



Defense Threat Reduction Agency
8725 John J. Kingman Road, MS
6201 Fort Belvoir, VA 22060-6201



DTRA-TR-18-70

TECHNICAL REPORT

Robustness Analysis and Anomaly Detection of Interdependent Physical and Social Networks

Distribution Statement A. Approved for public release, distribution is unlimited.

September 2018

HDTRA1-10-1-0120

Tarek Abdelzaher et al.

Prepared by:
University of Illinois at Urbana
620 East John Street
Champaign, IL 61820

DESTRUCTION NOTICE:

Destroy this report when it is no longer needed.
Do not return to sender.

PLEASE NOTIFY THE DEFENSE THREAT REDUCTION
AGENCY, ATTN: DTRIAC/ RD-NTF, 8725 JOHN J. KINGMAN ROAD,
MS-6201, FT BELVOIR, VA 22060-6201, IF YOUR ADDRESS
IS INCORRECT, IF YOU WISH IT DELETED FROM THE
DISTRIBUTION LIST, OR IF THE ADDRESSEE IS NO
LONGER EMPLOYED BY YOUR ORGANIZATION.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE	3. DATES COVERED (From - To)		
4. TITLE AND SUBTITLE			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION / AVAILABILITY STATEMENT					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)

UNIT CONVERSION TABLE

U.S. customary units to and from international units of measurement*

U.S. Customary Units	Multiply by Divide by [†]	International Units
Length/Area/Volume		
inch (in)	2.54 × 10 ⁻²	meter (m)
foot (ft)	3.048 × 10 ⁻¹	meter (m)
yard (yd)	9.144 × 10 ⁻¹	meter (m)
mile (mi, international)	1.609 344 × 10 ³	meter (m)
mile (nmi, nautical, U.S.)	1.852 × 10 ³	meter (m)
barn (b)	1 × 10 ⁻²⁸	square meter (m ²)
gallon (gal, U.S. liquid)	3.785 412 × 10 ⁻³	cubic meter (m ³)
cubic foot (ft ³)	2.831 685 × 10 ⁻²	cubic meter (m ³)
Mass/Density		
pound (lb)	4.535 924 × 10 ⁻¹	kilogram (kg)
unified atomic mass unit (amu)	1.660 539 × 10 ⁻²⁷	kilogram (kg)
pound-mass per cubic foot (lb ft ⁻³)	1.601 846 × 10 ¹	kilogram per cubic meter (kg m ⁻³)
pound-force (lbf avoirdupois)	4.448 222	newton (N)
Energy/Work/Power		
electron volt (eV)	1.602 177 × 10 ⁻¹⁹	joule (J)
erg	1 × 10 ⁻⁷	joule (J)
kiloton (kt) (TNT equivalent)	4.184 × 10 ¹²	joule (J)
British thermal unit (Btu) (thermochemical)	1.054 350 × 10 ³	joule (J)
foot-pound-force (ft lbf)	1.355 818	joule (J)
calorie (cal) (thermochemical)	4.184	joule (J)
Pressure		
atmosphere (atm)	1.013 250 × 10 ⁵	pascal (Pa)
pound force per square inch (psi)	6.984 757 × 10 ³	pascal (Pa)
Temperature		
degree Fahrenheit (°F)	[T(°F) - 32]/1.8	degree Celsius (°C)
degree Fahrenheit (°F)	[T(°F) + 459.67]/1.8	kelvin (K)
Radiation		
curie (Ci) [activity of radionuclides]	3.7 × 10 ¹⁰	per second (s ⁻¹) [becquerel (Bq)]
roentgen (R) [air exposure]	2.579 760 × 10 ⁻⁴	coulomb per kilogram (C kg ⁻¹)
rad [absorbed dose]	1 × 10 ⁻²	joule per kilogram (J kg ⁻¹) [gray (Gy)]
rem [equivalent and effective dose]	1 × 10 ⁻²	joule per kilogram (J kg ⁻¹) [sievert (Sv)]

* Specific details regarding the implementation of SI units may be viewed at <http://www.bipm.org/en/si/>.

[†] Multiply the U.S. customary unit by the factor to get the international unit. Divide the international unit by the factor to get the U.S. customary unit.

Abstract

This project developed models and algorithms for analyzing and improving the robustness of *information flow* in interdependent physical and social networks in counter-WMD scenarios. While much research in the DTRA portfolio focused on modeling the propagation of *failure cascades* through interconnected networks, a point of departure in this project was to recognize that networks (of humans or computer-controlled physical assets), unlike physical systems often used to model them, are made of intelligent entities whose ability to control damage or mitigate cascade propagation is *not* an inherent property but is rather determined in large part by the amount of *information* they have about the unfolding event. With that in mind, the project investigated techniques to enhance *observability* – the ability to correctly estimate physical state – with a focus on exploiting social media. The project developed a science of observability of physical events from social media. This science allows *modeling interconnected cyber-physical and social networks*, where physical phenomena interact with social observers generating both physical effects and information artifacts. Research outcomes included results such as extraction of pertinent information about the state of the physical world from signals that transpire on social media, analysis of reliability of individual sources and information items, algorithms for role discovery, and foundations for social data fusion for purposes of estimation of physical state, and detection, localization, and tracking of physical events. We also developed *scalable tools for heterogeneous network analysis* based on scalable distributed algorithms for tensor manipulation and machine learning, directed at solving the aforementioned problems.

Chapter 1: Introduction

This project developed models and algorithms for analyzing and improving the robustness of *information flow* in interdependent physical and social networks in counter-WMD scenarios. We focused on human sources as the originators of information and the agents for its dissemination, complementing other means of observation (such as physical sensors). A significant example of human-sourced information lies in social media posts, offering the means to investigate algorithms for information extraction from such a source. *The project developed analytical foundations and algorithms for understanding and enhancing observability – the ability to correctly estimate physical state – by exploiting social media data.*

While the project used social media extensively to evaluate algorithms for information extraction, the underlying mathematical foundations of observability are broadly applicable to other human-sourced data as well, such as human-generated reports and data originating from mobile crowd-sensing apps. The following contributions were attained:

- Robust information recovery and fact-finding (Chapter 2): Robustness was investigated in the context of both physical and social networks, leading to a book on analytic foundations of fact-finding.
- Physical event discovery (Chapter 3): Tools were developed for discovery of physical events from their social media signatures.
- Damage assessment (Chapter 4): Algorithms were developed that estimate physical damage from exaggerated, imprecise, and generally unreliable social media feeds.
- Role discovery (Chapter 5): Techniques were developed for identifying roles that nodes play on social networks from analysis of network topology in the neighborhood of those nodes.
- Scalable infrastructure support (Chapter 6): Libraries were developed for effective and scalable mathematical manipulation of large graphs, such as that necessary to carry out the aforementioned algorithms on large data sets.
- Other (Chapter 7): A variety of other contributions were developed in the general area of mining large social networks.

The following chapters describe the above contributions, respectively.

Chapter 2: Robust Information Recovery and Fact-finding

The first thrust of the project focused on information recovery in the aftermath of a mass-destruction event. Recovering an information delivery capability in the aftermath of a mass-destruction event entails solving two fundamental problems that pertain to physical and social networks, respectively; namely: (i) design of content delivery capabilities that are robust to resource outages and bottlenecks, and (ii) design of truth discovery algorithms from noisy, unreliable, and conflicting (social) network data to increase robustness to noise and misinformation.

Robust Information-maximizing Physical Networks

We developed content delivery services for the aftermath of mass-destruction events that enable survivors and first responders to post information (for example, broadcast descriptions of damages and images of the aftermath) in the absence of a functional communication infrastructure. These protocols exploited mobile devices, handling opportunistic forwarding (when such devices came in contact) and in-network storage. A novel feature of these protocols was that they maximized information content delivered subject to resource constraints. This was accomplished by assigning priorities to content objects for forwarding and replacement that depend on the degree of similarity (or dissimilarity) among them, such that a measure of entropy is maximized. Entropy-maximizing prioritization aims at reducing semantic redundancy (such as that between pictures of the same scene at the same location taken from slightly different angles). This is in contrast to redundancy among identical objects and among time series data. We demonstrated that, in resource constrained networks, our protocol significantly improved a measure of information coverage, thereby decreasing communication latency, reducing consumed bandwidth, and improving situation

understanding at the same time. We also explored the role of caching and content pre-fetching policies on mobile nodes in maximizing the rate of information delivery. A new cache replacement policy, called diversity caching, was designed that maximized a measure of information served, as opposed to cache hit rate. This policy was transitioned to BBN, where it was re-implemented on their composite networks testbed for demonstration to government clients. Experiences from running diversity caching at BBN demonstrated that exploitation of semantic relations between stored content items to prioritize delivery and content replacement increases the amount of information delivered, while minimizing both end-to-end latency and resources consumed. Furthermore, we developed an approach for hierarchical content naming that allows applications to encode relevant content semantics into content names, and allows network services, such as caching, to exploit hierarchical naming for information-maximizing delivery.

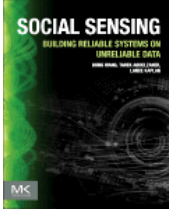
The above mechanisms aimed at providing support for content dissemination in resource-challenged environments, where communication infrastructure is destroyed. It paved the way for research on exploiting information posted in the aftermath of disasters, such that situation understanding is maximized.

Fact-finding from Human Observations

Information-maximizing networks make the best use of scarce physical resources in delivering key information from survivors, first-responders, and other social sources in the aftermath of a mass-destruction event. Unlike the case with well-calibrated and well-tested infrastructure sensors, humans are less reliable, and the likelihood that reported observations are correct is often unknown a priori. Approaches for robust information delivery were therefore augmented with truth discovery techniques that separate fact from rumor. Given a network of human sources of unknown reliability who together supply observations that cannot be immediately verified, we addressed the question of whether one can determine, in an analytically founded manner, the probability that a given observation is true. We developed an optimal solution to the above truth discovery problem. Optimality, in the sense of maximum likelihood estimation, was attained by solving an expectation maximization problem that returns the best hypothesis regarding the correctness of each observation. The approach was shown to outperform state of the art fact-finding heuristics, as well as simple baselines such as majority voting. We then developed algorithms for assessing credibility of information taking into account the topology of the social network over which the information propagates. We demonstrated that taking this topology into account can significantly enhance our ability to improve quality of extracted information in the face of noise, rumors, and misinformation that accompanies the chaos of mass destruction events. The algorithm was extended to operate on streaming data. New inference tools were developed to fill-in missing information items by leveraging known correlations between them. The approach was tested on several Twitter-based data sets from hurricanes, chemical attacks, acts of terrorism, nuclear disasters, and other real-life events showing that it attains a high level of accuracy in identifying correct information and suppressing less reliable inputs.

Our truth discovery work, mentioned above, confirmed the observation that independence between sources is a key property to reduce error in extracted information. If sources are not independent (e.g., repeating what they heard from others), information recovery is prone to rumor propagation and misinformation. To address this problem, we developed algorithms for unbiased source selection that attempt to maximize independence between the selected sources to support unbiased information recovery from social networks. Source selection was complemented by algorithms for rumor detection. We developed an approach for assessing the likelihood that a piece of information is in fact a rumor, in the absence of data provenance information. We also developed techniques for unbiased sampling of diverse opinions.

With additional funding (including a 6.2 transition effort) our fact-finding techniques were incorporated into a tool, called Apollo, designed to help analysts attain situation awareness from noisy Twitter data feeds. The tool automatically identifies credible information and stitches it into a storyline describing the event. Apollo, developed with joint funding from ARL and DTRA, was demonstrated at the DTRA Seminar Series on April 29th, 2014. The idea of using weak indicators to infer facts with a quantified degree of reliability was generalized to a sensing paradigm, called *social sensing*. One of the highlights of the project was the publication of a book on social sensing (with Morgan Kaufmann) by Dong Wang, Tarek Abdelzaher, and Lance Kaplan, entitled "Social Sensing: Building Reliable Systems on Unreliable Data," in 2015.



The book focused on describing the mathematical foundations for reliability guarantees in social sensing. More specifically, it developed analytic means for computing accurate error bounds in cases where information about the physical world is collected from social sources whose reliability is unknown. It contains examples of exploiting Twitter to reliably reconstruct relevant world state. Besides the formulation and solution of the basic social sensing problem in the aforementioned book, several advances were made on progressively more advanced versions of social sensing problems. These include:

- *Social sensing of time-varying phenomena*: The challenge of obtaining reliable information from noisy and conflicting data collected from unreliable social network sources was extended to cases, where the state in question changes dynamically. A challenge in designing the estimator was to properly handle conflicting data, when reported state of the same variable changes over time. Such a conflict could either be attributed to lack of source reliability or to an actual state change in the physical world. A maximum-likelihood estimator was developed to reconstruct the most probable state trajectories of different binary state variables using the noisy social observations as input.
- *Social sensing of interdependent phenomena*: Techniques were developed for exploiting physical dependencies between observed variables to improve the accuracy of social sensing. Physical phenomena are interdependent. Hence, the bits mentioned above are not independent variables but rather are correlated. When different phenomena are described on social networks, much noise is introduced due to lack of reliability of sources, as well as other misinformation, propaganda and errors. We demonstrated how knowledge of dependencies between different “bits” can improve estimation accuracy of ground truth from noisy observations. Of particular interest was to develop techniques that can be used at scale, where the dependencies among variables are extensive.
- *Scalable social sensing*: Finally, to support algorithms such as the above, common primitives were extracted and implemented in a scalable manner on server clusters. Specifically, a key primitive identified was one of clustering data by arbitrary distance functions. A service was developed that performs such clustering at scale, thereby allowing social sensing to be performed in real-time on very large amounts of data.

Chapter 3: Physical Event Discovery

Techniques for localizing physical events from their social media signatures were developed. A hallmark of the developed techniques has been the exploitation of source location affinity. Each source, by virtue of their location and interests is more likely to make references to a finite number of locations with higher probability. By identifying and exploiting such location affinities of sources on social networks, we are better able to determine locations of events described by these sources. The work naturally led to follow-up research aiming to understanding human mobility. While there has been fruitful research on modeling human mobility using tracking data (e.g., GPS traces), the recent growth of geo-tagged social media (GeoSM) brought new opportunities to this task because of its sheer size and multi-dimensional nature. Nevertheless, obtaining quality mobility models from the highly sparse and complex GeoSM data remained a challenge. We proposed GMOVE, a group-level mobility modeling method using GeoSM data. Our insight was that the GeoSM data usually contains multiple user groups, where the users within the same group share significant movement regularity. User grouping and mobility modeling are therefore two intertwined tasks: (1) better user grouping offers better within-group data consistency and thus leads to more reliable

mobility models; and (2) better mobility models serve as useful guidance that helps infer the group a user belongs to. GMOVE thus alternates between user grouping and mobility modeling. Real-life data sets demonstrated that GMOVE effectively generated meaningful group-level mobility models.

We then used geo-tagged social media (GTSM) data streams to conduct recency-aware urban activity modeling and worked out a new method, ReAct, that processed continuous streams and obtained recency-aware urban activity models on the fly. ReAct embeds all the regions, hours, and keywords into the same latent space to capture their correlations. Our experiments on the geo-tagged tweet streams in two major cities showed that ReAct significantly outperformed existing methods for location and activity retrieval, and was orders of magnitudes faster.

The idea that social networks respond predictably to physical state leads to the next natural question; namely, if physical state was independently measured and an anomalous state was observed, can anomalies observed in physical state be explained by correlating them with anomalies simultaneously detected in social network information feeds? The above work led to information gain metrics that identify keyword combinations whose frequency of occurrence has undergone an unusual change (measured in information gain) within an observation window, compared to their normal statistics. For example, a keyword combination such as “forest” and “fire” might be normally absent from tweets, but is being observed with high frequency in a given window, making it a high-information-gain item. Statistical analysis techniques were developed to identify high-information gain keyword combinations over sliding windows. Tweets containing such keyword combinations were then ranked by the corresponding information gain. These tweets were analyzed for location references and matched to locations where physical sensor anomalies were observed, hence normally explaining these anomalies. The approach was applied to explain the root causes of several traffic anomalies on California freeways, among other scenarios. For example, Figure 3 depicts a radiation sensor anomaly observed on February 18th, 2014, by a sensor near the Fukushima reactor.

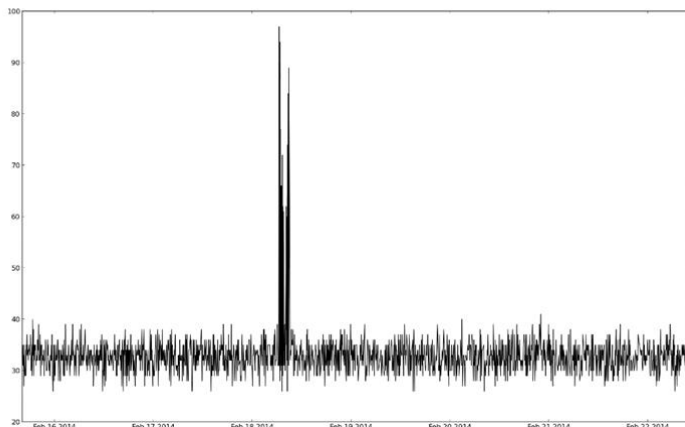


Figure 3. A radiation anomaly near the Fukushima reactor on February 18th, 2014.

The following tweet was identified as the explanation:

福島第一原子力発電所 5号機タービン建屋地下1階における漏えい検知器の動作について | 東京電力 <http://t.co/qbmpi49M2t>

Translating this tweet (from Japanese), it was determined that the tweet was about a report from TEPCO (the reactor owner and decommissioning operator), stating that a radioactive leak was detected in the first floor of Unit 5 in the basement of the turbine building of the Fukushima Daiichi Nuclear Power Station. The example illustrates that utility of social networks as means of explaining unusual phenomena detected by physical sensors.

Chapter 4: Damage Assessment

We also exploited social networks for facilitating *damage assessment* in the aftermath of critical events such as those involving use of weapons of mass destruction (WMDs). Damage assessment requires a sufficient degree of observability into the evolving physical state (such as the state of propagation of a physical damage cascade triggered by the event). Often accurate state assessment is hard to attain due to impaired access to “ground zero” and insufficient resources. The advent of social networks, such as Twitter, offers an opportunity to observe events through the social lens of survivors and their online networks, some of which may remain partially operable after the event. The goal of our effort was to investigate the degree to which physical state can be reconstructed from information propagated on such social networks.

As a running case study, we focused on analysis of gas shortage around New York City in the aftermath of hurricane Sandy. This hurricane hit the city at the end of October 2012. It was the second costliest hurricane on record in the US (after Katrina) at the time and the worst in 2012. A severe gas shortage ensued. The All Hazards Consortium logged the status of hundreds of gas stations in the affected area during the month of November, archiving daily gas availability at each station. In addition, our team collected tweets from the area (using the Twitter API to download tweets with the keyword “gas”). Sentiment analysis was performed on the tweets to determine if they are positive or negative. This analysis offered an opportunity to plot the average sentiment regarding the gas crisis as reflected on Twitter over time. Let us call this measured sentiment, the Twitter *social response*. The research question posed was one of understanding how social response evolves as a function of actual damage, and the degree to which damage evolution may be reconstructed from observing social response. Three sub-questions were then answered:

- First, what model best expresses the relation between damage and Twitter social response? Thinking of the total number of gas stations that are out of gas on a given day as input and social response as output of an auto-regressive moving average (ARMA) model with delay, is response affected more by the input terms (i.e., by actual damage) or by past output terms (i.e., previous sentiment)? What is the order of the model and how big is the delay term? In other words, how far back do terms that affect current sentiment go? Is the model linear or not? A linear model would indicate that humans react to different levels of damage in a consistent manner. Non-linear response would indicate that human behavior changes. For example, in a state of panic, social response might become disproportionate to actual damage. Hence, the structure of the model gives interesting insights into the nature of human response to physical emergencies, observed on social networks.
- Second, can the actual physical damage evolution be inferred from the observed social response on a social network like Twitter? The answer depends very much on the shape of the model discussed in the above bullet and the amount of noise present in the raw data. Given the model and noise terms, what is the accuracy of physical damage reconstruction? Moreover, the order of the model and the magnitude of the delay term affect how soon damage can be estimated. If sentiment is a delayed reaction then it carries less information on the present and more on the past.
- Third, does the evolution of sentiment on the social network actually affect damage propagation itself? In our example, the propagation of a gas outage follows supply and demand dynamics, where supply has been impaired by damage attributed to the disaster. Are these dynamics consistent with demand curves observed prior to the disaster, or does the demand curve needed to explain outage dynamics change after the disaster as a function of the sentiment? The latter would indicate that the sentiment does affect the evolution of damage propagation by altering the demand profile.

The highlights of the findings were as follows:

- Analyzing the parameters of the best-fitting ARMA Model that relates actual damage (i.e., outage expressed as percentage of gas stations out of gas on a given day) to social response (i.e., the prevalence of the negative Twitter sentiment about the crisis on a given day), it is seen that the social response is best explained by a non-linear model that is a combination of two linear ones. In the beginning, social response tracks average damage, leading to a model that acts as a low-pass filter. Later, social response becomes more of an auto-regressive function of itself. In other words, sentiment feeds on itself and becomes more weakly correlated to actual damage. The non-linearity

suggests a transition from rational response to a “panic” mode, where sentiment becomes inflated by what social media carry, more so than by physical observations.

- Having understood the non-linear model relating actual damage to social response, it becomes possible to estimate actual damage from the social response curve alone, as observed on Twitter. In other words, given the evolution of sentiment on Twitter, it is possible to estimate the evolution of the gas outage footprint over time. Figure 2 shows the recorded Twitter sentiment curve, denoted by *Act. S*, where higher values denote a more prevalent negative sentiment (and where 1 means 100% prevalence). It also shows the actual damage, denoted by *Act. D*, which denotes the percentage of gas stations out of gas as a function of time. It then compares the actual damage to damage estimated from the Twitter social response curve using a few versions of our models discussed above (denoted by *Est. D*). The comparison demonstrates a remarkable correspondence between the actual and estimated damage curves. The horizontal axis in the figure is the number of days elapsed since hurricane landfall. The figure demonstrates the feasibility of estimating damage footprint from social response.

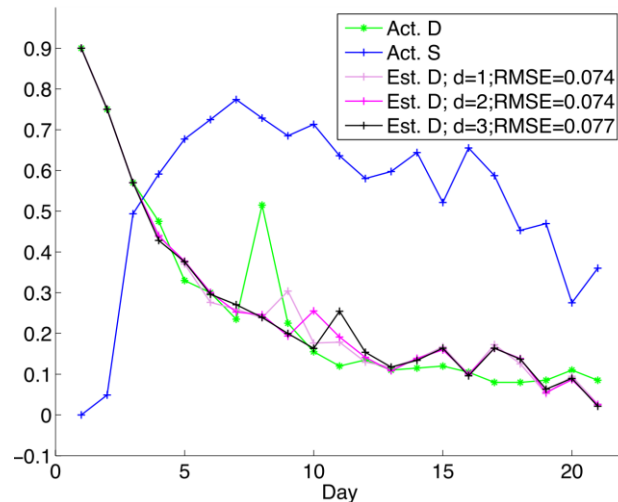


Figure 2. Estimating actual physical damage (*Act. D*) from social response (*Act. S*) in the aftermath of hurricane Sandy. Estimated damage curves (*Est. D*) match actual damage well.

- As might be expected, the physical damage and social response are indeed two interacting cascades. The social response exacerbates damage by artificially inflating demand, causing a larger outage footprint. In turn, the larger outage footprint exacerbates the sentiment, leading to a more prevalent negative social response.

While the above description aggregated all response and all damage into two scalar numbers, the work also considered spatial extensions of the model, where both damage and response were broken up by area (e.g., New York versus New Jersey) and relations were found between the resulting vectors. Models that best predict sentiment in an area were found to be more a function of worst neighbor damage, as opposed to local damage.

Chapter 5: Role Discovery

As mentioned above, in recovering information from social networks, a key concern lies in understanding dependence between sources (for purposes of unbiased source selection). This is accomplished by analysis of the social network graph. Recovered information quality is maximized when the sources selected are independent. Understanding the social network reduces to the problem of estimating social ties between pairs of individuals. We took an axiomatic approach to the problem of social tie strength estimation. Starting from a list of axioms, which a measure of tie strength must satisfy, we characterized functions that satisfy all the axioms. We then showed that there is a range of tie-strength measures that satisfy this characterization. A measure of tie strength induces a ranking on the edges of the social network

(and on the set of neighbors for every person). We showed that for applications where the ranking, and not the absolute value of the tie strength, is the important thing about the measure, the axioms are equivalent to a natural partial order. Once edges in the social network are characterized (as described above), we are further able to determine nodes with special roles in the network, based on the local topology. For example, we may want to identify leaders, followers, rumor spreaders, and information hubs. Intuitively, given a social network, two nodes belong to the same role if they have similar structural behavior. We developed RoIX (Role eXtraction), a scalable (linear in the number of edges), unsupervised learning approach for automatically extracting structural roles from general network data. We demonstrated the effectiveness of RoIX on several network-mining tasks.

Role discovery leads to the natural next question of network similarity. Given a set of k networks, possibly with different sizes and no overlaps in nodes or links, how can we quickly assess similarity between them? For example, is an emerging network of individuals, resources, and transactions similar to one that might indicate preparations for a WMD attack? What features should best characterize such a network? Is there a set of social theories that when represented by a small number of descriptive, numerical features, effectively serve as a “signature” for the network? Having such signatures will simplify a wealth of graph mining and social network analysis tasks, including clustering, outlier detection, and visualization. We introduced a novel, effective, and scalable method for solving the above problem. Another objective addressed was controlling propagation over large graphs. A key issue in WMD scenarios is to study properties of effect propagation over different types of graphs. Controlling the dissemination of an entity (e.g., meme, virus, etc) on a large graph is an interesting problem in many disciplines. Examples include epidemiology, computer security, and marketing, to name a few. We asked: which edges should we add or delete in order to speed-up or contain dissemination? We introduced effective and scalable algorithms to solve these dissemination problems and conducted a theoretical study and experiments on real topologies of varying sizes to demonstrate the effectiveness and scalability of our approaches.

Chapter 6: Scalable Infrastructure Support

Analysis of large networks of the sort mentioned above (e.g., for role discovery on Twitter) requires scalable algorithms. This motivated our research on scalability. We developed algorithms for scalable tensor (multi-dimensional array) decomposition which is an important data mining tool with applications including clustering, text mining, role discovery, and anomaly detection. We developed GigaTensor, a scalable tensor decomposition algorithm on MapReduce. GigaTensor can handle billion-length modes, and at least 100 times more scalable than the state of the art algorithm. We used GigaTensor to analyze a very large real world knowledge base tensor, and found interesting concepts, refinements, and use-cases. We also developed ParCube, a sampling based fast tensor decomposition algorithm which produces sparse factors. We used ParCube to find port scanning attacks on network traffic data, and to find anomalous activities in social network postings data. Finally, we developed scalable algorithms to accomplish common network mining tasks such as link prediction, clustering, outlier detection, ranking, and periodicity mining.

We also studied a distributed (Hadoop) implementation of ‘stochastic gradient descent’ analysis methods for tensors and coupled matrix-tensor factorizations. This led to one of the most scalable methods for the task. We also studied the use of tensor analysis for computer network security, in a honey-pot setting. Our tensor based method was able to group similar attacks (similar by timestamp, and/or exploit, and/or botnet used), and thus help sys-admins understand better how the attackers behave. We also developed a tensor-inspired method to spot temporal communities in real, time-evolving graphs. Continuing with the scalability objective, we studied the problem of determining the proper aggregation granularity for longitudinal (a.k.a., time-evolving or dynamic) network data for purposes of summarizing large amounts of information, when edges are added to the network in a streaming fashion. Longitudinal networks are often used to study topics such as dissemination, change detection, evolution of communities, or network growth. Aggregation lengths are often arbitrarily chosen based on intuition or convenience. The same network may be aggregated per-day in one study, per-week in another study, and per-month in yet another. It is unclear whether these interval lengths are optimal for the tasks being considered. We described a novel algorithmic framework, called ADAGE, which identifies the appropriate variable-length intervals depending on the structural properties of the graph and the tasks under consideration.

Chapter 7: Other Contributions

A significant number of other isolated contributions were developed over the course of the project. Those include linking named entities in Web text with heterogeneous information networks, personalized entity recommendation in heterogeneous information networks, phrase mining for construction of a topic hierarchy, identifying key phrases in unstructured text, entity recognition, among others. Details are omitted from the final report, but can be found in the individual publications.

Appendix A: Selected Awards

- **Jiawei Han:** *Excellence in Graduate and Professional Teaching Award*, Univ. of Illinois at Urbana-Champaign, May 2011
- **Tina Eliasi-Rad:** Third Place Prize, *Antivirus Research Grant/Gift Winner*, sponsored by pcantivirusreviews.com, \$300, 2011 (<http://www.pcantivirusreviews.com/Gifts-and-Grants/>)
- **2012 KDD Best Student Paper Award:** Yizhou Sun, Brandon Norick, Jiawei Han, Xifeng Yan, Philip S. Yu, Xiao Yu, "Integrating Meta-Path Selection with User-Guided Object Clustering in Heterogeneous Information Networks," at 2012 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'12), Beijing, China, Aug. 2012.
- **2012 KDD Best Poster Award:** Yizhou Sun, Brandon Norick, Jiawei Han, Xifeng Yan, Philip S. Yu, Xiao Yu, "Integrating Meta-Path Selection with User-Guided Object Clustering in Heterogeneous Information Networks," at 2012 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'12), Beijing, China, Aug. 2012
- **2012 JCDL Vannevar Bush Best Paper Award:** Hongbo Deng, Jiawei Han, Michael R. Lyu and Irwin King, "Modeling and Exploiting Heterogeneous Bibliographic Networks for Expertise Ranking," Proc. 2012 ACM/IEEE Joint Conf. on Digital Libraries (JCDL'12), Washington, D.C., June 2012.
- **ICDM 10-Year High Impact Paper Award:** for the paper "gSpan: Graph-Based Substructure Pattern Mining", by Xifeng Yan and Jiawei Han, published at Proc. 2002 Int. Conf. on Data Mining (ICDM'02), Maebashi, Japan, Dec. 2002, awarded at 2010 IEEE Int. Conf. on Data Mining, Vancouver, Canada, Dec. 2011.
- **Tina Eliasi-Rad:** Promoted to *Associate Professor with Tenure* effective July 1, 2012.
- **Tarek Abdelzaher,** *Outstanding Technical Achievement and Leadership Award*, IEEE Technical Committee of Real-time Systems, December 2012
- **Tina Eliassi-Rad,** *Best Interdisciplinary Paper Award*, The 21st ACM Conference on Information and Knowledge Management (CIKM'12), Maui, Hawaii, October 2012
- Our ICDM 2012 paper was selected as one of the best papers for possible publication in the journal of *Knowledge and Information Systems*: Hyungsul Kim, Yizhou Sun, Julia Hockenmaier, and Jiawei Han, "ETM: Entity Topic Models for Mining Documents Associated with Entities," Proc. of 2012 IEEE Int. Conf. on Data Mining (ICDM'12), Brussels, Belgium, Dec. 2012, pp. 349-358.
- Our ICDM 2013 paper was invited to the journal of "Knowledge and Information Systems" as one of the best papers: Chi Wang, Marina Danilevsky, Jialu Liu, Nihit Desai, Heng Ji, and Jiawei Han, "Constructing Topical Hierarchies in Heterogeneous Information Networks," Proc. 2013 IEEE Int. Conf. on Data Mining (ICDM'13), Dallas, TX, Dec. 2013.

- **KDD CUP 2013 Award:** Our team won the 2nd place of Track 2 competition on “Author Disambiguation in Bibliographic Data” with the techniques described in Jialu Liu, Kin Hou Lei, Jeffery Yufei Liu, Chi Wang, and Jiawei Han, “Ranking-Based Name Matching for Author Disambiguation in Bibliographic Data,” Proc. of KDD Cup 2013 Workshop (KDDCUP’13), Chicago, IL, Aug. 2013.
- **2013 ACM SIGKDD Doctoral Dissertation Award:** It was received by Yizhou Sun, for her Ph.D. thesis on “Mining heterogeneous information networks”.
- **NSF Graduate Fellowship:** Ph.D. student: Brandon Norick (2013-2014).
- **NSF Graduate Fellowship:** Ph.D. student: Ahmed El-Kishky (2013-2014). IBM Ph.D. Fellowship: Ph.D student: Quanquan Gu (2013-2014).
- **Grand Prize in Yelp Dataset Challenge,** for the paper: Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, Jiawei Han, “Mining Quality Phrases from Massive Text Corpora”, in Proc. of 2015 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD’15), Melbourne, Australia, May 2015
- **Best paper award for ICAC 2015:** Our paper on “SocialTrove” won the best paper award at ICAC 2015
- **Best of SDM 2015:** Selected as one of the best papers in the conference and invited to journal “Statistical Analysis and Data Mining (SADM),” special issue “Best of SDM 2015” for the paper: Jialu Liu, Chi Wang, Jing Gao, Quanquan Gu, Charu Aggarwal, Lance Kaplan, and Jiawei Han, “GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data”, in Proc. of 2015 SIAM Int. Conf. on Data Mining (SDM’15), Vancouver, Canada, Apr. 2015
- **Best of EDBT 2015:** Selected as one of the best papers in the conference and invited to journal ACM Transactions on Database Systems (TODS) as “Best of EDBT 2015” for the paper: Jonathan Kuck, Honglei Zhuang, Xifeng Yan, Hasan Cam, and Jiawei Han, “Query-Based Outlier Detection in Heterogeneous Information Networks”, in Proc. of 2015 Int. Conf. on Extending Database Technology (EDBT’15), Brussels, Belgium, Mar. 2015
- **CRA Outstanding Undergraduate Researcher Award:** Jiawei Han’s undergraduate student Urvashi Khandelwal receives 2015 CRA (Computer Research Association) Outstanding Undergraduate Researcher Award, January 2015
- **NDSEG Graduate Fellowship:** National Defense Science and Engineering Graduate Fellowship: Ph.D. student: Ahmed El-Kishky (2015-2017)
- **NSF Graduate Fellowship:** Ph.D. student: Brandon Norick (2014-2015)
- Han’s student Xiang Ren received 2017 **David J. Kuck Outstanding M.S. Thesis Award** from the Department of Computer Science at the University of Illinois at Urbana-Champaign, May 2017
- Han’s Ph.D. student Jingbo Shang receives 2017 **Google PhD Fellowship**, the sole recipient in the category of Structured Data and Database Management on North America
- Han’s PhD student Xiang Ren was selected by ACM SIGKDD as one of the students to receive a travel scholarship to attend the ACM 50th Celebration of the Turing Award (2017)
- **2016 UIUC Computer Science Distinguished Alumni Educator Award:** Han’s ex-PhD student Yizhou Sun (now assistant professor at CS, UCLA)
- **2016 Siebel Scholar:** Han’s M.S. student Wenqi He was honored as one of six Siebel Scholar award winners at UIUC in Sept. 2016
- Han’s Ph.D. student Xiang Ren receives 2016 **Google PhD Fellowship**, the sole recipient in the category of Structured Data and Database Management internationally

- ECML/PKDD 2015: Best Student Paper Runner-Up award for the paper: Chao Zhang, Shan Jiang, Yucheng Chen, Yidan Sun, and Jiawei Han, "Fast Inbound Top-K Query for Random Walk with Restart", in Proc. of 2015 European Conf. on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECMLPKDD'15), Porto, Portugal, Sept. 2015.
- SIGMOD'15: won **Grand Prize in Yelp Dataset Challenge**, 2015 for the paper: Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, Jiawei Han, "Mining Quality Phrases from Massive Text Corpora", in Proc. of 2015 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'15), Melbourne, Australia, May 2015
- SDM'15: Selected as one of the best papers in the conference and invited to journal "Statistical Analysis and Data Mining (SADM)" special issue "Best of SDM 2015" for the paper: Jialu Liu, Chi Wang, Jing Gao, Quanquan Gu, Charu Aggarwal, Lance Kaplan, and Jiawei Han, "GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data", in Proc. of 2015 SIAM Int. Conf. on Data Mining (SDM'15), Vancouver, Canada, Apr. 2015
- EDBT'15: Selected as one of the best papers in the conference and invited to journal "ACM Transactions on Database Systems (TODS)" as "Best of EDBT 2015" for the paper: Jonathan Kuck, Honglei Zhuang, Xifeng Yan, Hasan Cam, and Jiawei Han, "Query-Based Outlier Detection in Heterogeneous Information Networks", in Proc. of 2015 Int. Conf. on Extending Database Technology (EDBT'15), Brussels, Belgium, Mar. 2015
- **KDD dissertation award (runner up)**, for Prof. Evangelos Papalexakis, 2017
- Paper among the "Best of Conference" in ICDM 2016: Kijung Shin, Tina Eliassi-Rad, Christos Faloutsos. CoreScope: Graph Mining Using k-Core Analysis - Patterns, Anomalies and Algorithms. ICDM 2016: 469-478.

Appendix B: Research Transitions

- Jialu Liu, Jingbo Shang, and Jiawei Han, [Phrase Mining from Massive Text and Its Applications](#), Synthesis Lectures on Data Mining and Knowledge Discovery, Morgan & Claypool Publishers, 2017.
- Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, Jiawei Han, "[Mining Quality Phrases from Massive Text Corpora](#)", in Proc. of 2015 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'15), Melbourne, Australia, May 2015 (**won Grand Prize in Yelp Dataset Challenge, 2015**)
- Xiang Ren, Ahmed El-Kishky, Chi Wang, and Jiawei Han, "[Automatic Entity Recognition and Typing from Massive Text Corpora: A Phrase and Network Mining Approach](#)" (conference tutorial), 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
- Xiang Ren, Zeqiu Wu, Wenqi He, Meng Qu, Clare Voss, Heng Ji, Tarek Abdelzaher and Jiawei Han, "[CoType: Joint Extraction of Typed Entities and Relations with Knowledge Bases](#)", in Proc. of 2017 World-Wide Web Conf. (WWW'17), Perth, Australia, Apr. 2017.
- Co-PI Eliassi-Rad has received funding from Northeastern University's Global Resilience Institute on maritime container anomaly detection. In addition, co-PI Eliassi-Rad is in talks with National Geospatial-Intelligence Agency (NGA), which has shown interest in technology transfer from this project on the topic of maritime container anomaly detection.
- Abdelzaher received 6.2 funding for a project called MINI-DASS whose goal is to match mission-informed needed information to discoverable available sensing sources. The work was largely based on the social sensing framework described in this report.
- Abdelzaher received a \$4M DARPA award to develop a simulator of social media under the SocialSim program. The work leverages social sensing models described above.

Appendix C: Result Dissemination

- Yizhou Sun, Jiawei Han, Xifeng Yan, and Philip S. Yu, "Mining Knowledge from Interconnected Data: A Heterogeneous Information Network Analysis Approach," **Tutorial**, *Int. Conf. on Very Large Data Bases (VLDB'12/PVLDB)*, Aug. 2012.
- Yizhou Sun, Jiawei Han, Xifeng Yan, and Philip S. Yu, "On the Power of Heterogeneous Information Networks," **Tutorial**, *Int. Conf. on Advances in Social Network Analysis and Mining (ASONAM'12)*, Aug. 2012.
- Jiawei Han, Yizhou Sun, Xifeng Yan, and Philip S. Yu, "Mining Knowledge from Data: An Information Network Analysis Approach," **Tutorial**, *IEEE Int. Conf. on Data Engineering ICDE'12*, Arlington, VA, Apr. 2012.
- Jiawei Han, "Mining Heterogeneous Information Networks," *KDD Summer School*, Aug. 2012.
- Jiawei Han, "Data Mining with Social and Trajectory Data: Urban Computing in the Big Data Age," *CCF Advanced Disciplines Lectures*, Aug. 2012.
- Jiawei Han, "A Meta Path-Based Approach for Similarity Search and Mining of Heterogeneous Information Networks," **Tutorial**, *Workshop on Algorithms for Modern Massive Data Sets (MMDS)*, Stanford Univ., CA, July 2012.
- Jiawei Han, "Mining Information Networks," *DHS-UIUC Summer School*, June 2012.
- Jiawei Han, "Mining Heterogeneous Information Networks: The Next Frontier," **Keynote Talk**, *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'12)*, Aug. 2012.
- Jiawei Han, "Construction of Web-Based, Service-Oriented Information Networks: A Data Mining Perspective," **Keynote Talk**, *Int. Conf. on Web-Age Information Management (WAIM'12)*, Aug. 2012.
- Jiawei Han, *invited or distinguished seminars at*: Univ. of Notre Dame (Dec. 2011), Univ. of Michigan (Dec. 2011), UCLA (Jan. 2012), Univ. of Central Florida (Feb. 2012), Microsoft Research Asia (Aug. 2012), Renmin Univ. (May/Aug. 2012), Tsinghua Univ. (May 2012), Baidu Inc. (Aug. 2012).
- Tina Eliasi-Rad, and Christos Faloutsos "Discovering Roles and Anomalies in Graphs: Theory and Applications," **Tutorial**, *SIAM SDM 2012*.
- Tina Eliasi-Rad, *invited talks or distinguished seminars at* Lawrence Livermore National Laboratory, Network Security Innovation Center (April 2012), Lawrence Livermore National Laboratory, Center for Applied Scientific Computing (April 2012), Northern New Jersey Junior Science and Humanities Symposium (March 2012), Northeastern University (December 2011), MIT, CSAIL (December 2011), IBM Watson, Hawthorne (November 2011). Lehigh University (November 2011), UCLA IPAM (October 2011).
- C. Faloutsos, "Large Graph Mining - Patterns, Tools and Cascade Analysis," *KDD 2012 Summer School*, Beijing, China, Aug. 2012.
- C. Faloutsos and U. Kang, "Managing and Mining Large Graphs: Patterns and Algorithms," **Tutorial**, *SIGMOD 2012*, Scottsdale, AZ, May 2012.
- T. Abdelzaher, "Fact-finding and Other Challenges in Social Sensing," **Invited Talk**, *IBM Research, Hawthorne, NY*, Aug. 2011
- T. Abdelzaher, "Challenges in Human-centric Sensor Networks," *ARTIST Summer School*, Aix-les-Bains, France, Sept. 2011
- T. Abdelzaher, "Social Sensing Challenges for a Smarter Planet," **Keynote Talk**, *2nd Annual ELLIIT Workshop*, Lund, Sweden, Oct. 2011

- T. Abdelzaher, "Social Sensing Services and Applications," **Invited Talk**, *ETH Zurich*, Zurich, Switzerland, Nov. 2011
- T. Abdelzaher, "Analytics of Social Sensing," **Invited Talk**, *Swedish Institute of Computer Science (SICS)*, Stockholm, Sweden, Nov. 2011
- T. Abdelzaher, "Social Sensing Services and Applications," **Invited Talk**, *Ericsson Labs*, Lund, Sweden, Nov. 2011
- T. Abdelzaher, "Analytics of Social Sensing," **Invited Talk**, *Uppsala University*, Uppsala, Sweden, Dec. 2011
- T. Abdelzaher, "Social Sensing: Challenges with Humans in the Loop," **Invited Talk**, *University of Nebraska, Lincoln*, Lincoln, NE, Jan. 2012
- T. Abdelzaher, "Smart Cities: Playground for Cooperating Objects," **Panel**, *CONET 2012*, CPS Week, Beijing, China, April 2012
- T. Abdelzaher, "Challenges, Opportunities and Future Directions for Mobile Sensing," **Panel**, *Mobile Sensing*, CPS Week, Beijing, China, April 2012
- T. Abdelzaher, "Information-centric Networking: A Research Frontier for Cyber-physical Systems," **Keynote Talk**, *2nd International Workshop on Cyber-Physical Networking Systems (CPNS 2012)*, Macau, China, June 2012
- T. Abdelzaher, "Research Questions in Social Sensing," **Keynote Talk**, *IEEE ICDCS*, Macau, China, June 2012
- T. Eliassi-Rad and C. Faloutsos, "Discovering Roles and Anomalies in Graphs: Theory and Applications," **Tutorial**, *ECML PKDD Conference*, September 2013. Slides available at: <http://eliassi.org/>
- S. Papadimitriou and T. Eliassi-Rad, "Mining Data from Mobile Devices," **Tutorial**, *19th ACM SIGKDD Conference*, August 2013. Website: <http://mobilemining.clusterhack.net/>
- Tim Weninger, and Jiawei Han, "Information Network Analysis and Extraction on the World Wide Web," **Tutorial**, *Int. Conf. on the World Wide Web (WWW'13)*, Rio de Janeiro, Brazil, May 2013.
- Manish Gupta, Jing Gao, Charu Aggarwal, and Jiawei Han, "Outlier Detection for Temporal Data," **Tutorial**, *SIAM Data Mining Conf. (SDM'13)*, Austin, TX, May 2013.
- Tim Weninger and Jiawei Han, "Exploring Structure and Content on the Web: Extraction and Integration of the Semi-Structured Web," **Tutorial**, *Intl. Conf. on Web Search and Data Mining (WSDM'13)*, Rome, Italy, February 2013.
- B. Aditya Prakash and Christos Faloutsos, "Understanding and Managing Cascades in Large Graphs," **Tutorial**, *ECML/PKDD 2012*, Bristol and Tutorial at *VLDB 2012*, Istanbul
- Yizhou Sun, Jiawei Han, Xifeng Yan, and Philip S. Yu, "Mining Knowledge from Interconnected Data: A Heterogeneous Information Network Analysis Approach," **Tutorial**, *Int. Conf. on Very Large Data Bases (VLDB'12/PVLDB)*, August 2012.
- Yizhou Sun, Jiawei Han, Xifeng Yan, and Philip S. Yu, "On the Power of Heterogeneous Information Networks," **Tutorial**, *Int. Conf. on Advances in Social Network Analysis and Mining (ASONAM'12)*, August 2012.
- Spiros Papadimitriou and Tina Eliassi-Rad, "Data from Mobile Devices: A Survey of Smart Sensing and Analytics," **Tutorial**, *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, IL, August 2013.
- Manish Gupta, Jing Gao, Charu Aggarwal and Jiawei Han, "Outlier Detection for Graph Data," **Tutorial**, *IEEE/ACM Int. Conf. on Social Networks Analysis and Mining (ASONAM'13)*, Niagara Falls, Canada, August 2013.

- Tina Eliassi-Rad, Christos Faloutsos, "Discovering Roles and Anomalies in Graphs: Theory and Applications," **Tutorial**, *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD'13)*, Prague, Czech Republic, September 2013.
- Tarek Abdelzaher, "Cyber-physical Systems with Humans in the Loop," **Tutorial**, *Artist Summer School*, Aix-les-Bains, France, September 2013.
- Manish Gupta, Jing Gao, Charu Aggarwal, Jiawei Han, "Outlier Detection for Temporal Data," **Tutorial**, *ACM International Conference on Information and Knowledge Management (CIKM 2013)*, San Francisco, CA, October 2013.
- Danai Koutra, Tina Eliassi-Rad, Christos Faloutsos, "Node Similarity, Graph Similarity and Matching: Theory and Applications," **Tutorial**, *SIAM International Conference on Data Mining (SDM'14)*, Philadelphia, PA, April 2014.
- Lidan Wang, Jimmy Lin, Donald Metzler and Jiawei Han, "Learning to Efficiently Rank on Big Data," **Tutorial**, *Int. World-Wide Web Conf. (WWW'14)*, Seoul, Korea, April 2014.
- Jiawei Han and Chi Wang, "Mining Latent Entity Structures from Massive Unstructured and Interconnected Data," **Tutorial**, *ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'14)*, Snowbird, UT, June 2014.
- Jiawei Han, Chi Wang, Ahmed El-Kishky, "Bringing Structure to Text: Mining Phrases, Entity Concepts, Topics, and Hierarchies," **Tutorial**, *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'14)*, New York, NY, August 2014.
- Jiawei Han, "From Social Networks to Heterogeneous Social and Information Networks: A Data Mining Perspective," **Keynote Talk**, *SNAKDD Int. Workshop on Social Network Mining and Analysis (SNAKDD'13)*, Chicago, IL, August 2013.
- Tarek Abdelzaher, "Quality of Information: A Network Science Perspective," **Invited Talk**, *Army Research Laboratories, Adelphi, MD*, August 2013.
- Tarek Abdelzaher, "Quality of Information: A Network Science Perspective," **Invited Talk**, *Army Research Laboratories, Aberdeen Proving Ground, MD*, August 2013.
- Jiawei Han, "Challenging Problems for Scalable Mining of Heterogeneous Social and Information Networks", **Keynote Talk**, *SIGKDD Int. Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications (BigMine'13)*, Chicago, IL, August 2013.
- Tina Eliassi-Rad, "Measuring Tie Strength in Implicit Social Networks," **Invited Talk**, *ACM SIGKDD Workshop on Mining and Learning from Graphs (MLG 2013)*, August 2013.
- Tarek Abdelzaher, "Cyber-physical Systems with Humans in the Loop," **Invited Speaker**, *Artist Summer School*, Aix-les-Bains, France, September 2013.
- Jiawei Han, "Mining and Exploring Semi-Structured, Heterogeneous Social and Information Networks", **Invited Talk**, *5th Int. Workshop on Network Theory: Network Science Meets the Science of Teams*, Northwestern Univ., Chicago, IL, October 2013.
- Tarek Abdelzaher, "Quality of Information in Social Sensing," **Keynote Talk**, *ITA Annual Fall Meeting*, Palisades, NY, October 2013.
- Tarek Abdelzaher, "Social Sensing: Making Reliable Observations from Unreliable Data," **Invited Talk**, *Syracuse University*, Syracuse, NY, October 2013.
- Tina Eliassi-Rad, "Measuring Tie Strength in Implicit Social Networks," **Invited Talk**, *Rutgers DIMACS, Workshop on Statistical Analysis of Network Dynamics and Interactions*, November 2013.

- Tarek Abdelzaher, "Data Analytics for Human-centric Cyber-physical Systems," **Invited Talk**, *Global Innovation Festival, Daegu Gyeongbuk Institute of Science and Technology (DGIST)*, Daegu, Korea, November 2013.
- Tina Eliassi-Rad, University of Notre Dame, "Discovering Roles in Graphs: Algorithms and Applications," **Invited Talk**, *Department of Computer Science and Engineering, Notre Dame, IN*, December 2013.
- Tarek Abdelzaher, "Cyber-physical Systems in Social Spaces: A Data Reliability Perspective," **Invited (Award) Talk**, *IEEE Real-time Systems Symposium*, Vancouver, Canada, December 2013.
- Jiawei Han, "Taming the Complexity of Information and Networks for Future Wars", **Invited Talk**, *Army Research Strategy Planning Meeting on Information at the Tactical Edge*, organized by ARL, Potomac, Maryland, December 2013.
- Tarek Abdelzaher, "Social Sensing: Making Reliable Observations from Unreliable Data," **Invited Talk**, *Hong Kong Polytechnic University*, Hong Kong, December 2013.
- Tina Eliassi-Rad, "Some Advances in Graph Mining: Theory, Algorithms, and Applications," **Invited Talk**, *University of California Davis, Department of Computer Science*, Davis, CA, January 2014.
- Tina Eliassi-Rad, "Advances in Graph Mining," **Invited Talk**, *Lawrence Livermore National Laboratory, Center for Applied Scientific Computing*, Livermore, CA, January 2014.
- Tina Eliassi-Rad, "Advances in Graph Mining," **Invited Talk**, *Facebook*, Menlo Park, January 2014.
- Tina Eliassi-Rad, *UCLA IPAM Workshop on Mathematics of Social Learning*, **Invited Talk**, January 2014.
- Christos Faloutsos, "Large Graph Mining - Patterns, Explanations and Cascade Analysis," **Invited Talk**, *NCSU Lecture Series*, February 2014.
- Tarek Abdelzaher, "Social Sensing: Making Reliable Observations from Unreliable Data," **Keynote Talk**, *Danish Academy of Technical Sciences (Big Data Seminar)*, Aalborg, Denmark, February 2014.
- Christos Faloutsos, "Large Graph Mining - Patterns, Explanations and Cascade Analysis," **Invited Talk**, *Case Western Reserve University, EECS Seminar Series*, March 2014.
- Tina Eliassi-Rad, "Discovering Roles in Graphs: Algorithms and Applications," **Invited Talk**, *Spring Workshop on Mining and Learning (SMiLe)*, Oostende, Belgium, March 2014.
- Christos Faloutsos, "Mining Large, Dynamic Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Keynote Talk**, *WWW*, April 2014.
- Tarek Abdelzaher, "Real-time Open Source Analysis of WMD Events: The Syria Chemical Attacks and Other Selected Case-studies," **Invited Talk**, *DTRA Seminar Series*, April 2014.
- Tina Eliassi-Rad, "Discovering Roles in Graphs: Algorithms and Applications," **Invited Talk**, *SIAM SDM Workshop on Mining Networks and Graphs: A Big Data Analytics Challenge*, Philadelphia, PA, April 2014.
- Tina Eliassi-Rad, "Structural Features Threaten Privacy across Social Graphs" **Invited Talk**, *Charles River Workshop on Private Analysis of Social Networks*, Boston University, Boston, MA, May 2014.
- Jiawei Han, "Construction, Exploration and Mining of Information Networks", **Invited Talk**, *Yahoo! 2014 Big Thinker forum*, California, May 2014.

- Tina Eliassi-Rad, "Discovering Roles in Graphs: Algorithms and Applications," **Invited Talk**, *International School and Conference on Network Science (NetSci'14)*, Berkeley, CA, June 2014.
- Jiawei Han, "Construction and Mining of Heterogeneous Information Networks: Will It Be a Key to Web-Aged Information Management and Mining?," **Keynote Talk**, *Int. Conf. on Web-Age Information Management (WAIM'14)*, Macau, China, June 2014.
- Tarek Abdelzaher, "The Power of Analytics," **Invited Talk**, *International Conference on Smart Computing (SmartComp)*, Hong Kong, November 2014
- Tarek Abdelzaher, "Anomaly Determination and Situation Understanding," **Invited Talk**, *IBM Research*, Yorktown Heights, NY, December 2014
- Tarek Abdelzaher, "Resource Management in Self-Aware Computing Systems," **Invited Talk**, *Dagstuhl Seminar*, Dagstuhl, Germany, January 2015
- Tarek Abdelzaher, "Cyber-physical Systems in Social Spaces," **Invited Talk**, *Microsoft Research*, Redmond, WA, April 2015
- Tarek Abdelzaher, "Social Network Signal Processing," **Keynote Talk**, *International Conference on Distributed Computing in Sensor Systems (DCoSS)*, Fortaleza, Brazil, June 2015
- Christos Faloutsos, "Mining Large Graphs," **Invited Talk**, *Distinguished Lecture Series, Anniversary Edition, University of Toronto*, November 2014
- Christos Faloutsos, "Mining Large Graph," **Invited Talk**, *Distinguished Lecture Series, Univ. of Illinois Chicago*, April 2015
- Christos Faloutsos, "Mining Large Graphs," **Invited Talk**, *Distinguished Lecture Series, Univ. of Southern California*, April 2015
- Tina Eliassi-Rad, "Some Advances in Graph Mining: Theory, Algorithms, and Applications," **Invited Talk**, *University of Leuven*, Leuven, Belgium, July 2014
- Tina Eliassi-Rad, distinguished seminars at the University of Antwerp, Antwerp, Belgium, July 2014, Yahoo! NYC, New York, NY August 2014, Yahoo!, Sunnyvale, CA, August 2014, MIT Lincoln Laboratory's 5th Annual Graph Exploitation Symposium (GraphEx), Dedham, MA, August 2014, Northeastern University, College of Computer and Information Science, Boston, MA, December 2014, The 9th Annual Workshop for Women in Machine Learning (WiML), Montreal, Canada, December 2014, NYC Women in Machine Learning and Data Science Meetup, New York, NY, December 2014, University of California, Santa Cruz, Department of Computer Science, January 2015, and the University of California, Santa Barbara, Department of Computer Science, January 2015
- Tina Eliassi-Rad, SIAM SDM Workshop on Mining Networks and Graphs: A Big Data Analytic Challenge, **Panel** on "Future Challenges in Mining Large Networks", Vancouver, Canada, April 2015
- Tina Eliassi-Rad, International School and Conference on Network Science (NetSci), Satellite on Higher-Order Models in Network Science (HONS), Zaragoza, Spain, June 2015
- Tina Eliassi-Rad, International School and Conference on Network Science (NetSci), Satellite on Arts, Humanities, and Complex Networks, Panel on "Scientific Cultural Analysis", Zaragoza, Spain, June 2015
- Tina Eliassi-Rad, Institute for Humanitarian Informatics (iHi), **Invited Participant** at Workshop, Georgetown University, Washington, DC, June 2015.
- Tina Eliassi-Rad, CyberInfrastructure for NETWORK science (CINET) Workshop, Virginia Tech, Blacksburg, VA, July 2015

- Jiawei Han, "Construction and Mining of Heterogeneous Information Networks from Text Data," **Keynote Talk**, *Annual meeting of the Association for Computational Linguistics (ACL'15)*, Beijing, China, July 2015
- Jiawei Han, "Construction and Mining of Semi-Structured Heterogeneous Information Networks," **Keynote Talk**, *International School and Conference on Network Science (NetSci-x-2015)*, Rio de Janeiro, Brazil, Jan. 2015
- Jiawei Han, "Taming Interconnect and Unstructured Data: Construction and Mining of Heterogeneous Information Networks," **Distinguished Seminar**, *Case-Western Reserve University*, Nov. 2014
- Jiawei Han, "Construction, Exploration, and Mining of Semi-Structured, Heterogeneous Information Networks," **Tutorial**, *Military Sensing Symposium (MSS) National Symp. on Sensor Data Fusion (NSSDF)*, Springfield, VA, Oct. 2014
- Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han, "Truth Discovery and Crowdsourcing Aggregation: A Unified Perspective," **Tutorial**, *Int. Conf. on Very Large Data Bases (VLDB'15)*, KohalaCoast, Hawaii, Sept. 2015
- Xiang Ren, Ahmed El-Kishky, Chi Wang, and Jiawei Han, "Automatic Entity Recognition and Typing from Massive Text Corpora: A Phrase and Network Mining Approach," **Tutorial**, *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15)*, Sydney, Australia, Aug. 2015
- Jiawei Han, Heng Ji and Yizhou Sun, "Successful Data Mining Methods for NLP," **Tutorial**, *Annual Meeting of the Association for Computational Linguistics and Int. Joint Conf. on Natural Language Processing (ACL-IJCNLP 2015)*, Beijing, China, July 2015
- Christos Faloutsos, Yasushi Sakurai and Yasuko Matsubara "Mining and Forecasting of Big Time-series data," **Tutorial**, *ACM SIGMOD*, Melbourne, AU, June 2015.
- Christos Faloutsos, Alex Beutel and Leman Akoglu, "Graph-Based User Behavior Modeling: From Prediction to Fraud Detection," **Tutorial**, *KDD 2015*, Sydney, Australia, Aug. 2015.
- Danai Koutra, Tina Eliassi-Rad, Christos Faloutsos, "Node and Graph Similarity: Theory and Applications," **Tutorial**, *14th IEEE International Conference on Data Mining (ICDM'14)*, Shenzhen, China, December 2014.
- Spiros Papadimitriou and Tina Eliassi-Rad, "Mining Mobility Data," **Tutorial**, *24th International World Wide Web Conference (WWW'15)*, Florence, Italy, May 2015.
- Jingbo Shang, Xiang Ren, Meng Jiang, and Jiawei Han, "Mining Entity-Relation-Attribute Structures from Massive Text Data," **Tutorial**, *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'17)*, Halifax, Nova Scotia, Canada, August 2017.
- Xiang Ren, Meng Jiang, Jingbo Shang and Jiawei Han, "Building Structured Databases of Factual Knowledge from Massive Text Corpora," **Tutorial**, *ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'17)*, June 2017.
- Xiang Ren, Meng Jiang, Jingbo Shang and Jiawei Han, "Constructing Structured Information Networks from Massive Text Corpora," **Tutorial**, *World-Wide Web Conf. (WWW'17)*, Perth, Australia, Apr. 2017.
- Chao Zhang, Quan Yuan, and Jiawei Han, "Bringing Semantics to Spatiotemporal Data Mining: Challenges, Methods, and Applications," **Tutorial**, *IEEE Int. Conf. on Data Engineering (ICDE'17)*, San Diego, California, Apr. 2017.
- Zhenhui (Jessie) Li, Fei Wu, and Jiawei Han, "Trajectory Data Mining", **Tutorial**, *IEEE Int. Conf. on Big Data (BigData'16)*, Washington, D.C., Dec. 2016
- Meng Jiang, Peng Cui, Jiawei Han, "Data-Driven Behavioral Analytics: Observations, Representations and Models", **Tutorial**, *ACM Int. Conf. on Knowledge and Information Management (CIKM'16)*, Indianapolis, IN, Oct. 2016

- Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han, "Mining Reliable Information from Passively and Actively Crowdsourced Data," **Tutorial**, *ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD'16)*, San Francisco, CA, Aug. 2016
- Jiawei Han, "Construction and Mining of Text-Rich Heterogeneous Information Networks", International **Summer School** on Web Science and Technology (WebST'16), Bilbao, Spain, July 18-22, 2016
- Xiang Ren, Ahmed El-Kishky, Heng Ji, Jiawei Han, "Automatic Entity Recognition and Typing in Massive Text Data", **Tutorial**, *ACM SIGMOD Conf. on Management of Data (SIGMOD'16)*, San Francisco, CA, July 2016
- Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han, "Towards Veracity Challenge in Big Data," **Tutorial**, *SIAM Data Mining Conf. (SDM'16)*, Miami, FL, May 2016
- Xiang Ren, Ahmed El-Kishky, ChiWang, and Jiawei Han, "Automatic Entity Recognition and Typing in Massive Text Corpora," **Tutorial**, *Int. World-WideWeb Conf. (WWW'16)*, Montreal, Canada, April 2016
- Jing Gao, Qi Li, Bo Zhao, Wei Fan, and Jiawei Han, "Truth Discovery and Crowdsourcing Aggregation: A Unified Perspective," **Tutorial**, *Int. Conf. on Very Large Data Bases (VLDB'15)*, Kohala Coast, Hawaii, Sept. 2015
- Xiang Ren, Ahmed El-Kishky, Chi Wang, and Jiawei Han, "Automatic Entity Recognition and Typing from Massive Text Corpora: A Phrase and Network Mining Approach," **Tutorial**, *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15)*, Sydney, Australia, Aug. 2015
- Jiawei Han, Heng Ji and Yizhou Sun, "Successful Data Mining Methods for NLP," **Tutorial**, *Annual Meeting of the Association for Computational Linguistics and Int. Joint Conf. on Natural Language Processing (ACL-IJCNLP 2015)*, Beijing, China, July 2015
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Invited Distinguished Seminar**, *IBM Almaden Research Center*, California, August 2017
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Invited Distinguished Seminar**, *Technische Univ. of Eindhoven*, Netherlands, July 2017
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Invited Talk**, *Univ. of Edinburgh*, Scotland, July 2017
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Keynote Talk**, *Int. Database Engineering & Applications Symposium (IDEAS'17)*, Bristol, England, July 2017
- Jiawei Han, "Construction of Structured Networks from Massive Text Data: A Data-Driven Approach," **Tutorial**, *U.S. Army Research Lab*, Aberdeen, June 2017
- Jiawei Han, "Exploring the Power of Embedding in Construction and Mining of Multi-Genre Networks," **Tutorial**, *U.S. Army Research Lab*, Adelphi, June 2017
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Keynote Talk**, *Int. FLAIRS Conf.*, Marco Island, FL, May 2017
- Jiawei Han, "Construction of Structured Heterogeneous Networks from Massive Text Data," **Keynote Talk**, *ACM SIGMOD 2017 Workshop on Network Data Analytics (NDA)*, Chicago, IL, May 2017
- Jiawei Han, "Exploring Text Mining for Health Data Analytics," **Invited Talk**, *Illinois Health Data Analytics Summit*, Champaign, IL, May 2017
- Jiawei Han, "Automated Mining of Phrases and Structures from Massive Text Corpora," **Invited Talk**, *Facebook Inc., Menlo Park*, CA, March 2017

- Jiawei Han, "Exploring Text Mining for Bio-Data Annotation and Retrieval," **Invited Talk**, *Medical School, Stanford University*, Stanford, CA, March 2017
- Jiawei Han, "Networks Everywhere or Nowhere: On Construction of Semi-Structured Heterogeneous Networks from Massive Text Data," **Invited Talk**, *Beckman Institute, UIUC*, Feb. 2017
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Keynote Talk**, *MORS Emerging Techniques Special Meeting (METSM)*, Alexandria, VA, Virginia, Dec. 6, 2016.
- Jiawei Han, "Finding Truth from Multiple Conflicting Sources: A Probabilistic Network-Based Approach," **Invited Talk**, *Network Science Seminar Series at DTRA (Defense Threat Reduction Agency)*, Fort Belvoir, Virginia, Dec. 7, 2016.
- Jiawei Han, "On the Power of Big Data: Mining Structures from Massive, Unstructured Text Data," **Keynote Talk**, *IEEE Int. Conf. on Big Data (BigData'16)*, Washington, D.C., Dec. 5-8, 2016.
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Keynote Talk**, *NSERC CREATE DAV Camp: Data Analytics & Visualization Research Showcase and Bootcamp (CreateDAV-2016)*, Toronto, Canada, Nov. 7, 2016.
- Jiawei Han, "From Data to Knowledge: A Data-to-Network-to-Knowledge (D2N2K) Paradigm," **Keynote Talk**, *International Summer School on Web Science and Technology (WebST'16)*, Bilbao, Spain, July 18-22, 2016
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Invited Talk**, *LinkedIn Inc.*, June 2016
- Jiawei Han, "Construction and Mining of Heterogeneous Information Networks from Massive Text Data," **Distinguished Seminar**, *Arizona State Univ.*, May 2016
- Jiawei Han, "Mining Structures from Massive Bio-Text Data: A Data-Driven Approach," **Invited Talk**, *Univ. of Wisconsin-Madison*, April 2016
- Jiawei Han, "Recent Progress on Mining Bio-Text Data: A Data-Driven Approach," **Invited Talk**, *NIH BD2K Consortium Teleconference Meeting*, March 2016
- Jiawei Han, "Mining Structures from Massive Text Data: A Data-Driven Approach," **Invited Talk**, *Google Inc.*, March 2016
- Jiawei Han, "Structuring and Mining of Heterogeneous Information Networks," **Invited Talk**, *Workshop on Incomplete Network Data (WIND)*, *Sandia National Lab*, Livermore, CA, March 2016
- Jiawei Han, "Networks Everywhere: On Construction of Semi-Structured Heterogeneous Networks from Massive Text Data," **Keynote Talk**, *Int. School and Conference on Network Science (NetSci-x-2016)*, Wroclaw, Poland, Jan. 2016
- Jiawei Han, "Construction and Mining of Semi-Structured Heterogeneous Information Networks," **Keynote Talk**, *IEEE International Symposium on Multimedia (IEEE-ISM'15)*, Miami, Florida, Dec. 2015
- Jiawei Han, "Towards Construction and Mining of Biological Information Networks from Text Data," **Distinguished Seminar**, *UCLA*, Dec. 2015
- Jiawei Han, "Turning Big Data to Big Knowledge: A Data-to-Network-to-Knowledge Paradigm," **Distinguished Seminar**, *RPI, Wuhan Univ., Fudan Univ.*, Oct. 2015
- Jiawei Han, "Construction and Mining of Heterogeneous Information Networks from Text Data," **Invited DHS E-Seminar**, *Rutgers Univ.*, Sept. 2015
- Jiawei Han, "Towards Construction and Mining of Biological Information Networks from Text Data," **Invited Talk**, *Mayo Clinic Individualizing Medicine*, Sept. 2015

- Jiawei Han, "Constructing Structured Heterogeneous Information Networks from Massive Unstructured Text Data," **Invited Talk**, *Boeing, Inc.*, Sept. 2015.
- Jiawei Han, "From Data to Knowledge: Construction and Mining of Heterogeneous Information Networks," **Distinguished Seminar**, *Univ. of New South Wales*, Sydney, Australia, Aug. 2015
- Jiawei Han, "Construction of Structured Networks from Massive Text Data," **Invited Talk**, *Microsoft Research Asia*, Beijing, China, July 2015
- Jiawei Han, "From Data to Knowledge: Construction and Exploration of Heterogeneous Information Networks," **Invited Talk**, *ADL (Advanced Distinguished Lecture)*, Beijing, China, July 2015.
- Jiawei Han, "From Data to Knowledge: Construction and Exploration of Heterogeneous Information Networks," **Invited Talk**, *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences*, Beijing, China, July 2015
- Jiawei Han, "Construction and Mining of Heterogeneous Information Networks from Text Data," **Keynote Talk**, *Annual Meeting of the Association for Computational Linguistics (ACL'15)*, Beijing, China, July 2015
- Jiawei Han, "Mining Anomalies in Heterogeneous Networks: Query-Based, Static and Dynamic Network Outliers," **Invited Talk**, *Information Trust Inst., UIUC*, Jan. 2015
- Jiawei Han, "Construction and Mining of Semi-Structured Heterogeneous Information Networks," **Keynote Talk**, *Int. International School and Conference on Network Science (NetSci-x-2015)*, Rio de Janeiro, Brazil, Jan. 2015.
- Christos Faloutsos, "Mining Large Graphs," **Distinguished Seminar**, *Univ. of Illinois, Chicago, Distinguished Lecture Series*, April 2015
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Distinguished Seminar**, *Univ. of Southern California, Distinguished Lecture Series*, April 2015
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Distinguished Seminar**, *Arizona State University, Distinguished Guest Lecture Series*, Nov. 2015
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Distinguished Seminar**, *York University Distinguished Guest Lecture Series*, February, 2016
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Distinguished Seminar**, *George Mason University Distinguished Lecture Series*, March 2016
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Distinguished Seminar**, *Distinguished Lecture Series, UCSD*, Oct. 2016
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Distinguished Seminar**, *Distinguished Lecture Series, UCR*, April 2017
- Christos Faloutsos, "Mining Large Graphs: Patterns, Cascades, Fraud Detection, and Algorithms," **Keynote Talk**, *ICML'16, NYC*, June 2016
- Christos Faloutsos, Yasushi Sakurai, and Yasuko Matsubara, "Mining and Forecasting of Big Time-series data," **Tutorial**, *ACM SIGMOD, Melbourne, AU*, June 2015
- Christos Faloutsos, Alex Beutel, and Leman Akoglu, "Graph-Based User Behavior Modeling: From Prediction to Fraud Detection," **Tutorial**, *KDD 2015, Sydney, Australia*, Aug. 2015
- Yasushi Sakurai, Yasuko Matsubara, and Christos Faloutsos, "Mining Big Time-series Data on the Web," **Tutorial**, *WWW'16, Montreal, Canada*, April 2016
- Yasushi Sakurai, Yasuko Matsubara, and Christos Faloutsos, "Smart Analytics for Big Time-series Data," **Tutorial**, *KDD 2017, Halifax, Nova Scotia, Canada*, Aug. 13-17, 2017

- Meng Jiang, Srijan Kumar, VS Subrahmanian, and Christos Faloutsos, “Data-Driven Approaches towards Malicious Behavior Modeling,” **Tutorial**, *KDD 2017*, Halifax, Nova Scotia, Canada, Aug. 13-17, 2017
- Tina Eliassi-Rad, talks and seminars at the Computational Social Science Summer School, July 2017, International School and Conference on Network Science (NetSci), Machine Learning in Network Science Satellite, June 2017, International School on Network Science (NetSci), June 2017, New England Machine Learning (NEML) Day, May 2017, Harvard Medical School, Channing Network Science Seminar, February 2017, University of Chicago, Computational Social Science and Public Policy Colloquium, January 2017, Military Operations Research Society, Emerging Techniques Special Meeting (METSM), December 2016, GESIS Computational Social Science Winter Symposium, Cologne, Germany, November 2016, 8th International Conference on Social Informatics (SocInfo), Bellevue, WA, November 2016, MIT, Machine Learning Seminar, Cambridge, MA, October 2016, The Institute for Quantitative Social Science (IQSS), Applied Statistics Workshop, Cambridge, MA, October 2016, University of Massachusetts Amherst, Computational Social Science Institute, Amherst, MA, October 2016, Two Sigma Labs, New York, NY, July 2016, University of Antwerp’s Data Science Summit, Venice, Italy, September 2016, Conference on Complex Systems (CCS), Amsterdam, Netherlands, September 2016, UCLA IPAM Culture Analytics Culminating Workshop, Lake Arrowhead, CA, June 2016, UCLA IPAM Workshop III: Cultural Patterns: Multiscale Data-driven Models, Los Angeles, CA, May 2016,
- Tina Eliassi-Rad, IEEE ICDM **Panel** on “How We Can/Should Handle the Offline vs. Online Data Gap?”, Barcelona, Spain, December 2016
- Tina Eliassi-Rad, IEEE ICDM PhD Forum, “Top Ten List of Things That I’ve Learned Advising PhD Students,” Barcelona, Spain, December 2016
- Tina Eliassi-Rad, German Center for Research and Innovation (GCRI), Panel Moderator for “Big Data – Small Devices,” New York, NY, March 2016

Bibliography

1. Zhijun Yin, Liangliang Cao, Jiawei Han, Chengxiang Zhai, and Thomas Huang, “LPTA: A Probabilistic Model for Latent Periodic Topic Analysis,” *Proc. 2011 IEEE Int. Conf. on Data Mining (ICDM’11)*, Vancouver, Canada, Dec. 2011.
2. U Kang and Christos Faloutsos, “Beyond ‘Caveman Communities’: Hubs and Spokes for Graph Compression and Mining,” *IEEE International Conference on Data Mining (ICDM) 2011*, Vancouver, Canada.
3. Md Yusuf Sarwar Uddin, Hongyan Wang, Fatemeh Saremi, Guo-Jun Qi, Tarek Abdelzaher and Thomas Huang, “PhotoNet: A Similarity-aware Picture Delivery Service for Situation Awareness,” *IEEE Real-time Systems Symposium (RTSS)*, Vienna, Austria, December 2011.
4. Tao Wang, Jennifer Neville, Brian Gallagher, Tina Eliassi-Rad, “Correcting Bias in Statistical Tests for Network Classifier Evaluation,” *ECML/PKDD (3) 2011*: 506-521
5. Jennifer Neville, Brian Gallagher, Tina Eliassi-Rad, Tao Wang, “Correcting evaluation bias of relational classifiers with network cross validation,” *Knowl. Inf. Syst.* 30(1): 31-55 (2012) 2011
6. Leman Akoglu, Duen Horng Chau, U Kang, Danai Koutra, and Christos Faloutsos, “OPAvion: Mining and visualization in large graphs,” *ACM SIGMOD Conference 2012*, Scottsdale, AZ, USA.
7. Hongbo Deng, Jiawei Han, Michael R. Lyu and Irwin King, “Modeling and Exploiting Heterogeneous Bibliographic Networks for Expertise Ranking,” *Proc. 2012 ACM/IEEE Joint Conf. on Digital Libraries (JCDL’12)*, Washington, D.C., June 2012. (**Vannevar Bush Best Paper Award**)
8. Manish Gupta, Jing Gao, Yizhou Sun, and Jiawei Han, “Integrating Community Matching and Outlier Detection for Mining Evolutionary Community Outliers,” *Proc. of 2012 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD’12)*, Beijing, China, Aug. 2012
9. Manish Gupta, Jing Gao, Yizhou Sun, and Jiawei Han, “Community Trend Outlier Detection Using Soft Temporal Pattern Mining,” *Proc. of 2012 European Conf. on Machine Learning and*

Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD'12), Bristol, UK, Sept. 2012.

10. Mangesh Gupte, Tina Eliassi-Rad, "Measuring tie strength in implicit social network," *ACM Web Science* 2012.
11. Keith Henderson, Brian Gallagher, Tina Eliassi-Rad, Hanghang Tong, Sugato Basu, Leman Akoglu, Danai Koutra, Christos Faloutsos, Lei Li, "RoIX: Structural role extraction & mining in large graphs," *ACM KDD* 2012.
12. U Kang, Duen Horng Chau, and Christos Faloutsos, "PEGASUS: Mining Billion-Scale Graphs in the Cloud," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* 2012, Kyoto, Japan
13. U Kang, Evangelos Papalexakis, Abhay Harpale, and Christos Faloutsos, "GigaTensor: Scaling Tensor Analysis Up By 100 Times, Algorithms and Discoveries," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)* 2012, Beijing, China.
14. U Kang, Hanghang Tong, Jimeng Sun, Ching-Yung Lin, and Christos Faloutsos, "GBASE: An Efficient Analysis Platform for Large Graphs," *VLDB Journal*, 2012.
15. Danai Koutra, Evangelos E. Papalexakis, Christos Faloutsos, "TensorSplat: Spotting Latent Anomalies in Time," *16th Panhellenic Conference on Informatics with international participation (PCI)*, Piraeus, Greece, Oct. 2012.
16. Zhenhui Li, Jingjing Wang, and Jiawei Han, "Mining Periodicity for Sparse and Incomplete Event Data," *Proc. of 2012 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'12)*, Aug. 2012.
17. Lu Liu, Feida Zhu, Meng Jiang, Jiawei Han, Lifeng Sun, and Shiqiang Yang, "Mining diversity on social media networks," *Multimedia Tools and Applications*, 56(1): 179-205 (2012)
18. Evangelos E. Papalexakis, Alex Beutel, Peter Steenkiste, "Network Anomaly Detection using Co-clustering," *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM* 2012, Istanbul, Turkey.
19. Evangelos E. Papalexakis, Christos Faloutsos, Nikolaos D. Sidiropoulos, "ParCube: Sparse Parallelizable Tensor Decompositions," *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases 2012 (ECML-PKDD)*, Bristol, United Kingdom.
20. Eunsoo Seo, Prasant Mohapatra, Tarek F. Abdelzaher, "Identifying Rumors and their Sources in Social Networks," *SPIE Defense, Security, and Sensing*, Baltimore, Maryland, April 2012.
21. Mani Srivastava, Tarek Abdelzaher, Boleslaw K. Szymanski, "Human-centric Sensing," *Philosophical Transactions of the Royal Society, special issue on Wireless Sensor Networks*, Vol. 370, No. 1958, pp. 176-197, January 2012. (invited).
22. Mudhakar Srivatsa, Sihyung Lee, and Tarek Abdelzaher, "Mining Diverse Opinions," in *Proc. Military Communications Conference (MILCOM '12)*, Orlando, FL, October 2012.
23. Yizhou Sun, Charu C. Aggarwal, and Jiawei Han, "Relation Strength-Aware Clustering of Heterogeneous Information Networks with Incomplete Attributes", *Proc. 2012 Int. Conf. on Very Large Data Bases (VLDB'12/PVLDB)*, Istanbul, Turkey, Aug. 2012.
24. Yizhou Sun and Jiawei Han, "Mining on Heterogeneous Information Networks: Principles and Methodologies," Morgan & Claypool Publishers, 2012.
25. Yizhou Sun, Jiawei Han, Charu C. Aggarwal, and Nitesh Chawla, "When Will It Happen? Relationship Prediction in Heterogeneous Information Networks," *Proc. 2012 ACM Int. Conf. on Web Search and Data Mining (WSDM'12)*, Seattle, WA, Feb. 2012.
26. Yizhou Sun, Brandon Norick, Jiawei Han, Xifeng Yan, Philip S. Yu, and Xiao Yu, "Integrating Meta-Path Selection with User Guided Object Clustering in Heterogeneous Information Networks," *Proc. of 2012 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'12)*, Aug. 2012 (**Best Student Paper Award**)
27. Lu-An Tang, Quanquan Gu, Xiao Yu, Jiawei Han, Thomas La Porta, Alice Leung, Tarek Abdelzaher, and Lance Kaplan, "IntruMine: Mining Intruders in Untrustworthy Data of Cyber-Physical Systems," *Proc. 2012 SIAM Int. Conf. on Data Mining (SDM'12)*, Anaheim, CA, April 2012.
28. Lu-An Tang, Xiao Yu, Sangkyum Kim, Jiawei Han, Yizhou Sun, Wen-Chih Peng, Hector Gonzalez, Sebastian Seith, "Multidimensional Analysis of Atypical Events in Cyber-Physical Data," *Proc. 2012 IEEE Int. Conf. on Data Engineering (ICDE'12)*, Arlington, VA, Apr. 2012.

29. Hanghang Tong, Spiros Papadimitriou, Christos Faloutsos, Philip S. Yu, Tina Eliassi-Rad, "Gateway finder in large graphs: problem definitions and fast solutions," *Information Retrieval* 15(3-4): 391-411, 2012
30. Md Yusuf S Uddin, Md Tanvir Al Amin, Hieu Le, Tarek Abdelzaher, Boleslaw Szymanski, Tommy Nguyen, "On Diversifying Source Selection in Social Sensing," In Proc. *9th International Conference on Networked Sensing Systems*, Antwerp, Belgium, June 2012.
31. Chi Wang, Jiawei Han, Qi Li, Xiang Li, Wen-Pin Lin, and Heng Ji, "Learning Hierarchical Relationships among Partially Ordered Objects with Heterogeneous Attributes and Links," Proc. *2012 SIAM Int. Conf. on Data Mining (SDM'12)*, Anaheim, CA, April 2012.
32. Dong Wang, Lance Kaplan, Tarek Abdelzaher, Charu Aggarwal, "On Scalability and Robustness Limitations of Real and Asymptotic Confidence Bounds in Social Sensing," *9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Seoul, Korea, June 2012
33. Dong Wang, Hieu Le, Lance Kaplan, Tarek Abdelzaher, "On Truth Discovery in Social Sensing: A Maximum Likelihood Estimation Approach," In Proc. *11th ACM/IEEE Conference on Information Processing in Sensor Networks (IPSN)*, April 2012.
34. Jingjing Wang and Bhaskar Prabhala, "Periodicity Based Next Place Prediction," Proc. of *Workshop on Mobile Data Challenge* by Nokia, Newcastle, UK, June 2012. (**Nokia Challenge Competition: 2nd place on next location prediction**)
35. Bo Zhao, Benjamin I. P. Rubinstein, Jim Gemmell, and Jiawei Han, "A Bayesian Approach to Discovering Truth from Conflicting Sources for Data Integration," Proc. *2012 Int. Conf. on Very Large Data Bases (VLDB'12/PVLDB)*, Istanbul, Turkey, Aug. 2012.
36. Xiao Yu, Yizhou Sun, Brandon Norick, Tiancheng Mao, and Jiawei Han, "User Guided Entity Similarity Search Using Meta-Path Selection in Heterogeneous Information Networks," In Proc. of *2012 Int. Conf. on Information and Knowledge Management (CIKM'12)*, Maui, Hawaii, Oct. 2012, pp. 2025-2029.
37. H. Tong, B.A. Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos, "Gelling, and Melting, Large Graphs by Edge Manipulation," In Proc. of the *21st ACM CIKM Conference*, October 2012. (acceptance rate: 13%; *best interdisciplinary paper award*)
38. Hyungsul Kim, Yizhou Sun, Julia Hockenmaier, and Jiawei Han, "ETM: Entity Topic Models for Mining Documents Associated with Entities," In Proc. of *2012 IEEE Int. Conf. on Data Mining (ICDM'12)*, Brussels, Belgium, Dec. 2012, pp. 349-358. (selected as one of the best papers for possible publication in the journal "Knowledge and Information Systems")
39. Fatemeh Saremi, Praveen Jayachandran, Forrest Landola, Yusuf Sarwar, Tarek Abdelzaher, Aylin Yener, "On Schedulability and Time Composability of Multisensor Data Aggregation Networks," In Proc. *15th International Conference on Information Fusion*, Singapore, July 2012.
40. Quanquan Gu and Jiawei Han, "Experimental Design on Graphs via Generalization Error Bound Minimization," In Proc. of *2012 IEEE Int. Conf. on Data Mining (ICDM'12)*, Brussels, Belgium, Dec. 2012, pp. 882-887.
41. Quanquan Gu, Tong Zhang, Chris Ding, and Jiawei Han, "Selective Labeling via Error Bound Minimization," In Proc. of *2012 Neural Information Processing Systems Conf. (NIPS'12)*, Lake Tahoe, NV, Dec. 2012.
42. Manish Gupta, Jing Gao, Yizhou Sun, and Jiawei Han, "Community Trend Outlier Detection Using Soft Temporal Pattern Mining," In Proc. of *2012 European Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD'12)*, Bristol, UK, Sept. 2012, pp. 692-708.
43. K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu, L. Akoglu, D. Koutra, C. Faloutsos, and L. Li, "RoIX: Structural Role Extraction and Mining in Large Networks," In Proc. of the *18th ACM SIGKDD Conference*, August 2012.
44. Tarek Abdelzaher and Dong Wang, "Analytic Challenges in Social Sensing," *The Art of Wireless Sensor Networks*, Springer, (expected in) 2013.
45. M. Berlingerio, D. Koutra, T. Eliassi-Rad, C. Faloutsos, "Network Similarity via Multiple Social Theories," In Proc. of the *5th IEEE/ACM ASONAM Conference*, August 2013.
46. Marina Danilevsky, Chi Wang, Fangbo Tao, Son Nguyen, Gong Chen, Nihit Desai, and Jiawei Han, "AMETHYST: A System for Mining and Exploring Topical Hierarchies in Information Networks," (system demo) In Proc. of *2013 ACM SIGKDD Int. Conf. on Knowledge Discovery*

- and Data Mining (KDD'13), Chicago, IL, Aug. 2013.
47. Hongbo Deng, Jiawei Han, Hao Li, Heng Ji, Hongning Wang, and Yue Lu, "Exploring and Inferring User-User Pseudo-Friendship for Sentiment Analysis with Heterogeneous Networks," In *Proc. of 2013 SIAM Data Mining Conf. (SDM'13)*, Austin, TX, May 2013.
 48. William Dron, Alice Leung, Md Uddin, Shiguang Wang, Tarek Abdelzaher, Ramesh Govindan, John Hancock, "Information-maximizing Caching in Ad Hoc Networks with Named Data Networking," In *Proc. 2nd IEEE International Workshop on Network Science (NSW)*, West Point, NY, April, 2013
 49. S. Gilpin, T. Eliassi-Rad, I. Davidson, "Guided Learning for Role Discovery (GLRD): Framework, Algorithms, and Applications," In *Proc. of the 19th ACM SIGKDD Conference*, August 2013. (acceptance rate: 17.4%)
 50. Quanquan Gu, Charu Aggarwal, Jialu Liu, and Jiawei Han, "Selective Sampling on Graphs for Classification," In *Proc. of 2013 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, IL, Aug. 2013.
 51. Quanquan Gu, Charu Aggarwal and Jiawei Han, "Unsupervised Link Selection in Networks," In *Proc. 2013 Int. Conf. on Artificial Intelligence and Statistics (AISTAT'13)*, Scottsdale, AZ, Apr. 2013.
 52. Quanquan Gu and Jiawei Han, "Clustered Support Vector Machine," In *Proc. 2013 Int. Conf. on Artificial Intelligence and Statistics (AISTAT'13)*, Scottsdale, AZ, Apr. 2013.
 53. Manish Gupta, Jing Gao, Xifeng Yan, Hasan Cam, and Jiawei Han, "On Detecting Association-Based Clique Outliers in Heterogeneous Information Networks," In *Proc. of 2013 IEEE/ACM Int. Conf. on Social Networks Analysis and Mining (ASONAM'13)*, Niagara Falls, Canada, Aug. 2013
 54. H. Hang, X. Wei, M. Faloutsos, and T. Eliassi-Rad, "Entelecheia: Detecting P2P Botnets in their Waiting Stage," In *Proc. of the 12th IEEE IFIP Networking Conference*, May 2013.
 55. Danai Koutra, Joshua Vogelstein, Christos Faloutsos, "DeltaCon: A Principled Massive-Graph Similarity Function," In *Proc. SDM 2013*, Austin, Texas, May 2013.
 56. H. Liu, S. Gu, C. Pan, W. Zheng, S. Li, S. Hu, S. Wang, D. Wang, T. Amin, Z. Xie, R. Govindan, C. Aggarwal, A. Barnoy, T. Abdelzaher, "Metis: Inference-based Information Extrapolation from Participatory Sensing in Disaster Response Applications," Submitted to ACM Sensys 2013
 57. Jialu Liu, Chi Wang, Jing Gao, and Jiawei Han, "Multi-View Clustering via Joint Nonnegative Matrix Factorization," In *Proc. of 2013 SIAM Data Mining Conf. (SDM'13)*, Austin, TX, May 2013.
 58. Jialu Liu, Chi Wang, Marina Danilevsky, and Jiawei Han, "Large-Scale Spectral Clustering on Graphs," In *Proc. of 2013 Int. Joint Conf. on Artificial Intelligence (IJCAI'13)*, Beijing, China, August 2013.
 59. Yasuko Matsubara, Lei Li, Evangelos E. Papalexakis, David Lo, Yasushi Sakurai, Christos Faloutsos, "F-Trail: Finding Patterns in Taxi Trajectories," In *Proc. PAKDD 2013*, Gold Coast, Queensland, Australia, April 2013
 60. Misael Mongiovi, Petko Bogdanov, Razvan Ranca, Ambuj K. Singh, Evangelos E. Papalexakis, Christos Faloutsos, "NetSpot: Spotting Significant Anomalous Regions on Dynamic Networks," In *Proc. SDM 2013*, Austin, Texas, May 2013
 61. Evangelos E. Papalexakis, Leman Akoglu, Dino Ienco, "Do more Views of a Graph help? Community Detection and Clustering in Multi-Graphs," In *Proc. Fusion 2013*, Istanbul, Turkey, 2013
 62. Guo-Jun Qi, Charu C. Aggarwal, Jiawei Han, and Thomas Huang, "Mining Collective Intelligence in Groups," In *Proc. of 2013 Int. Conf. on World Wide Web (WWW'13)*, Rio de Janeiro, Brazil, May 2013.
 63. Lu-An Tang, Xiao Yu, Quanquan Gu, Jiawei Han, Alice Leung, and Thomas La Porta, "Mining Lines in the Sand: On Trajectory Discovery From Untrustworthy Data in Cyber-Physical System," In *Proc. of 2013 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, IL, Aug. 2013.
 64. Fangbo Tao, Kin Hou Lei, Jiawei Han, ChengXiang Zhai, Xiao Cheng, Marina Danilevsky, Nihit Desai, Bolin Ding, Jing Ge, Heng Ji, Rucha Kanade, Anne Kao, Qi Li, Yanen Li, Cindy Xide Lin, Jialiu liu, Nikunj Oza, Ashok Srivastava, Rod Tjoelker, Chi Wang, Duo Zhang, and Bo Zhao, "EventCube: Multi-Dimensional Search and Mining of Structured and Text Data," (system demo), In *Proc. of 2013 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, IL, Aug. 2013.

65. Md Yusuf S Uddin, Hossein Ahmadi, Tarek Abdelzaher, Robin Kravets. "Inter-contact Routing for Energy-constrained Disaster Response Networks," accepted to *IEEE Transaction on Mobile Computing*, 2013.
66. Dong Wang, Tarek Abdelzaher, Lance Kaplan, and Charu Aggarwal, "Recursive Fact-finding: A Streaming Approach to Truth Estimation in Crowdsourcing Applications," In *Proc. 33rd International Conference on Distributed Computing Systems (ICDCS)* Philadelphia, PA, July 2013.
67. Dong Wang, Md Tanvir Amin, Tarek Abdelzaher, Lance Kaplan, Charu C. Aggarwal, Shen Li, Siyu Gu, Hengchang Liu, Chenji Pan, XinLei Wang, Prasant Mohapatra, Raghu Ganti, Hieu Le, Boleslaw K. Szymanski, "Humans as Unreliable Binary Sensors with Uncertain Provenance: Model, Performance, and Limitations," Submitted to Sensys 2013.
68. Chi Wang, Marina Danilevsky, Nihit Desai, Yinan Zhang, Phuong Nguyen, Thrivikrama Taula, and Jiawei Han, "A Phrase Mining Framework for Recursive Construction of a Topical Hierarchy," In *Proc. of 2013 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, IL, Aug. 2013.
69. Dong Wang, Lance Kaplan, Tarek Abdelzaher, "Maximum Likelihood Analysis of Conflicting Observations in Social Sensing," *ACM Transactions on Sensor Networks*, Accepted in 2013.
70. Dong Wang, Lance Kaplan, Tarek Abdelzaher, and Charu Aggarwal, "On Credibility Estimation Tradeoffs in Assured Social Sensing," *IEEE Journal On Selected Areas in Communication (JSAC)*, Vol. 31, Issue 6, pp. 1026 - 1037, June 2013.
71. Hyungsul Kim, Xiang Ren, Yizhou Sun, Chi Wang, and Jiawei Han, "Semantic Frame-Based Document Representation for Comparable Corpora," *Proc. IEEE Int. Conf. on Data Mining (ICDM'13)*, Austin, TX, Dec. 2013.
72. Long T. Le, Tina Eliassi-Rad, Foster Provost, Lauren Moores, "Hyperlocal: Inferring Location of IP Addresses in Real-time Bid Requests for Mobile Ads," in *Proc. 6th ACM SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN'13)*, Orlando, FL, November 2013.
73. Scott Deean Chen, Ying-Yu Chen, Jiawei Han, and Pierre Moulin, "A Feature-Enhanced Ranking-Based Classifier for Multimodal Data and Heterogeneous Information Networks," *Proc. IEEE Int. Conf. on Data Mining (ICDM'13)*, Austin, TX, Dec. 2013.
74. Dong Wang, Tarek Abdelzaher, Lance Kaplan, Raghu Ganti, "Exploitation of Physical Constraints for Reliable Social Sensing," in *Proc. Real-Time Systems Symposium (RTSS)*, Vancouver, Canada, December 2013.
75. Shiguang Wang, Shaohan Lu, Shen Li, Hengchang Liu, Md Uddin, and Tarek Abdelzaher, "MINERVA: Information-Centric Programming for Social Sensing," In *Proc. International Conference on Computer Communications and Networks (ICCCN)*, Nassau, Bahamas, July 2013.
76. Chi Wang, Hongning Wang, Jialu Liu, Ming Ji, Lu Su, Yuguo Chen, Jiawei Han, "On the Detectability of Node Grouping in Networks," In *Proc. of 2013 SIAM Data Mining Conf. (SDM'13)*, Austin, TX, May 2013.
77. Tina Eliassi-Rad, "Social Order in Online Social Networks," In *Encyclopedia of Social Network Analysis and Mining (ESNAM)*, Eds: R. Alhajj and J. Rokne, Springer, 2013.
78. Sean Gilpin, Tina Eliassi-Rad, Ian Davidson, "Guided Learning for Role Discovery (GLRD): Framework, Algorithms, and Applications," *Proc. 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'13)*, Chicago, IL, August 2013.
79. Raghu Ganti, Mudhakar Srivatsa, Hengchang Liu, and Tarek Abdelzaher "Spatio-Temporal Spread of Events in Social Networks: A Gas Shortage Case Study," *Military Communications Conference (MILCOM)*, San Diego, CA, November 2013.
80. Michele Berlingerio, Danai Koutra, Tina Eliassi-Rad, Christos Faloutsos, "Network Similarity via Multiple Social Theories," *Proc. 5th IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'13)*, Niagara Falls, Canada, August 2013.
81. Md Tanvir Al Amin, Tarek Abdelzaher, Dong Wang, Boleslaw Szymanski, "Crowd-sensing with Polarized Sources," In *Proc. 10th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, Marina Del Rey, CA, May 2014.
82. Miguel Araujo, Spiros Papadimitriou, Stephan Gunnemann, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Evangelos Papalexakis, Danai Koutra, "Com2: Fast Automatic Discovery of

- Temporal (Comet) Communities,” In Proc. *PAKDD*, Tainan, Taiwan (Best student paper award - runner up), 2014.
83. Alex Beutel, Abhimanu Kumar, Evangelos E. Papalexakis, Partha Pratim Talukdar, Christos Faloutsos, Eric P. Xing, “FlexiFaCT: Scalable Flexible Factorization of Coupled Tensors on Hadoop,” In Proc. *SIAM International Conference on Data Mining (SDM)*, Philadelphia, PA, April 2014.
 84. Prasanna Giridhar, Md Tanvir Amin, Tarek Abdelzaher, Lance Kaplan, Jemin George, and Raghu Ganti, “ClariSense: Clarifying Sensor Anomalies using Social Network Feeds,” In Proc. *6th International Workshop on Information Quality and Quality of Service for Pervasive Computing (IQ2S)*, Budapest, Hungary, March 2014.
 85. Siyu Gu, Chenji Pan, Hengchang Liu, Shen Li, Shaohan Hu, Lu Su, Shiguang Wang, Dong Wang, Tanvir Amin, Ramesh Govindan, Charu Aggarwal, Raghu Ganti, Mudhakar Srivatsa, Amotz Barnoy, Peter Terlecky, Tarek Abdelzaher “Data Extrapolation in Social Sensing for Disaster Response,” In Proc. *10th IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, Marina Del Rey, CA, May 2014.
 86. Quanquan Gu, Tong Zhang, Jiawei Han, “Batch-Mode Active Learning via Error Bound Minimization,” Proc. *Conf. on Uncertainty in Artificial Intelligence (UAI)*, Quebec City, Quebec, Canada, July 2014.
 87. Manish Gupta, Jing Gao, Charu Aggarwal, and Jiawei Han, “Outlier Detection for Temporal Data,” *Morgan & Claypool Publishers*, 2014.
 88. Manish Gupta, Jing Gao, Charu C. Aggarwal, and Jiawei Han, “Outlier Detection for Temporal Data: A Survey,” Accepted by *IEEE Trans. on Knowledge and Data Engineering*, (to appear), 2014.
 89. Sue E. Kase, Elizabeth K. Bowman, Md Tanvir Al Amin, Tarek Abdelzaher, “Exploiting social media for Army operations: Syrian civil war use case,” *SPIE Defense, Security, and Sensing*, Baltimore, Maryland, May 2014.
 90. Ching-Hao Mao, Chung-Jung Wu, Kuo-Chen Lee, Evangelos E. Papalexakis, Christos Faloutsos, “MalSpot: Multi Malicious Network Behavior Patterns Analysis,” in Proc. *PAKDD*, Tainan, Taiwan, 2014.
 91. Xiang Ren, Jialu Liu, Xiao Yu, Urvashi Khandelwal, Quanquan Gu, Lidan Wang, and Jiawei Han, “ClusCite: Effective Citation Recommendation by Information Network-Based Clustering,” Proc. *ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'14)*, New York, NY, Aug. 2014.
 92. Xiang Ren, Yujing Wang, Xiao Yu, Jun Yan, Zheng Chen, and Jiawei Han, “Heterogeneous Graph-Based Intent Learning with Queries, Web Pages and Wikipedia Concepts,” Proc. *ACM Int. Conf. on Web Search and Data Mining (WSDM'14)*, New York City, NY, Feb. 2014.
 93. Wei Shen, Jiawei Han, and Jianyong Wang, “A Probabilistic Model for Linking Named Entities in Web Text with Heterogeneous Information Networks,” Proc. *ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'14)*, Snowbird, UT, June 2014.
 94. Sucheta Soundarajan, Tina Eliassi-Rad, Brian Gallagher, “A Guide to Selecting a Network Similarity Method,” in Proc. *SIAM International Conference on Data Mining (SDM)*, Philadelphia, PA, April 2014.
 95. Dong Wang, Tarek Abdelzaher, Lance Kaplan, “Surrogate Mobile Sensing,” *IEEE Communications Magazine*, Accepted in 2014.
 96. Dong Wang, Tarek Abdelzaher, Dan Roth, Clare Voss, Lance Kaplan, Stephen Tratz, Jamal Laoudi, Douglas Briesch, “Provenance-assisted Classification in Social Networks,” *IEEE Journal of Selected Topics in Signal Processing (J-STSP)*, Accepted in 2014.
 97. Dong Wang, Tanvir Amin, Shen Li, Tarek Abdelzaher, Lance Kaplan, Siyu Gu, Chenji Pan, Hengchang Liu, Charu Aggrawal, Raghu Ganti, XinLei Wang, Prasant Mohapatra, Boleslaw Szymanski, and Hieu Le, “Humans as Sensors: An Estimation Theoretic Perspective,” In Proc. *13th International Conference on Information Processing in Sensor Networks (ACM/IEEE IPSN)*, Berlin, Germany, April 2014.
 98. Dong Wang, Lance Kaplan, Tarek Abdelzaher, “Maximum Likelihood Analysis of Conflicting Observations in Social Sensing,” *ACM Transactions on Sensor Networks*, Volume 10 Issue 2, January 2014.

99. Chi Wang, Xueqing Liu, Yanglei Song, and Jiawei Han, "Scalable Moment-based Inference for Latent Dirichlet Allocation," Proc. *European Conf. on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECMLPKDD'14)*, Nancy, France, Sept. 2014.
100. Xiao Yu, Xiang Ren, Yizhou Sun, Quanquan Gu, Bradley Sturt, Urvashi Khandelwal, Brandon Norick, and Jiawei Han, "Personalized Entity Recommendation: A Heterogeneous Information Network Approach," Proc. *ACM Int. Conf. on Web Search and Data Mining (WSDM'14)*, New York City, NY, Feb. 2014.
101. Chao Zhang, Jiawei Han, Lidan Shou, Jiajun Lu, and Thomas La Porta, "Splitter: Mining Fine Grained Sequential Patterns in Semantic Trajectories," Proc. *Int. Conf. on Very Large Data Bases (VLDB'14)*, Hangzhou, China, Sept. 2014.
102. Manish Gupta, Jing Gao, Charu C. Aggarwal, and Jiawei Han, "Outlier Detection for Temporal Data: A Survey," *IEEE Trans. on Knowledge and Data Engineering*, 26(9):2250-2267, 2014.
103. Honglei Zhuang, Jing Zhang, George Brova, Jie Tang, Hasan Cam, Xifeng Yan and Jiawei Han, "Mining Query-Based Subnetwork Outliers in Heterogeneous Information Networks," in Proc. of 2014 *IEEE Int. Conf. on Data Mining (ICDM'14)*, Shenzhen, China, Dec. 2014.
104. Huan Gui, Yizhou Sun, Jiawei Han, and George Brova, "Modeling Topic Diffusion in Multi-Relational Bibliographic Information Networks," in Proc. of 2014 *ACM Int. Conf. on Information and Knowledge Management (CIKM'14)*, Shanghai, China, Nov. 2014.
105. Chi Wang, Jialu Liu, Nihit Desai, Marina Danilevsky, and Jiawei Han, "Constructing Topical Hierarchies in Heterogeneous Information Networks," *Knowledge and Information Systems (KAIS)*, 2014
106. Dian Yu, Hongzhao Huang, Taylor Cassidy, Heng Ji, Chi Wang, Shi Zhi, Jiawei Han and Clare Voss, "The Wisdom of Minority: Unsupervised Slot Filling Validation based on Multi-dimensional Truth-Finding with Multi-layer Linguistic Indicators," in Proc. of 2014 *Int. Conf. on Computational Linguistics (Coling'14)*, Dublin, Ireland, Aug. 2014
107. Chao Zhang, Jiawei Han, Lidan Shou, Jiajun Lu, and Thomas La Porta, "Splitter: Mining Fine Grained Sequential Patterns in Semantic Trajectories," Proc. of 2014 *Int. Conf. on Very Large Data Bases (VLDB'14)*, Hangzhou, China, Sept. 2014.
108. Md Tanvir Al Amin, Shen Li, Muntasir Raihan Rahman, Panindra Tumkur Seetharamu, Tarek Abdelzaher, Indranil Gupta, Mudhakar Srivatsa, Raghu Ganti, Reaz Ahmed, "SocialTrove: A Self-summarizing Storage Service for Social Sensing," In Proc. The 12th *IEEE International Conference on Autonomic Computing (ICAC)*, Grenoble, France, July 2015.
109. Prasanna Giridhar, Shiguang Wang, Tarek Abdelzaher, Jemin George, Lance Kaplan, Raghu Ganti, "Joint Localization of Events and Sources in Social Networks," In Proc. *International Conference on Distributed Computing in Sensor Systems (DCoSS)*, Fortaleza, Brazil, June 2015.
110. Shiguang Wang, Dong Wang, Lu Su, Lance Kaplan, Tarek Abdelzaher, "Towards Cyber-physical Systems in Social Spaces: The Data Reliability Challenge," In Proc. *IEEE Real-time Systems Symposium (RTSS)*, Rome, Italy, December 2014.
111. Tina Eliassi-Rad, "Social Order in Online Social Networks," *Encyclopedia of Social Network Analysis and Mining*, Springer, October 2014: 1918-1920.
112. Priya Govindan, Tina Eliassi-Rad, Jin Xu, Shawndra Hill, Chris Volinsky, "Threatening Privacy across Social Graphs: A Structural Features Approach," In Proc. of the 2014 *IEEE International Conference on Data Mining (ICDM) Workshops*, Shenzhen, China, December 2014.
113. Dong-Anh Nguyen, Tarek Abdelzaher, Xuan-Hong Dang, Raghu Ganti, Ambuj Singh, Mudhakar Srivatsa, "On Critical Event Observability using Social Networks: A Disaster Monitoring Perspective," *Military Communications Conference (MILCOM)*, Baltimore, MD, October 2014.
114. Dong Wang, Tarek Abdelzaher, Lance Kaplan, "Surrogate Mobile Sensing," *IEEE Communications Magazine*, Vol. 52, No. 8, August 2014.
115. Long T. Le, Tina Eliassi-Rad, "Measuring Coverage and Divergence of Reading Behaviors Among Friends," In Proc. the 1st *ACM SIGKDD Workshop on Data Science for News Publishing (NewsKDD'14)*, New York, NY, August 2014.
116. Priya Govindan, Sucheta Soundarajan, Tina Eliassi-Rad, "Finding the Most Appropriate Auxiliary Data for Social Graph De-anonymization," In Proc. the 1st *ACM SIGKDD Workshop on Data Ethics*, New York, NY, August 2014.
117. Xiang Ren, Jialu Liu, Xiao Yu, Urvashi Khandelwal, Quanquan Gu, Lidan Wang, and Jiawei Han, "ClusCite: Effective Citation Recommendation by Information Network-Based Clustering," Proc.

- 2014 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'14), New York, NY, Aug. 2014.
118. Miguel Araujo, Stephan Guennemann, Spiros Papadimitriou, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Evangelos Papalexakis, Danai Koutra, "Discovery of 'comet' communities in temporal and labeled graphs (Com2)" *Knowledge and Information Systems Journal*, Springer, May 2015.
 119. Evangelos E. Papalexakis, Christos Faloutsos, Nikolaos D. Sidiropoulos, "ParCube: Sparse Parallelizable CANDECOMP-PARAFAC Tensor Decomposition," *ACM Transactions on Knowledge Discovery from Data*, 2015
 120. Shaohan Hu, Shen Li, Shuochao Yao, Lu Su, Ramesh Govindan, Reginald Hobbs, Tarek F. Abdelzaher, "On Exploiting Logical Dependencies for Minimizing Additive Cost Metrics in Resource-Limited Crowdsensing," In Proc. *International Conference on Distributed Computing in Sensor Systems (DCoSS)*, Fortaleza, Brazil, June 2015.
 121. Shiguang Wang, Lu Su, Shen Li, Shuochao Yao, Shaohan Hu, Lance Kaplan, Tanvir Amin, Tarek Abdelzaher, Hongwei Wang, "Scalable Social Sensing of Interdependent Phenomena," In Proc. *14th ACM/IEEE Conference on Information Processing in Sensor Networks (IPSN)*, Seattle, WA, April 2015.
 122. Prasanna Giridhar, Tarek Abdelzaher, Jemin George, Lance Kaplan, "Physical Event Localization from Social Network Data," In Proc. *7th International Workshop on Information Quality and Quality of Service for Pervasive Computing (IQ2S)*, St. Louis, Missouri, USA, March 2015. (Invited)
 123. Senjuti Basu Roy, Tina Eliassi-Rad, Spiros Papadimitriou, "Fast Best-Effort Search on Graphs with Multiple Attributes," *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 27(3):755-768, 2015.
 124. Véronique Van Vlasselaer, Leman Akoglu, Tina Eliassi-Rad, Monique Snoeck, Bart Baesens: Guilt-by-Constellation, "Fraud Detection by Suspicious Clique Memberships," In Proc. of the *48th Annual Hawaii International Conference on System Sciences (HICSS'15)*, Kauai, HI, January 2015.
 125. Keith Henderson, Brian Gallagher, Tina Eliassi-Rad, "EP-MEANS: An Efficient Nonparametric Clustering of Empirical Probability Distributions," In Proc. of the *30th ACM SIGAPP Symposium On Applied Computing (SAC'15)*, Salamanca, Spain, April 2015.
 126. Long T. Le, Tina Eliassi-Rad, Hanghang Tong, "MET: A Fast Algorithm for Minimizing Propagation in Large Graphs with Small Eigen-Gaps," In Proc. of the *2015 SIAM International Conference on Data Mining (SDM'15)*, Vancouver, British Columbia, Canada, April 2015.
 127. Véronique Van Vlasselaer, Cristián Bravo, Olivier Caelen, Tina Eliassi-Rad, Leman Akoglu, Monique Snoeck, Bart Baesens, "APATE: A Novel Approach for Automated Credit Card Transaction Fraud Detection using Network-Based Extensions," *Decision Support Systems*, Elsevier, May 2015.
 128. Véronique Van Vlasselaer, Tina Eliassi-Rad, Leman Akoglu, Monique Snoeck and Bart Baesens, "AFRAID: Fraud Detection via Active Inference in Time-Evolving Social Networks," Proceedings of the *2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'15)*, Paris, France, August 2015 (to appear).
 129. Ahmed El-Kishky, Yanglei Song, Chi Wang, Clare R. Voss, Jiawei Han, "Scalable Topical Phrase Mining from Text Corpora", *PVLDB* 8(3): 305 - 316, 2015. (Also, in Proc. 2015 Int. Conf. on Very Large Data Bases (VLDB'15), Kohala Coast, Hawaii, Sept. 2015)
 130. Qi Li, Yaliang Li, Jing Gao, Lu Su, Bo Zhao, Murat Demirbas, Wei Fan, and Jiawei Han, "A Confidence-Aware Approach for Truth Discovery on Long-Tail Data," *PVLDB* 8(4): 425-436, 2015) (Also, in Proc. 2015 Int. Conf. on Very Large Data Bases (VLDB'15), KohalaCoast, Hawaii, Sept. 2015)
 131. Shi Zhi, Jiawei Han, and Quanquan Gu, "Robust Classification of Information Networks by Consistent Graph Learning," in Proc. of 2015 *European Conf. on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECMLPKDD'15)*, Porto, Portugal, Sept. 2015.
 132. Chao Zhang, Shan Jiang, Yucheng Chen, Yidan Sun, and Jiawei Han, "Fast Inbound Top-K Query for Random Walk with Restart," in Proc. of 2015 *European Conf. on Machine Learning*

- and Principles and Practices of Knowledge Discovery in Databases (ECMLPKDD'15)*, Porto, Portugal, Sept. 2015.
133. Chao Zhang, Yu Zheng, Xiuli Ma, Jiawei Han, "Assembler: Efficient Discovery of Spatial Coevolving Patterns in Massive Geosensory Data," in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
 134. Xiang Ren, Ahmed El-Kishky, Chi Wang, Fangbo Tao, Clare R. Voss, Heng Ji, Jiawei Han, "ClusType: Effective Entity Recognition and Typing by Relation Phrase-Based Clustering," in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
 135. Chaney, D. Blei, T. Eliassi-Rad. "A Probabilistic Model for Using Social Networks in Personalized Item Recommendation", *Proceedings of the 9th ACM Recommender Systems Conference (RecSys'15)*, Vienna, Austria, September 2015.
 136. V. Van Vlasselaer, T. Eliassi-Rad, L. Akoglu, M. Snoeck, B. Baesens. "AFRAID: Fraud Detection via Active Inference in Time-evolving Social Networks", *Proceedings of the 7th IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'15, Industrial Track)*, Paris, France, August 2015.
 137. L.T. Le, T. Eliassi-Rad, H. Tong. "MET: A Fast Algorithm for Minimizing Propagation in Large Graphs with Small Eigen-Gaps", *Proceedings of the 2015 SIAM International Conference on Data Mining (SDM'15)*, Vancouver, British Columbia, Canada, April 2015.
 138. K. Henderson, B. Gallagher, T. Eliassi-Rad. "EP-MEANS: An Efficient Nonparametric Clustering of Empirical Probability Distributions", *Proceedings of the 30th ACM SIGAPP Symposium on Applied Computing (SAC'15)*, Salamanca, Spain, April 2015.
 139. V. Van Vlasselaer, L. Akoglu, T. Eliassi-Rad, M. Snoeck, B. Baesens. "Guilt-by-Constellation: Fraud Detection by Suspicious Clique Memberships", *Proceedings of the 48th Annual Hawaii International Conference on System Sciences (HICSS'15)*, Kauai, HI, January 2015.
 140. Fenglong Ma, Yaliang Li, Qi Li, Minghui Qui, Jing Gao, Shi Zhi, Lu Su, Bo Zhao, Heng Ji, and Jiawei Han, "FaitCrowd: Fine Grained Truth Discovery for Crowdsourced Data Aggregation," in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
 141. Jingjing Wang, Wenzhu Tong, Hongkun Yu, Min Li, Xiuli Ma, Haoyan Cai, Tim Hanratty, and Jiawei Han, "Mining Multi-Aspect Reflection of News Events in Twitter: Discovery, Linking and Presentation", in Proc. of 2015 IEEE Int. Conf. on Data Mining (ICDM'15), Atlantic City, NJ, Nov. 2015
 142. Ahmed El-Kishky, Yanglei Song, Chi Wang, Clare R. Voss, and Jiawei Han, "Scalable Topical Phrase Mining from Text Corpora", *PVLDB 8(3)*: 305 - 316, 2015. Also, in Proc. 2015 Int. Conf. on Very Large Data Bases (VLDB'15), Kohala Coast, Hawaii, Sept. 2015.
 143. Shi Zhi, Jiawei Han, and Quanquan Gu, "Robust Classification of Information Networks by Consistent Graph Learning", in Proc. of 2015 European Conf. on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECMLPKDD'15), Porto, Portugal, Sept. 2015.
 144. Chao Zhang, Shan Jiang, Yucheng Chen, Yidan Sun, and Jiawei Han, "Fast Inbound Top-K Query for Random Walk with Restart", in Proc. of 2015 European Conf. on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECMLPKDD'15), Porto, Portugal, Sept. 2015. **(Received the Best Student Paper Runner-Up award at ECML/PKDD 2015)**
 145. Chao Zhang, Yu Zheng, Xiuli Ma, Jiawei Han, "Assembler: Efficient Discovery of Spatial Coevolving Patterns in Massive Geosensory Data", in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
 146. Xiang Ren, Ahmed El-Kishky, Chi Wang, Fangbo Tao, Clare R. Voss, Heng Ji, Jiawei Han, "ClusType: Effective Entity Recognition and Typing by Relation Phrase-Based Clustering", in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
 147. Chi Wang, Xueqing Liu, Yanglei Song, Jiawei Han, "Towards Interactive Construction of Topical Hierarchy: A Recursive Tensor Decomposition Approach", in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015

148. Shi Zhi, Bo Zhao, Wenzhu Tong, Jing Gao, Dian Yu, Heng Ji, and Jiawei Han, "Modeling Truth Existence in Truth Discovery", in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
149. Honglei Zhuang, Aditya Parameswaran, Dan Roth, Jiawei Han, "Debiasing Crowdsourced Batches", in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
150. Xiang Ren, Ahmed El-Kishky, Chi Wang, and Jiawei Han, "Automatic Entity Recognition and Typing from Massive Text Corpora: A Phrase and Network Mining Approach" (conference tutorial), 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
151. Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, Jiawei Han, "Mining Quality Phrases from Massive Text Corpora", in Proc. of 2015 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'15), Melbourne, Australia, May 2015 (**won Grand Prize in Yelp Dataset Challenge, 2015**)
152. Fangbo Tao, Bo Zhao, Ariel Fuxman, Yang Li, Jiawei Han, "Leveraging Pattern Semantics for Constructing Entity Taxonomies in Enterprises", in Proc. of 2015 Int. Conf. on World-Wide Web (WWW'15), Florence, Italy, May 2015
153. Huan Gui, Ya Xu, Anmol Bhasin, Jiawei Han, "Network A/B Testing: From Sampling to Estimation", in Proc. of 2015 Int. Conf. on World-Wide Web (WWW'15), Florence, Italy, May 2015
154. Jialu Liu, Chi Wang, Jing Gao, Quanquan Gu, Charu Aggarwal, Lance Kaplan, and Jiawei Han, "GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data", in Proc. of 2015 SIAM Int. Conf. on Data Mining (SDM'15), Vancouver, Canada, Apr. 2015 (**selected as one of the best papers in the conference and invited to journal Statistical Analysis and Data Mining (SADM) special issue "Best of SDM 2015"**)
155. Mengting Wan, Yunbo Ouyang, Lance Kaplan, Jiawei Han, "Graph Regularized Meta-path Based Transductive Regression in Heterogeneous Information Network", in Proc. of 2015 SIAM Int. Conf. on Data Mining (SDM'15), Vancouver, Canada, Apr. 2015
156. Wei Feng, Chao Zhang, Wei Zhang, Jiawei Han, Jianyong Wang, Charu Aggarwal, Jianbin Huang, "STREAMCUBE: Hierarchical Spatio-temporal Hashtag Clustering for Event Exploration over the Twitter Stream", in Proc of 2015 IEEE Int. Conf on Data Engineering (ICDE'15), Seoul, Korea, Apr. 2015
157. Jonathan Kuck, Honglei Zhuang, Xifeng Yan, Hasan Cam, and Jiawei Han, "Query-Based Outlier Detection in Heterogeneous Information Networks", in Proc. of 2015 Int. Conf. on Extending Database Technology (EDBT'15), Brussels, Belgium, Mar. 2015 (**selected as one of the best papers in the conference and invited to journal ACM Transactions on Database Systems (TODS) as Best of EDBT 2015**)
158. Jialu Liu, Charu Aggarwal and Jiawei Han, "On Integrating Network and Community Discovery", in Proc. of 2015 Int. Conf. on Web Search and Data Mining (WSDM'15), Shanghai, China, Feb. 2015, pp.117-126
159. Lu-An Tang, Xiao Yu, Quanquan Gu, Jiawei Han, Guofei Jiang, Alice Leung, and Thomas La Porta, "A Framework of Mining Trajectories from Untrustworthy Data in Cyber-Physical System", ACM Transactions on Knowledge Discovery from Data (TKDD), 9(3):16:1-16:35, 2015
160. Wei Shen, Jianyong Wang, Jiawei Han, "Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions", IEEE Transactions on Knowledge and Data Engineering (TKDE), 27(2): 443-460, 2015
161. Yong Yang, Lu Su, Mohammad Maifi Hasan Khan, Michael LeMay, Tarek F. Abdelzaher, Jiawei Han, "Power-Based Diagnosis of Node Silence in Remote High-End Sensing Systems", ACM Transactions on Sensor Networks (TOSN), 11(2): 33:1-33:33, 2015
162. Chi Wang and Jiawei Han, Mining Latent Entity Structures, Synthesis Lectures on Data Mining and Knowledge Discovery, Morgan & Claypool Publishers, 2015.
163. Chi Wang, Xueqing Liu, Yanglei Song, Jiawei Han, "Towards Interactive Construction of Topical Hierarchy: A Recursive Tensor Decomposition Approach," in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015
164. Yaliang Li, Qi Li, Jing Gao, Lu Su, Bo Zhao, Wei Fan, and Jiawei Han, "On the Discovery of Evolving Truth," in Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, Aug. 2015

165. Chenguang Wang, Yangqiu Song , Ahmed El-Kishky, Dan Roth, Ming Zhang, Jiawei Han, "Incorporating World Knowledge to Document Clustering via Heterogeneous Information Networks," in Proc. of *2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15)*, Sydney, Australia, Aug. 2015
166. Honglei Zhuang, Aditya Parameswaran, Dan Roth, Jiawei Han, "Debiasing Crowdsourced Batches," in Proc. of *2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15)*, Sydney, Australia, Aug. 2015
167. Chenguang Wang, Yangqiu Song, Dan Roth, Chi Wang, Jiawei Han, Heng Ji, and Ming Zhang, "Constrained Information-Theoretic Tripartite Graph Clustering to Identify Semantically Similar Relations," in Proc. *2015 Int. Joint Conf. on Artificial Intelligence (IJCAI'15)*, Buenos Aires, Argentina, July 2015
168. C. Bravo, V. Van Vlasselaer, O. Caelen, T. Eliassi-Rad, L. Akoglu, M. Snoeck, B. Baesens. "APATE: A Novel Approach for Automated Credit Card Transaction Fraud Detection Using Network-based Extensions", *Decision Support Systems*, 75: 38-48, 2015.
169. S. Basu Roy, T. Eliassi-Rad, S. Papadimitriou. "Fast Best-Effort Search on Graphs with Multiple Attributes", *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 27(3): 755-768, 2015.
170. Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, Jiawei Han, "Mining Quality Phrases from Massive Text Corpora," in Proc. of *2015 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'15)*, Melbourne, Australia, May 2015 (won Grand Prize in Yelp Dataset Challenge, 2015)
171. Mengting Wan, Yunbo Ouyang, Lance Kaplan, Jiawei Han, "Graph Regularized Meta-path Based Transductive Regression in Heterogeneous Information Network," in Proc. of *2015 SIAM Int. Conf. on Data Mining (SDM'15)*, Vancouver, Canada, Apr. 2015
172. Jialu Liu, Chi Wang, Jing Gao, Quanquan Gu, Charu Aggarwal, Lance Kaplan, and Jiawei Han, "GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data," in Proc. of *2015 SIAM Int. Conf. on Data Mining (SDM'15)*, Vancouver, Canada, Apr. 2015 (selected as one of the best papers in the conference and invited to journal Statistical Analysis and Data Mining (SADM) special issue "Best of SDM 2015")
173. Jonathan Kuck, Honglei Zhuang, Xifeng Yan, Hasan Cam, and Jiawei Han, "Query-Based Outlier Detection in Heterogeneous Information Networks," in Proc. of *2015 Int. Conf. on Extending Database Technology (EDBT'15)*, Brussels, Belgium, Mar. 2015 (selected as one of the best papers in the conference and invited to journal ACM Transactions on Database Systems (TODS) as Best of EDBT 2015)
174. Jialu Liu, Charu Aggarwal and Jiawei Han, "On Integrating Network and Community Discovery," in Proc. of *2015 Int. Conf. on Web Search and Data Mining (WSDM'15)*, Shanghai, China, Feb. 2015, pp.117-126
175. Wei Shen, Jianyong Wang, Jiawei Han, "Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions," *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 27(2): 443-460 (2015)
176. Huan Gui, Jialu Liu, Fangbo Tao, Meng Jiang, Brandon T. Norick, and Jiawei Han, "Large-Scale Embedding Learning in Heterogeneous Event Data", in Proc. of *2016 Int. Conf. on Data Mining (ICDM'16)*, Barcelona, Spain, Dec. 2016.
177. Xiang Ren, Wenqi He, Meng Qu, Lifu Huang, Heng Ji, and Jiawei Han, "AFET: Automatic Fine-Grained Entity Typing by Hierarchical Partial-Label Embedding", in Proc. of *2016 Conf. on Empirical Methods in Natural Language Processing (EMNLP'16)*, Austin, TX, Nov. 2016
178. Fangbo Tao, Honglei Zhuang, Chi Wang Yu, Qi Wang, Taylor Cassidy, Lance Kaplan, Clare Voss, Jiawei Han, "Multi-Dimensional, Phrase-Based Summarization in Text Cubes", *Data Eng. Bulletin* 39(3), Sept. 2016, pp. 74-84.
179. Chao Zhang, Keyang Zhang Quan Yuan, Luming Zhang, Tim Hanratty, and Jiawei Han, "GMove: Group-Level Mobility Modeling Using Geo-Tagged Social Media", in Proc. of *2016 ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD'16)*, San Francisco, CA, Aug. 2016
180. Meng Jiang, Christos Faloutsos, Jiawei Han, "CatchTartan: Representing and Summarizing Dynamic Multicontextual Behaviors", in Proc. of *2016 ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD'16)*, San Francisco, CA, Aug. 2016

181. Xiang Ren, Wenqi He, Meng Qu, Clare R. Voss, Heng Ji, Jiawei Han, "Label Noise Reduction in Entity Typing by Heterogeneous Partial-Label Embedding", in Proc. of 2016 ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD'16), San Francisco, CA, Aug. 2016
182. Mengting Wan, Xiangyu Chen, Lance Kaplan, Jiawei Han, Jing Gao, Bo Zhao, "From Truth Discovery to Trustworthy Opinion Discovery: An Uncertainty-Aware Quantitative Modeling Approach", in Proc. of 2016 ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD'16), San Francisco, CA, Aug. 2016
183. Chao Zhang, Guangyu Zhou, Quan Yuan, Honglei Zhuang, Yu Zheng, Lance Kaplan, Shaowen Wang, Jiawei Han, "GeoBurst: Real-time Local Event Detection in Geo-tagged Tweet Stream", in Proc. of 2016 ACM SIGIR Conf. on Research & Development in Information Retrieval (SIGIR'16), Pisa, Italy, July 2016
184. Jingjing Wang, Changsung Kang, Yi Chang and Jiawei Han, "Learning Hostname Preference to Enhance Search Relevance", in Proc. of 2016 Int. Joint Conf. on Artificial Intelligence (IJCAI'16), New York City, NY, July 2016
185. Huan Gui, Jiawei Han, Quanquan Gu, "Towards Faster Rates and Oracle Property for Low-Rank Matrix Estimation", in Proc. of 2016 Int. Conf. on Machine Learning (ICML'16), New York City, NY, June 2016.
186. Xiang Ren, Ahmed El-Kishky, Heng Ji, and Jiawei Han, "Automatic Entity Recognition and Typing in Massive Text Data", (Conference Tutorial) 2016 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'16), San Francisco, CA, June 2016
187. Jingbo Shang, Wenzhu Tong, Jian Peng, and Jiawei Han, "DPClass: An Effective but Concise Discriminative Patterns-Based Classification Framework", in Proc of 2016 SIAM Int. Conf. on Data Mining (SDM'16), Miami, FL, May 2016
188. Jingbo Shang, Jian Peng, and Jiawei Han, "MACFP: Maximal Approximate Consecutive Frequent Pattern Mining under Edit Distance", in Proc of 2016 SIAM Int. Conf. on Data Mining (SDM'16), Miami, FL, May 2016
189. Jialu Liu, Xiang Ren, Jingbo Shang, Taylor Cassidy, Clare Voss and Jiawei Han, "Representing Documents via Latent Keyphrase Inference", in Proc. of 2016 Int. World-Wide Web Conf. (WWW'16), Montreal, Canada, April 2016
190. Min Li, Jingjing Wang, Wenzhu Tong, Hongkun Yu, Xiuli Ma, Yucheng Chen, Haoyan Cai, Jiawei Han, "EKNOT: Event Knowledge from News and Opinions in Twitter", Proc. of AAAI Conf. on Artificial Intelligence (AAAI'16) (system demo), Phoenix, AZ, Feb. 2016
191. Jingjing Wang, Min Li, Jiawei Han and Xiaolong Wang, "Modeling Check-in Preferences with Multidimensional Knowledge: A Minimax Entropy Approach", in Proc. of 2016 Int. Conf. on Web Search and Data Mining (WSDM'16). San Francisco, CA. Feb. 2016.
192. V. Van Vlasselaer, T. Eliassi-Rad, L. Akoglu, M. Snoeck, B. Baesens. "GOTCHA! Network-based Fraud Detection for Social Security Fraud", *Management Science* (INFORMS), pp. 1-21, July 2016.
193. C. Chen, H. Tong, B.A. Prakash, T. Eliassi-Rad, M. Faloutsos, C. Faloutsos. "Eigen-optimization on Large Graphs by Edge Manipulation", *ACM Transactions on Knowledge Discovery in Data* (TKDD), 10(4), Article 49, June 2016.
194. C. Chen, H. Tong, B.A. Prakash, C.E. Tsourakakis, T. Eliassi-Rad, C. Faloutsos, D.H. Chau. "Node Immunization on Large Graphs: Theory and Algorithms", *IEEE Transactions on Knowledge and Data Engineering* (TKDE), 28(1): 113-126, 2016.
195. K. Shin, T. Eliassi-Rad, C. Faloutsos: "Patterns and Anomalies in k-Cores of Real-World Graphs with Applications", *Knowledge and Information Systems* (KAIS), 2017 (to appear).
196. S. Soundarajan, T. Eliassi-Rad, B. Gallagher, A. Pinar. epsilon-WGX: Adaptive Edge Probing for Enhancing Incomplete Networks. *Proceedings of the 9th International ACM Web Science Conference* (WebSci'17), Troy, NY, June 2017
197. K. Shin, T. Eliassi-Rad, C. Faloutsos. "CoreScope: Graph Mining Using k-Core Analysis--Patterns, Anomalies, and Algorithms", *Proceedings of the 16th IEEE International Conference on Data Mining* (ICDM'16), Barcelona, Spain, December 2016. (**Invited for fast-track journal publication as among best in conference**)
198. S. Soundarajan, T. Eliassi-Rad, B. Gallagher, A. Pinar. "MaxReach: Reducing Network Incompleteness through Node Probes", *Proceedings of the 2016 IEEE/ACM International*

- Conference on Advances in Social Networks Analysis and Mining (ASONAM'16)*, San Francisco, CA, August 2016.
199. P. Govindan, S. Soundarajan, T. Eliassi-Rad, C. Faloutsos. "NimbleCore: A Space-efficient External Memory Algorithm for Estimating Core Numbers", *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'16)*, San Francisco, CA, August 2016.
 200. S. Basu Roy, T. Eliassi-Rad, S. Papadimitriou. "Fast Best-Effort Search on Graphs with Multiple Attributes", *Proceedings of the 32nd IEEE International Conference on Data Engineering (ICDE'16)*, TKDE Poster Track, Helsinki, Finland, May 2016.
 201. S. Soundarajan, A. Tamersoy, E.B. Khalil, T. Eliassi-Rad, D.H. Chau, B. Gallagher, K. Roundy. "Generating Graph Snapshots from Streaming Edge Data", *Proceedings of the 25th International World Wide Web Conference (WWW'16)*, Poster Track, Montreal, Canada, April 2016.
 202. Chao Zhang, Keyang Zhang, Quan Yuan, Fangbo Tao, Luming Zhang, Tim Hanratty, Jiawei Han, "ReAct: Online Multimodal Embedding for Recency-Aware Spatiotemporal Activity Modeling", In Proc. of 2017 ACM SIGIR Conf. on Research & Development in Information Retrieval (SIGIR'17), Tokyo, Japan, Aug. 2017
 203. Chao Zhang, Dongming Lei, Quan Yuan, Honglei Zhuang, Lance Kaplan, Shaowen Wang, Jiawei Han, "GeoBurst+: Effective and Real-Time Local Event Detection in Geo-Tagged Tweet Streams", accepted by ACM Transactions on Intelligent Systems and Technology (ACM TIST), 2017
 204. Jialu Liu, Jingbo Shang, and Jiawei Han, *Phrase Mining from Massive Text and Its Applications, Synthesis Lectures on Data Mining and Knowledge Discovery*, Morgan & Claypool Publishers, 2017.
 205. Xiang Ren, Zeqiu Wu, Wenqi He, Meng Qu, Clare Voss, Heng Ji, Tarek Abdelzaher and Jiawei Han, "CoType: Joint Extraction of Typed Entities and Relations with Knowledge Bases", in Proc. of 2017 World-Wide Web Conf. (WWW'17), Perth, Australia, Apr. 2017.
 206. Chao Zhang, Keyang Zhang, Quan Yuan, Haoruo Peng, Yu Zheng, Tim Hanratty, Shaowen Wang and Jiawei Han, "Regions, Periods, Activities: Uncovering Urban Dynamics via Cross-Modal Representation Learning", in Proc. of 2017 World-Wide Web Conf. (WWW'17), Perth, Australia, Apr. 2017.
 207. Xiang Ren, Yuanhua Lv, Kuansan Wang and Jiawei Han, "Comparative Document Analysis for Large Text Corpora", in Proc. of 2017 ACM Int. Conf. on Web Search and Data Mining (WSDM'17), Cambridge UK, Feb. 2017