q-MST: an example of cryptocurrency market	35
<i>Stanisław Drożdż, Jarosław Kwapien and Marcin Wątorek</i>	
Percolation in heterogeneous spatial networks with long-range interactions	38
<i>Guy Amit, Dana Vaknin Ben Porath, Sergey Buldyrev and Amir Bashan</i>	
Validity of a one-dimensional reduction of dynamical systems on networks	41
<i>Prosenjit Kundu, Hiroshi Kori and Naoki Masuda</i>	
Resonating power flows across highly renewable electricity networks	44
<i>Mathias Beltoft and Martin Greiner</i>	
Exact solution of heterogeneous Markovian SI epidemics on networks	45
<i>Massimo Achterberg and Piet Van Mieghem</i>	
Robustness of Noisy Quantum Networks	48
<i>Bruno Coutinho, William J. Munro, Kae Nemoto and Yasser Omar</i>	
Fixation probability of switching networks	51
<i>Jnanajyoti Bhaumik and Naoki Masuda</i>	
The Frechet distribution in drone networks	54
<i>Maria Raftopoulou, Remco Litjens and Piet Van Mieghem</i>	
Analysis of Perturbed Networks	57
<i>Andras Farago</i>	
Floquet Theory for Spreading Dynamics over Periodically Switching Networks	60
<i>Maisha Islam Sejunti, Naoki Masuda and Dane Taylor</i>	
Fire Together, Wire Together?	64
<i>Gail Gilboa Freedman</i>	

VII

Maximum information entropy production for network dynamics modelling	67
<i>Noam Abadi, Diego Garlaschelli and Franco Ruzzenenti</i>	
Quantifying Biobank Impact	71
<i>Rodrigo Dorantes Gilardi, John Michael Gaziano and Albert-László Barabási</i>	
Extremists of a Feather: Homophily in Violent Extremist Networks	75
<i>Anina Schwarzenbach and Michael Jensen</i>	
Extended abstract: Analysis of the competition among viral strains using a temporal interaction-driven contagion model	78
<i>Alex Abbey, Yuval Shahar and Osnat Mokryn</i>	
In temporal networks correlation lags are informative of approaching bifurcations	81
<i>Giulio Tirabassi and Cristina Masoller</i>	
Outlier mining in high-dimensional datasets based on Jensen-Shannon distances and graph structure analysis	84
<i>Alex S. O. Toledo, Riccardo Silini, Laura Carpi and Cristina Masoller</i>	
From time series to networks in R with the ts2net package	87
<i>Leonardo Ferreira</i>	
Territorial Development as an innovation driver: a complex network approach	90
<i>Francesco De Nicolò, Alfonso Monaco, Giuseppe Ambrosio, Loredana Bellantuono, Roberto Cilli, Angela Lombardi, Ester Pantaleo, Sabina Tangaro, Flaviano Zandonai, Nicola Amoroso and Roberto Bellotti</i>	
Enabling Unmanned Aerial Vehicles Operations over Cities	93
<i>William Roberto de Paiva and Andre Franceschi De Angelis</i>	
The Global Network of Embodied Nitrogen Flows - Across Countries and Sectors	96
<i>Kaan Hidiroglu and Franco Ruzzenenti</i>	
Disentangling activity-aware human flows reveals the hidden functional organization of urban systems	101
<i>Riccardo Gallotti, Giulia Bertagnolli and Manlio De Domenico</i>	
Functional alignments in brain connectivity networks	104
<i>Ruaridh Clark, Mario Parra, Keith Smith, Javier Escudero, Agustin Ibanez, William McGeown and Malcolm Macdonald</i>	

Redundancy in the Structure and Dynamics of Complex Systems	107
<i>Luis M. Rocha</i>	
Influence of heterogeneous age-group contact patterns on critical vaccination rates for herd immunity to SARS-CoV-2	109
<i>Joan Saldana and Caterina Scoglio</i>	
The valuation of information by dynamic decentralised criticality measures in complex data flow networks	112
<i>Yaniv Proselkov, Manuel Herrera, Marco Perez Hernandez, Ajith Kumar Parlikad and Alexandra Brintrup</i>	
Evaluating network membership with an application to the European Research Programmes	115
<i>Matteo Cinelli, Giovanna Ferraro and Antonio Iovanella</i>	
Disentangling the Growth of Web3 Blockchain-based Networks by Graph Evolution Rules	118
<i>Alessia Galdeman, Maria Pia Chiatante, Matteo Zignani and Sabrina Gaito</i>	
Does heterogeneity slow down the diffusion of new products?	121
<i>Gadi Fibich and Amit Golan</i>	
Attribute-aware Community Events in Feature-rich Dynamic Networks . .	124
<i>Andrea Failla, Federico Mazzoni and Salvatore Citraro</i>	
Distance Dependent Bayesian Community Detection	127
<i>Isabelle Beaudry, Marcelo Mendoza, Naim Bro, Fabian Sepúlveda and Hernán Sarmiento</i>	
The Flow of Corporate Control in the Global Ownership Network	130
<i>Takayuki Mizuno, Shohei Doi and Shuhei Kurizaki</i>	
Executive women’s networks: the affinity rule of social capital	133
<i>Gabriela Contreras, Ruth Mateos de Cabo and Ricardo Gimeno</i>	
N-actors conflict model on scale-free networks	136
<i>Eduardo Jacobo-Villegas, Bibiana Obregón-Quintana and Lev Guzmán-Vargas</i>	
Correlation tensor spectra of XRP transaction networks	139
<i>Abhijit Chakraborty, Tetsuo Hatsuda and Yuichi Ikeda</i>	
Stochastic approximation of Adaptive Voter models	142
<i>Emmanuel Kravitzch, Vineeth S. Varma, Yezekael Hayel and Antoine O. Berthet</i>	
Estimation of flow trajectories in a multi-lines network	145
<i>Guillaume Guex, Romain Loup and François Bavaud</i>	

IX

COVID-19 impact on international trade	148
<i>Célestin Coquidé, José Lages, Leonardo Ermann and Dima L Shepelyansky</i>	
Network structures of a centralised and a decentralized market. A direct comparison	151
<i>Sylvain Mignot and Annick Vignes</i>	
COVID-19 Spreading Dynamics via Air Transportation Network: A Case Study in China	155
<i>Xinyue Chen, Shuai Liu and Xiaoqian Sun</i>	
Delving into Individual Interactions in the IETF Time Evolving Social Graph	159
<i>Matthew Barnes, Richard Clegg and Vincenzo Nicosia</i>	
Centrality in the Macroeconomic Multi-Network Explains the Spatio-Temporal Evolution of Country Per-Capita Income	162
<i>Giorgio Fagiolo and Davide Samuele Luzzati</i>	
Community recovery from temporal and real-valued network data	165
<i>Konstantin Avrachenkov, Maximilien Drevetton and Lasse Leskelä</i>	
Structural analysis of SARS-CoV-2 Spike Protein variants through Graph Embedding.	168
<i>Pietro Hiram Guzzi, Ugo Lomoio, Barbara Puccio and Pierangelo Veltri</i>	
Hierarchical communities in complex networks	172
<i>Michael Schaub, Jiaze Li and Leto Peel</i>	
Detectability of hierarchical communities in networks	175
<i>Leto Peel and Michael Schaub</i>	
Unveiling the higher-order organization of multivariate time series	178
<i>Andrea Santoro, Federico Battiston, Giovanni Petri and Enrico Amico</i>	
Estimating Affective Polarization on a Social Network	180
<i>Marilena Hohmann and Michele Coscia</i>	
Exploring Glioblastoma multiforme through Network Analysis highlights topologically critical nodes	183
<i>Apurva Badkas, Sébastien De Landtsheer and Thomas Sauter</i>	
Vaccine hesitancy in Twitter (spanish language)	186
<i>Israel Morales-Trigueros, Carlos Adolfo Piña García and Ruben Flores Gonzalez</i>	

Exploring topics in LDA models through Statistically Validated Networks: directed and undirected approaches	189
<i>Alessandro Albano, Mariangela Sciandra and Antonella Plaia</i>	
Interplay between network structure and function in connectome-inspired Reservoir Computing	192
<i>Valeria D'Andrea, Michele Puppim and Manlio De Domenico</i>	
Assessing the Role of Influencers in Tweet Propagation	195
<i>Joel Nilsson, Alberto Zenere and Claudio Altafini</i>	
Vector-based Pedestrian Navigation in Cities	198
<i>Christian Bongiorno, Yulun Zhou, Marta Kryven, David Theurel, Alessandro Rizzo, Paolo Santi, Joshua Tenenbaum and Carlo Ratti</i>	
Mainshock Identification by community detection techniques	201
<i>Aditi Seal, Swarandeeep Sahoo and Prosanta Kumar Khan</i>	
Comparison of metrics for measuring the editor scatteredness and article complexity on Wikipedia	204
<i>Fumiko Ogushi</i>	
The Evolution of the Covid-19 Vaccine Debate on Twitter	206
<i>Arnaldo Santoro, Alessandro Galeazzi, Teresa Scantamburlo and Fabiana Zollo</i>	
Non-Monotonic Dynamics in Node Ranking	209
<i>Shahar Somin, Yaniv Altshuler, Alex Pentland and Erez Shmueli</i>	
Excess closure in a multilayer population-scale social network	213
<i>Eszter Bokanyi, Eelke Heemskerk, Yuliia Kazmina and Frank Takes</i>	
Epidemic evolution in a dynamic multilayer network to study the spread of antibiotic resistance	216
<i>Paola Stolfi, Davide Vergni, Emiliano Mancini and Filippo Castiglione</i>	
Model-independent hyperbolic embedding of directed networks	219
<i>Bianka Kovács and Gergely Palla</i>	
Law and Success: A Complexity Perspective	222
<i>Michal Shur-Ofry</i>	
The distribution of cover times of random walks on random regular graphs	225
<i>Ofer Biham, Ido Tishby and Eytan Katzav</i>	
Cities As Complex Neural Networks: A Case Study For House Price Estimation	228
<i>Julien Randon-Furling and Sarah Soleiman</i>	

Directed percolation in temporal networks	231
<i>Arash Badie-Modiri, Abbas K. Rizi, Márton Karsai and Mikko Kivelä</i>	
Systemic Risk in Decentralized Finance (DeFi)	233
<i>Stefan Kitzler, Christian Diem, Pietro Saggese, Bernhard Haslhofer and Stefan Thurner</i>	
Opinion polarization as a natural consequence of beliefs being interrelated	236
<i>Anna Zafeiris</i>	
Graph Partitions in Chemistry	240
<i>Ioannis Michos and Vasilios Raptis</i>	
Non-monotonic transients to synchrony in Kuramoto networks and electrochemical oscillators	244
<i>Oleh Omel'Chenko, Jorge Luis Ocampo-Espindola and Istvan Z. Kiss</i>	
Partial synchronization in neural networks: chimeras and beyond	247
<i>Anna Zakharova</i>	
Effects of online social capital on academic performance	248
<i>Agusti Canals, Josep Cobarsí-Morales and Eva Ortoll</i>	
Correlation Networks measure Technology Evolution on Stack Overflow . .	251
<i>Conrad Borchers and Fabian Braesemann</i>	
Universality and phase transitions in contracting networks	254
<i>Eytan Katzav, Ido Tishby, Barak Budnick and Ofer Biham</i>	
Multivariate permutation entropy via the Cartesian graph product to analyse two-phase flow	257
<i>John Stewart Fabila-Carrasco, Chao Tan and Javier Escudero</i>	
Resilience of Coupled Social and Technical Networks in Open Source Software	260
<i>Johannes Wachs and William Schueller</i>	
Relatedness, complexity, and growth: The relevance of the value-added approach in CEE countries	263
<i>Tamás Sebestyén, Zoltán Elekes and Erik Braun</i>	
Networks of Causal Relationships in Financial Markets	267
<i>Vladimir Boginski, Oleg Shirokikh, Grigory Pastukhov, Alexander Semenov, Sergiy Butenko, Alexander Veremyev and Eduardo Pasiliao</i>	

Structural biases in university rankings: a complex network approach to bridge the gap	270
<i>Loredana Bellantuono, Alfonso Monaco, Nicola Amoroso, Vincenzo Aquaro, Marco Bardoscia, Annamaria Demarinis Loiotile, Angela Lombardi, Sabina Tangaro and Roberto Bellotti</i>	
From the connectome to action: emergent dynamics in a robotic model based on the connectome of the nematode <i>C. elegans</i>	273
<i>Pablo M. Gleiser, Carlos E. Valencia Urbina and Sergio A. Cannas</i>	
Social Networks in Times of Turkey's Currency Crisis	276
<i>N. Gizem Bacaksizlar Turbic, Melih Can Yardı, Özgem Elif Acar, Oğuz Gürerk, Ali Hürriyetöglü, Tolga Etgü and Erdem Yörük</i>	
Solitary, cluster and chimera states in globally coupled nonlinear oscillators	280
<i>Leonhard Schülen, Alexander Gerdes, Matthias Wolfrum and Anna Zakharova</i>	
The impact of diversity on group performance	283
<i>Fabian Baumann, Agnieszka Czaplicka and Iyad Rahwan</i>	
Reaction impurity prediction using a data mining approach	286
<i>Adarsh Arun, Zhen Guo, Simon Sung and Alexei Lapkin</i>	
Interspecific competition affects the structural stability of mutualistic networks	289
<i>Johan Dubbeldam</i>	
Testing the dynamics of online polarization	291
<i>Matteo Cinelli, Carlo Valensise and Walter Quattrociocchi</i>	
Propagation of disruptions in supply networks of essential goods: A population-centered perspective of systemic risk	294
<i>William Schueller, Christian Diem, Melanie Hinterplattner, Johannes Stangl, Beate Conrady, Markus Gerschberger and Stefan Thurner</i>	
Networks for smoking dynamics	297
<i>Adarsh Prabhakaran, Valerio Restocchi and Benjamin Goddard</i>	
Double-Threshold Models for Network Influence Propagation	300
<i>Vladimir Boginski, Alexander Semenov, Alexander Veremyev and Eduardo Pasiliao</i>	
An information-theoretic approach to hypergraph psychometrics	303
<i>Daniele Marinazzo, Fernando Rosas, Massimo Stella, Renzo Comolatti, Nigel Colenbier, Sebastiano Stramaglia and Yves Rosseel</i>	

Graph clustering of large trade network: case study of Russian wood industry	306
<i>Nikita Sinitsyn, Mikhail Elutin, Andrei Korotkov, Artem Kushlevich, Roman Menyalshikov and Dmitriy Rusakov</i>	
Network theory as a novel tool to analyze cardiac arrhythmia	311
<i>Sirius Fuenmayor, Enid Van Nieuwenhuyse, Sander Hendrickx, Robin Van den Abeele, Arthur Santos Bezerra, Sebastien Knecht, Mattias Duytschaever, Daniele Marinazzo and Nele Vandersickel</i>	
Noise Transmission in Layered Complex Networks	316
<i>Melvyn Tyloo</i>	
A comparative analysis of multiplex phonological and orthographic networks	319
<i>Pablo A. Lara-Martinez, Bibiana Obregon-Quintana and Lev Guzman-Vargas</i>	
A Path-Based Approach to Analyzing the Global Liner Shipping Network	322
<i>Timothy LaRock, Mengqiao Xu and Tina Eliassi-Rad</i>	
The orthoBackbone: an evolutionarily-conserved backbone of genetic networks	325
<i>Rion Brattig Correia, Paulo Navarro Costa and Luis M. Rocha</i>	
The metric backbone preserves community structure and is a primary transmission subgraph in contact networks	328
<i>Rion Brattig Correia, Alain Barrat and Luis M. Rocha</i>	
Influence maximization in Boolean networks	331
<i>Thomas Parmer, Filippo Radicchi and Luis Rocha</i>	
Analysis of the Structure and Dynamics of European Flight Networks . . .	333
<i>Matteo Milazzo, Federico Musciotto, Salvatore Miccichè and Rosario Nunzio Mantegna</i>	
Low-order descriptors of high-order dependencies	336
<i>Tomas Scagliarini, Davide Nuzzi, Yuri Antonacci, Luca Faes, Fernando E. Rosas, Daniele Marinazzo and Sebastiano Stramaglia</i>	
Ecological Networks and State Shifts in the Earth's Biosphere	341
<i>Sabin Roman and Francesco Bertolotti</i>	
The Unconstrained LFR Benchmark	342
<i>Bojan Evkoski, Igor Mozetic and Petra Kralj Novak</i>	
The Russian invasion of Ukraine through the lens of ex-Yugoslavian Twitter	345
<i>Bojan Evkoski, Igor Mozetic, Petra Kralj Novak and Nikola Ljubesic</i>	

Link prediction in blockchain online social networks with textual information	348
<i>Manuel Dileo, Cheick Tidiane Ba, Matteo Zignani and Sabrina Gaito</i>	
A Spinglass Model of Video Recorded Street Violence	351
<i>Jeroen Bruggeman, Don Weenink and Bram Mak</i>	
What deep learning can bring to two decades of correlations, hierarchies, networks and clustering in financial markets	355
<i>Gautier Marti</i>	
Ecological validation of soil food-web robustness for managed grasslands .	358
<i>Letizia Stella Di Mauro, Alessandro Pluchino, Erminia Conti and Christian Mulder</i>	
Birth-Death Processes Reproduce the Epidemic Footprint	360
<i>Gerrit Großmann and Michael Backenköhler</i>	
Robustness of Social Cohesion in Growing Groups.....	363
<i>Enrico Maria Fenoaltea, Fanyuan Meng, Run-Ran Liu and Matus Medo</i>	
Understanding the Evolution of Reddit in Temporal Networks induced by User Activity	366
<i>Nora Marcoux, Pedro G Lind, Johannes Langguth, Andreas Huber and Daniel Thilo Schroeder</i>	
Dynamics of <i>Drosophila melanogaster</i> social interaction networks	370
<i>Milan Petrovic, Ana Filošević Vujnović and Ana Meštrović</i>	
Quantifying complexity and similarity of chess openings using online chess communities data	373
<i>Giordano De Marzo and Vito Dp Servedio</i>	
Weighted network motifs as random walk patterns	376
<i>Francesco Picciolo, Franco Ruzznenti, Petter Holme and Rossana Mastrandrea</i>	
Which sparrows are making a summer? The role of conformism and network structure in shaping cooperative behavior.	379
<i>Tamás Sebestyén and Emese Braun</i>	
A novel representation for higher order networks	383
<i>Tomaso Aste and Yuanrong Yuanrong</i>	
Application of Non-negative Matrix and Tensor Factorizations to Money Flows in Economic Networks.....	386
<i>Yoshi Fujiwara</i>	

Consensus clustering using projective distance between partitions	389
<i>Boris Mirkin and Andrei Sitnikov</i>	
Inferring the transmission dynamics of Avian Influenza from news and environmental data	392
<i>Nejat Arınik, Roberto Interdonato, Mathieu Roche and Maguelonne Teisseire</i>	
Optimizing robustness of modular networks based on reinforced nodes . . .	395
<i>Yael Kfir-Cohen, Dana Vaknin Ben Porath and Shlomo Havlin</i>	
Nonlinear machine learning pattern recognition and bacteria-metabolite multilayer network analysis of perturbed gastric microbiome	399
<i>Claudio Duran and Carlo Vittorio Cannistraci</i>	
Universal growth of social groups	401
<i>Ana Vranić, Jelena Smiljanić and Marija Mitrović Dankulov</i>	
Ensemble of Opinion Dynamics Models to Understand the Role of the Undecided about Vaccines	404
<i>Jacopo Lenti and Giancarlo Ruffo</i>	
A technique for inducing multiconsensus exploiting network symmetries in multi-agent systems	407
<i>Cinzia Tomaselli, Lucia Valentina Gambuzza, Francesco Sorrentino and Mattia Frasca</i>	
Edge-snapping in complex networks with limited resources	410
<i>Alessandra Corso, Lucia Valentina Gambuzza, Pietro De Lellis and Mattia Frasca</i>	
Shedding a light on the ESG risk factor: a multiplex approach	413
<i>Davide Stocco and Francesca Grassetti</i>	
Regional Trade Network of the EU27 for Medical Products in the Fight against the Pandemic: Policy Implications for Self Sufficiency in Critical Inputs	417
<i>Semanur Soygyigit and Sheri Markose</i>	
Academic Support Network Reflects Doctoral Experience and Productivity	421
<i>Ozgur Can Seckin and Onur Varol</i>	
Calculation of Information Cascade Final Size Distributions	425
<i>Kazumasa Oida</i>	
Formation and Development of Echo Chambers on Twitter under the Pandemic	429
<i>Hiroshi Iyetomi and Wataru Souma</i>	

Real-world data in rheumatoid arthritis: patient similarity networks as a tool for clinical evaluation of disease activity	432
<i>Ondřej Janča, Eva Kriegová, Pavel Horák, Martina Skácelová, Jakub Savara, Anna Petráčková and Miloš Kudělka</i>	
The Price of a Skill: Why the Value of a Skill Depends on its “Neighbours”	435
<i>Fabian Stephany, Ole Teutloff, Georg von Richthofen and Hendrik Send</i>	
Topological-temporal properties of evolving high order networks	438
<i>Alberto Ceria and Huijuan Wang</i>	
Revealing Beyond-Pairwise Functional Connectivity Structure via Multivariate Entropy Decomposition of Human Neuroimaging Data	441
<i>Thomas Varley, Maria Pope, Joshua Faskowitz and Olaf Sporns</i>	
Spatio-Temporal social network: a Case Study of Children’s Behaviour in the Schoolyard	446
<i>Maedeh Nasri, Yung-Ting Tsou, Mitra Baratchi, Alexander Koutamanis, Sarah Giest and Carolien Rieffe</i>	
Users susceptibility in online social media	450
<i>Gianluca Nogara, Luca Luceri, Stefano Cresci and Silvia Giordano</i>	
Sandpiles in Networks with Variable Topology: reconnected 2D grid	453
<i>Victor Munoz and Alejandro Zamorano</i>	
Multivariate Information Theory Uncovers Synergistic Subsystems of the Human Cerebral Cortex	456
<i>Maria Pope, Thomas F. Varley, Joshua Faskowitz and Olaf Sporns</i>	
Is consensus robust against nodes removal in the Deffuant model during homophily scale-free growing?	460
<i>Yerali Gandica, Omid Roozmand, Floriana Gargiulo and Guillaume Deffuant</i>	
Network motifs of the human temporal cortex and their computational implications	463
<i>Yangfan Peng, Antje Bjelde, Pau Vilimelis Aceituno, Benjamin Grewe and Jörg Geiger</i>	
Blending Machine Learning and mechanistic models to learn ecological time series	464
<i>Victor Boussange, Pau Vilimelis Aceituno and Loïc Pellissier</i>	
Methods Evaluation - Missing Data in Age-stratified Contacts Predictions	465
<i>Ramona Ottow</i>	

XVII

The extremely modular structure of growing hyperbolic networks	469
<i>Gáspár Sámuel Balogh, Bianka Kovács and Gergely Palla</i>	
Analyzing the commentator network within the French YouTube Environment	472
<i>Kurt Maxwell Kusterer, Sylvain Mignot and Annick Vignes</i>	
Local Assortativity in Weighted and Directed Complex Networks	476
<i>Uta Pigorsch and Marc Sabek</i>	
The adjacent possible and the dynamics of meaning	479
<i>Peter Persoon and Pantelis Koutroumpis</i>	
Political polarization: Persistent hatred or issue-dependent sidings?	483
<i>Letizia Iannucci, Ali Faqeeh, Ted Hsuan Yun Chen and Mikko Kivela</i>	
Network Specialization: A Topological Mechanism for the Emergence of Cluster Synchronization	486
<i>Benjamin Webb, Erik Hannesson, Jordan Sellers and Ethan Walker</i>	
Building co-morbidity networks via Bayesian network reconstruction	491
<i>Guillermo Romero Moreno, Valerio Restocchi and Jacques Fleuriot</i>	
The way to webify the financial system in China—from individual credit reporting	494
<i>Jing Zhang</i>	
The sample size value in Network Medicine: an application of gene co-expression networks	496
<i>Joaquim Aguirre-Plans, Bingsheng Chen, Deisy Morselli Gysi and Albert-Laszlo Barabasi</i>	
Sparsification of deep neural networks before training using complex network analysis	499
<i>Cameron Taylor, Qihang Yao, Mustafa Burak Gurbuz, Shreyas Malakarjun Patil and Constantine Dovrolis</i>	
Does modular function result in modular structure in sparse deep neural networks?	503
<i>Shreyas Malakarjun Patil, Loizos Michael and Constantine Dovrolis</i>	
Friendship Modulates Hierarchical Relations in Public Elementary Schools	507
<i>Cristian Candia, Melanie Oyarzun and Carlos Rodriguez-Sickert</i>	
Analysis of Oscillatory Time series using the Visibility Graph method . . .	510
<i>Elizabeth Garcés and Víctor Muñoz</i>	

Persistent homology to analyse disruptions of functional and effective brain connectivity	513
<i>Jaroslav Hlinka, Anna Pidnebesna and Luigi Caputi</i>	
Aberrant change in brain network flexibility during the performance of Theory of Mind task in schizophrenia patients	515
<i>Narges Chinichian, Ahmad Ehyaei, Pablo Reinhardt and Henrik Walter</i>	
Exploring the energy complexity nexus in the production space with network theory	530
<i>Franco Ruzzenenti, Noam Abadi, Kas Jansma and Rossana Mastrandrea</i>	
Detecting network hyper-motifs in real networks and their emergent properties	533
<i>Miri Adler and Ruslan Medzhitov</i>	

Exploring topics in LDA models through Statistically Validated Networks: directed and undirected approaches

Alessandro Albano¹, Mariangela Sciandra¹, and Antonella Plaia¹

Department of Economics, Business and Statistics, University of Palermo, Viale delle Scienze,
Edificio 13, Palermo, Italy.

alessandro.albano@unipa.it, mariangela.sciandra@unipa.it,
antonella.plaia@unipa.it

1 Introduction

Probabilistic topic models are machine learning tools for processing and understanding large text document collections. Among the different models in the literature, Latent Dirichlet Allocation (LDA) [1] has turned out to be the benchmark of the topic modelling community. The key idea is to represent text documents as random mixtures over latent semantic structures called *topics*. Each topic follows a multinomial distribution over the vocabulary words. In order to understand the result of a topic model, researchers usually select the top- n (*essential words*) words with the highest probability given a topic and look for meaningful and interpretable semantic themes [2], [3], [4].

This work proposes a new method for exploring topics in LDA models, using Statistically Validated Networks (SVNs). The main idea of the proposed method is to consider co-occurrence between essential words as a measure of association. Two different approaches, called *undirected* and *directed* are proposed. Firstly, the symmetrical association between two words is taken into account, i.e. how many times two words are found in the same sentence. Conversely, in the directed approach, the order in which the words are in the sentence is also considered.

We use hypothesis testing to assess whether the co-occurrence between two words can be attributed to the chance or if these links carry relevant information about the structure of topics. Specifically, textual data is represented as a *bipartite network* in which one set of nodes is made by sentences, and the other set of nodes is made by a list of essential words associated with a given topic. A link between a word and a sentence is set if the word belongs to that sentence. Therefore, the projection of the bipartite network on the set of words results in a word-co-occurrence network [5].

Note that the directed approach produces a *directed network* while the undirected one an *undirected network*. Indeed, a directed link from one word to another may be validated, but not the other way around. The two methods are applied to a real dataset, highlighting the differences.

2 Method

Let us consider a corpus of N sentences and an LDA model providing a list of K latent topics, each described by an ordered list of essential words $w_1, \dots, w_i, \dots, w_n$. For each

topic, an *undirected* and a *directed* SVN are constructed by performing a battery of statistical tests against the null hypothesis of random co-occurrence, one for each pair of words. Let w_i and w_j be two essential words in a topic, their co-occurrences are denoted as:

- Symmetric X_{ij} : the number of sentences in which w_i and w_j appear together;
- Anti-symmetric from w_i to w_j , $X_{i \rightarrow j}$: the number of sentences in which w_i precedes w_j ;
- Anti-symmetric from w_j to w_i , $X_{j \rightarrow i}$: the number of sentences in which w_j precedes w_i ,

note that $X_{ij} = X_{i \rightarrow j} + X_{j \rightarrow i}$.

In the undirected approach, we aim at statistically validating only the symmetric co-occurrences X_{ij} against a null hypothesis of random co-occurrence, While in the directed method, we test the statistical significance of $X_{i \rightarrow j}$ and $X_{j \rightarrow i}$. The null hypothesis must account for the heterogeneity of the considered words, that is, the total number of times they individually appear in the text, N_i and N_j , respectively. The probability distribution that describes the random co-occurrence is the hypergeometric distribution. Indeed, assuming that the actual co-occurrences of these words (that may be symmetric or antisymmetric depending on the model) is N_{ij} , then the probability that a value larger than or equal to N_{ij} is observed by chance, according to the null hypothesis, is:

$$p_v(N_{ij}|N_i, N_j, N) = \sum_{X=N_{ij}}^{\min(N_i, N_j)} \frac{\binom{N_i}{X} \binom{N-N_i}{N_j-X}}{\binom{N}{N_j}}. \quad (1)$$

where parameters N_i and N_j naturally allow for the incorporation of the aforementioned heterogeneity of words in the null hypothesis.

To claim that the number of co-occurrences, N_{ij} , between words is too large to be consistent with the null hypothesis of random co-occurrences, we shall set a threshold α of statistical significance. Since we are facing multiple and dependent comparisons, we needed to use a correction method for multiple hypothesis testing; we opted for the more conservative Bonferroni correction [6].

3 Results

We applied our methods to a corpus of articles from the New York Times¹, setting 50 topics and considering ten essential words for each topic. Directed and undirected SVNs were obtained, and the Betweenness centrality [7] and the PageRank [8] measures were computed to spot the most influential words in each topic. Fig. 1 reports the SVNs related to one of the estimated topics. Note that the directed approach produces more sparse networks. The reason is twofold: i) the total number of symmetric co-occurrences X_{ij} splits into two smaller ones $\{X_{i \rightarrow j}, X_{j \rightarrow i}\}$ being more challenging to be validated; ii) twice as many tests are performed in the directed approach; thus the

¹<https://www.kaggle.com/nzalake52/new-york-times-articles>

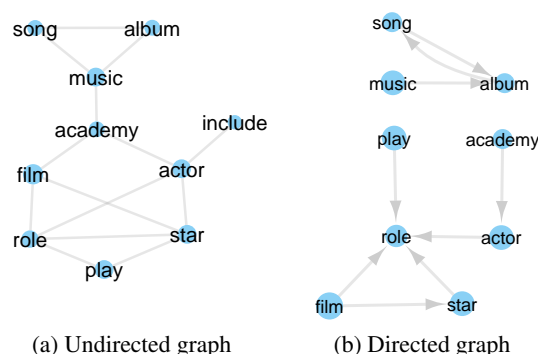


Fig. 1. SVN of one the estimated topics from the New York Times dataset.

Bonferroni correction is stronger. The undirected network shows only one big component, including all the essential words. It is important to stress that the word “academy” plays a fundamental role in being a bridge, thus having high centrality degree. Its role is less relevant in the directed network since their antisymmetric co-occurrences with the words “music”, and “film” are not significant. As a matter of fact, the directed network has two components corresponding to two sub-topics within the main one. The word “role” become the most influential term in the main component.

References

1. David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
2. David Newman, Sarvnaz Karimi, and Lawrence Cavedon. External evaluation of topic models. In *in Australasian Doc. Comp. Symp., 2009*. Citeseer, 2009.
3. David Mimno, Hanna Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 262–272, 2011.
4. Jey Han Lau, David Newman, and Timothy Baldwin. Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 530–539, 2014.
5. Dmitry Paranyushkin. Identifying the pathways for meaning circulation using text network analysis. *Nodus Labs*, 26, 2011.
6. Rupert G. Miller. Simultaneous statistical inference. springer-verlag, 1981.
7. Albert-László Barabási. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375, 2013.
8. Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.