

Azadeh Nematzadeh

School of Informatics and Computing, Indiana University
azadeh.n@gmail.com
812-567-5097
<http://azadeh.github.io>

- EDUCATION** *PhD candidate, Complex systems and network analysis*
School of Informatics and Computing, Indiana University, expected summer 2017.
Thesis: Information Diffusion and Behavior in Groups of Networked Individuals
Visiting graduate student researcher
Instituto Gulbenkian de Ciencia, Oeiras, Portugal, 2009-2010.
I applied machine learning techniques to solve bioinformatics research questions.
Graduate student, Information security
School of Informatics and Computing, Indiana University, 2008-2009.
I proposed a framework to assure the privacy of online health systems.
Master of Science, Information security
Amirkabir University, Tehran, Iran, 2008.
Thesis: A Calculus for Trust Management in Semantic Web
Bachelor of Science, Computer engineering
Shahid Beheshti University, Tehran, Iran, 2005.
Thesis: An Ontology-based Ranking Method to Retrieve Documents based on their Relevancy Score
- SKILLS** Proficient in **Python, Pig, MySQL, L^AT_EX**.
Experience with **Spark, D3, ggplot, R, C++, C#, Gephi, Bash script**.
Technical Skills: Network Science, Data Science, Statistics, Machine Learning
- MAJOR AWARD** *LinkedIn economic graph challenge*, One of eleven data science projects selected to analyze LinkedIn Data, 2015.
Project: Forecasting Large-Scale Industrial Evolution.
- PATENTS** Yong-Yeol Ahn, Azadeh Nematzadeh, Ian Wood, Jaehyuk Park, Yizhi Jing, and Michael Conover. Pat. US 15/229,956 (Pending).
Title: “Data Flow Based Feature Vector Clustering”.
- SELECTED DATA SCIENCE PROJECTS** **Project: Forecasting Large-Scale Industrial Evolution.**
- DataSet: LinkedIn data, S&P 500
 - Description: We construct a global labor flow network to reveal the industry on the hierarchical organization of firms and migration of people. We quantify the impact of geography and industry on the job mobility across scales. Furthermore, we demonstrate that organic, geo-industrial clusters of firms discovered from the network better capture industrial sectors that grow or decline together than traditional categories of firms based on industry or geography.
 - Technologies: Hadoop, Pig, Python, Spark, Gephi

- Techniques: PageRank, Network visualization, Network analysis, Clustering, Community detection, Regression analysis, Entropy measures, Prediction, Feature engineering, Log odds ratio informative Dirichlet prior, etc.

Project: Information Overload in Twitch’s Chat Channels.

- Data Set: Twitch chat data
- Description: I conducted a case study to analyze how communication patterns change when people experience information overload. In this case study, I examined conversations among users on Twitch’s platform in its IRC chat channels. I measured the structural and textual features of communications such as user output, user interaction, and information content per message across a wide range of information loads.
- Technologies: Python, R
- Techniques: Text mining, NLP, Feature engineering, Entropy measures, Regression analysis, Multilevel statistics, etc.

Project: Consumer Choices in Yelp.

- Data Set: Yelp, census, Google geocoding
- Description: I have studied how people make choices and explore in a case study based on Yelp restaurant review data. I have shown that it is possible to predict the flux of restaurant reviewers between two cities based on the gravity law of human mobility. I have also built a prediction model to predict how Yelp users will rate the restaurants based on the Yelp users’ review history, and their social contact review history.
- Technologies: Python, MySql
- Techniques: Network analysis, Word2Vec, Item-item recommendation system, Prediction, Similarity measures, Regression analysis, Mobility analysis, Hypothesis testing, etc.

Project: Measuring Agenda Change in Political Discourse.

- Data Set: Congress Data, Twitter
- Description: I developed a computational approach to track temporal changes of (1) issue ownership and (2) issue frames. In the former application, I identified the critical words come to characterize the policy focus of each political group at different time steps. In the latter application, I concentrated on the rate and pattern of frame turnover in the parliamentary discourse on controversial policy issues.
- Description: I studied temporal changes in patterns of users’ collective behaviors on the Twitter’s microblogging network. I analyzed the change in topological and statistical characteristics of the retweet network for two popular political hashtags (#tcot, and #p2). I also analyzed how the participation of users in the discussion about these hashtags and their interactions with others changed over the time.
- Technologies: Python, MySql
- Techniques: Crawler, Text mining, Feature engineering, Entropy measure, Concept drift, Time series analysis, clustering, statistical analysis, etc.

Project: Prevalence and Identification of Auto Scam on Craigslist.

- Data Set: Craigslist ads
- Description: Craigslist ads are viewed by millions of Internet users each month, making it an attractive target for fraudsters and miscreants. Unsurprisingly, it has even been labeled

a “cesspool of crime.” I studied automobile scam on Craigslist. I performed feature engineering to find features distinguishing scam from good ads. Using these features I show that an SVM based classifier can differentiate between scam and trustworthy ads with 99% accuracy.

- Technologies: Python, MySQL
- Techniques: Crawler, Text mining, Feature engineering, Classification, statistical analysis.

Project: BioCreative Competition.

- Data Set: PubMed Data
- Technologies: Python
- Techniques: Crawler, Text mining, Feature engineering, Classification, statistical analysis.

**SELECTED
PUBLICATIONS**

Information overload in group communication of Twitch.

Under review in epj data science journal, 2016, **Nematzadeh**, et. al.

Cooperative and competitive spreading dynamics on the human connectome.

J. Neuron, 2015, Bratislav, Betzel, **Nematzadeh**, et. al.

Optimal network modularity for information diffusion.

Physical review letters, 2014, **Nematzadeh**, et. al.

Experience design framework for securing large scale ICTS.

DRS, 2014, **Nematzadeh**, et. al.

A linear classifier and a statistical approach to method extraction in the PPI literature.

BMC bioinformatics, 2011, Loureno, Conover, Wong, **Nematzadeh**, et. al.

Testing extensive use of NER tools in article classification and a statistical approach for method interaction extraction in the PPI literature.

BioCreative III Workshop 2010, Loureno, Conover, Wong, **Nematzadeh**, et. al.

Threat analysis of online health information system.

International Conference on Pervasive Technologies, 2010, **Nematzadeh**, et. al.

Privacy concerns of semantic web.

ITNG 2008 IEEE, 2008, **Nematzadeh**, et. al.

A cheating model for cellular automata-based secret sharing schemes.

J. configurations, 2007, Jafarpour, **Nematzadeh**, et. al.

Modeling architectural access control with UML.

IPM, 2007, **Nematzadeh**, et. al.

Orank: An ontology based system for ranking documents.

International Journal of Computer Science, 2006, Shamsfard, **Nematzadeh**, et. al.

**PRESENTED
POSTERS**

Optimal network modularity for fast global information diffusion

Network Frontier Workshop, 2015.

Effect of network community in information diffusion

Network Frontier Workshop, 2013.

How do issue frames evolve in political discourse? A computational analysis of the presence and magnitude of shifts between policy frames

PolMeth XXIX. 29th Annual for Political Methodology, 2012.

Classification approach toward Craigslist scam detection

Graduate cohort workshop, 2011.

**RECENT
PROFESSIONAL
EXPERIENCE**

Research Assistant, Under supervision of Alessandro Flammini, Center for complex networks and systems research, Fall 2013-Spring 2015.

Description: I utilize mathematical and computational methods to study complex social systems. The list of my recent works includes studying the role of network structure in driving diffusion of novel ideas or technologies, computationally probing the coarse-grained wired structure of the human brain, studying the emergence of new communication patterns in groups subject to information overload, and modeling the interplay between items popularity and friends recommendations in determining global patterns of information consumption.

Research Assistant, Under supervision of Luis Rocha, Gulbenkian institute, 2009-2010.

Description: I focused on building an agent-based model of the t-cell regulatory model in the immune system. I used this model to perform binary classification. I evaluated this model using Enron spam email data. I also participated in a Biocreative competition, that was about building a classification to detect protein-protein interaction documents using PubMed data.

Research Assistant, Under supervision of Jean Camp, School of Informatics, 2009-2010.

Description: worked on analysis privacy and security threats of online health system. I proposed a framework similar to DRM technology to preserve the security and privacy of online health systems. I also worked on the E.T.H.O.S (Ethical Technology in the Homes of Seniors) project over the summer where we designed a secure system to help senior citizens get the healthcare they need. During this period, I started researching bio-inspired models to detect trustworthy and untrustworthy content.

Solution Engineer, Nokia Siemens Network, 2007-2008.

Description: I created engineering solutions for mobile networks; system installation, integration, and maintenance.

Software Engineer, Niro Research Institute, 2006.

Description: I worked as a software analyst, designer, and developer.

Software Engineer, 2005.

Description: I built software and performed system analysis and design. For instance, I helped build a database and a web base software inventory and accounting system.

Student Researcher University MultiMedia research lab, 2004.

Description: I studied an agent-based language.

**RECENT
TEACHING
EXPERIENCE**

Lecturer, Search Informatics, Indiana University, Fall 2016.

I taught student how to build a functioning web-based search engine in Python.

Guest Lecturer, Mathematical Modeling of Complex Systems, Fall 2016.

Lecturer, Multimedia Systems, Shahid Rajaei University, Tehran, Iran, Spring 2007.

Associate Instructor, Math Foundation, Indiana University, Spring-Summer 2016.

Associate Instructor, Information Representation, Indiana University, Spring 2011- Spring 2013.

Associate Instructor, Organizational Informatics of Security, Indiana University, Spring 2009.

Associate Instructor, Analytical Foundation of Security, Indiana University, Fall 2009.

Associate Instructor, Artificial Intelligence, Shahid Beheshti University, Spring 2004 - Fall 2006.

Associate Instructor, Principles of Computer and Programming, Shahid Beheshti University.

RECENT TALKS I presented my work on information overload in IC2S2, 2016.

I presented my work on the various dimensions of human exploratory behaviors in The Cognitive Lunch colloquium, Indiana University psychology and brain science department, 2015.

I presented my work on modularity and information diffusion in the foundation of Complex Network session of CCS2015, September 2015.

I presented my work on the various dimensions of human exploratory behaviors in CCS2015, September 2015.

I presented my work on the various dimensions of human exploratory behaviors in 2015 Computational Social Science Summit, May 2015.

I presented my work on modularity and information diffusion in Community Detection and Co-Evolving Networks session of Complenet, March 2015.

I presented my work on statistical analysis of Yelp data in a statistical consulting class (taught by Stephanie Dickinson), Indiana University, 2014.

I presented my work on the effect of network community in information diffusion, The Cognitive Lunch colloquium, Indiana University psychology and brain science department, 2014.

ACADEMIC SERVICES & VOLUNTEER WORKS

Reviewer of Journal of the Royal Society Interface (2015), IC2S2 (2015), Scientific Reports (2014-2015), and Web of Science Conference (2014).

Openhatch workshop , Mentor, Fall. 2013-2014.

Complex System reading group, Organized the complex system reading group, Indiana University, 2015-2016.

Complex System PhD student meeting, member of the organizing committee, Indiana University, 2011-2012.

NCWIT Award , Review and score applications for the NCWIT Award for Aspirations in Computing, 2011.

IT association, One of the main organizers of the IT group at Shahid Beheshti University, Activities of this group included presenting lectures about some different aspects of IT (Data Mining, Intelligent Agents, eCommerce, ITS), 2004.

Robocup instructor, Instructor of two Robocup-leagues preparation classes for pre-selected high-school students, 2006.

OTHER AWARDS

CCS2015, Travel grant to participate in the CCS 2015 conference.

Computational Social Science Summit, Travel grant for the Summit 2015 conference.

Grad-Cohort award, Graduate cohort workshop, Boston, 2011.

NSF scholar award, Doctoral consortium of PETRA conference. Greece, 2010.

WISE 2010, Summer school sponsored by the Team for Research in Ubiquitous Secure Technology (TRUST), CMU, 2010.

WISE 2009, Summer school sponsored by the Team for Research in Ubiquitous Secure Technology (TRUST), Berkeley University, 2009.