# REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188



| 1. REPORT DATE (DD-MM-YYYY) | 2. REPORT TYPE | 3. DATES COVERED (From - To) |
|---|---|---|
| 18-11-2021 | Final Report | 18-Aug-2016 - 17-Aug-2021 |

**4. TITLE AND SUBTITLE**
Final Report: Forecasting Emergent Phenomena with Human Computer Collaboration

**5a. CONTRACT NUMBER**
W911NF-16-1-0524

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**
611102

**6. AUTHORS**

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAMES AND ADDRESSES**
Rensselaer Polytechnic Institute
110 8th Street

Troy, NY 12180 -3522

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)**
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

**10. SPONSOR/MONITOR'S ACRONYM(S)**
ARO

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**
67398-LS.60

**12. DISTRIBUTION AVAILIBILITY STATEMENT**
Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

**14. ABSTRACT**

**15. SUBJECT TERMS**

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 15. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Boleslaw Szymanski |
| UU | UU | UU | UU | | 19b. TELEPHONE NUMBER 518-276-6838 |

# RPPR Final Report
as of 19-Nov-2021

Agency Code: 21XD

Proposal Number: 67398LS **Agreement Number: W911NF-16-1-0524**
**INVESTIGATOR(S):**

**Name:** Boleslaw Szymanski
**Email:** szymab@rpi.edu
**Phone Number:** 5182766838
**Principal:** Y

Organization: **Rensselaer Polytechnic Institute**
Address: 110 8th Street, Troy, NY 121803522
Country: USA
DUNS Number: 002430742 EIN: 141340095
**Report Date:** 17-Nov-2021 Date Received: 18-Nov-2021
**Final Report** for Period Beginning 18-Aug-2016 and Ending 17-Aug-2021
**Title:** Forecasting Emergent Phenomena with Human Computer Collaboration
**Begin Performance Period:** 18-Aug-2016 **End Performance Period:** 17-Aug-2021
**Report Term:** 0-Other
Submitted By: Boleslaw Szymanski Email: szymab@rpi.edu
Phone: (518) 276-6838


**STEM Degrees:** 25 **STEM Participants:** 25

**Major Goals:** Our ultimate goal of this research is to utilize data crowd-sourced from the experts to gain a quantitative model of the state of global risk networks that may provide some actionable insights for the U.S. government and the U.S. Army. The global risk network by definition is composed of risks of varying complexity, potential economical damage, and probability of activating. Hence, the need for risk mitigation processes may vary from difficult to extremely complicated for various risks and impose various costs. The risk activation probabilities constantly evolve. Accordingly, our quantification of the activation probabilities and their impacts could provide a timely and invaluable guide to any cost-benefit analysis involved in the design of policies or strategies aimed at global risk minimization. The insights provided by our model may also enable the domain experts to provide tailor-made recommendations for the pertinent risks to curb the likelihood of systemic failures. The effects of such recommendations can also be tested using our model.
Accordingly, the final contribution of the research is to provide insights for the current state of the risks, guidance for curbing the probability of global failures and the tool for evaluating the proposed solutions. Guidance may include optimal risk network control with minimum control energy and minimum cost of deterring activation of risks. This in turn, may ultimately lead to the continuous control of the network of threats and an early response system to activating risks to reliably limit the disruptive effects of emergent phenomena. Last year, we accomplished the important step in this direction by developing a data driven minimum total cost continuous control system for global risks. It includes two phases: the first called reactive control deactivates currently active risks, followed by the proactive control that continuously deters new activation of risks. This year we have developed a new methodology for supervised Learning of the global risk network activation from media event reports. We also created and validated the software tool for creating a risk knowledge graph directly from the annotated risk events crawled from Wikipedia. We also demonstrated possible applications of our methodology to the current pandemic, showing that different approaches taken by different countries vary vastly on their efficiency.
Our current work aims at creating a methodology for incorporating lessons that we learned about the best human and machine partnerships for global risk networks in a generalized approach to threat networks on increasingly fine-grained levels of specificity such that the model can go from global, to regional, to national levels of analysis. Since the most dangerous form of risk failures are cascades of risk activation, and in view of current pandemics, we work on locating the source of spreading in many applications, from risk networks, to epidemic first patients, to misinformation sources. We improved currently existing algorithms for such detection. In the last year of the grant, we have considered an intriguing and promising direction of looking at the value of silence present in the parts of a network, to which spread has not yet reached. We obtained improved results. We finished this research and completed the paper for submission in the near future. Among risks that undermine the stability of nations and regions are social disruption. Accordingly, in the last year of the grant, we studied dynamics processes in complex social networks, ranging from the different kinds of society polarization to cultural completion and resilience. Finally,

the third topic addressed in the last year of the grant focused on an important issue of properly identifying communities in the complex social networks, especially those with incomplete information like covert terrorist networks. We completed this research and already published one paper and submitted the second paper for publication. Thanks to the high novelty of both papers, one is already the source for U.S. provisional patent, and the other is under consideration for a patent by the Rensselaer Office of Intellectual Property Innovation.

**Accomplishments:** We grouped our paper published or accepted in the reporting period into four clusters mentioned in the Accomplishment Sections. They are listed below.

I. Risk network control

1. "From Data to Complex Network Control of Airline Flight Delays," Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, and Boleslaw K. Szymanski, Scientific Reports 11:18715, Sept. 21, 2021, https://doi.org/10.1038/s41598-021-98112-7

We demonstrate how to design control of the dynamical system of airline flight delays directly from the date about airline flight performance collected by the government. We show that the optimal control significantly lowers the cost of airline flight delays for airlines, airports and passengers.

2. "Heuristic assessment of choices for risk network control," Christopher Brissette, Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, Boleslaw K. Szymanski, Scientific Reports 11:7645, April 7, 2021. https://doi.org/10.1038/s41598-021-85432-x

We present a possible application of our risk network control methodology to global pandemics. The results show that different combinations of countermeasures incur drastically different costs.

II. Dynamics Processes in Complex Networks

3. "Social Networks through the Prism of Cognition," Radoslaw Michalski, Boleslaw K. Szymanski, Przemysław Kazienko, Christian Lebiere, Omar Lizardo, Marcin Kulisiewicz, Complexity, 2021:4963903, Jan. 8, 2021, doi.org/10.1155/2021/4963903

We demonstrate that mimicking human memory via detailed cognitive models improves predicting evolution of social networks.

4. “Message sorting: a mechanism that hinders the system’s discriminative power,” Diego F.M. Oliveira, Kevin S. Chan, Physica A, submitted 2020.

We propose an agent-based model to study the effects of message sorting on the diffusion of low- and high-quality information in online social networks. We introduce a model that shows our approach to sorting messages intensifies the exposure of high-quality information increasing the overall system’s quality while preserving its diversity.

5. "Become a better you: Correlation between the change of research direction and the change of scientific performance," Xiaoyao Yu, Boleslaw K. Szymanski, Tao Jia, Journal of Informetrics 15(3):101193, Aug. 2021, https://doi.org/10.1016/j.joi.2021.101193

Here, we quantify the change of research direction in one’s early and late career. The direction change is associated with the chance to increase scientific impact, but noy with the productivity change; this change is uncorrelated with diversity or switching among topics. In conclusion it is beneficial to promote multi-disciplinary collaborations.

6. "A partial knowledge of friends of friends speeds social search," Amr Elsisy, Boleslaw K. Szymanski, Jasmine A. Plum, Miao Qi, Alex Pentland, PLoS ONE 16(8):e0255982, Aug. 19, 2021, https://doi.org/10.1371/journal.pone.0255982

The success of people in forwarding efficiently knowing only personal connections is still not fully explained. To study this problem, we emulate it on a real location-based social network, Gowalla that provides explicit information about friends and temporal locations of each user. We use it to conduct a massive computational experiment to

establish new necessary and sufficient conditions for achieving social search efficiency. The results show that only the distribution of friendship edges and the partial knowledge of friends of friends are essential and sufficient for the efficiency of social search.

7. "Nuclear reaction network unveils novel reaction patterns based on stellar energies," Chunheng Jiang, Boleslaw K Szymanski, Jie Lian, Shlomo Havlin and Jianxi Gao, New Journal of Physics 23 :083035, Aug.19, 2021, https: //doi.org/10.1088/1367-2630/ac1a3d

We assemble a nuclear reaction network in which a node represents a nuclide, and a link represents a direct reaction between nuclides. Interestingly, the degree distribution of the nuclear network is bimodal, unusual in nature. We discover three rules that govern the structure pattern of this network: (i) linking choices determine reaction-types, (ii) network distances between the reacting nuclides are short, and (iii) each node in- and out-degrees are close to each other. Incorporating these three rules unveils the underlying nuclear reaction patterns hidden in a large and dense nuclear reaction network.

III. Social Network Polarization

8. “Creation, Evolution, and Dissolution of Groups in Social Networks,” Ashwin Bahulkar, James Flamino, Boleslaw K. Szymanski, Kevin Chan, and Omar Lizardo, Scientific Reports 11:17470, September 1, 2021, https://doi.org/10. 1038/s41598-021-96805-7

We demonstrate that group dynamics is driven by members attempting to maximize their benefits from group interactions. This leads to across group polarization. We built a highly predictive analytical model for group dynamics to validate our conclusions.

9. "Polarization and tipping points," Michael W. Macy, Manqing Ma, Daniel R. Tabin, Jianxi Gao, and Boleslaw K. Szymanski, PNAS, June 2021, in press.

Our study was motivated by a highly disturbing puzzle. Confronted with a deadly global pandemic that threatened not only massive loss of life but also the collapse of our medical system and economy, why were we unable to put partisan divisions aside and unite in a common cause. We used a computational model to search for an answer in the phase transitions of political polarization. The model reveals asymmetric hysteresis trajectories with tipping points that are hard to predict and that make polarization extremely difficult to reverse once the level exceeds a critical value.

10. "Shifting Polarization and Twitter News Influencers between two U.S. Presidential Elections," James Flamino, Alessandro Galeazzi, Stuart Feldman, Michael W. Macy, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Alexandre Bovet, Hernan A. Makse, Boleslaw K. Szymanski, Science Advances, submitted Nov. 2021.

Social media are decentralized and interactive, which has transformed the dynamics of political communication. Using nearly a billion tweets, we measure the change between the 2016 and 2020 U.S. Presidential elections in the composition and polarization of Twitter influencers and the diffusion of political information among such influencers and those they influence. We identify ``influencers'' by their ability to spread information and classify them as affiliated with a media organization, a political organization, or unaffiliated. Most of the top influencers were affiliated with media organizations in both elections, yet 10 percent were replaced in 2020 by unaffiliated influencers or by those affiliated with center or right-orientated political organizations. From 2016 to 2020, polarization increased among influencers and among those whom they influence. Only a quarter of the top influencers in 2016 retained their status in 2020, showing high susceptibility to turnover.

IV. Challenges of Community Detection in Complex Networks

11. "A Network Generator for Covert Network Structures," Information Science, vol. 584, pp.387-398, https://doi. org/10.1016/j.ins.2021.10.066, online Nov. 6. 2021.

We design a novel method of generating random networks with varying variance and edge weight that are statistically similar to the original network but are void of privacy information. These networks enable a new structure analytics by finding community structure that is most resilient to perturbations.

12. "Minimizing Uncertainty Cost of Network Structure from Noisy Data," Amr Elsisy, Boleslaw K. Szymanski,

Social Networks, submitted Oct. 2021.

We use entropy measures to establish limits of the community structure uncertainty. For this structure we find a network with the minimum uncertainty about its communities. Finally, we introduce the novel measure of the cost of uncertainty and introduce an efficient heuristic that constructs a community structure with the minimal expected cost of uncertainty.

**Training Opportunities:** During the reporting period, we provided training opportunities in the STEM areas in research relevant to DoD. We trained one postdoctoral researcher, Dr. Diego Fregolente Mendes de Oliveira, who stayed 33 months and then joined the Department of Mathematics and The Statistical and Applied Mathematical Sciences Institute at the University of North Carolina, Chapel Hill, NC. We also graduate eight students in the reported period. Two Computer Science Ph.D.'s joined Google, Inc., one Computer Science Ph.D. joined Facebook, Inc. and three Physics Ph.D. accepted postdoc positions at CMU and at the IBM Research Center in Boston, MA, and in the Network Science and Technology Center at RPI. We also graduated five MS students in Computer Science. One of them joined a college as a teaching faculty, another joined Ph.D. program at Harvard University while the remaining three accepted offers from industry. We also involved 12 undergraduates in Computer Science research for various periods ranging from one to three semesters. They left with B.Sc. in Computer Science degrees and joined Ph.D. programs at other universities.

# RPPR Final Report
as of 19-Nov-2021

**Results Dissemination:** Results are divided into the topics used in the Accomplishment Section. Several planned dissemination events were canceled due to the pandemic.

I. Risk network control

1. Prof. Boleslaw Szymanski presented an invited talk titled “From Data to Complex Network Control With Applications to COVID-19 Pandemic and Airline Flight Delays,” by Gyorgy Korniss, Jianxi Gao, Xiang Niu, Chunheng Jiang Boleslaw K. Szymanski (RPI) at the 11th Polish Symposium on Physics in Economy and Social Sciences, FENS2021, online, on July 2, 2021,

II. Dynamics Processes in Complex Networks

2. Prof. Tarek Abdelzaher presented a talk titled “The Curse of Asymmetric Polarization Dynamics in the Age of Information Overload,” by Chao Xu, Jinyang Li, Dachun Sun, Ruijie Wang, Tarek Abdelzaher, Jesse Graham, and Boleslaw Szymanski at the International Workshop on Social Sensing (SocialSens 2021): Special Edition on Information Operations on Social Media on June 7, 2021.

3. Prof. Janusz Holyst presented an invited talk titled “Multiplicity of social attributes enhances Heider interactions in social triads,” by Piotr Górski, Joanna Linczuk, Raissa D’Souza, Boleslaw Szymanski, Klavdiya Bochenina and Janusz Holyst, at the 11th Polish Symposium on Physics in Economy and Social Sciences, FENS2021, online, on July 2, 2021,

4. Prof. Radoslaw Michalski presented a talk titled “Cognition-driven Temporal Social Networks,” by Radoslaw Michalski, Boleslaw K. Szymanski, Przemyslaw Kazienko, Christian Lebiere, Omar Lizardo, and Marcin Kulisiewicz at Networks 2021 Conference, online, July 6, 2021.

III. Social Network Polarization

5. Prof. Alexandre Bovet presented a talk titled “Twitter Influencers and Increased Media Polarization during two U. S. Presidential Elections,” by Alexandre Bovet, Stuart Feldman, James Flamino, Alessandro Galeazzi, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Hernán A. Makse, and Boleslaw K. Szymanski at the PCNET, A Satellite Workshop, at A Networks 2021 Conference, online on June 24, 2021.

6. Brendan Cross presented a talk titled “Twitter Influencers and Increased Media Polarization during two U.S. Presidential Elections,” by Alexandre Bovet, Stuart Feldman, James Flamino, Alessandro Galeazzi, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Hernán A. Makse, and Boleslaw K. Szymanski at the Computational Social Science Satellite Event at a Conference on Complex Systems CCS21, online on 27 October 2021.

IV. Challenges of Community Detection

7. Dr. Amr Elsisy presented at talk titled “A Synthetic Network Generator for Covert Network Analytics,” by Amr Elsisy, Aamir Mandviwalla, Boleslaw K. Szymanski, and Thomas Sharkey at the Communities in Networks, ComNets, A Satellite Workshop at Networks 2021 Conference, online on July 1, 2021.

8. Prof. Boleslaw Szymanski presented an invited talk titled “A Network Generator for Covert Network Structures,” by Amr Elsisy, Aamir Mandviwalla, Boleslaw K. Szymanski, and Thomas Sharkey at the USC 9/11 Symposium, at Los Angeles, CA, online on September 15, 2021.

**Honors and Awards:** Dr. Boleslaw Szymanski received:
(i) 2021 Network Science Society Service Award for his many service contributions to network science, including being a global strategic leader, fundraiser, and hands-on, on-the-ground institution builder. He is a third recipient of this award, after Dr. Albert Laszlo Barabasi, founder of the Society and Dr. Guido Caldarelli;
(ii) 2018 Life Fellow member of IEEE

**Protocol Activity Status:**

# RPPR Final Report
as of 19-Nov-2021

**Technology Transfer:** 1. Provisional U.S. patent: “A Method For Computing Likelihood Of Clustering For Network with Incomplete Information,” U.S. Patent Application No.: 63/168,216 submitted 03/30/2021, RPI Ref.: 2021-048, Law Office Ref.: 104640-100. (co-inventors: Boleslaw K. Szymanski, Amr Esisy, and Aamir Mandviwalla), assignee: RPI, Troy, NY., 03/31/2021

2. Invention Disclosure: "A method and Apparatus for Minimizing the Expected Cost of Uncertainty of Network Structure," under consideration for applying by the RPI Office of Intellectual Property Optimization, (co-inventors Boleslaw K. Szymanski and Amr Esisy), assignee: RPI, Troy, NY., 1/12/2021.

**PARTICIPANTS:**

**Participant Type:** PD/PI
**Participant:** Boleslaw K. Szymanski
**Person Months Worked:** 5.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Co PD/PI
**Participant:** Brian Uzzi
**Person Months Worked:** 1.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Postdoctoral (scholar, fellow or other postdoctoral position)
**Participant:** Diego Fregolente Oliveira
**Person Months Worked:** 15.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Xin Lin
**Person Months Worked:** 5.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Panagiotis Karampourniotis
**Person Months Worked:** 5.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Xiang Niu
**Person Months Worked:** 10.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Xiaoyan Lu
**Person Months Worked:** 5.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Stephen Dipple
**Person Months Worked:** 3.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Ashwin Bahulkar
**Person Months Worked:** 3.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Amr Elsisy
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** James Flamino
**Person Months Worked:** 1.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Partha Mukherjee
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Juntao Zhuang
**Person Months Worked:** 1.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)

**Participant:** Matthew Mohr
**Person Months Worked:** 5.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Daniel Tabin
**Person Months Worked:** 1.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Graduate Student (research assistant)
**Participant:** Bowen Gong
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Miao Qi
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Alex Lieberman
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Weiki: Dai
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Ted Louison
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Duo Jiang,
**Person Months Worked:** 2.00 **Funding Support:**

Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Alexander Meandzija
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** James Osika
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Ross DeVito
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Samuel Cohen
**Person Months Worked:** 4.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Andrew Smith
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Frederick Buchanan
**Person Months Worked:** 4.00 **Funding Support:**
Project Contribution:
National Academy Member: N

**Participant Type:** Undergraduate Student
**Participant:** Lake Yin
**Person Months Worked:** 2.00 **Funding Support:**
Project Contribution:
National Academy Member: N

# RPPR Final Report
as of 19-Nov-2021

**International Travel:**

POL 20 days

**International Collaboration:**

POL

**ARTICLES:**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Scientific Programming
Publication Identifier Type: Other Publication Identifier: REP-17-24177
Volume: 8 Issue: 1 First Page #: 2508
Date Submitted: 2/7/18 12:00AM Date Published: 2/6/18 5:00AM
Publication Location: on-line publisher site
**Article Title:** Fast and accurate detection of spread source in large complex networks
**Authors:** Robert Paluch, Xiaoyan Lu, Krzysztof Suchecki, Boles?aw K. Szymanski, Janusz A. Ho?yst
**Keywords:** information spread, complex networks, source detection
**Abstract:** Spread over complex networks is a ubiquitous process with increasingly wide applications. Locating spread sources is often important, e.g. finding the patient one in epidemics, or source of rumor spreading in social network. Pinto, Thiran and Vetterli introduced an algorithm (PTVA) to solve the important case of this problem in which a limited set of nodes act as observers and report times at which the spread reached them. PTVA uses all observers to find a solution. Here we propose a new approach in which observers with low quality information (i.e. with large spread encounter times) are ignored and potential sources are selected based on the likelihood gradient from high quality observers. Our Gradient GMLA reduces the PVTA complexity by at least N/log(N), where N is the network size. Extensive numerical tests performed on synthetic networks and real Facebook network demonstrate that for scale-free networks GMLA yields higher quality localization results than PTVA does.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Scientific Reports
Publication Identifier Type: DOI Publication Identifier: 10.1038/s41598-018-19726-y
Volume: 8 Issue: 1 First Page #:
Date Submitted: 2/4/18 12:00AM Date Published: 1/22/18 5:00AM
Publication Location: online
**Article Title:** Author Correction: Limits of Risk Predictability in a Cascading Alternating Renewal Process Model
**Authors:** Xin Lin, Alaa Moussawi, Gyorgy Korniss, Jonathan Z. Bakdash, Boleslaw K. Szymanski
**Keywords:** correction
**Abstract:** Correction to: Scientific Reports https://doi.org/10.1038/s41598-017-06873-x, published online 27 July 2017 The Acknowledgments section of this Article was incomplete and should read: “This work was partially supported by DTRA Award No. HDTRA1-09-1-0049, by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053 (the ARL Network Science Collaborative Technology Alliance), and by the Army Research Office under grant W911NF-16-1-0524. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies either expressed or implied of the U.S. Army Research Laboratory or the U.S. Government”.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Information Sciences
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.ins.2017.09.063
Volume: 424 Issue: First Page #: 55
Date Submitted: 2/4/18 12:00AM Date Published: 9/29/17 9:00AM
Publication Location: Amsterdam, Netherlands
**Article Title:** Adaptive modularity maximization via edge weighting scheme
**Authors:** Xiaoyan Lu, Konstantin Kuzmin, Mingming Chen, Boleslaw K. Szymanski
**Keywords:** community detection, modularity maximization, regression model
**Abstract:** Modularity maximization is the state-of-the-art methods for community detection. Yet it suffers from the resolution limit problem. To solve it, we expand the meaning of the edges that currently indicate only propensity for sharing the same community by allowing negative weights to indicate aversion for putting nodes into one community. We also present a novel regression model which assigns weights to the edges of a graph according to their local topological features to enhance the accuracy of modularity maximization algorithms. We construct artificial graphs based on the parameters sampled from a given unweighted network and train the regression model on ground truth communities of these artificial graphs in a supervised fashion. The extraction of local topological edge features is done in linear time. Experimental results on real and synthetic networks show that other community detection algorithms improve their performance significantly by finding communities in the weighted graphs.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Scientific Report
Publication Identifier Type: Other Publication Identifier:
Volume: 8 and 7 Issue: First Page #:
Date Submitted: 2/7/18 12:00AM Date Published: 1/22/18 10:00AM
Publication Location: internet on-line
**Article Title:** Limits of Risk Predictability in a Cascading Alternating Renewal Process Model
**Authors:** Xin Lin, Alaa Moussawi, Gyorgy Korniss, Jonathan Z. Bakdash & Boleslaw K. Szymanski
**Keywords:** risk network, limits of predictability, alternating renewal processes
**Abstract:** Most models ignore interconnectivity and interdependence of risks. We propose the Cascading Alternating Renewal Process (CARP) to forecast interconnected global risks. We also establish prediction precision as a function of input data size by using alternative long ground truth data generated by simulations of the CARP model with known parameters. We illustrate the approach on a model of fires in artificial cities assembled from basic city blocks with diverse housing. The results confirm that parameter recovery variance exhibits power law decay as a function of the length of available ground truth data. Using CARP, we also demonstrate estimation using a disparate dataset that also has dependencies: real-world prediction precision for the global risk model based on the World Economic Forum Global Risk Report. We conclude that the CARP model is an efficient method for predicting catastrophic cascading events with potential applications to emerging local and global interconnected risks.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** PLOS ONE
Publication Identifier Type: DOI Publication Identifier: 10.1371/journal.pone.0193007
Volume: 13 Issue: 3 First Page #:
Date Submitted: 8/18/18 12:00AM Date Published: 3/28/18 8:00AM
Publication Location: Internet
**Article Title:** Peer-to-peer lending and bias in crowd decision-making
**Authors:** Pramesh Singh, Jayaram Uparna, Panagiotis Karampourniotis, Emoke-Agnes Horvat, Boleslaw Szymar
**Keywords:** crowdfinancing, flat-world hypothesis, regression analysis
**Abstract:** We investigate the “flat-world” hypothesis that globalization leads to economic equality using crowdfinancing data for over 660,000 loans in 220 nations/territories made from 2005 to 2013. We find that peer-to-peer lending networks are decreasing flatness which is strongly associated with multiple variables: relatively stable patterns in the difference in the per capita GDP between borrowing and lending nations, ongoing migration flows from borrowing to lending nations worldwide, and the existence of a tie as a historic colonial and a spatial preference in lending for geographically proximal borrowers. To estimate the these patterns for future, we construct a network of borrower and lending nations based on the observed data, perturb the network and stochastically simulate policy and event shocks or regulatory shocks. The simulations project a drift towards rather than away from flatness, unless there is a loss of the top borrowing nations, which decreases flatness.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Applied Network Science
Publication Identifier Type: DOI Publication Identifier: 10.1007/s41109-018-0077-0
Volume: 3 Issue: 24 First Page #:
Date Submitted: 8/18/18 12:00AM Date Published: 8/1/18 4:00AM
Publication Location: Internet
**Article Title:** Evolution of threats in the global risk network
**Authors:** Xiang Niu, Alaa Moussawi, Gyorgy Korniss, Boleslaw K. Szymanski
**Keywords:** Global risk network, Cascading failures, Network evolution, Mean-field steady state
**Abstract:** With growing population and advancements in technology, the global economy steadily increases in size and complexity which exacerbates global vulnerabilities and may lead to unforeseen consequences such as global pandemics fueled by air travel, or cascading failures caused by the weakest link in a supply chain. So a quantitative understanding of the mechanisms driving global network vulnerabilities is urgently needed. Each year the World Economic Forum publishes an authoritative report on the state of the global economy and identifies dangerous risks. Using a Cascading Alternating Renewal Process approach to model these risks, we can answer critical questions about their evolution. We analyze the asymptotic state of risks (risk levels which would be reached in the long term if the risks were left unabated) given a snapshot in time; this elucidates the various challenges faced by the world community over time. We also investigate the influence exerted by each risk on others.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** IEEE Transactions on Computational Social Systems
Publication Identifier Type: DOI Publication Identifier: 10.1109/TCSS.2018.2857479
Volume: 5 Issue: 3 First Page #: 1
Date Submitted: 8/18/18 12:00AM Date Published: 8/17/18 8:00AM
Publication Location: Internet
**Article Title:** Scalable Prediction of Global Online Media News Virality
**Authors:** Xiaoyan Lu, Boleslaw K. Szymanski
**Keywords:** community detection, information cascades, online media, parallelization, supercomputer.
**Abstract:** News reports shape the public perception of the critical social, political, and economical events around the world yet predicting which news will go viral is a challenging task. We propose a scalable community-based probabilistic framework to model the spreading of news about events in online media. Our approach exploits the latent community structure in the global news media and uses the affiliation of the early adopters with a variety of communities to identify the events widely reported in the news at the early stage of their spread. The time complexity of our approach is linear in the number of news reports. This approach is also amenable to efficient parallelization and was parallelized for message passing paradigm and tested on the RPI AMOS, one of the fastest Blue Gene/Q supercomputers in the world. Compared to the exiting methods, the model improves early news virality predictability by 20% and its parallelization achieves orders of magnitude speedup.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Scientific Reports
Publication Identifier Type: Other Publication Identifier: arxiv.org/pdf/1803.02961.pdf
Volume: 9 Issue: First Page #: 5573
Date Submitted: 5/24/19 12:00AM Date Published: 3/8/18 10:00AM
Publication Location: Internet
**Article Title:** Influence Maximization for Fixed Heterogeneous Thresholds
**Authors:** Panagiotis. D. Karampourniotis, Boleslaw K. Szymanski, Gyorgy Korniss
**Keywords:** Influence Maximization, threshold model, variable threshold, spread in complex networks
**Abstract:** Influence Maximization is a NP-hard problem of selecting the optimal set of influencers in a network. Here, we propose two new node metrics to guide influence maximization: Balanced Index (BI) and Group Performance Index (GPI). The first is fast to compute and assigns top values to nodes with high resistance to adoption, and with large out-degree. The second metric measures performance of each node as an initiator when it is a part of randomly selected initiator set of which all teammate are assigned a score inversely proportional to the number of initiators causing the desired spread. We test them on the Linear Threshold Model with fixed and known thresholds. We also study the impact of network degree assortativity and threshold distribution on the cascade size for metrics including ours. The results demonstrate our two metrics deliver strong performance for influence maximization.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Physica A
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.physa.2019.03.0340378-4371
Volume: 525 Issue: First Page #: 657
Date Submitted: 5/24/19 12:00AM Date Published: 4/1/19 8:00AM
Publication Location: Internet
**Article Title:** The effects of trust and influence on the spreading of low and high quality information
**Authors:** Diego F. M. Oliveira, and Kevin S. Chan
**Keywords:** Networks, limited attention, competition, information
**Abstract:** In this work, we employ a minimal agent-based model to explore the mechanisms that regulate competition between memes that spread online. We investigate the case in which each piece of information has a quality, and the higher is the quality the higher are the chances of being transmitted. The model allows us to study the impact of influential nodes on the spreading behavior. We show that meme's quality does not guarantee virility, but there is a strong correlation between the meme's success and the influence of the agent who introduced it. When trust is incorporated into the model and the agents can decided whether or not to accept a meme, we show that both lifetime and popularity distributions have broad power-law tails indicating that only a few memes spread virally through the population reproducing perfectly the broad distributions obtained from empirical data.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 5-Submitted
**Journal:** Nature Communication
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 5/24/19 12:00AM Date Published: 5/24/19 8:00AM
Publication Location: internet
**Article Title:** Data-driven control of networked risks with minimal cost
**Authors:** Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, and Boleslaw K. Szymanski
**Keywords:** Risk networks, optimal control, global risks
**Abstract:** Global risks are intricately interconnected with one another through an underlying network that creates pathways along which dangerous and damaging events can rapidly and broadly spread, threatening modern societies, yet we lack control tools for diminishing the networked risks with minimal cost. Risk dynamics entangles both the cost of control input to drive active risks to inactivity and keep inactive risks in that state and the cost of damage caused by the currently risks until they are inactivated. This creates the trade-off between proactive and reactive risk control. We develop a mathematical framework using the Linear-Quadratic Regulator (LQR) control and validate it in both the annual global risk networks for the past 5 years and the networked domestic flight delay risks of U.S. airlines. Our analytical tools yield the minimal total cost of controlling the different types of networked risks, and informs the specific control strategies for networked risk management.
**Distribution Statement:** 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Applied Network Science
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 8/21/21 12:00AM Date Published: 5/2/19 8:00AM
Publication Location: internet
**Article Title:** Message sorting: a mechanism that hinders the system's discriminative power
**Authors:** Diego F. M. Oliveira, and Kevin S. Chan
**Keywords:** Message sorting, discriminative power, agent-based models
**Abstract:** Data-driven risk networks describe many complex system dynamics arising in fields such as epidemiology and ecology. They lack explicit dynamics and have multiple sources of cost, both of which are beyond the current scope of traditional control theory. We construct the global economy risk network by combining the consensus of experts from the World Economic Forum with risk activation data to define its topology and interactions. Many of these risks pose significant economic costs when active. We introduce a method for converting network interaction data into continuous dynamics to which we apply optimal control. We contribute the first method for constructing and controlling risk network dynamics based on empirically collected data. We simulate applying this method to control the spread of COVID-19 and show that the choice of risks through which the network is controlled has significant influence on both the cost of control and the total cost of keeping network stable.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Computer Science (Springer)
Publication Identifier Type: Other Publication Identifier: paper 8
Volume: Issue: First Page #:
Date Submitted: 8/28/20 12:00AM Date Published: 10/15/19 12:00PM
Publication Location: Berlin, Germany
**Article Title:** Supervised Learning of the Global Risk Network Activation from Media Event Reports
**Authors:** Xiang Niu, Gyorgy Korniss, Boleslaw K. Szymanski
**Keywords:** Global risk network, Knowledge graph, Auto-detection,Supervised learning
**Abstract:** The World Economic Forum (WEF) publishes annual reports of global risks highly impact the world's economy. Analyzing and modeling of evolution of risks is popular. Yet, few studies focus on validation of the global risk networks published by the WEF. Here, we fi?rst create a risk knowledge graph from the annotated risk events crawled from the Wikipedia. Then, we compare the dependencies of risks in the WEF and Wikipedia networks, and fi?nd that they share over 50% of their edges. Edges unique to each network signify the different perspectives of the experts and the public on global risks. To reduce the cost of manual annotation of events triggering risk activation, we build an auto-detection tool to filters out over 80% media reported unrelated events During fi?lterring, our tool also continuously learns relevant keywords from the event sentences. Using locations extracted from the risk knowledge graph, we find characteristics of spatial distributions of the global risk categories.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Information Sciences
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.ins.2020.03.082
Volume: 525 Issue: First Page #: 54
Date Submitted: 8/28/20 12:00AM Date Published: 7/1/20 4:00AM
Publication Location:
**Article Title:** Asymptotic resolution bounds of generalized modularity and multi-scale community detection
**Authors:** Xiaoyan Lu, Brendan Cross, Boleslaw K. Szymanski
**Keywords:** community detection, modularity maximization, resolution limit, stochastic block model, Bayes model selection
**Abstract:** There is no theory bounding the maximization performance in realistic networks where edges are heterogeneously distributed. Using the random graph properties, we establish asymptotic theoretical bounds on the resolution parameter for which the generalized modularity maximization performs well. From this new perspective, we explain the resolution limit of modularity maximization in a simple and straightforward way. Given a network produced by the stochastic block models, the communities for which the resolution parameter is larger than their densities are likely to be spread among multiple clusters, while for communities for which this parameter is smaller than their inter community edge density will be merged. No suitable resolution parameter exits when the intracommunity edge density in a subgraph is lower than the inter-community edge density in some other subgraph. For such networks, we propose a progressive agglomerative heuristic algorithm to detect significant communities.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** PLOS ONE
Publication Identifier Type: DOI Publication Identifier: 10.1371/journal.pone.0232888
Volume: 15 Issue: 5 First Page #:
Date Submitted: 8/28/20 12:00AM Date Published: 5/1/20 4:00AM
Publication Location:
**Article Title:** Modeling competitive evolution of multiple languages
**Authors:** Zejie Zhou, Boleslaw K. Szymanski, Jianxi Gao, Ming Tang
**Keywords:** language evolution, cultural diversity, language extinction
**Abstract:** There is no theory bounding the maximization performance in realistic networks where edges are heterogeneously distributed. Using the random graph properties, we establish asymptotic theoretical bounds on the resolution parameter for which the generalized modularity maximization performs well. From this new perspective, we explain the resolution limit of modularity maximization in a simple and straightforward way. Given a network produced by the stochastic block models, the communities for which the resolution parameter is larger than their densities are likely to be spread among multiple clusters, while for communities for which this parameter is smaller than their inter community edge density will be merged. No suitable resolution parameter exits when the intracommunity edge density in a subgraph is lower than the inter-community edge density in some other subgraph. For such networks, we propose a progressive agglomerative heuristic algorithm to detect significant communities.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Applied Network Science
Publication Identifier Type: DOI Publication Identifier: 10.1007/s41109-019-0238-9
Volume: 4 Issue: 1 First Page #:
Date Submitted: 8/28/20 12:00AM Date Published: 12/1/19 5:00AM
Publication Location:
**Article Title:** On community structure in complex networks; challenges and opportunities
**Authors:** Hocine Cherifi, Gergely Palla, Boleslaw K. Szymanski, Xiaoyan Lu
**Keywords:** community detection, stochastic block model, time evolving networks, Immunization, centrality, epidemic spreading
**Abstract:** Community structure is a very relevant feature encountered in numerous real-world applications of networked systems. Despite the tremendous effort of a large interdisciplinary community of scientists working on this subject over the past few decades to characterize, model, and analyze communities, more investigations are needed in order to better understand the impact of community structure and its dynamics on networked systems. Here, we first focus on generative models of communities in complex networks and their role in developing strong foundation for community detection algorithms. We discuss modularity and the use of modularity maximization as the basis for community detection. Then, we follow with an overview of the Stochastic Block Model and its different variants as well as inference of community structures from such models. Next, we focus on time evolving networks, Finally, we discuss immunization strategies essential for targeting the influential spreaders of epidemics.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Future Generation Computer Systems
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.future.2020.06.023
Volume: 112 Issue: First Page #: 1070
Date Submitted: 8/28/20 12:00AM Date Published: 11/1/20 4:00AM
Publication Location:
**Article Title:** Optimizing sensors placement in complex networks for localization of hidden signal source: A review
**Authors:** Robert Paluch, Lukasz G. Gajewski, Janusz A. Holyst, Boleslaw K. Szymanski
**Keywords:** information diffusion, sensors placement, source localization, node selection strategies
**Abstract:** As the world becomes more and more interconnected, misinformation, fake news and malware, can spread very fast practically anonymously. We need systems to suppress such spread. Currently, the most effective methods for source localization are based on sensors which provide the times at which they detect the spread. We investigate the problem of the optimal placement of such sensors in complex networks and propose a new graph measure, called Collective Betweenness (CB), which we compare against four other metrics. Extensive numerical tests show clear difference in performance of the investigated methods for real or scale-free synthetic networks versus narrow degree distribution networks. The former have a clear region for any given method’s dominance in contrast to the latter where the performance is less homogeneous. We find there are two methods that consistently stand out. High Variance Observers do well on low stochasticity spread while our CB is best on highly unpredictable spread.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Nature Scientific Reports
Publication Identifier Type: DOI Publication Identifier: 10.1038/s41598-021-96805-7
Volume: 11 Issue: 9 First Page #: 17470
Date Submitted: 9/4/21 12:00AM Date Published: 9/1/21 4:00AM
Publication Location: Internet
**Article Title:** Creation, Evolution, and Dissolution of Groups in Social Networks
**Authors:** Ashwin Bahulkar, James Flamino, Boleslaw K. Szymanski, Kevin Chan, and Omar Lizardo
**Keywords:** group dynamics, group polarization, utility of ingroup interactions,
**Abstract:** Understanding why people join, stay, or leave social groups is of great interest to the social sciences. Yet, current empirical studies on groups rarely focus on group dynamics for lack of data combing people's opinions and group membership. We extract hundreds of groups and thousands of decisions about groups' memberships from the empirical NetSense data. We discover two group dynamic properties. First, the group opinion homogeneity increases with time. Second, the satisfaction from ingroup interactions depends on the conformity of member's opinions to those of the group. We define the gain of utility that members accumulate from ingroup interactions. We design two utility-based models of group dynamics. Both make accurate predictions. They show that members maximizing utility give rise to the discovered properties.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Complexity
Publication Identifier Type: DOI Publication Identifier: 10.1155/2021/4963903
Volume: 2021 Issue: First Page #:
Date Submitted: 1/20/21 12:00AM Date Published:
Publication Location: on line
**Article Title:** Social Networks through the Prism of Cognition
**Authors:** Radoslaw Michalski, Boleslaw Karol Szymanski, PrzemysLaw Kazienko, Christian Lebiere, Omar Lizard
**Keywords:** social network, network dynamics, memory model, cognition
**Abstract:** Human relations are driven by social events - people interact, exchange information, share knowledge and emotions, and gather news from mass media. These events leave traces in human memory, the strength of which depends on cognitive factors such as emotions or attention span. Each trace continuously weakens over time unless another related event activity strengthens it. Here, we introduce a novel Cognition-driven Social Network (CogSNet) model that accounts for cognitive aspects of social perception. The model explicitly represents each social interaction as a trace in human memory with its corresponding dynamics. The strength of the trace is the only measure of the influence that the interactions had on a person. For validation, we apply our model to NetSense data on social interactions among university students. The results show that CogSNet significantly improves the quality of modeling of human interactions in social networks.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Nature Scientific Reports
Publication Identifier Type: DOI Publication Identifier: 10.1038/s41598-021-85432-x
Volume: 11 Issue: 4 First Page #: 7645
Date Submitted: 8/21/21 12:00AM Date Published: 4/7/21 4:00AM
Publication Location: internet
**Article Title:** Heuristic assessment of choices for risk network control
**Authors:** Christopher Brissette, Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, Boleslaw K. Szymanski
**Keywords:** network control, pandemics, data-driven networks, global risk networks
**Abstract:** Data-driven risk networks describe many complex system dynamics arising in fields such as epidemiology and ecology. They lack explicit dynamics and have multiple sources of cost, both of which are beyond the current scope of traditional control theory. We construct the global economy risk network by combining the consensus of experts from the World Economic Forum with risk activation data to define its topology and interactions. Many of these risks pose significant economic costs when active. We introduce a method for converting network interaction data into continuous dynamics to which we apply optimal control. We contribute the first method for constructing and controlling risk network dynamics based on empirically collected data. We simulate applying this method to control the spread of COVID-19 and show that the choice of risks through which the network is controlled has significant influence on both the cost of control and the total cost of keeping network stable.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 4-Under Review
**Journal:** Expert System with Applications
Publication Identifier Type: Other Publication Identifier: arXiv:2008.04445v1
Volume: Issue: First Page #:
Date Submitted: 8/28/20 12:00AM Date Published:
Publication Location:
**Article Title:** Synthetic Network Generator for Covert Network Analytics
**Authors:** Amr Elsisy, Aamir Mandviwall, Boleslaw K. Szymanskia, Thomas Sharkey
**Keywords:** social networks, random weighted network generator, network structure stability, covert networks, hierarchical networks
**Abstract:** We introduce here a network generator for synthetic networks that are statistically similar to a real network but void of personal information about its members. The generator uses statistical data about a real or imagined covert organization network. It generates randomized instances of the Stochastic Block model of the network groups but preserves this network organizational structure. Their direct use o is to train on them the research and analytical tools for finding structure and dynamics of covert networks. Since these synthetic networks differ from each other, they can be used as a new source for network analytics. They provide alternative interpretations of the data about the original network. The analysts can find community structures which are frequent, therefore stable under perturbations. They may also analyze how how the stability changes with the strength of perturbation. For covert networks, the analysts can quantify statistically expected outcomes of interdiction.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Chaos; An Interdisciplinary Journal of Nonlinear Science
Publication Identifier Type: DOI Publication Identifier: 10.1063/5.0004983
Volume: 30 Issue: 6 First Page #: 063112
Date Submitted: 8/28/20 12:00AM Date Published: 6/1/20 4:00AM
Publication Location:
**Article Title:** Scaling laws and dynamics of hashtags on Twitter
**Authors:** Hongjia H. Chen, Tristram J. Alexander, Diego F. M. Oliveira, Eduardo G. Altmann
**Keywords:** hashtag frequency, twitter, scaling law
**Abstract:** Here we quantify the statistical properties and dynamics of the frequency of hashtag use on Twitter. Looking at the collection of all hashtags used in a period of time, we identify the scaling laws underpinning the hashtag frequency distribution (Zipf's law), the number of unique hashtags as a function of sample size (Heaps' law), and the fluctuations around expected values (Taylor's law). While these scaling laws appear to be universal, in the sense that similar exponents are observed irrespective of when the sample is gathered, the volume and nature of the hashtags depends strongly on time, with the appearance of bursts at the minute scale, fat-tailed noise, and long-range correlations. We quantify this dynamics by computing the Jensen-Shannon divergence between hashtag distributions obtained t times apart and we find that the speed of change decays roughly as 1/4. Our findings are Our findings are based on the analysis of 3.5 billion hashtags used between 2015 and 2016.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 5-Submitted
**Journal:** Physica A
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 8/21/21 12:00AM Date Published:
Publication Location:
**Article Title:** Message sorting: a mechanism that hinders the system's discriminative power
**Authors:** Diego F. M. Oliveira, Kevin S. Chan
**Keywords:** networks, competition, limited attention, information load.
**Abstract:** In this work, we propose an agent-based model to study the effects of message sorting on the diffusion of low- and high-quality information in online social networks. We investigate the case in which each piece of information has a numerical proxy representing its quality, and the higher the quality, the greater are the chances of being transmitted further in the network. The model allows us to study how sorting information in the agent’s attention list according to their quality, node’s influence and popularity affect the overall system’s quality, diversity and discriminative power. We compare the three scenarios with a baseline model where the information is organized in a first-in first out manner. Our results indicate that such approach intensifies the exposure of high-quality information increasing the overall system’s quality while preserving its diversity. However, it deceases significantly the system’s discriminative power.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 4-Under Review
**Journal:** PNAS USA
Publication Identifier Type: DOI Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 3/4/21 12:00AM Date Published:
Publication Location:
**Article Title:** Emergence of New Influencer Elites Challenges Traditional Polarization Analysis
**Authors:** Stuart Feldman, James Flamino, Alessandro Galeazz, Brendan Cross, Zhenkun Zhou, Matteo Serafino,
**Keywords:** Political elites,| Political polarization, Misinformation, Influencers, ction, Social Media
**Abstract:** Analyses of polarization frequently distinguish “mass” and “elite”, but the traditional meaning of elite may no longer be the most relevant one in the age of social media. We propose to consider a new category of “influencers” in the current and future environments. The standard definitions of elites focus either on a small self-reinforcing grouping of political, economic, corporate, and communications leaders (like traditional media), or on the larger group of people who are even mildly directly active in politics. However, the Internet and social media have democratized access to publishing, and moved the focus from ability to publish to ability to influence. The new influencers come from a variety of backgrounds and are increasingly important in shaping society and politics. Here, we analyze differences between the traditional elites and new influencers, and their respective influence techniques used as political and diplomatic strategies.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 2-Awaiting Publicat
**Journal:** PNAS
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 8/21/21 12:00AM Date Published:
Publication Location:
**Article Title:** Polarization and Tipping Points
**Authors:** Michael W. Macya, Manqing Ma, Daniel R. Tabin, Jianxi Gao, and Boleslaw K. Szymanski
**Keywords:** Polarization, Phase transition,Tipping points, Hysteresis dynamics
**Abstract:** Research has documented increasing partisan division and extremist positions that are more pronounced among political elites than among voters. Attention has now begun to focus on how polarization might be attenuated. We use a general model of opinion change to see if the self-reinforcing dynamics of influence and homophily may be characterized by tipping points that make reversibility problematic. The model applies to a legislative body or other small, densely connected organization but does not assume country-specific institutional arrangements that would obscure the identification of fundamental regularities in the phase transitions. Agents in the model have initially random locations in a multidimensional issue space consisting of membership in one of two equal-sized parties and positions on ten issues. Agents then update their issue positions by moving closer to nearby neighbors and farther from those with whom they disagree, depending on the agents’ tolerance of disagreement.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** PLOS ONE
Publication Identifier Type: DOI Publication Identifier: 10.1371/journal.pone.0255982
Volume: 16 Issue: 8 First Page #: e0255982
Date Submitted: 8/21/21 12:00AM Date Published: 8/19/21 4:00AM
Publication Location: Internet
**Article Title:** A partial knowledge of friends of friends speeds social search
**Authors:** Amr Elsisy, Boleslaw K. Szymanski, Jasmine A. Plum, Miao Qi, Alex Pentland, Jaros?aw Jankowski
**Keywords:** social search, social routing, Milgram small world experiment
**Abstract:** Milgram empirically showed that people knowing only connections to their friends could locate any person in the U.S. in a few steps. Social network topology enables a node aware of its full routing to find an arbitrary target in even fewer steps. Yet, the success of people in forwarding efficiently knowing only personal connections is still not fully explained. We emulate it on a real location-based social network, Gowalla. It provides explicit information about friends and temporal locations of each user useful for studies of human mobility. We use it to conduct a massive computational experiment to establish new necessary and sufficient conditions for social search efficiency. Only the distribution of friendship edges and the partial knowledge of friends of friends are essential and sufficient for the efficiency of social search. Using a knowledge that about friends of friends is strongly nonlinear and most efficient with small fraction of it is used.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** New Journal of Physics
Publication Identifier Type: DOI Publication Identifier: 10.1088/1367-2630/ac1a3d
Volume: 23 Issue: 8 First Page #: 083035
Date Submitted: 8/21/21 12:00AM Date Published: 8/19/21 4:00AM
Publication Location: Internet
**Article Title:** Nuclear reaction network unveils novel reaction patterns based on stellar energies
**Authors:** Chunheng Jiang, Boleslaw K Szymanski, Jie Lian, Shlomo Havlin, Jianxi Gao
**Keywords:** nuclear reaction, network science, bimodal degree distribution, machine learning
**Abstract:** Milgram empirically showed that people knowing only connections to their friends could locate any person in the U.S. in a few steps. Social network topology enables a node aware of its full routing to find an arbitrary target in even fewer steps. Yet, the success of people in forwarding efficiently knowing only personal connections is still not fully explained. We emulate it on a real location-based social network, Gowalla. It provides explicit information about friends and temporal locations of each user useful for studies of human mobility. We use it to conduct a massive computational experiment to establish new necessary and sufficient conditions for social search efficiency. Only the distribution of friendship edges and the partial knowledge of friends of friends are essential and sufficient for the efficiency of social search. Using a knowledge that about friends of friends is strongly nonlinear and most efficient with small fraction of it is used.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Journal of Informetrics
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.joi.2021.101193
Volume: 15 Issue: 3 First Page #: 101193
Date Submitted: 8/21/21 12:00AM Date Published: 7/1/21 4:00AM
Publication Location: Inernet
**Article Title:** Become a better you: Correlation between the change of research direction and the change of scientific performance
**Authors:** Xiaoyao Yu, Boleslaw K. Szymanski, Tao Jia
**Keywords:** science of science, scientific productivity, evolution of research interests
**Abstract:** Scientists deciding their research agenda collectively shape contemporary science. Here we ask a unexplored question: how is a scientist's change of research agenda affects scientific performance. Using publication records of over 14,000 authors in physics, we quantitatively measure the extent of research direction and performance changes of individuals. We identify a strong positive correlation between the direction change and impact change. Scientists changing direction produce works with increased scientific impact compared to their past ones, but also have a higher growth rate of scientific impact. Yet, the direction change is not associated with productivity change. Those who stay in familiar topics do not publish faster than those who venture out into new fields. We show that the probability of switching topics is uncorrelated to the diversity of research agenda. Our finding sheds light on a range of problems in the career development of individual scientists.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Information Sciences
Publication Identifier Type: DOI Publication Identifier: 10.1016/j.ins.2021.10.066
Volume: 584 Issue: First Page #: 387
Date Submitted: 11/16/21 12:00AM Date Published: 1/1/22 5:00AM
Publication Location: online
**Article Title:** A network generator for covert network structures
**Authors:** Amr Elsisy, Aamir Mandviwalla, Boleslaw K. Szymanski, Thomas Sharkey
**Keywords:** Random weighted network generator, Network structure stability, Covert networks, Hierarchical networks
**Abstract:** We focus on organizational structures in covert networks, such as criminal or terrorist networks. Their members engage in illegal activities and attempt to hide their association and interactions with these networks. We introduce a novel method of rewiring covert networks parameterized by the edge connectivity standard deviation and statistically similar to themselves and to the original network. The higher-level organizational structures are modeled as a multi-layer network while the lowest level uses the Stochastic Block Model. Such synthetic networks provide alternative structures for data about the original network. Using them, analysts can find structures that are frequent, therefore stable under perturbations. In experiments, we generate many synthetic networks from two covert networks. Only a few structures of synthetics networks repeat, with the most stable ones shared by 18% of all synthetic networks making them strong candidates for the ground truth structure.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
**Journal:** Scientific Reports
Publication Identifier Type: DOI Publication Identifier: 10.1038/s41598-021-98112-7
Volume: 11 Issue: 1 First Page #:
Date Submitted: 11/16/21 12:00AM Date Published: 9/1/21 4:00AM
Publication Location: online
**Article Title:** From data to complex network control of airline flight delays
**Authors:** Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, Boleslaw K. Szymanski
**Keywords:** complex networks, network control, delayed flights
**Abstract:** Many critical complex systems and networks are continuously monitored, creating vast volumes of data describing their dynamics. We introduce a multidisciplinary framework using network science and control theory to accomplish these goals. We demonstrate its use on a meaningful example of a complex network of U.S. domestic passenger airlines aiming to control flight delays. Using the real data on such delays, we build a flight delay network for each airline. Analyzing these networks, we uncover and formalize their dynamics. We use this formalization to design the optimal control for the flight delay networks. The results of applying this control to the ground truth data on flight delays demonstrate the low costs of the optimal control and significant reduction of delay times, while the costs of the delays unabated by control are high. Thus, the introduced here framework benefits the passengers, the airline companies and the airports.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 5-Submitted
**Journal:** Science Advances
Publication Identifier Type: Other Publication Identifier: arXiv:2111.02505
Volume: Issue: First Page #:
Date Submitted: 11/18/21 12:00AM Date Published:
Publication Location:
**Article Title:** Shifting Polarization and Twitter News Influencers between two U.S. Presidential Elections
**Authors:** James Flamino, Alessandro Galeazzi, Stuart Feldman, Michael W. Macy, Brendan Cross, Zhenkun Zho
**Keywords:** political polarization, Tweeter influencers, spread of opinions, complex networks
**Abstract:** Social media are decentralized and interactive, which has transformed the dynamics of political communication. Using nearly a billion tweets, we measure the change between the 2016 and 2020 U.S. Presidential elections in the composition and polarization of Twitter influencers and the diffusion of political information among such influencers and those they influence. We identify “influencers” by their ability to spread information and classify them into those affiliated with a media organization, a political organization, or unaffiliated. Most of the top influencers were affiliated with media organizations during both elections, yet 10 percent were replaced in 2020 by unaffiliated influencers or by those affiliated with center or right-orientated political organizations. From 2016 to 2020, polarization increased among influencers and among those whom they influence. Only a quarter of the top influencers in 2016 retained this status in 2020, indicating high susceptibility to turnover.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**Publication Type:** Journal Article Peer Reviewed: Y **Publication Status:** 5-Submitted
**Journal:** Social Networks
Publication Identifier Type: Other Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 11/18/21 12:00AM Date Published:
Publication Location:
**Article Title:** Minimizing Uncertainty Cost of Network Structure from Noisy Data
**Authors:** Amr Elsisy and Boleslaw K. Szymanski
**Keywords:** Noisy sata, community Structure uncertainty, cost of uncertainty
**Abstract:** Community structures often used for the network analysis. Yet, the prevalence of noise in data massively collected for large networks and intentionally obfuscated in covert networks distorts edges and community structures of such networks. To address this challenge, we repeatedly rewire a given network with a community structure created from the noisy data while controlling the standard deviation of the edges' weights and connectivity creating the original network variants compatible with the noisy data. We use several entropy-based metrics to identify the lowest uncertainty community structure among them. However, in many applications, e.g., maintaining a complex system or disrupting a criminal or terrorist organization, the uncertainty cost is more essential than uncertainty itself. We introduce a novel heuristic that creates a community structure with the lowest expected cost of uncertainty that can be defined by the user as a function of difference between two community structures.
**Distribution Statement:** 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

**CONFERENCE PAPERS:**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)
Date Received: 28-Aug-2017 Conference Date: 29-May-2017 Date Published: 29-May-2017
Conference Location: Orlando / Buena Vista, FL, USA
**Paper Title:** Predicting Viral News Events in Online Media
**Authors:** Xiaoyan Lu, Boleslaw Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** 6th International Conference on Complex Networks and Their Applications Complex Networks
Date Received: 05-Feb-2018 Conference Date: 01-Dec-2017 Date Published: 02-Jan-2018
Conference Location: Lyon, France
**Paper Title:** Evolution of the Global Risk Network Mean-Field Stability Point
**Authors:** Xiang Niu, Alaa Moussawi, Noemi Derzsy, Xin Lin, Gyorgy Korniss and Boleslaw K. Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** 2018 International Workshop on Social Sensing (SocialSens)
Date Received: 20-Aug-2018 Conference Date: 17-Apr-2018 Date Published: 17-Apr-2018
Conference Location: Orlando, FL
**Paper Title:** The Future of Deception: Machine-Generated and Manipulated Images, Video, and Audio?
**Authors:** Jonathan Z. Bakdash, Char Sample, Monica Rankin, Murat Kantarcioglu, Jennifer S. Holmes, Sue Kase
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Conference on Complex Systems
Date Received: 04-Oct-2019 Conference Date: 04-Oct-2019 Date Published: 04-Oct-2019
Conference Location: Singapure
**Paper Title:** Detection of spread source in complex networks
**Authors:** Robert Paluch, Lukasz Gajewski, Xiaoyan Lu, Krzysztof Suchecki, Boleslaw Szymanski, and Janusz Ho
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Conference on Complex Systems
Date Received: 04-Oct-2019 Conference Date: 04-Oct-2019 Date Published: 04-Oct-2019
Conference Location: Singapure
**Paper Title:** Finding the source of local information cascades in complex networks
**Authors:** Krzysztof Suchecki, Boleslaw K. Szymanski, and Janusz A. Holyst
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** APS March Meeting
Date Received: 24-May-2019 Conference Date: 09-Mar-2018 Date Published:
Conference Location: Los Angeles, CA
**Paper Title:** CARP Model for Multi-Risk Dynamics
**Authors:** Alaa Moussawi, Xiang Niu, Noemi Derzsy, Xin Lin, Gyorgy Korniss, and Boleslaw Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Conference on Complex Systems
Date Received: 21-Aug-2021 Conference Date: 04-Oct-2019 Date Published: 04-Oct-2019
Conference Location: Singapure
**Paper Title:** Competition and spreading of low and high quality information in online social networks
**Authors:** Diego Fregolent Mendes de Oliveira, and Kevin S. Chan
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Seventh Int. Conference on Complex Networks & Their Applications
Date Received: 24-May-2019 Conference Date: 13-Dec-2018 Date Published: 13-Dec-2018
Conference Location: Cambridge, U.K.
**Paper Title:** The effects of trust and influence on the spreading of low and high quality information
**Authors:** Diego F. M. Oliveira, and Kevin S. Chan
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** The Tenth Annual Conference on Complex Networks, Complenet 2019
Date Received: 24-May-2019 Conference Date: 18-Mar-2019 Date Published: 21-Mar-2019
Conference Location: Tarragona, Spain
**Paper Title:** Sorting information by quality as a mechanism to hinder the system’s discriminative power
**Authors:** Diego F. M. Oliveira, and Kevin S. Chen
Acknowledged Federal Support: **Y**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** ISODS IV Satellite, NetSci 2019
Date Received: 07-Jun-2019 Conference Date: 28-May-2019 Date Published: 28-May-2019
Conference Location: Burlington, VT
**Paper Title:** Control of Risk Networks: Minimizing Time and Energy
**Authors:** Xiang Niu, Jianxi Gao, and Boleslaw K. Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** IEEE ICCCN, 2019
Date Received: 04-Oct-2019 Conference Date: 30-Jul-2019 Date Published: 27-Sep-2019
Conference Location: Valencia, Spain
**Paper Title:** A Reaction-Based Approach to Information Cascade Analysis
**Authors:** James Flamino and Boleslaw K. Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Conference on Complex System
Date Received: 04-Oct-2019 Conference Date: 02-Oct-2019 Date Published: 04-Oct-2019
Conference Location: Singapore, Singapore
**Paper Title:** Finding Efficient Spreaders for Information Diffusion in Social Networks
**Authors:** Boleslaw Szymanski, Panagiotis Karampourniotis, Gyorgy Korniss
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** IX Conference onComplex Networks and Their Applications
Date Received: 04-Mar-2021 Conference Date: 04-Dec-2020 Date Published: 22-Jan-2021
Conference Location: Madrid, Spain
**Paper Title:** Learning Parameters for Balanced Index Influence Maximization
**Authors:** Manqing Ma., Gyorgy Korniss, Boleslaw K. Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** 9th Conference on Complex Networks and Their Applications
Date Received: 21-Aug-2021 Conference Date: 02-Dec-2020 Date Published: 05-Feb-2021
Conference Location: Madrit, Spain, remote participation
**Paper Title:** Learning Parameters for Balanced Index Influence Maximization
**Authors:** Manqing Ma, Gyorgy Korniss, and Boleslaw K. Szymanski
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published
**Conference Name:** Complex Networks X
Date Received: 21-Aug-2021 Conference Date: 19-Mar-2019 Date Published: 06-Mar-2019
Conference Location: Tarragona, Spain
**Paper Title:** Spread and control of misinformation with heterogeneous agents
**Authors:** Pedro Cisneros-Velarde and Diego F. M. Oliveira and Kevin S. Chan
Acknowledged Federal Support: **Y**

**DISSERTATIONS:**

# RPPR Final Report
as of 19-Nov-2021

**Publication Type:** Thesis or Dissertation
**Institution:** Rnesselaer Polytechnic Institute
Date Received: 14-Aug-2019 Completion Date: 6/6/19 4:35PM
**Title:** The Dynamics in Opinion and Global Risk Networks: Modeling, Discovering and Control
**Authors:** Xiang Niu
Acknowledged Federal Support: **Y**

**Publication Type:** Thesis or Dissertation
**Institution:** Rensselaer Polytechnic Institute
Date Received: 16-Nov-2021 Completion Date: 11/10/21 1:21PM
**Title:** DYNAMICS AND STABILITY OF EDGES AND COMMUNITIES IN SOCIAL NETWORKS
**Authors:** Amr Elsisy
Acknowledged Federal Support: **Y**

**Publication Type:** Thesis or Dissertation
**Institution:** Rensselaer Polytechnic Institute
Date Received: 16-Nov-2021 Completion Date: 11/16/21 1:18AM
**Title:** ANALYZING AND MODELING HUMAN BEHAVIORAL DYNAMICS IN SOCIAL NETWORKS
**Authors:** James,Flamino
Acknowledged Federal Support: **Y**

**PATENTS:**

**Intellectual Property Type:** Patent Date Received: **21-Aug-2021**
**Patent Title:** A method for computing likelihood of clustering of network with with incomplete information (provisional)
**Patent Abstract:**
**Patent Number:** 63/168,216
Patent Country: USA
Application Date: 30-Mar-2021 Application Status: 1
Date Issued:

**Partners**

,

2016-2021 Collaboration sponsored by EU Maria Curie Sklodowska Program to support collaboration between EU a

I certify that the information in the report is complete and accurate:
Signature: Boleslaw K. Szymanski
Signature Date: 11/18/21 10:04AM

# W911NF1610524 : Forecasting Emergent Phenomena with Human Computer Collaboration

**Reporting Period:** AUG 18, 2016 to AUG 17, 2021

**Date Received:**

**Submitter:** Boleslaw Szymanski

---



---

## Major Goals

Our ultimate goal of this research is to utilize data crowd-sourced from the experts to gain a quantitative model of the state of global risk networks that may provide some actionable insights for the U.S. government and the U.S. Army. The global risk network by definition is composed of risks of varying complexity, potential economical damage, and probability of activating. Hence, the need for risk mitigation processes may vary from difficult to extremely complicated for various risks and impose various costs. The risk activation probabilities constantly evolve. Accordingly, our quantification of the activation probabilities and their impacts could provide a timely and invaluable guide to any cost-benefit analysis involved in the design of policies or strategies aimed at global risk minimization. The insights provided by our model may also enable the domain experts to provide tailor-made recommendations for the pertinent risks to curb the likelihood of systemic failures. The effects of such recommendations can also be tested using our model. Accordingly, the final contribution of the research is to provide insights for the current state of the risks, guidance for curbing the probability of global failures and the tool for evaluating the proposed solutions. Guidance may include optimal risk network control with minimum control energy and minimum cost of deterring activation of risks. This in turn, may ultimately lead to the continuous control of the network of threats and an early response system to activating risks to reliably limit the disruptive effects of emergent phenomena. Last year, we accomplished the important step in this direction by developing a data driven minimum total cost continuous control system for global risks. It includes two phases: the first called reactive control deactivates currently active risks, followed by the proactive control that continuously deters new activation of risks. This year we have developed a new methodology for supervised Learning of the global risk network activation from media event reports. We also created and validated the software tool for creating a risk knowledge graph directly from the annotated risk events crawled from Wikipedia. We also demonstrated possible applications of our methodology to the current pandemic, showing that different approaches taken by different countries vary vastly on their efficiency. Our current work aims at creating a methodology for incorporating lessons that we learned about the best human and machine partnerships for global risk networks in a generalized approach to threat networks on increasingly fine-grained levels of specificity such that the model can go from global, to regional, to national levels of analysis. Since the most dangerous form of risk failures are cascades of risk activation, and in view of current pandemics, we work on locating the source of spreading in many applications, from risk networks, to epidemic first patients, to misinformation sources. We improved currently existing algorithms for such detection. In the last year of the grant, we have considered an intriguing and promising direction of looking at the value of silence present in the parts of a network, to which spread has not yet reached. We obtained improved results. We finished this research and completed the paper for submission in the near future. Among risks that undermine the stability of nations and regions are social disruption. Accordingly, in the last year of the grant, we studied dynamics processes in complex social networks, ranging from the different kinds of society polarization to cultural

completion and resilience. Finally, the third topic addressed in the last year of the grant focused on an important issue of properly identifying communities in the complex social networks, especially those with incomplete information like covert terrorist networks. We completed this research and already published one paper and submitted the second paper for publication. Thanks to the high novelty of both papers, one is already the source for U.S. provisional patent, and the other is under consideration for a patent by the Rensselaer Office of Intellectual Property Innovation.

---

**Accomplishments Under Goals**

We grouped our paper published or accepted in the reporting period into four clusters mentioned in the Accomplishment Sections. They are listed below. I. Risk network control

1. "From Data to Complex Network Control of Airline Flight Delays," Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, and Boleslaw K. Szymanski, Scientific Reports 11:18715, Sept. 21, 2021, https://doi.org/10.1038/s41598-021-98112-7

We demonstrate how to design control of the dynamical system of airline flight delays directly from the date about airline flight performance collected by the government. We show that the optimal control significantly lowers the cost of airline flight delays for airlines, airports and passengers.

2. "Heuristic assessment of choices for risk network control," Christopher Brissette, Xiang Niu, Chunheng Jiang, Jianxi Gao, Gyorgy Korniss, Boleslaw K. Szymanski, Scientific Reports 11:7645, April 7, 2021. https://doi.org/10.1038/s41598-021-85432-x

We present a possible application of our risk network control methodology to global pandemics. The results show that different combinations of countermeasures incur drastically different costs.

II. Dynamics Processes in Complex Networks

3. "Social Networks through the Prism of Cognition," Radoslaw Michalski, Boleslaw K. Szymanski, Przemysław Kazienko, Christian Lebiere, Omar Lizardo, Marcin Kulisiewicz, Complexity, 2021:4963903, Jan. 8, 2021, doi.org/10.1155/2021/4963903

We demonstrate that mimicking human memory via detailed cognitive models improves predicting evolution of social networks.

4. “Message sorting: a mechanism that hinders the system’s discriminative power,” Diego F.M. Oliveira, Kevin S. Chan, Physica A, submitted 2020.

We propose an agent-based model to study the effects of message sorting on the diffusion of low- and high-quality information in online social networks. We introduce a model that shows our approach to sorting messages intensifies the exposure of high-quality information increasing the overall system’s quality while preserving its diversity. 5. "Become a better you: Correlation between the change of research direction and the change of scientific performance," Xiaoyao Yu, Boleslaw K. Szymanski, Tao Jia, Journal of Informetrics 15(3):101193, Aug. 2021, https://doi.org/10.1016/j.joi.2021.101193

Here, we quantify the change of research direction in one’s early and late career. The direction change is associated with the chance to increase scientific impact, but noy with the productivity change; this change is uncorrelated with diversity or switching among topics. In conclusion it is beneficial to promote multi-disciplinary collaborations.

6. "A partial knowledge of friends of friends speeds social search," Amr Elsisy, Boleslaw K. Szymanski, Jasmine A. Plum, Miao Qi, Alex Pentland, PLoS ONE 16(8):e0255982, Aug. 19, 2021,

https://doi.org/10.1371/journal.pone.0255982

The success of people in forwarding efficiently knowing only personal connections is still not fully explained. To study this problem, we emulate it on a real location-based social network, Gowalla that provides explicit information about friends and temporal locations of each user. We use it to conduct a massive computational experiment to establish new necessary and sufficient conditions for achieving social search efficiency. The results show that only the distribution of friendship edges and the partial knowledge of friends of friends are essential and sufficient for the efficiency of social search.

7. "Nuclear reaction network unveils novel reaction patterns based on stellar energies," Chunheng Jiang, Boleslaw K Szymanski, Jie Lian, Shlomo Havlin and Jianxi Gao, New Journal of Physics 23 :083035, Aug.19, 2021, https://doi.org/10.1088/1367-2630/ac1a3d

We assemble a nuclear reaction network in which a node represents a nuclide, and a link represents a direct reaction between nuclides. Interestingly, the degree distribution of the nuclear network is bimodal, unusual in nature. We discover three rules that govern the structure pattern of this network: (i) linking choices determine reaction-types, (ii) network distances between the reacting nuclides are short, and (iii) each node in- and out-degrees are close to each other. Incorporating these three rules unveils the underlying nuclear reaction patterns hidden in a large and dense nuclear reaction network.

III. Social Network Polarization

8. “Creation, Evolution, and Dissolution of Groups in Social Networks,” Ashwin Bahulkar, James Flamino, Boleslaw K. Szymanski, Kevin Chan, and Omar Lizardo, Scientific Reports 11:17470, September 1, 2021, https://doi.org/10.1038/s41598-021-96805-7 We demonstrate that group dynamics is driven by members attempting to maximize their benefits from group interactions. This leads to across group polarization. We built a highly predictive analytical model for group dynamics to validate our conclusions.

9. "Polarization and tipping points," Michael W. Macy, Manqing Ma, Daniel R. Tabin, Jianxi Gao, and Boleslaw K. Szymanski, PNAS, June 2021, in press.

Our study was motivated by a highly disturbing puzzle. Confronted with a deadly global pandemic that threatened not only massive loss of life but also the collapse of our medical system and economy, why were we unable to put partisan divisions aside and unite in a common cause. We used a computational model to search for an answer in the phase transitions of political polarization. The model reveals asymmetric hysteresis trajectories with tipping points that are hard to predict and that make polarization extremely difficult to reverse once the level exceeds a critical value.

10. "Shifting Polarization and Twitter News Influencers between two U.S. Presidential Elections," James Flamino, Alessandro Galeazzi, Stuart Feldman, Michael W. Macy, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Alexandre Bovet, Hernan A. Makse, Boleslaw K. Szymanski, Science Advances, submitted Nov. 2021.

Social media are decentralized and interactive, which has transformed the dynamics of political communication. Using nearly a billion tweets, we measure the change between the 2016 and 2020 U.S. Presidential elections in the composition and polarization of Twitter influencers and the diffusion of political information among such influencers and those they influence. We identify ``influencers" by their ability to spread information and classify them as affiliated with a media organization, a political organization, or unaffiliated. Most of the top influencers were affiliated with media organizations in both elections, yet 10 percent were replaced in 2020 by unaffiliated influencers or by those affiliated with center or right-orientated political organizations. From 2016 to

2020, polarization increased among influencers and among those whom they influence. Only a quarter of the top influencers in 2016 retained their status in 2020, showing high susceptibility to turnover.

IV. Challenges of Community Detection in Complex Networks

11. "A Network Generator for Covert Network Structures," Information Science, vol. 584, pp.387-398, https://doi.org/10.1016/j.ins.2021.10.066, online Nov. 6. 2021. We design a novel method of generating random networks with varying variance and edge weight that are statistically similar to the original network but are void of privacy information. These networks enable a new structure analytics by finding community structure that is most resilient to perturbations.

12. "Minimizing Uncertainty Cost of Network Structure from Noisy Data," Amr Elsisy, Boleslaw K. Szymanski, Social Networks, submitted Oct. 2021.

We use entropy measures to establish limits of the community structure uncertainty. For this structure we find a network with the minimum uncertainty about its communities. Finally, we introduce the novel measure of the cost of uncertainty and introduce an efficient heuristic that constructs a community structure with the minimal expected cost of uncertainty.

---

**Plans Next Period**

---

**Results Dissemination**

Results are divided into the topics used in the Accomplishment Section. Several planned dissemination events were canceled due to the pandemic.

I. Risk network control 1. Prof. Boleslaw Szymanski presented an invited talk titled “From Data to Complex Network Control With Applications to COVID-19 Pandemic and Airline Flight Delays,” by Gyorgy Korniss, Jianxi Gao, Xiang Niu, Chunheng Jiang Boleslaw K. Szymanski (RPI) at the 11th Polish Symposium on Physics in Economy and Social Sciences, FENS2021, online, on July 2, 2021,

II. Dynamics Processes in Complex Networks 2. Prof. Tarek Abdelzaher presented a talk titled “The Curse of Asymmetric Polarization Dynamics in the Age of Information Overload,” by Chao Xu, Jinyang Li, Dachun Sun, Ruijie Wang, Tarek Abdelzaher, Jesse Graham, and Boleslaw Szymanski at the International Workshop on Social Sensing (SocialSens 2021): Special Edition on Information Operations on Social Media on June 7, 2021.

3. Prof. Janusz Holyst presented an invited talk titled “Multiplicity of social attributes enhances Heider interactions in social triads,” by Piotr Górski, Joanna Linczuk, Raissa D’Souza, Boleslaw Szymanski, Klavdiya Bochenina and Janusz Holyst, at the 11th Polish Symposium on Physics in Economy and Social Sciences, FENS2021, online, on July 2, 2021,

4. Prof. Radoslaw Michalski presented a talk titled “Cognition-driven Temporal Social Networks,” by Radoslaw Michalski, Boleslaw K. Szymanski, Przemyslaw Kazienko, Christian Lebiere, Omar Lizardo, and Marcin Kulisiewicz at Networks 2021 Conference, online, July 6, 2021.

III. Social Network Polarization

5. Prof. Alexandre Bovet presented a talk titled “Twitter Influencers and Increased Media Polarization during two U.S. Presidential Elections,” by Alexandre Bovet, Stuart Feldman, James

Flamino, Alessandro Galeazzi, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Hernán A. Makse, and Boleslaw K. Szymanski at the PCNET, A Satellite Workshop, at A Networks 2021 Conference, online on June 24, 2021.

6. Brendan Cross presented a talk titled “Twitter Influencers and Increased Media Polarization during two U.S. Presidential Elections,” by Alexandre Bovet, Stuart Feldman, James Flamino, Alessandro Galeazzi, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Hernán A. Makse, and Boleslaw K. Szymanski at the Computational Social Science Satellite Event at a Conference on Complex Systems CCS21, online on 27 October 2021. IV. Challenges of Community Detection

7. Dr. Amr Elsisy presented at talk titled “A Synthetic Network Generator for Covert Network Analytics,” by Amr Elsisy, Aamir Mandviwalla, Boleslaw K. Szymanski, and Thomas Sharkey at the Communities in Networks, ComNets, A Satellite Workshop at Networks 2021 Conference, online on July 1, 2021.

8. Prof. Boleslaw Szymanski presented an invited talk titled “A Network Generator for Covert Network Structures,” by Amr Elsisy, Aamir Mandviwalla, Boleslaw K. Szymanski, and Thomas Sharkey at the USC 9/11 Symposium, at Los Angeles, CA, online on September 15, 2021.

---

**Honors and Awards**

Dr. Boleslaw Szymanski received: (i) 2021 Network Science Society Service Award for his many service contributions to network science, including being a global strategic leader, fundraiser, and hands-on, on-the-ground institution builder. He is a third recipient of this award, after Dr. Albert Laszlo Barabasi, founder of the Society and Dr. Guido Caldarelli; (ii) 2018 Life Fellow member of IEEE

---

**Training Opportunities**

During the reporting period, we provided training opportunities in the STEM areas in research relevant to DoD. We trained one postdoctoral researcher, Dr. Diego Fregolente Mendes de Oliveira, who stayed 33 months and then joined the Department of Mathematics and The Statistical and Applied Mathematical Sciences Institute at the University of North Carolina, Chapel Hill, NC. We also graduate eight students in the reported period. Two Computer Science Ph.D.’s joined Google, Inc., one Computer Science Ph.D. joined Facebook, Inc. and three Physics Ph.D. accepted postdoc positions at CMU and at the IBM Research Center in Boston, MA, and in the Network Science and Technology Center at RPI. We also graduated five MS students in Computer Science. One of them joined a college as a teaching faculty, another joined Ph.D. program at Harvard University while the remaining three accepted offers from industry. We also involved 12 undergraduates in Computer Science research for various periods ranging from one to three semesters. They left with B.Sc. in Computer Science degrees and joined Ph.D. programs at other universities.

---

**Technology Transfer**

1. Provisional U.S. patent: “A Method For Computing Likelihood Of Clustering For Network with Incomplete Information,” U.S. Patent Application No.: 63/168,216 submitted 03/30/2021, RPI Ref.: 2021-048, Law Office Ref.: 104640-100. (co-inventors: Boleslaw K. Szymanski, Amr Esisy, and Aamir Mandviwalla), assignee: RPI, Troy, NY., 03/31/2021

2. Invention Disclosure: "A method and Apparatus for Minimizing the Expected Cost of Uncertainty of Network Structure," under consideration for applying by the RPI Office of Intellectual Property Optimization, (co-inventors Boleslaw K. Szymanski and Amr Esisy), assignee: RPI, Troy, NY., 1/12/2021.

---

**Participants**

| Name | Role | Person Months |
|---|---|---|
| Uzzi, Brian | Co PD/PI | 1 |
| Bahulkar, Ashwin | Graduate Student (research assistant) | 3 |
| Dipple, Stephen | Graduate Student (research assistant) | 3 |
| Elsisy, Amr | Graduate Student (research assistant) | 2 |
| Flamino, James | Graduate Student (research assistant) | 1 |
| Gong, Bowen | Graduate Student (research assistant) | 2 |
| Karampourniotis, Panagiotis | Graduate Student (research assistant) | 5 |
| Lin, Xin | Graduate Student (research assistant) | 5 |
| Lu, Xiaoyan | Graduate Student (research assistant) | 5 |
| Mohr, Matthew | Graduate Student (research assistant) | 5 |
| Mukherjee, Partha | Graduate Student (research assistant) | 2 |
| Niu, Xiang | Graduate Student (research assistant) | 10 |
| Tabin, Daniel | Graduate Student (research assistant) | 1 |
| Zhuang, Juntao | Graduate Student (research assistant) | 1 |
| Szymanski, Boleslaw | PD/PI | 5 |
| Oliveira, Diego | Postdoctoral (scholar, fellow or other postdoctoral position) | 15 |
| Buchanan, Frederick | Undergraduate Student | 4 |
| Cohen, Samuel | Undergraduate Student | 4 |
| Dai, Weiki: | Undergraduate Student | 2 |
| DeVito, Ross | Undergraduate Student | 2 |
| Jiang,, Duo | Undergraduate Student | 2 |
| Lieberman, Alex | Undergraduate Student | 2 |
| Louison, Ted | Undergraduate Student | 2 |
| Meandzija, Alexander | Undergraduate Student | 2 |
| Osika, James | Undergraduate Student | 2 |
| Qi, Miao | Undergraduate Student | 2 |
| Smith, Andrew | Undergraduate Student | 2 |
| Yin, Lake | Undergraduate Student | 2 |