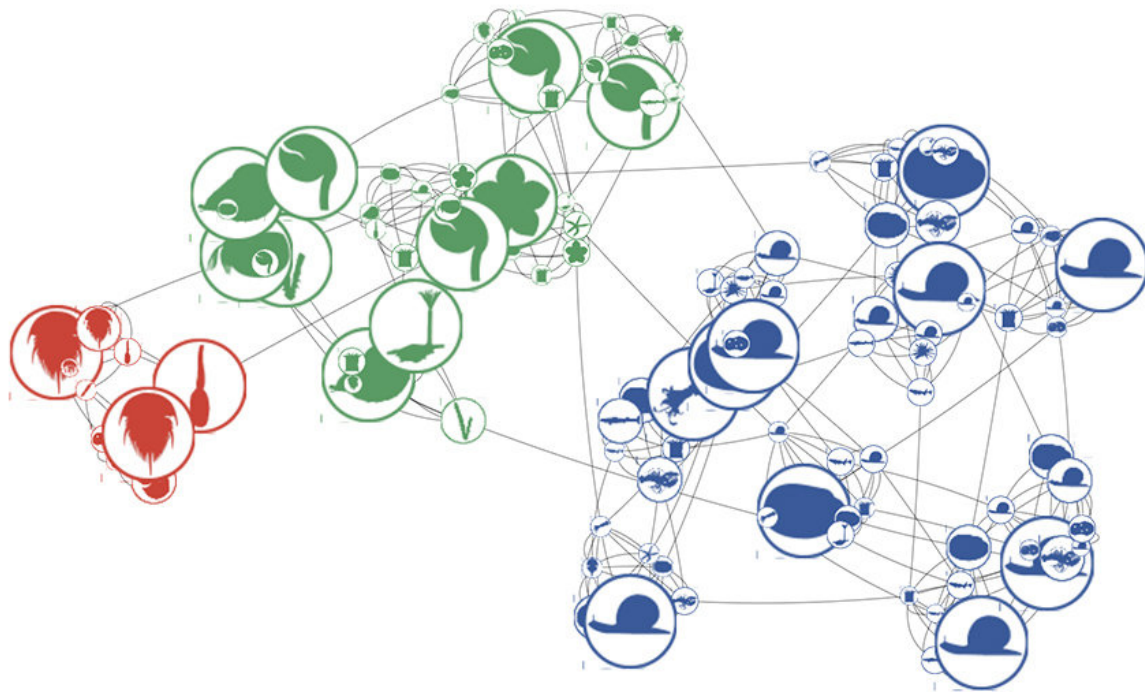# Applying network analysis to natural history: Technique popularized through social media ranks impact of extinctions

April 23 2018

Network analysis of marine fossil records over the past 541 million years revealed the ecological impact of mass extinctions and may help us anticipate the consequences of a 'sixth mass extinction.' Credit: Rensselaer Polytechnic Institute

A team of researchers is using network analysis techniques - popularized through social media applications - to find patterns in Earth's natural history, as detailed in a paper published today in the *Proceedings of the National Academy of Sciences* (*PNAS*). By using network analysis to search for communities of marine life in the fossil records of the Paleobiology Database, the team, including researchers at Rensselaer Polytechnic Institute, was able to quantify the ecological impacts of major events like mass extinctions and may help us anticipate the consequences of a "sixth mass extinction."

"Network analysis can transform into a digestible form databases so huge that it's impossible to look at substantial portions of the data altogether," said Peter Fox, a Tetherless World Constellation Chair and professor of earth and environmental science, computer science, and cognitive science at Rensselaer.

"The power of our approach is that multidimensional data embedded in the network can inform and discover trends in the data, turning an endless grid of numbers into a picture that reveals multiple relationships at a glance."

The team's approach offers new perspective on the ecological impacts of present-day species extinctions, said Drew Muscente, a postdoctoral research fellow at Harvard University and lead author on the paper. Given the rate of species disappearances over the past few centuries, many scientists suspect that Earth is in the midst of the sixth mass extinction.

"The fossil record contains evidence of repeated mass extinctions. Data on how ancient communities of organisms changed during these events can help us understand the potential consequences of the present-day biodiversity crisis," said Muscente. "Our work shows that this crisis, regardless of what you call it, may irreparably alter communities of

organisms and their ecosystems in some surprising ways, which can't be predicted with other methods."

One picture that emerges from the analysis is a ranking of the ecological impact of major events, with the Great Ordovician Biodiversification Event having the largest effect on ecology, followed in descending order by the Permian-Triassic, Cretaceous-Paleogene, Devonian, and Triassic-Jurassic mass extinctions. The analysis shows that the Ordovician mass extinction may have had less ecological impact than previously estimated, and likewise, the significance of the Devonian extinction may be under-appreciated.

Fox and Rensselaer researchers Anirudh Prabhu, Hao Zhong, and Ahmed Eleish joined lead author Muscente and Andrew Knoll of Harvard University, and Michael B. Meyer and Robert Hazen of the Carnegie Institution for Science on the research, which expands upon a suite of earlier work applying network analysis to mineralogy data. Their work is funded with a three-year grant from the W.M. Keck Foundation.

"The groundbreaking work reported in this article illustrates how next-generation data analytics created for one domain can transform other fields of study," said Professor Curt M. Breneman, dean of the Rensselaer School of Science. "This provides a look ahead into the impact of data-driven science in the 21st century."

Social network analysis can be used to identify groups of friends, transmission of disease, and extremist groups by identifying communities of people - whose common attributes like location, interests, or gender reveal their association absent an outright declaration - on social media. Just as network analysis reveals communities of people, researchers can use network analysis of Earth and life science databases to discover associations of ancient organisms (e.g., species and genera) that lived in the past and learn something about how those

"paleocommunities" changed over time, said Fox.

In earlier work, the team applied network analysis to a mineralogical database. Each mineral recorded was defined by as many as 17 attributes - aspects like chemical composition, mode of formation, location - and the results, as published in American Mineralogist, predicted the existence of 1,500 as yet undiscovered minerals, at least 14 of which have since been found. Recent work on network analysis of mineralogical data has also been published in American Mineralogist and the International Journal of Geo-Information.

In the PNAS paper, titled "Quantifying ecological impact of mass extinctions with network analysis of fossil communities," the team tackled the Paleobiology Database, a "non-governmental, nonprofit public resource of paleontological data." The database contains data on the locations, ages, environments, and affinities of fossils, representing more than 350,000 ancient taxa preserved at more than 190,000 fossil collection sampling points around the world during the last 600 million years of Earth history. The team restricted their dataset to occurrences of fossils of marine animals that lived in the Phanerozoic Eon, the interval of time that began with the explosion of animal life 541 million years ago and continues to the present day.

In the authors' networks, each fossil taxon (e.g., order, family, or genus) becomes a "node," which can be visualized in network graphs as a dot. The nodes are connected to each other if those organisms lived together and were fossilized at the same sites in the past. This approach results in the organization of nodes into clusters, which represent ancient communities of marine animals and can be identified using computational and statistical methods. Because taxa (and communities) originate and become extinct over time, geologic age manifests as an implicit aspect of network structure. In the graphs, taxa and communities that lived at different times in Earth history are distributed throughout

the networks, with the distances between nodes being directly related to the time span separating their ages. Overall, the resulting network diagram portrays aspects like the density of the network at different time periods, the degree of centrality of nodes and groups of nodes, the number of connections between nodes, and more. To represent even more attributes, the team is moving into representation in three-dimensions and virtual reality. The result, said Fox, is "a very substantial fraction of the Paleobiology Database in a single graph."

The approach lends itself to new discoveries, not just about the fossils themselves, but "about the correspondence between the fossils and the environment they lived in," said Fox. In the PNAS result, researchers use the fossil record to quantify the impacts of extinctions, but by combining data from the fossil record with information from the mineral record, the researchers hope that network analysis could lead to other insights into the evolution of the Earth system, for instance, how life and environments changed in response to atmospheric oxygenation or the change from nitrogen-rich to nitrogen-poor conditions.

"When we combine this work, we'll have a multi-layer network where we can look at the correspondence between the fossil network and the mineral network," said Fox. "And that's never been done before."

Research on network analysis of Earth and life science databases fulfills The New Polytechnic, an emerging paradigm for higher education which recognizes that global challenges and opportunities are so great they cannot be adequately addressed by even the most talented person working alone. Rensselaer serves as a crossroads for collaboration—working with partners across disciplines, sectors, and geographic regions—to address complex global challenges, using the most advanced tools and technologies, many of which are developed at Rensselaer. Research at Rensselaer addresses some of the world's most pressing technological challenges—from energy security and sustainable

development to biotechnology and human health. The New Polytechnic is transformative in the global impact of research, in its innovative pedagogy, and in the lives of students at Rensselaer.

More information: A. D. Muscente el al., "Quantifying ecological impacts of mass extinctions with network analysis of fossil communities," *PNAS* (2018). www.pnas.org/cgi/doi/10.1073/pnas.1719976115

Provided by Rensselaer Polytechnic Institute