Judges' Commentary: Measuring Network Influence and Impact

Chris Arney
Dept. of Mathematical Sciences
U.S. Military Academy
West Point, NY 10996
`david.arney@usma.edu`

Kathryn Coronges
Department of Behavioral Sciences and Leadership
U.S. Military Academy
West Point, NY

Tina Hartley
Dept. of Mathematical Sciences
U.S. Military Academy
West Point, NY

Jessica Libertini
Dept. of Mathematics
Virginia Military Institute
Lexington, VA

Introduction

The topic area for this year's Interdisciplinary Contest in Modeling (ICM) was network science. Network science will continue to be the topical area for one of next year's ICM problems. However, there will also be a second ICM problem, involving human-environment interactions in the areas of environmental science, including climatology, food security, and geography. For teams that want to organize early for next year's contest, prepare by studying network modeling or environmental science and assemble a team with one of those subjects in mind.

The UMAP Journal 35 (2–3) (2014) 249–269.

The ICM continues to be an opportunity for teams to tackle challenging real-world problems that require a wide breadth of understanding in multiple academic subjects and skill in modeling interdisciplinary phenomena. These kinds of interdisciplinary study and modeling are included in the definition of network science. The complexity of ICM problems, along with the short duration of the contest, requires effective communication and coordination among team members. One of the most challenging issues for ICM teams is to organize and collaborate effectively to use each team member's skills and talents to tackle the diverse nature of ICM problems. Teams that resolve this organizational challenge and co-operate well often submit 20-page solutions that rise to the higher levels of ICM awards.

The Problem Statement

One of the techniques to determine influence of academic research is to build and measure properties of citation or co-author networks. Co-authoring a manuscript usually connotes a strong influential connection between researchers.

One of the most famous academic co-authors was the 20th-century mathematician Paul Erdős, who had over 500 co-authors and published over 1,400 technical research papers.

It is ironic (or perhaps not!) that Erdős is also one of the influencers in building the foundation for the emerging interdisciplinary science of networks, particularly, through his publication with Alfred Rényi of the paper "On random graphs" [1959].

Erdős's role as a collaborator was so significant in the field of mathematics that mathematicians often measure their closeness to Erdős through analysis of Erdős's amazingly large and robust co-author network (see Grossman [2014]).

The unusual and fascinating story of Paul Erdős as a gifted mathematician, talented problem solver, and master collaborator is provided in many books and online Websites (e.g., O'Connor and Robertson [2000]). Perhaps his itinerant lifestyle, frequently staying with or residing with his collaborators, and giving much of his money to students as prizes for solving problems, enabled his co-authorships to flourish and helped build his astounding network of influence in several areas of mathematics.

To measure such influence as Erdős produced, there are network-based evaluation tools that use co-author and citation data to determine impact factor of researchers, publications, and journals. Some of these are *Science Citation Index*, H-factor, Impact factor, Eigenfactor, etc. Google Scholar is also a good data tool to use for network influence or impact data collection and analysis. Your team's goal for ICM 2014 is to analyze influence and impact in research networks and other areas of society.

We summarize the tasks for the teams in this year's ICM problem:

1. Build the co-author network of the 511 Erdos1 co-authors. This will take some skilled data extraction and modeling efforts to obtain nodes (the Erdős co-authors) and their links (connections with one another as co-authors). There are over 18,000 lines of raw data in Erdos1 file, but many of them will not be used since they are links to people outside the Erdos1 network. Analyze the properties of this network.

2. Develop influence measure(s) to determine who in this Erdos1 network has significant influence within the network. Consider who has published important works or connects important researchers within Erdos1.

3. Another type of influence measure might be to compare the significance of a research paper by analyzing the important works that follow from its publication. Choose a set of foundational papers in the emerging field of network science (a possible set was provided). Use these papers to develop a model to determine their relative influence. Which of the papers of the set do you consider is the most influential in network science and why?

4. Implement your algorithm on a completely different set of network influence data.

5. Discuss the science and utility of modeling influence and impact within networks. Could individuals, organizations, nations, and society use influence methodology to improve relationships, conduct business, and make wise decisions?

A Short Historical Reflection on Paul Erdős and the Erdős Co-author Network

Paul Erdős's creative research advanced graph theory, combinatorics, discrete mathematics, and number theory and laid foundations for the applied subjects of computer and network science. He excelled at modeling number systems and graphical structures, and determining their properties. He worked on many of the most important problems in these fields; and, through tireless effort and amazing skill, he became the most prolific and eccentric mathematician of modern times. He published over 1,500 scholarly papers. Paul Hoffman, who wrote Erdős's biography *The Man Who Loved Only Numbers*, wrote, "Erdős's style was one of intense curiosity, a style he brought to everything he confronted" [2012, 21]. The ICM hopes to develop that trait of curiosity in its contestants.

Because of Erdős's extensive collaborations, mathematicians began tracking and counting his collaborators. A special network numbering system was devised such that a person who collaborated by publishing a paper with Erdős was given an Erdős number of 1. Collaboration (publishing) with an Erdős 1 author gave a mathematician an Erdős number of 2, and so on. If there is no chain of co-authorships connecting someone with Erdős, the person's Erdős number is infinite. The result of this effort, using the Erdős Number Project site (`http://www.oakland.edu/enp/`) [Grossman 2014] and data through the MathSciNet service of the American Mathematical Society's *Mathematical Reviews*, is an elaborate collaboration graph of the mathematics research community that captures the connections of over 400,000 authors. Elaborate records and a myriad of statistics are kept on the collaborations and connection record of Erdős. Today the Erdős Number Project Website provides all sorts of trivia, such as the data in **Table 1** on co-author connections to Erdős [Grossman 2014]. The table shows the number of people with Erdős number 1, 2, 3, . . . , according to the electronic data from MathSciNet (slightly different than other data sources).

Table 1.
Numbers of mathematicians with particular Erdős numbers.

Erdős number	Number of mathematicians
0	1
1	504
2	6593
3	33,605
4	83,642
5	87,760
6	40,014
7	11,591
8	3,146
9	819
10	244
11	68
12	23
13	5

Thus, at the moment when this table was tabulated, the median Erdős number was 5, the mean 4.65, and the standard deviation 1.21. In addition to these 268,000 people with finite Erdős number, there are about 50,000 published mathematicians who have collaborated with others but have an infinite Erdős number, and 84,000 who have never published joint work (and therefore also have an infinite Erdős number).

Erdős lived from 1913 to 1996 and spent six decades living out of two tattered suitcases. He would show up on the doorsteps of his colleagues prepared to do work, and they would accommodate him. After working through a problem or two and writing a paper, he would move on to the next research station, hopefully to confront the next problem and find its solution.

Erdős received many awards, which allowed him the freedom to travel and the money to pay student solvers of his legendary challenge problems. His challenge problems were often easy to state but difficult to solve. These numerous cash giveaways to student problem-solvers made Erdős's campus visits special events. Another quotation about Erdős from Hoffman's book reinforces an interdisciplinary perspective: "For Erdős, mathematics was a glorious combination of sciences and art." [1998, 27]

Judges' Criteria

The panel of judges was impressed by the modeling of many teams. Many papers were rich in network modeling methodology and modeling creativity. To ensure that the individual judges assessed submissions on the same criteria, a judging rubric and guide was developed. The general framework used to evaluate submissions is described below. The main thrust of ICM problem-grading is finding and evaluating modeling that includes good science and leads to measurable outcomes and a viable solution.

Executive Summary

It was important in the summary that students succinctly and clearly explained the highlights of their submissions. The executive summary should contain brief descriptions of both the problem and the bottom-line results. One mark of better papers was a summary with a well-connected and concise description of the methodology, results and recommendations.

Modeling

Well-defined measures of influence and impact were needed to build a viable model. Many teams started with standard network centrality measures and modified them to produce influence or impact effects. Other teams calculated other measures from clustering, community building, and dynamic measures. In this problem, teams needed to develop viable influence measure(s) to determine who in the network has significant influence within the network. For some teams, influence was a scalar value; others established a multidimensional vector with several components.

Many teams used network analysis software packages such as Gephi, ORA, Pajek, and UICNET for both calculations and visualizations. In many cases, the resulting mathematical analysis included statistical measures. Some teams used the explicit structures of networks or graphs to determine classic nodal measures and properties. In such cases, critical assumptions such as the directionality of influence and weights of connections within the network led to viable network models.

Better papers discussed the differences in co-author and citation networks by explaining that a co-author network is not directed and a citation network model is directional. Similarly, the Erdős co-author network is now static (nearly 20 years after his death); but the citation network is dynamic, with new citations to papers occurring frequently.

No matter the modeling framework, the assumptions needed for these models and the careful and appropriate development of these models were important in evaluating the quality of the solutions. The better submissions explicitly discussed why key assumptions were made and how these assumptions affected model development.

Stronger submissions presented a balanced mix of mathematics and prose rather than a series of equations and parameter values without explanation. One major discriminator was the use or misuse of arbitrary parameters without any explanation or analysis. Establishing and explaining parameter values in models are at least as significant as making and validating assumptions.

Perhaps the most challenging aspect of this problem was determining the topic and data collection of the test application dataset. Collecting good data where their influence measure was appropriate was a challenge. The judges recognized this challenge and rewarded papers with strong datasets that produced viable network models.

Science

The ICM modelers discussed the science of influence at many levels. Some teams did effective background research and analysis of this aspect of the problem, included elements of their scientific analysis, and described how their model fit into the science of influence. In this case, powerful scientific analysis was performed best by making strong, insightful connections between the precise mathematical measures that teams created and the abstract notions of influence that they produced from social theory.

No matter what level of modeling was performed by the teams, the interdisciplinary nature of this problem was revealed in the science requirements and the background investigation performed by the teams. The ICM students were exposed to the nature of influence in information theory in performing their background research, and the team reports required proper documentation of the team's research sources.

Data/Validity/Sensitivity

Filtering over the 18,000 lines of data to extract the 511 co-author nodes and nearly 1,700 links was a challenge for some teams, as was collecting data for their own test application. Sensitivity analysis to determine the effects of assumptions and data validity were empowering for some of the teams. Sensitivity analysis is especially important for highly-structured and powerful data-rich models such as networks. Some network structures are highly robust and flexible, while others are more fragile and highly sensitive to data errors or changes. While this sensitivity analysis is a challenging element of network modeling, it was important to address this modeling issue in the report. Teams that did this well quickly rose to the top of judges' evaluations.

Strengths/Weaknesses

Discussion of the strengths and weaknesses of the models is where students demonstrate their deeper understanding of what they have created. The utility of a model fades quickly if team members do not understand the limitations or constraints of their assumptions or the implications of their methodology. Networks are non-reductive, complex structures and, therefore, the strengths and weakness are often hidden from direct view or full control of the modeler. Some of the better reports presented these elements despite these challenges.

Communication/Visuals/Charts

To clearly explain solutions, teams used multiple modes of expression including diagrams and graphs, and—for this contest—clearly written English. A report that could not be understood did not progress to the final rounds of judging. Judges were often well informed through the amazing array of powerful charts and graphs that explained both models and results. The graphics shown in **Figures 1–3** provide a glimpse of the richness of this kind of presentation.

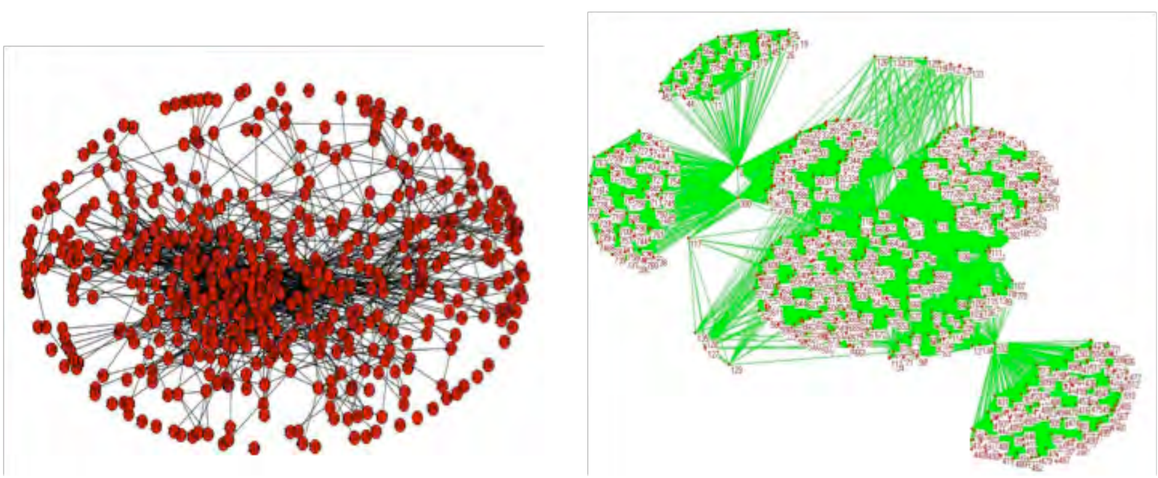


Figure 1. Many teams provided informative network graphs to show the entire co-authors' network model. The graphic on the left is from Team 25425 (Beijing University of Posts and Telecommunications). The graphic on the right is from Team 30407 (Hong Kong Polytechnic University).

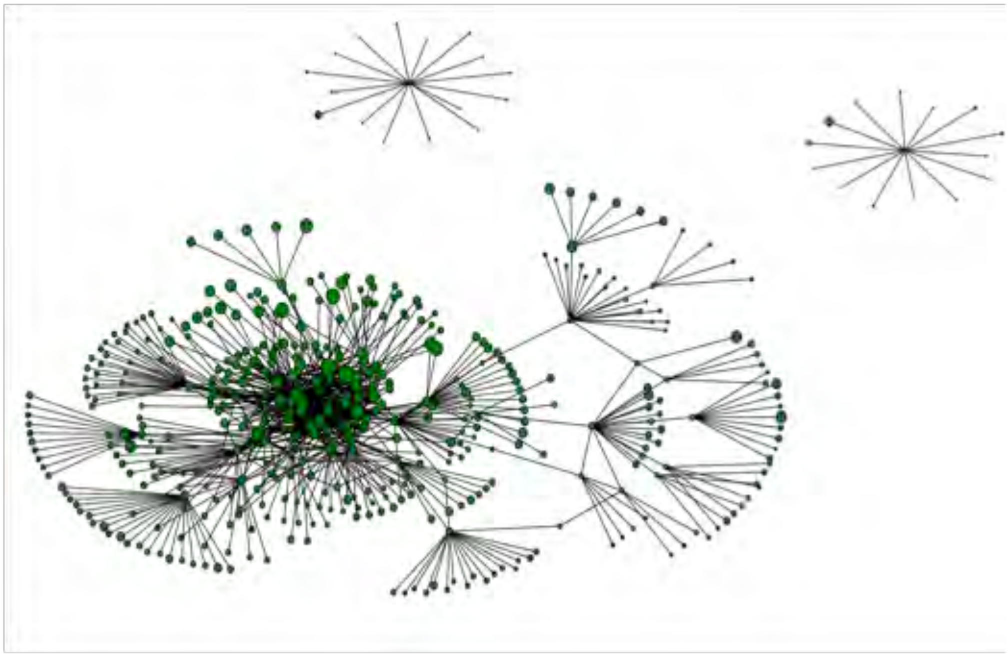


Figure 2. Some reports contained elaborate co-citation network diagrams, like this one from Team 31227 (Humboldt State University).

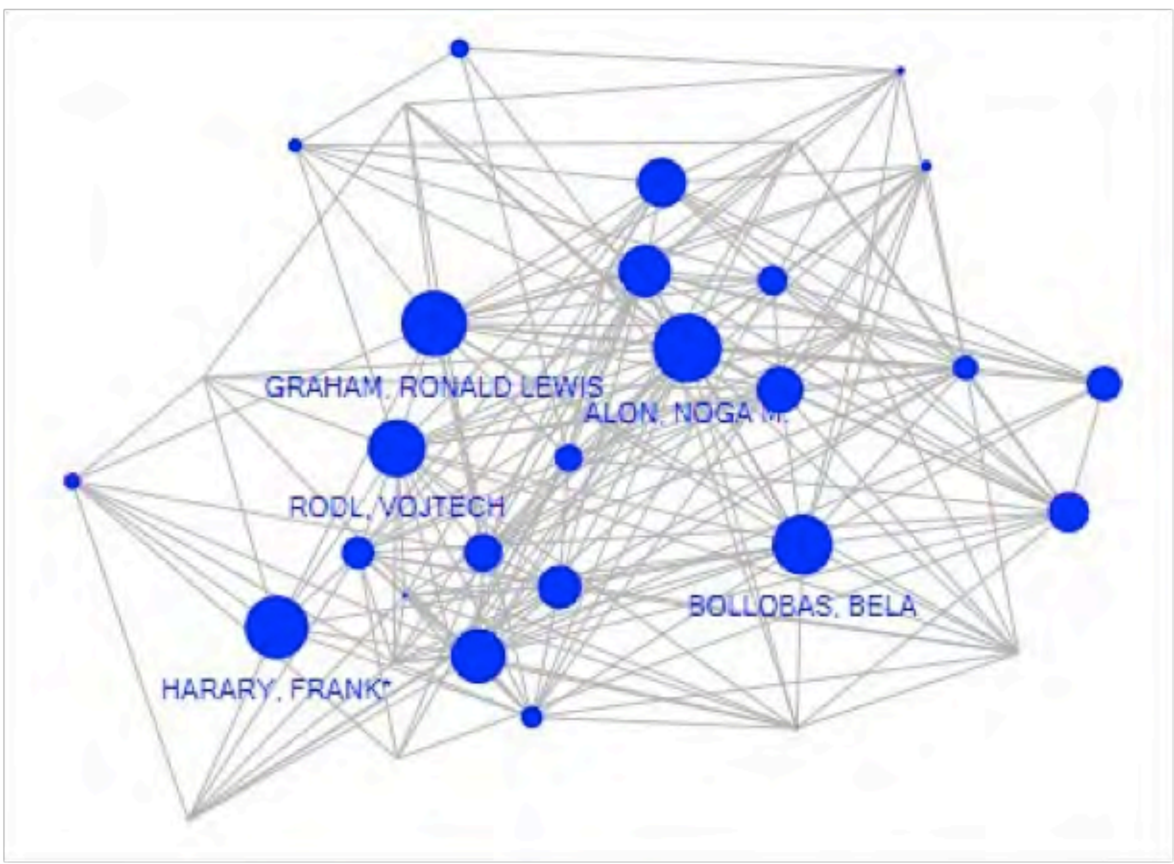


Figure 3. Other graphics zoomed in on significant parts of the co-author network to show details of the most influential co-authors, like this one from the report by Team 26715 (Peking University).

Discussion of the Outstanding Papers

Despite the common dataset and tasks, many different approaches were used by ICM teams to model various aspects of the problem. As a result, the submissions were varied and interesting. Overall, the basic modeling was often sound, creative, and powerful. Those papers that did not reach final judging generally suffered from various shortcomings. Some lacked clear explanation of the structure of their model. They provided some details but not a complete description of their model and its purpose. Others failed to connect their mathematical models to the aspects and basic elements of the science of influence. In general, incomplete or awkward communication was the most significant discriminator in determining which papers reached the final judging stage.

Although the six Outstanding papers used different methodologies, they all addressed the problem in a comprehensive way. These Outstanding papers were generally well written and presented clear explanations of their modeling procedures. In several of the Outstanding papers, a unique or innovative approach distinguished them from the rest of the finalists. Others were noteworthy for either the thoroughness of their modeling or the significance of their results. Summaries of the six Outstanding papers follow.

Central University of Finance and Economics, China: "Influence Measures in Networks"

The team from Central University of Finance and Economics gave their report the title "Influence Measures in Networks" to reflect their focused and quite thorough investigation of network proximity as a proxy of influence among network members. They rightly point to the limitations of some of the traditional network metrics, namely centrality measures, in their inability to efficiently handle link weights and to account for the whole network structure. This group instead combined the Shapley approach, a concept developed in cooperative game theory, with a cohesion measure (KPP-POS, developed by Stephen Borgatti) to capture influence effects. The limitation of this measure is that it can be applied only to undirected networks.

The team's approach to directed networks was somewhat less novel, making modest modification to the frequently used PageRank measure.

The team's approach to combine conventional social network analytic methods with the less obvious, game-theoretic Shapley approach showed a breadth and depth in their handling of this problem. Challenged with developing a more appropriate and meaningful influence measure, they modified Borgatti's cohesion measure to enable inclusion of weighted edges, and more notably, used the Shapley method to account for the rank order of nodes.

As a nice introduction to their analysis, the team provided descriptive information about the networks, such as the degree distribution, path length, and clustering. They also showed that their new metric of influence gives different results than conventional betweenness centrality values. They chose to validate their Shapley-cohesion measure on a weighted network that they built by selecting actors who have collaborated with the popular British actor, Jude Law. Unfortunately, on that dataset, their measure did not perform differently from betweenness centrality. The Jude Law network may have been too small to enable detection of difference among these metrics.

The team's approach to directed networks was one that was used by many ICM teams. However, this team successfully modified the network data to overcome some of the limitations of the PageRank method. Specifically,

- they incorporated additional papers into the citation network to augment the data specified in the problem; and
- they weighted the papers as a function of the number of times they were cited by these endogenous nodes.

The schematic of their model is provided in **Figure 4**.

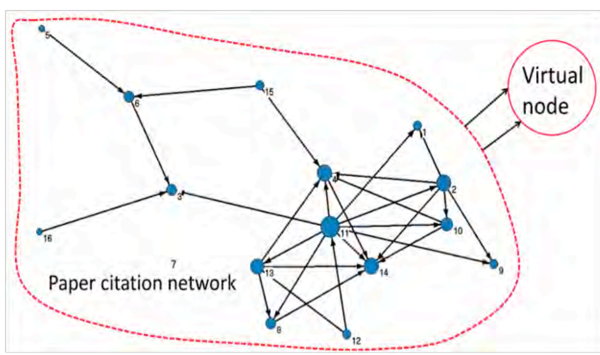


Figure 4. Network model supplemented with additional co-citation papers.

With their modified or optimized PageRank measure, the team found influence rankings among publications that differed from the standard PageRank measure. However, they did not investigate whether those papers that jumped in rank were any more meaningful than the standard rank. In addition, the team did not deal with the dimension of time—ignoring the fact that articles could not be cited by articles that appeared earlier in time, or that articles that were cited more recently or over longer periods of time were probably more influential.

Judges were impressed with the ability to combine the conventional network metrics with game-theoretic approaches. Further, the team showed thoughtfulness about applying network science concepts to a problem of social influence. This paper was well written and contained graphic results such as that given in **Figure 5** to show the nodal degree distribution between the Erdos1 network and a similar-sized random network.

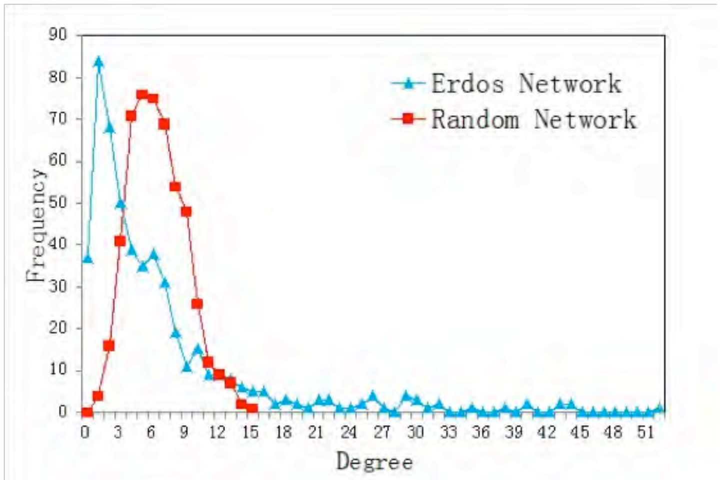


Figure 5. Nodal degree distribution of the Erdős co-author network and a similar-size random graph.

National University of Defense Technology, China: "The Research of Influence Based on the Characteristic of a Network"

The team from National University of Defense Technology took two very different approaches in analyzing the graphs, depending on whether or not the graph was directed or not. For the undirected co-author network, this team explained and computed many of the established network metrics for each node, including degree, betweenness, closeness, and eigenvector centralities. Rather than use all these metrics, the team provided a clear argument for first identifying those with the most authority (those who had published the most number of times with Erdős), and then ranking these authoritative co-authors based only on eigenvector centrality.

To qualitatively validate the results produced by this algorithm, the team did online searches to learn about the careers of each of the top five mathematicians on their results list. For the citation network, this team gave a visual representation of the network, laid out on a temporal axis. The inclusion of time in the network visualization was an excellent example of how something relatively simple can really make a difference in the ease of interpretation of the results.

For the analysis, this team's approach to the citation network relied on the fact that the resulting network is a directed acyclic graph. This team then attempted to identify the most influential paper by examining four different centrality measures, only to discover that their results were inconsistent.

From there, they determined that a topological sorting algorithm would sufficiently decrease runtime compared to a matrix multiplication method such as PageRank.

Leveraging the topological structure of the graph, the team defined and computed a contribution coefficient that took both the number of citations and the timing of those citations into consideration. They discussed how self-citation could influence their results, along with giving a modified model and a sensitivity analysis based on a range of values for their self-contribution coefficient.

Lastly, this team applied their algorithms to construct a directed ownership network of 500 U.S. media corporations and to identify the top media companies. After performing their analysis, they validated their results by looking at published business rankings. Following this, the team provided an insightful list of potential applications and a discussion of the benefits of using network science in business and military decision-making.

The judges were impressed by this paper's strong links between the theory, applications, and mathematics. The visualizations of their networks provided meaningful insights into their analysis as shown in the network graph in **Figure 6**. They calculated many of the standard network measures; but instead of consolidating all of them, they presented a convincing argument for only using certain measures. Additionally, these students demonstrated that they understood the most predictable paths for solving the problems, and they showed how they could use inherent network properties to improve upon the more obvious choices. The judges were very impressed by this team's qualitative approach to validating their results through internet research.

Southeast University, China (INFORMS Winner): "Who are the 20%?"

This team developed a model that they called the Relation Distance Model, which utilized three standard centrality measures (degree centrality, betweenness centrality, and closeness centrality) to construct a normalized centrality measure vector for each node in the network. The Euclidean distance from each node's centrality measure vector to an ideal vector was computed and used to determine the most influential nodes in the network. The team did a nice analysis of the results of their model, and then used eigenvector centrality results as a means to validate the results of their model, which exposed a limitation of the model.

To analyze the citation network, the team developed a model that they called the Authority-Popularity Model. Each paper or node was weighted based on the impact factor of the publishing journal. The model then used the Modified PageRank Algorithm to determine a value representing the importance of each paper in the network, which the team classified as its

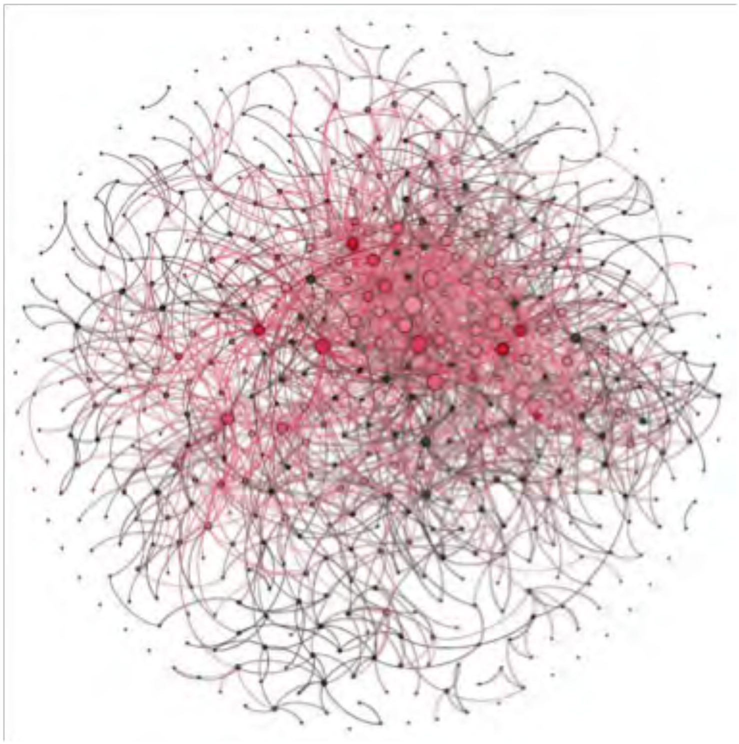


Figure 6. Structure of the Erdos1 network as presented in the report from the Outstanding team from the National University of Defense Technology, China.

level of authority or "depth of influence." Each node was also assigned a value based on the number of times per year that it was cited, which the team classified as its level of popularity or "width of influence." These two scores were then plotted on a log scale to visually segregate those papers that had both high authority and high popularity.

The Authority-Popularity Model was then applied to a co-star network that consisted of 15 movie stars and the links between them. The team recognized that this network would have weighted links (based on the rating of the movie) rather than weighted nodes, and adapted the model accordingly.

A particularly impressive feature of this paper was the strong use of visuals and graphics to clearly present the models and display results in a meaningful way. **Figure 7** is an example. In addition, judges were impressed with the thorough development of each of the models, the analysis of the effects of significant parameters, and the candid discussion of the models' strengths and weaknesses. Finally, the judges appreciated the team's use of their model to propose an innovative method for a researcher to quickly "boost" their influence in order to co-author with a leading figure in network science using the Pareto principle, or 80–20 rule.

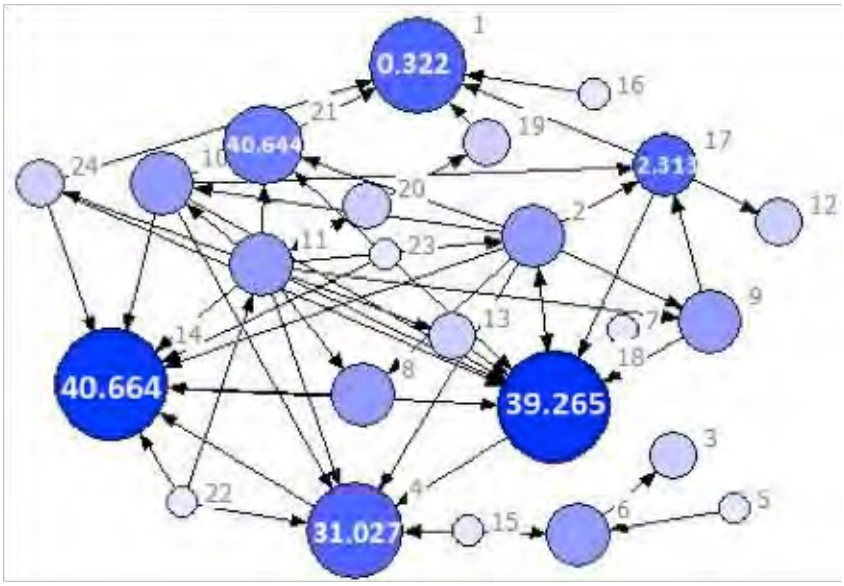


Figure 7. Portrayal of the co-citation network by the Outstanding team from Southeast University, China.

Tsinghua University, China: "Who is the Hidden Champion in a Network?"

This team performed standard network analysis on the Erdős co-author network, considering four standard centrality measures: degree centrality, eigenvector centrality, closeness centrality, and betweenness centrality. They did a very nice job of explaining each centrality measure and interpreting its meaning. The team evaluated the citation network both as a directed and as an undirected graph. While recognizing that the citation network is a directed network, they initially transformed it using symmetric relationships to find two related undirected networks, which they called a co-bibliography network and a co-citation network. The team then evaluated these related networks using the same centrality measures as they used for the co-author network. Next, the team analyzed the citation network as a directed graph, using applicable centrality measures such as the Modified Katz centrality.

The team then applied the methods that they used for the Erdős co-author network analysis to two different data sets: Chinese pop singers and Chinese movie actors. The network for singers had singers as vertices with an edge between two singers if they recorded a song written by the same songwriter. In the network for actors, an edge represents the fact that two actors are in a movie together. Upon completing this analysis, the team then developed a new approach: constructing a bipartite graph of singers-songwriters and actors-films, and using network centrality measures to rank the popularity of each set.

The judges were impressed by the team's clear presentation of the problem and the report's thorough and well-explained analysis. The executive summary and introduction were extremely well written, and the paper concluded with a clear discussion of the strengths and weaknesses of the team's analysis. One innovative aspect of the paper was the team's development and utilization of an algorithm called *stress majorization* to produce a graph of the network, which minimizes distance between vertices that are connected to present an optimal visual depiction of the network.

Tsinghua University, China: "A Three-Dimensional Network Impact Analysis Model Based on Centralizing, Connecting and Spreading Characteristics"

This team from Tsinghua University, as indicated in the title, realized that the task of identifying the most influential node in a network depends on the definition of influence. This team divided the concept of influence into three characteristics:

- centralizing characteristics, which aim to identify the nodes with the most central location in the topology of the network;
- connecting characteristics, which aim to identify the nodes whose positions are crucial to the connectedness of the whole network; and
- spreading characteristics, which aim to identify the nodes that are most capable of promoting flow of information through the network.

For the majority of these characteristics, the team carefully selected established measures from the field of network science and thoughtfully explained how each measure was relevant to its assigned family of characteristics. Specifically, the measures of degree centrality, eigenvector centrality, and page rank were chosen for the centralizing characteristics, while the measures of betweenness centrality, clustering coefficient, and a node removal method for evaluating total loss were chosen for the connecting characteristics. After finding only one established measure, that of closeness centrality, for their spreading characteristics, the team presented the clear development of two new network measures, *spreading breadth index* and *spreading depth index*, both of which factor time into the flow of information through the network.

After calculating all nine of these measures for each node, the team used principal component analysis for each family of characteristics, reducing the measure to a 3-vector, which could be reduced further to a scalar by applying weights based on the goal of the analysis. This approach was then applied to the co-author network, the paper citation network, and the users' comments on a Website for movies that are popular in China. Ultimately, the vectors for the most influential co-authors were presented visually on a three-dimensional graph. The team also presented very meaningful visualizations for their co-author and citation networks.

While many teams used similar sets of established network measures, the judges were impressed by this team's understanding and intuition about how each of these established metrics measured different aspects of influence. When they were not able to find established metrics that measured the elements that they were interested in capturing, this team developed new metrics and wrote detailed explanations of their measurement process. Additionally, this team presented a meaningful way to reduce all of these measures into a scalar, allowing ease of comparison and ranking of nodes. This talented team did a very strong job of connecting the mathematics to the more abstract meaning of influence through clear written prose. Additionally, their paper made excellent use of tables, charts, and graphical representations of their networks to convey results. See **Figure 8** for an example of the team's graphics to display the citation network.

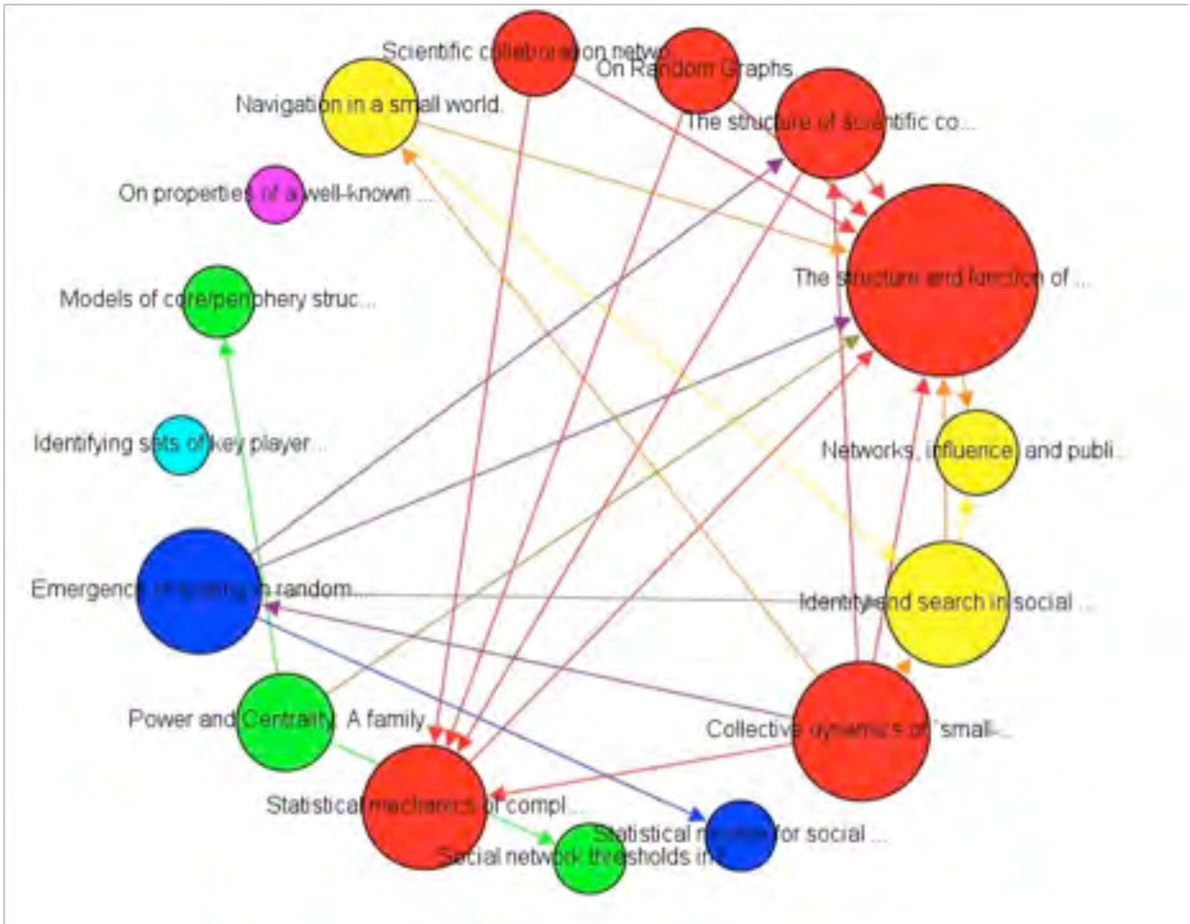


Figure 8. Network of fundamental papers' co-citations, by the Outstanding team from Tsinghua University, China.

The judges realize that given the timeline of the competition, not every team will have an opportunity to tackle masterfully all elements of the problem. This team appears to have done some nice analysis to identify the most influential user based on comments on a Website for popular movies in China; however, this section of the report was not as strongly presented as the other applications. Additionally, although the team did a nice job of explaining some factors that may contribute to the sensitivity of their model, they did not follow through with any computational results in their report.

Xidian University, China: "Methods of Measuring Influence Using Network Model"

The team from Xidian University did excellent work in

- laying out a set of criteria that one should consider in the evaluation of influence,
- how they mapped out these criteria to their specific approach, and finally,
- translating the algorithm components to meaning in a social context.

They developed two new measures of centrality to account for network influence: *importance degree* (combines degree centrality and clustering coefficient) and *influence degree* (weights PageRank with importance degree measure). They focus on these two dimensions of influence, which they attempt to describe: "[importance degree] reflects the researcher's ability to contribute to the ... network by contacting other researchers, while [influence degree]... shows [how] the researcher is affected by ... her/his partners and [how they] can ... [assert] her/his overall influence."

They validate their model with a network of actors who have collaborated with the Chinese movie star Tony Leung Chiu Wai. Analysis of importance degree and influence degree show that these metrics reflect different dimensions of social influence.

Importance degree ranks nodes by combining degree centrality (number of links) and clustering coefficient (links among neighbors), which the team turn into a piecewise function to deal with time of collaboration. In their piecewise function, they account for the year of collaboration between each author and Erdős. They identify the group who collaborated with Erdős in his earlier years as "old researchers," for whom they note "their frequent and early cooperation help them develop and grow in the collaborative network...." One of the most innovative aspects of this team's solution was their thorough handling of the dimension of time on influence.

The team powerfully presents a comparison of the two time periods to examine the changes over time. They found that mathematicians with large differences in their influence before and after their collaboration with Erdős are "young researchers" (those with later collaboration dates). While these researchers were less integrated in the network because they joined the collaboration late, they were successful in creating connections with high-influence researchers, thus drastically improving their influence metrics.

Authors who lose influence over time are those who have many collaborators in the Erdős 2 network (2 steps away from Erdős) but do not integrate into the Erdős 1 network. Judges were uncertain about some elements of the piecewise equation, making this aspect of the paper difficult to judge.

This team found a creative solution to identify social influence. Their algorithms enabled them to analyze these data, providing useful insights about the influence dynamics. For example, they suggest that a researcher can enhance their influence degree by cooperating with highly-influential researchers, even if they are themselves low-influence individuals; and high-influence individuals will lose influence over time if they are unable to collaborate with the core community. Crisp, clear graphical displays such as that in **Figure 9** helped the presentation of this paper.

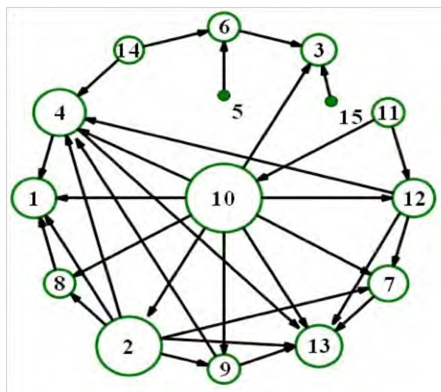


Figure 9. Influence links in the co-citation network, from the paper by the Outstanding team from Xidian University, China.

Conclusion

Among the 1,028 papers, there were many strong and innovative submissions that made judging both exciting and challenging. It was very gratifying to see so many students with the ability to combine modeling, science and effective communication skills in order to understand such large, complex datasets and build viable network models for their analysis.

Recommendations for Future Participants

- **Answer the problem.** Weak papers sometimes do not address a significant part of the problem. Outstanding teams often cover all the bases and then go beyond for some aspects of the problem.
- **Manage your time.** Every year there are submissions that do an outstanding job on one aspect of the problem, then "run out of gas" and are

unable to complete their solution. Outstanding teams have a plan and adjust as needed to submit a complete solution.

- **Coordinate your plan.** It is obvious in weaker papers that the work and writing was spilt between group members, then pieced together into the final report. For example, the output from one model or one step in a process doesn't match the input for the next model or a section appears in the paper that does not fit with the rest of the report. The more your team can coordinate the efforts of its members and integrate the writing, the stronger your final submission will be.
- **Do more than just model.** The model itself is not the solution. Some weak papers present a strong model, then stop. Outstanding teams use their models to produces results, understand the problem and recommend or produce a solution.
- **Explain what you are doing and why.** Weaker submissions tend to use too many equations and too few words. Problem approaches appear out of nowhere. Outstanding teams explain what they are doing and why.

References

Erdős, P., and A. Rényi. 1959. On random graphs, I. *Publicationes Mathematicae* 6: 290–297.

Grossman, Jerry. 2014. The Erdős Number Project. `http://www.oakland.edu/enp/`. Accessed 1 May 2014.

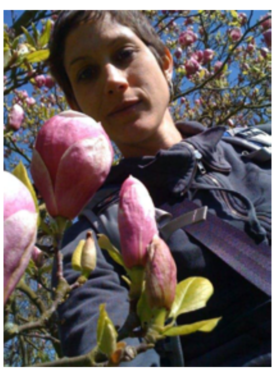
"This page was last updated on April 29, 2014 (but subpages may have been updated more recently). However, the lists of coauthors and the various other statistics on this site are updated about once every five years. The current version was posted on October 20, 2010 and includes all information listed inMathSciNet through mid-2010. The next update will probably occur around 2015."

Hoffman, Paul. 1998. *The Man Who Loved Only Numbers: The Story of Paul Erdős and the Search for Mathematical Truth*. New York: Hyperion.

O'Connor, J.J., and E.F. Robertson. 2000. Paul Erdős. `http://www-history.mcs.st-and.ac.uk/Biographies/Erdos.html`.

About the Authors

Chris Arney graduated from the U.S. Military Academy and served as an intelligence officer in the U.S. Army. His academic studies resumed at Rensselaer Polytechnic Institute with an M.S. (computer science) and a Ph.D. (mathematics). He spent most of his 30-year military career as a mathematics professor at West Point, before becoming Dean of the School of Mathematics and Sciences and Interim Vice President for Academic Affairs at the College of Saint Rose in Albany, NY. Chris then moved to Research Triangle Park, NC, where he served in various positions in the Army Research Office. His technical interests include mathematical modeling, cooperative systems, pursuit-evasion modeling, robotics, artificial intelligence, military operations modeling, and network science; his teaching interests include using technology and interdisciplinary problems to improve undergraduate teaching and curricula. He is the founding director of COMAP's Interdisciplinary Contest in Modeling (ICM). In August 2009, he rejoined the faculty at West Point as the Network Science Chair and Professor of Mathematics.



Kate Coronges received a Master's in Public Health and a Ph.D. from the University of Southern California in Human Health Behavior. She was an Assistant Professor in the Dept. of Behavioral Sciences and Leadership and a Research Fellow in the Network Science Center at the U.S. Military Academy for four years. Currently, she works as a Program Manager at the Army Research Office. Her research interests focus on the role of social and organizational network structures, and the dynamics of these networks, in communication patterns and performance of teams, groups, and societies. She is active in shaping and building momentum towards specific domains of social science basic research important in military settings, ranging from small-team dynamics (such as shared decision-making and collective intelligence), to belief and behavior propagation in groups and communities, to public and global policy (to include energy, education, information security and health care systems), and methodological challenges involved with modeling multidimensional and multifunctional systems.

Tina Hartley is an Academy Professor at the U.S. Military Academy and an active-duty officer. Her Ph.D. is in computational mathematics from George Mason University. Tina began her military career as an Air Defense Artillery Officer, and also served as an Operations Research Analyst. She is currently the Director of the Core Mathematics Program at the U.S. Military Academy. She has been an ICM judge for the past six years.

Jessica Libertini started her career as an engineer, earning a B.S. and an M.S. in mechanical engineering from Johns Hopkins University and Rensselaer Polytechnic Institute, respectively. She spent nine years at General Dynamics working on projects ranging from the design of submarines to the development of a multinational layered missile defense system. After earning her Ph.D. in applied mathematics from Brown University in 2008, Jessica left industry and began her academic career at the U.S. Military Academy, where she held the positions of Assistant Professor and National Research Council Fellow. While there, Jessica used her engineering background to motivate students to address large, open-ended, and meaningful questions, both in the classroom and as the coach of the competitive mathematics team. She has since served on the faculty at the University of Rhode Island and is currently an Assistant Professor of Applied Mathematics at Virginia Military Institute. Given her background in engineering, her research spans a wide variety of mathematical approaches to modeling, analyzing, and simulating medical, military, and physical applications; she also participates in the scholarship of teaching and learning, focusing on undergraduate pedagogy and the elements of a successful transition from high school to college.