CENTER FOR
COMPLEXITY
& BIOSYSTEMS
University of Milan

ISSUE #11 / SEPTEMBER, 2023
NEWSLETTER

COMPLEXITYBIOSYSTEMS.IT
COMPLEXITY@UNIMI.IT

PAOLO BOLDI
Professor of computer science and CC&B member

How to measure importance within a social network

Complex system analysis is an interdisciplinary field that calls for the collaboration of experts from various domains and employs of a broad array of mathematical models and computational tools. A significant portion of the research in this field exploits graph theory, depicting the system as a graph with nodes representing components and arcs corresponding to interactions.

This graph is then scrutinized using a multitude of techniques, encompassing traditional graph-theoretical algorithms, statistical analysis, and machine learning, among others. Depending on the specific use case, the graph can be either directed or undirected, according to whether the relations between nodes being symmetrical (like the friendship relation in Facebook) or asymmetrical (like the relation between followers/followees in Twitter).
One classical problem in graph analysis is understanding which nodes are important in a graph; this question is answered quantitatively by the so-called centrality indices, that aim at measuring how "central" is a node in a graph. The idea is of course not new, and was largely popularized in the context of sociology starting from the 50s, although the same problem was around already in the nineteenth century. Many solutions (in the form of centrality indices) have been proposed along the years, some of them being in a way or another related to the distances between the nodes (the so-called "geometric centralities", like closeness, harmonic centrality, degree centrality among others), some other related to forms of connectivity (like betweenness centrality), and finally some being more inspired by spectral analysis of some matrix (like PageRank or Seeley's index). Especially in computer science, spectral measures are sometimes preferred due partly to the fact that they can be approximated more efficiently and have a nicer interpretation.

These indices provide various notions of centrality, and often determine different importance ranking of nodes of the same graph; one tool that people often rely on to decide which centrality index to use in a specific context is to establish which properties (called, in this context, "axioms") indices do or do not satisfy.

An example of centrality axiom is "score monotonicity" and goes as follows: an index is "score monotone" if the score of a node always increases when a single arc is added toward that node. In Twitter parlance: a certain centrality index is score monotone if when I obtain a new follower, I always increase my score. An even more important axiom is "rank monotonicity": if a single arc is added toward a node, the node dominates (i.e., has a larger score than) the same set of nodes it used to dominate before adding the arc, and possibly some more. Some years ago, studying the most common centrality indices, we proved that they were almost all score and rank monotone on directed graphs. Again, in our Twitter use-case this means that: yes, having a new follower always increases your importance, no matter what.

Extending these results to the undirected (Facebook) realm has a number of technical hindrances. The natural undirected version of score monotonicity would be: if two nodes get connected, they always both increase their score (rank monotonicity is generalized similarly). Now the question is whether the centrality measures we normally use are score/rank monotone also in the undirected case. →

→ The somewhat surprising results we obtained is that most indices FAIL to be rank monotone. In Facebook parlance: no, having a new friend does not always increase your importance. This is totally counterintuitive, also because our results are not only theoretical, but can be found in real-world networks. Numerous (counter)examples were, for instance, found in the Hollywood co-actorship graph, using PageRank as centrality. If a famous actor like Meryl Streep were to make a new movie with some obscure actor, say, Yasuhiro Tsushima, not only this new relation would not be advantageous for both, but it turns out that it can be not worthwile for the less known actor! In a sort of a vampire-like behavior, Meryl Streep would suck away some more notoriety from working with Yasuhiro Tsushima, who would be instead harmed by this new co-actorship.

Among other things, these discoveries have an impact on the choice of centrality measures on undirected graphs, and highlight that different tools should be adopted when studying networks of different types. Tools that make sense in a certain context yield counterintuitive behaviors in other contexts.

New Study Reveals Potential Breakthrough in Triple-Negative Breast Cancer Prognosis

A recent study has investigated the prognostic and predictive potential of ARIADNE, a transcriptomic test for triple-negative breast cancer (TNBC) developed by researchers at the Center for Complexity and Biosystems of the University of Milan and commercialized by the spinoff Complexdata. The study was conducted by Caterina La Porta, from the Department of Environmental and Policy of the University of Milan and by Stefano Zapperi from the Department of Physics "Aldo Pontremoli" of the same University. Their findings have been published in the International Journal of Molecular Sciences and have the potential to change the way TNBC is diagnosed and treated in the future.

TNBC is a highly invasive and heterogeneous subtype of breast cancer that often has a high recurrence rate and poor outcome. The study found that ARIADNE was more effective than other common pathological indicators, such as grade, stage, and nodal status, in stratifying TNBC patients into groups with different disease-free survival statistics. Additionally, the study found that the classification provided by ARIADNE led to statistically significant differences in the rates of pathological complete response within the groups.
These findings offer hope for improved treatment and outcomes for TNBC patients. The development of prognostic and predictive markers like ARIADNE could lead to more targeted and effective treatment plans, ultimately improving the quality of life for those with TNBC.

"Triple-negative breast cancer is a challenging disease to treat, and we need better tools to predict patient outcomes and response to treatment. Our study suggests that ARIADNE may be a useful tool for clinicians in stratifying TNBC patients and selecting the most appropriate treatment plan for each individual" explains Caterina La Porta who coordinated the study. "The development of ARIADNE involved a unique combination of computational algorithms and expert knowledge in the field of breast cancer. We believe that our approach could be extended to other cancer subtypes and ultimately lead to the development of more accurate and personalized cancer diagnostics" – concludes Stefano Zapperi.

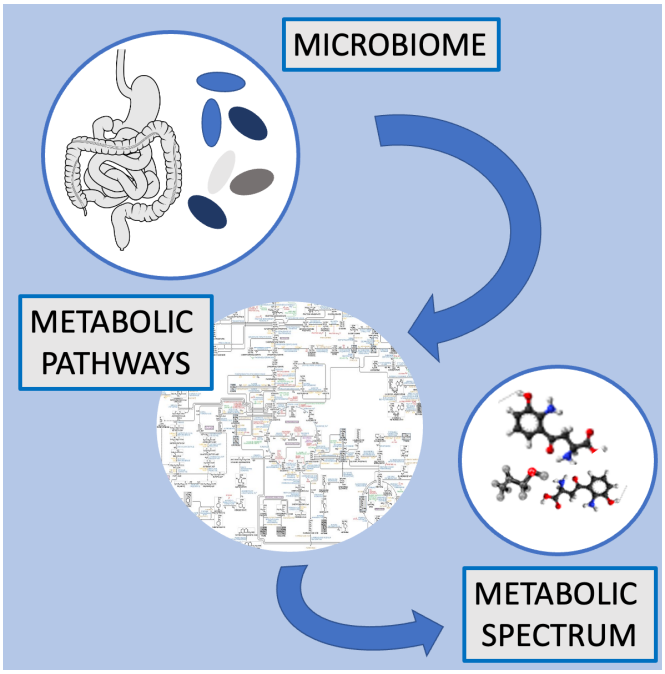
READ THE PAPER
https://www.mdpi.com/1422-0067/24/7/6422

Human microbiome: a novel quantitative method to identify its deregulation in relation to metabolic diseases

It is known that the metabolic activity of all the microorganisms that make up the human microbiome interacts with host metabolism, contributing to human health in a way that is not yet fully understood.

A group of researchers from the Center for Complexity and Biosystems at the University of Milan developed STELLA, a new computational method to derive the spectrum of metabolites associated with an individual's microbiome. The method was then applied to data obtained from the microbiome of patients with autism spectrum disorder and multiple sclerosis. The paper has just been published in the journal iScience.

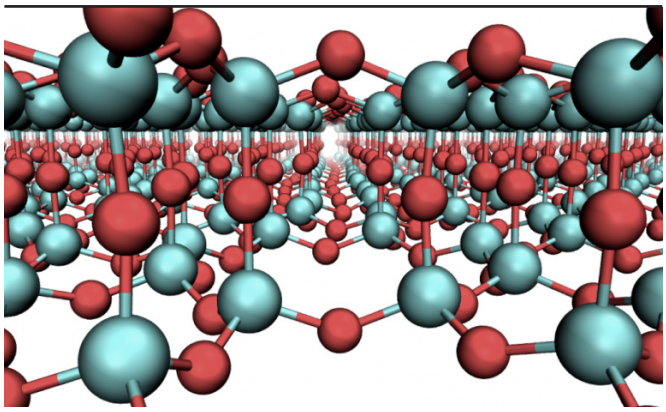
"STELLA integrates the known information on the metabolic pathways associated with each bacterial species and extracts from these the list of metabolic products of each singular reaction through an automatic analysis," says Caterina La Porta



professor of General Pathology in the Department of Environmental Science and Policy and coordinator of the study, " The great strength and innovation of Stella is that it allows by comparing the results of a single subject with the metabolic profile data of healthy subjects, to identify individual metabolic alterations." "The STELLA platform helps to identify new targets to make traditional therapies more effective using an integrated approach that addresses the complexity of the metabolic network," concludes Stefano Zapperi. professor at the "Aldo Pontremoli" Department of Physics at the University of Milan and co-author of the study.

READ THE PAPER
https://www.cell.com/iscience/fulltext/S2589-0042(22)02141-1

Predicting failure using AI



Being able to predict when and where a material will fracture is a key issue with important industrial implications in the area of device and component monitoring. Researchers from the Center for Complexity and Biosystems and the "Aldo Pontremoli" Department of Physics at the University of Milan in collaboration with colleagues from the Department of Materials Science at the Friedrich-Alexander University of Erlangen-Nuremberg in Germany, in a paper recently published in Nature Communications, have shown that through artificial intelligence it is possible to predict the fracture of silica glasses by examining their microstructure.
Thanks to recent advances in deep learning, it is possible to obtain accurate fracture predictions even for highly disordered solids such as glasses. Unfortunately, however,
the huge number of parameters used by artificial neural networks often makes a physical interpretation of the results impossible. This problem affects not only fracture prediction but is found in multiple applications of artificial intelligence. Researchers at the Center for Complexity and Biosystems applied a method to identify the areas of the microstructural image most used by the neural network for fracture prediction. This provided insight into what characteristics make a material more susceptible to fracture. "Neural networks are black boxes"-explains Stefano Zapperi, professor of theoretical matter physics and coordinator of the research-"and this is an important limitation in scientific research where the main purpose is to explain the origin of a phenomenon. Thanks to the method we used, it was possible to better understand what are the relevant aspects that determine the fracture of the material and thus obtain not only a prediction but also a greater fundamental understanding of the mechanics of glasses." "The strategy we developed lends itself to further applications"-adds Roberto Guerra, associate professor in the physics department and co-author of the paper-"such as to design disordered materials with better fracture toughness properties."

READ THE PAPER
https://www.nature.com/articles/s41467-022-30530-1

Three questions to...
Flavio Furia
Ph.D. student in Computer Science

Can you describe your current research activities?
Being a member of the LAW (Laboratory for Web Algorithmics), my main research field is graph theory. In particular, I focus on network centrality, which involves the study and analysis of functions used to quantitatively assess how "central" nodes are in graphs: we call them centrality indices. Graphs are used to model systems with entities linked one to another through some kind of relationship, but the "importance" of an entity is not uniquely defined: popularity, influence and reputation are just a few examples of notions we want to measure. In fact, over the years countless different strategies to compute node centrality have been proposed and choosing the best for a given scenario is not a trivial decision.
Currently, I'm studying how different centrality indices react to slight modifications of the underlying graph. For example, questions like "Does making new friends in a social network always increase your reputation?" have answers that are less obvious than they seem at first glance. We have shown that for some of the most used centrality indices making a new friend actually decreases your reputation. At the same time, we think that at least one of the two endpoints always benefits from the new relation, and our current experiments are rather promising in this direction.

What you think are the key challenges in your research field?
Graph theory is a very broad field: when dealing with graphs, every problem can be solved in many different ways, borrowing techniques from different areas of mathematics. In my opinion, it is crucial to look at problems from different points of view and with an open mind.
Speaking about network centrality, I think that one of the most interesting challenges is to participate in the process of building its theoretical foundations. Nowadays, centrality is deeply used in many real-world experiments, but different indices serve different purposes: I think that any "theoretical" result can actively

support the choice of the best centrality index for a given scenario.

Does your work have potential industrial applications?
Graphs can be used to model any system with entities related one to another. Indeed, countless industrial applications can benefit from research and discoveries in graph theory. For example, it is worth mentioning that initially network centrality was popular mainly among sociologists, but it turned out to be very useful in the context of link analysis. PageRank, the renowned algorithm initially used by Google Search to rank web pages, is nothing but a centrality index where the underlying graph is the Web, with nodes being web pages and arcs being links among them.

Three questions to...
Davide D'Ascenzo

Ph.D. student in Artificial Intelligence

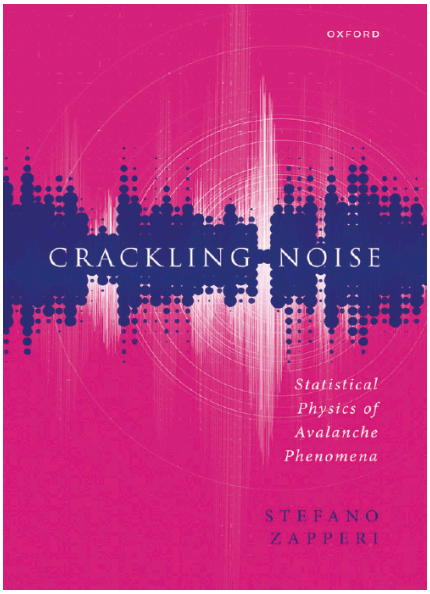
Can you describe your current research activities?
My PhD program is in Artificial Intelligence, so my main research is about learning systems. An exciting focus in the AI community is the role of networks (which I prefer to call graphs) either as building blocks of learning systems or as complex structures to analyze. I am also exploring other diverse topics with the hope, in the future, of bridging knowledge from various fields. In this sense, the CC&B is a great way to cross-contaminate different research areas.
In particular, I am currently working on understanding the theoretical principles of learning systems and practically analyzing graphs with these systems. An interesting practical problem that I recently encountered is the reconstruction of a graph with high symmetrical properties that, after a small perturbation, exhibits far fewer symmetries than before. Here, by perturbation, I mean the addition or deletion of a few edges from the graph. The meaning of symmetry is more nuanced, and it has to do with the similarities and repeating patterns you can find looking at the graph. The objective is to understand if it is possible to "correct" these perturbations through some kind of restoration process. You can think of it as an error-correcting mechanism for some types of graphs. This can be extremely helpful in the bioinformatics domain, where biological and chemical networks often show highly symmetrical structures but where the measurements are often noisy and perfect reconstruction is really hard.

What you think are the key challenges in your research field?
Artificial Intelligence, and more specifically, Deep Learning, is a huge research area that keeps growing every day. The main challenge is to understand how the different ideas of several research groups can be unified under a common mathematical framework. The current learning systems are built using trial-and-error approaches and are specifically tailored for certain problems. Understanding why these systems behave as they do or why some systems are less precise than others in certain tasks is an important step in improving this research field.

Does your work have potential industrial applications?
Both learning systems and graphs are ubiquitous. Learning systems are currently used in many different applications, including image recognition, object detection, translation, and many others. On our smartphones, we have several learning systems that have been trained to simplify our lives (for example, the auto-correction feature in mobile keyboards). Graphs are used to describe a lot of complex systems: social networks, biochemical compounds, transportation networks, etc. Any improvement, both theoretical and practical, can benefit society and the industries that are using such tools.

New book
Crackling Noise: Statistical Physics of Avalanche Phenomena

The response of materials and the functioning of devices is often associated with noise. In this book, Stefano Zapperi concentrates on a particular type of noise, known as crackling noise, which is characterized by an intermittent series of broadly distributed pulses. While representing a nuisance in many practical applications, crackling noise can also tell us something useful about the microscopic processes ruling the material's behavior.
The book is published by Oxford University Press. It provides a comprehensive overview of key concepts and theoretical models, explores the many applications of the theory of crackling noise in materials science and includes expansive discussions considering implications for the life sciences.



LINK TO THE BOOK
https://global.oup.com/academic/product/crackling-noise-9780192856951?cc=it&lang=en&

CC&B is a Coordinated Research Center at the University of Milan
Research within CC&B is supported by the European Research Council

UNIVERSITÀ DEGLI STUDI DI MILANO

European Research Council