sponsored by ARL

Network Science Collaborative Technology Alliance Newsletter

Welcome

Thank you for your interest in the U.S. Army Research Laboratory (ARL) **Network Science Collaborative Technology Alliance** (NS CTA). The NS CTA is a collaborative research alliance between ARL, other government researchers, and a consortium of 17 university and industry research organizations. The program began on 28 September, 2009, and was renewed for a second 5-year phase that began 28 September, 2014. The NS CTA unites researchers across organizations, technical disciplines, and research areas to address the critical technical challenges posed by the complex web of interacting networks within which the Army mission must be performed.

Our goal is to perform foundational cross-disciplinary research in network science, enabling greatly enhanced soldier performance and greatly enhanced speed and precision for complex military operations.

Network science is the study of the properties, models, and theories that apply to many varieties of networks, the understanding of how different genres of networks dynamically interact and co-evolve, and the use of this understanding in the analysis, prediction, design, and control of many varieties and systems of networks. To support this ambitious research program, the NS CTA has created shared, distributed, experimental resources throughout the Alliance as well as a distributed network science research facility, led from the ARL Network Science Research Laboratory (NSRL) in the ARL Adelphi Laboratory Center (ALC), MD. The NS CTA's collaborative research also makes extensive use of the ARL Open Campus facilities to support researcher rotations from throughout the consortium to ARL.

The NS CTA research program for Phase 2 is structured as five interdisciplinary technical areas (or *thrusts*), each of which unites Alliance expertise in social/cognitive networks, information networks, and communication networks. These areas are Co-evolution and Dynamics of Inter-genre Networks (Co-EDIN); Information Processing Across Networks for Decision-Making (IPAN); Quality of Information for Semantically-Adaptive Networks (QoI-SAN); Trust, Influencing, Modeling & Enhancing Human Performance (TIME); and the Science of Multi-genre Network Experimentation (Exp). The 22 current research tasks that comprise these five areas involve the contributions of 73 primary researchers, including 20 government researchers, 13 senior industry researchers, and 40 faculty members.

For this first newsletter of Phase 2, we have invited leads of our five technical areas to introduce their areas to our newsletter readers; future newsletters will highlight specific insights and achievements as they arise throughout the program.

Please see the program website **http://www.ns-cta.org/** for more information about the research efforts and current news and events.

This newsletter provides members of the public a quick sampling of what is happening around the NS CTA. These newsletters are archived on the NS CTA program website.

To subscribe or unsubscribe to our newsletter, please contact us at ns-cta-blog@ns-cta.org.

IEEE Fellowship

Prof. Aylin Yener of PSU has been named Fellow by the Institute of Electrical and Electronics Engineers (IEEE), as of January 1, 2015 for contributions to wireless communication theory and wireless information security. IEEE elevates less than one-tenth of one percent of its voting membership to Fellow status. Aylin Yener also has been elected to the Board of Governors of the IEEE Information Theory Society for the term 2015-2017.

News and tips on NS CTA Rotations at ALC

By John P. Hancock

Rotations and visits are an exciting topic at ARL this year. It dovetails into the substantial and fast moving Open Campus initiative that is now well underway. I am in a long-term, part-time rotation at ARL Adelphi Laboratory Complex (ALC), visiting ALC every Wednesday since August of 2014. In this article, I summarize the news and events and include some personal notes and practical tips that come with rotation experience.

ARL has a big vision for building an Open Campus (link: http://www.arl.army.mil/opencampus/). Open Campus is a new approach to facilitate in-person collaboration between ARL researchers and academic and industry partners in a laboratory environment. On December 9-10, 2014, an open house for Open Campus publicized the plans and vision. The NS CTA is an important part of Open Campus plans and a new Network Science Research Lab (NSRL) that embraces the Open Campus principles will be built out this year. The NSRL will offer collaboration, computation, and analysis motivations for researchers to visit and engage in rotations for network science.

I have been working with ARL in various ways since the late 90's, but the new Open Campus has brought new opportunities. In 2013, I observed Tarek Abdelzaher over a couple of weeks at ARL, which culminated in a seminar at ALC. He reprised this seminar at APG. I was impressed at the synergy and interaction that this rotation created; later in the summer of 2014 when ARL called for rotation candidates, Tarek's experience inspired me to volunteer for a longer term rotation that suited on-going experimentation projects.

Tarek had to sit at a table in a borrowed office and use telephone-based internet to get external connectivity. I can report a much improved visitor experience in my rotation. Brian Riviera and a few others were able to put me in an office, with a phone and Wi-Fi internet access, rapidly in my first rotation visit. The amazing thing was how this improved my experience at ARL.

ARL researchers are busy people who have multiple priorities; when I visit, I do not expect them to spend all day with me. I, also, have other responsibilities that I need to address daily. The availability of an office, desk, phone, and internet connection that I can use as needed for NS CTA and other business makes working effectively at ARL possible for me. In the past, I used to pass down-time in conference rooms, the corner of someone's office, or even the cafeteria; now I have a work-home at ARL that is useful beyond my regular rotation days—facilitating visit interactions with BBN and other ArtisTech staff working on integrated experimentation.

A lively collaboration between Dr. David Alberts of IDA, the NSRL and other X2 task researchers is one product of this rotation. Regular, in-person discussions have accelerated each phase of the project, from envisioning a campaign of experiments across ARL programs, to converging on an interface between network genre components, to planning specific experiments, to debugging irregularities in our multi-network simulation.

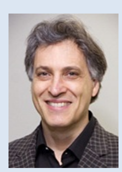
Focused, hands-on multi-month collaboration with Kelvin Marcus (ARL) and others for integration, deployment, and analysis of experiments facilitated another X2 effort that investigated how different virtualization schemes enable scale-up of different types of network interactions. We could not have accomplished the rapid integration and experimentation without the extended and recurring interaction that rotation made possible.

Just last week, the ARL NSRL team moved into the 2nd floor Open Campus space, so now we share that workspace home. In the Open Campus space, even researchers with escort-required badges can now walk down the hall, access team meeting spaces, and visit NSRL staff offices on the same floor. I am eagerly anticipating the new NSRL, which will be on the other side of the building later this year. The lab should do for hands-on experimentation collaboration what the office-side of Open Campus has done for us with planning and meeting collaboration.

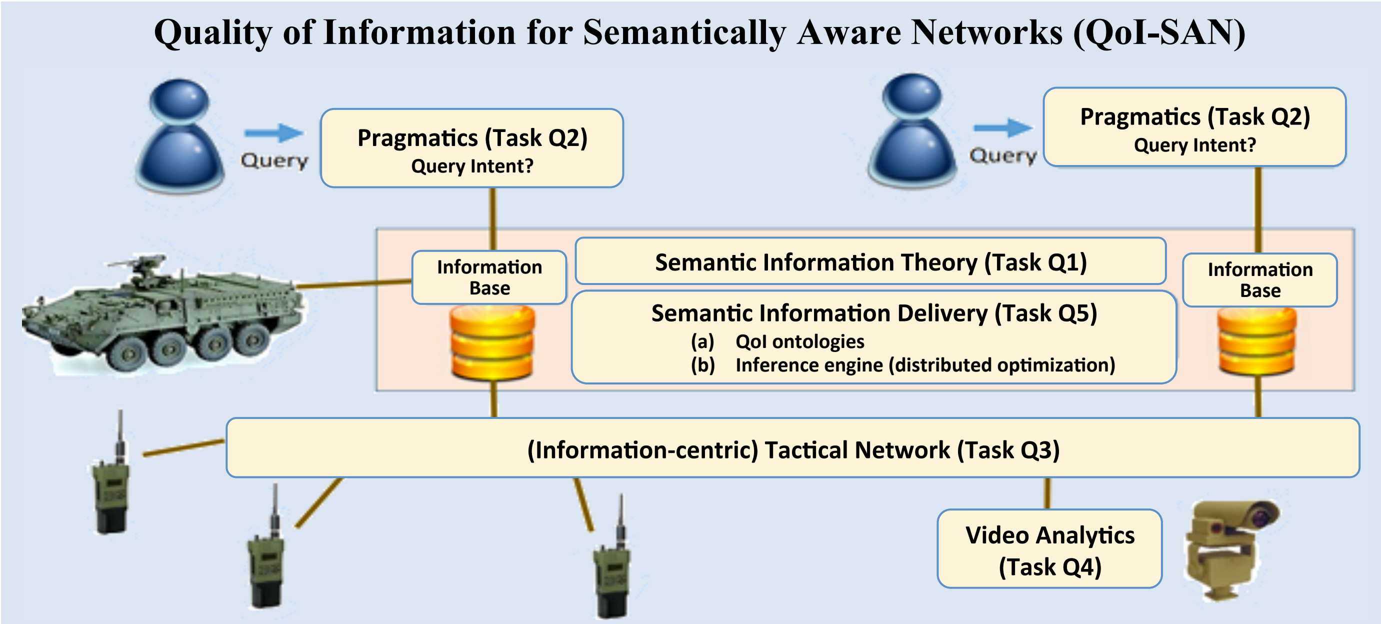
I don't have space here to fully describe the new NSRL, but let me preview some highlights. The new lab space will have a brand new set of advanced servers tied together with an upgraded virtual experiment environment. This environment offers NS CTA researchers access to user-specified configurations of virtualized server components and data. At first, this access will probably be limited to local access in the lab at ALC, but later there may be remote access as well. Laboratory terminals will be able to access the server farm and collaborative workspace to work with teams that include ARL researchers. New analytic, data, and unique NS simulation tools are also in the works and will become increasingly important as the NS CTA matures and our research produces transitionable products.

An exciting new level of collaboration between academic, industry, and ARL researchers is emerging with infrastructure, policy, and facilities to make it effective and enjoyable. I encourage others to consider jumping in to take advantage of the opportunity. I have learned a lot of practical lessons about life in rotation (such as: get to the cafeteria before 1pm for best entrée selection and dial 9-9 for an external phone call…), so if I can help anyone in the NS CTA with practical matters of rotation, feel free to contact me. See you on Wednesday.

Teacher of the Year



Prof. Brian Uzzi was named Teacher of the Year at Northwestern University.



The QoI-SAN thrust seeks to develop a foundational science to model, characterize, and control information delivered through multi-genre networks. This development is based on the semantics and context of information requests and requisite composite quality-of-information measures (both intrinsic and contextual). Our premise is that the overall *information capacity* of networks will be maximized by taking into account the meaning of the message (semantics) and the context in which the requested information is used (pragmatics). We use prior knowledge to drive information requirements or derive additional information and exploit the tradeoffs between data representation, the desired information quality, and communication capabilities.

In prior years we evolved the QoI research from network-centric algorithms that considered the meaning of information into an interdisciplinary mix of semantics, information, and resource allocation of information sources and networks. Last year, important advances were made in the areas of semantic-information theory and delivery, video analytics, and QoI-aware caching. In the upcoming year, we will expand the research efforts on the semantic aspects of QoI and the video analytics. We are starting to apply pragmatics to make a more natural linkage between queries and the information and QoI requirements they infer.

Semantic information theory addresses new network communication performance criteria that reflect the underlying semantics of the communicated information. In a key result this year, we conducted experiments with a source that produces semantic messages to develop a better understanding of how word meanings affect the semantic error performance in a noisy communication network. We formulated an index assignment scheme utilizing semantic distances based on lexical taxonomies as a distortion measure in a communication system. Our findings indicate the need for development of semantics-aware physical systems that allow for better integration of human factors and intelligence within complex systems design.

In a more practical setting for semantic information delivery, we explored a generalized infrastructure for describing relationships between information items. This infrastructure can be used to quantify the importance and diversity of information, thereby helping prioritize quality-aware information extraction from a tactical network. We focused on automatically extracting ontologies of information objects in mobile devices. These ontologies describe spatial and temporal relationships, as well as information quality attributes like resolution and content properties. These extracted ontologies can be queried in a language similar to SQL called SparQL, to provide a uniform method for performing semantic queries over information relationships.

We applied the idea of semantics to video-analytics. In one undertaking, we proposed a framework wherein a service receiving the content first solicits metadata (e.g., image and video features) from any device uploading content. By intelligently comparing this metadata with that associated with previously uploaded content, the service effectively identifies (and thus enables the suppression of) redundant content. We designed and implemented a framework, wherein with a small amount of metadata, redundancies in content that is generated by distributed nodes are detected with fairly high accuracy (70 %). In similar work, we augment stored images with meta-data related to the orientation, position, and other parameters of the built-in camera.

An important goal for the next five years is the synthesis and reconciliation of techniques into a unified, formal semantic information theory with a clear foundation relating semantic issues to communications needs. Our work in semantic information theory over the next year will explore semantic analogues to two capabilities developed by classical information theory: semantic source coding (optimal compression of semantic information) and semantic channel coding (reliable communication of semantic information). Our main hypothesis is that we can improve the fundamental limits of representation and communication significantly over classical semantics-agnostic results.

We will investigate the role of pragmatics in translating human intent to algorithmic understanding. Our work will further discern quality requirements for queries to produce answers based on the content of the message and based on knowledge about the information needs of the participants. Understanding these requirements will allow us to develop algorithms to trade-off required resources and achieve better quality based on the intent of human queries, coupled with contextual information ranging from command intent to immediate battlefield mission goals.

Another important topic in the upcoming year is characterizing the flow of information through multi-genre networks. We consider both the quality of the information as it propagates and the quantity of information that different network structures and types can support. We will develop foundational models, analytical tools, and experimental studies for characterizing information flow in multi-genre networks—including human networks, computer networks, and their combination.

Finally, we are addressing the delivery of semantic information in a more practical sense. We focus on information in general and video information in more detail. For video applications, we will develop a semantic querying framework over attributes of videos, taxonomize and characterize the semantic information content of videos and associated processing algorithms, and design quality-aware resource allocation and management algorithms that achieve the trade-offs between resource usage and quality delivered.

More generally, we will investigate the use of *ontological reasoning* to enable efficient, quality-aware, semantic information delivery (SID) to meet mission requirements by retrieving information of sufficient quality for decision-making from resource-constrained tactical networks. Our major scientific innovation will be in theoretical foundations and algorithms for *distributed performance optimization* in retrieving information from the tactical network. These algorithms leverage an understanding of *query semantics*, *context, and relations between content items*.

Research Awards

Prof. Heng Ji, Rensselaer Polytechnic Institute (RPI), received IBM Faculty and Bosch Research awards.

Lifetime Achievement

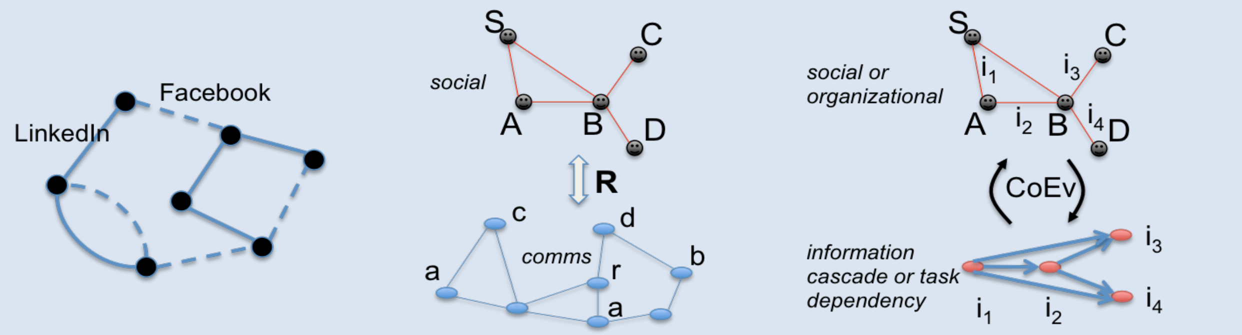
Prof. Marshall Scott Poole, University of Illinois at Urbana-Champaign, received the Joseph E. McGrath Award for Lifetime Achievement in the Study of Groups from the Interdisciplinary Network for Group Research (INGRoup).

Trust, Influence, Modeling, and Enhancing Human Performance (TIME)

The TIME thrust takes interdisciplinary research approaches to trust, influence, and modeling to inform enhancing human performance in social-cognitive networks. *Trust* is relevant to human performance in networked environments and achieving appropriate human trust in information and automation in intelligent systems. The i*nfluence* of both individuals and communities within and between cultures is key to understanding and shaping social dynamics and the spread of opinions. *Modeling* approaches include mathematical and computational models of human networks embedded in information and communication networks, scaling individual cognitive architectures to groups of human networks, and integration of disparate modeling approaches. As with other NS CTA components, TIME underwent deep transformation coming into Y6 and is now organized into five tasks:

- **Task T1, Information-based decision making and trust in networks.** T1 concentrates on development methods to measure and enhance human trust in decision-making contexts that use information from networked sources—human and automated systems alike.
- **Task T2, Social dynamics, opinion spreading, and influencing in social networks.** T2 leverages and expands the accomplishments of previous years of the NS CTA program. It investigates the role of individual-level social drivers and mechanisms in macro-level behavior of the system, including opinion spreading and influencing in social networks.
- **Task T3, Trust, influence and the enhanced human performance of multi-genre crowd networks.** This task studies technical advances to discover associations among outcome-based trust (repayment, commitment to ideals, and cooperation), network dynamics, and text and images of people. Mathematical advances and algorithms will help perfect the Army's use of crowd networks.
- **Task T4, Cognition and trust in composite networks**. T4 uses cognitive modeling, user modeling, and game-theoretical strategy analysis to research the relationships among human trust behavior, situation awareness, and performance. In this task, team members are building a realistic multi-genre model that involves humans, information sources, and intelligent agents.
- **Task T5, Group dynamics emerge from individuals: scaling up cognitive architectures and models of cognition.** T5 advances scientific knowledge about group dynamics, decision support technology, and their links to individual and group cognition. It brings together cognitive architectures, network science, and decision support technology.

Co-EDIN: Co-Evolution and Dynamics of Inter-genre Networks



Two networks embedded in the same "node space" *Two networks loosely embedded in different "node spaces"* *Two co-evolving networks embedded in different "node spaces"*

Inter-genre networks consist of multiple networks linked to each other at varying levels. Up through Year 5, the goal of the EDIN (Evolution and Dynamics of Inter-genre Networks) thrust was to mathematically characterize the rich interactions within inter-genre networks and the dynamics that occur within or across networks. In Year 6, EDIN has evolved to a new thrust—Co-EDIN. In Co-EDIN, the emphasis has matured into *co-evolution* in composite multi-genre networks.

Year 5 Accomplishments

Our key accomplishments as EDIN included the development of a theory of composite graphs for modeling dynamic interacting networks and developing efficient algorithms for their modification to meet a specific performance need; modeling and analysis of group behaviors using techniques ranging beyond traditional graph theory; development of a systematic, foundational theory of time-varying graphs; development of algorithms to predict mobility within a certain degree of accuracy; development of temporal community discovery algorithms; development of methods for controlling such networks so that they are driven to a desirable state; and the construction of a behavioral content-centric genotype of users in a social network. EDIN's Year 5 research was organized into seven tasks:

- **E1, Composite networks.** We studied theoretical aspects of *multiplex networks*, a system of multiple networks having different types of links but sharing a common set of nodes, which arise naturally in social networks and transportation networks. We were the first to analytically characterize the site and bond percolation thresholds of random multiplex networks, and also studied the problem of correlated edge dynamics. We developed guaranteed-factor approximation algorithms to identify the most critical alerts in communication/information composite networks. Finally, we developed advanced algorithmic tools for centralized and decentralized design of CG structures possessing desirable properties such as low messaging latency through command-and-control networks.
- **E2, Groups in networks.** This task developed a generative model for the inherent nature of collaboration structures (i.e., who works with whom, modeled as simplicial complexes). Such structures are different from the artifacts of collaboration (i.e., papers, movies). We also studied the structure and evolution of missed collaborations in large networks such as DBLP and IMDB. We explored functions that reflect the strength of a network of teams modeled as simplicial complexes. Finally, we developed efficient algorithms for opinion sampling, essentially finding the minimum subset of vertices in a node-weighted graph/hypergraph such that the information learned on this subset is sufficient to infer the information for the entire network.
- **E3, Foundations of time-varying networks.** We introduced the notion of "slice trees" to study the problem of compressing time-varying graphs from the perspective of dynamics on edge presence as well as real valued node/edge attributes (e.g., opinions in social networks and average speeds on road networks). We developed algorithms for time-varying networks for fast computation (at the cost of some tolerable error) of reachability and k-connectivity over time as well as for their temporal augmentation. We proposed Com2, a novel and fast incremental tensor-decomposition approach that can discover both transient and periodic ("Comet") communities. We received a Best Student paper Runner-up Award for this work. We studied and proposed an efficient node recruitment strategy for unlabeled, undirected networks based on the notion of compounding excess degrees. We received a Best Paper Award for this.

- **E4, Prediction of location and network evolution.** We investigated the interplay of user demographic and mobile social behaviors in a large-scale countryside mobile network. Significant efforts were spent on studying the next-place prediction problem. While using user IDs and coordinates alone yields 53% prediction accuracy, this accuracy was improved to over 70% using two approaches: using explicit semantic labels and discovering latent semantic trajectories. We also developed a method to mine spatio-temporal periodic patterns in traffic data and used these periodic behaviors to summarize a huge road network.
- **E5, Extrapolating network properties and behaviors.** We proposed the first method that can perform axiomatic, model-free forecasts of complex stochastic processes. We developed a new popularity competition model that explains how multiple social networks evolve, compete, and die over time (e.g., rise of Facebook at the expense of MySpace). We also developed methods for recruiting the "right" individuals in a network without access to the network topology or node labels—ENGAGE (Expected Node Gain Algorithm under Graph knowledge Evolution.
- **E6, Controllability of complex networks.** We studied distinct schemes for partial control, i.e., controlling a subset of target nodes in a complex network, and were able to rank order control schemes (e.g., random vs. local) in various regimes of degree heterogeneity. Results were used to develop a theory of control in layered networks. We also analytically and numerically explored the *eigen-spectrum* of controlling and observing complex networks for different fractions of driver nodes, leading to the discovery of several fundamental laws.
- **E7, Content-specific network behavior or Social Genome.** We performed an empirical study of the mean-field adoption rate of hashtags in different topic backbone networks. We found statistically significant evidence that hashtags diffuse fastest on their corresponding topic networks, thus a novel hashtag can potentially be classified into a topic by how it relatively spreads on each topic network. We also developed tools to quantify the randomness in the activity sequence, and studied how "social genotypes" evolve over time (and developed methods for estimating uncertainty), and how collective behaviors evolve in networks signifying research interests.

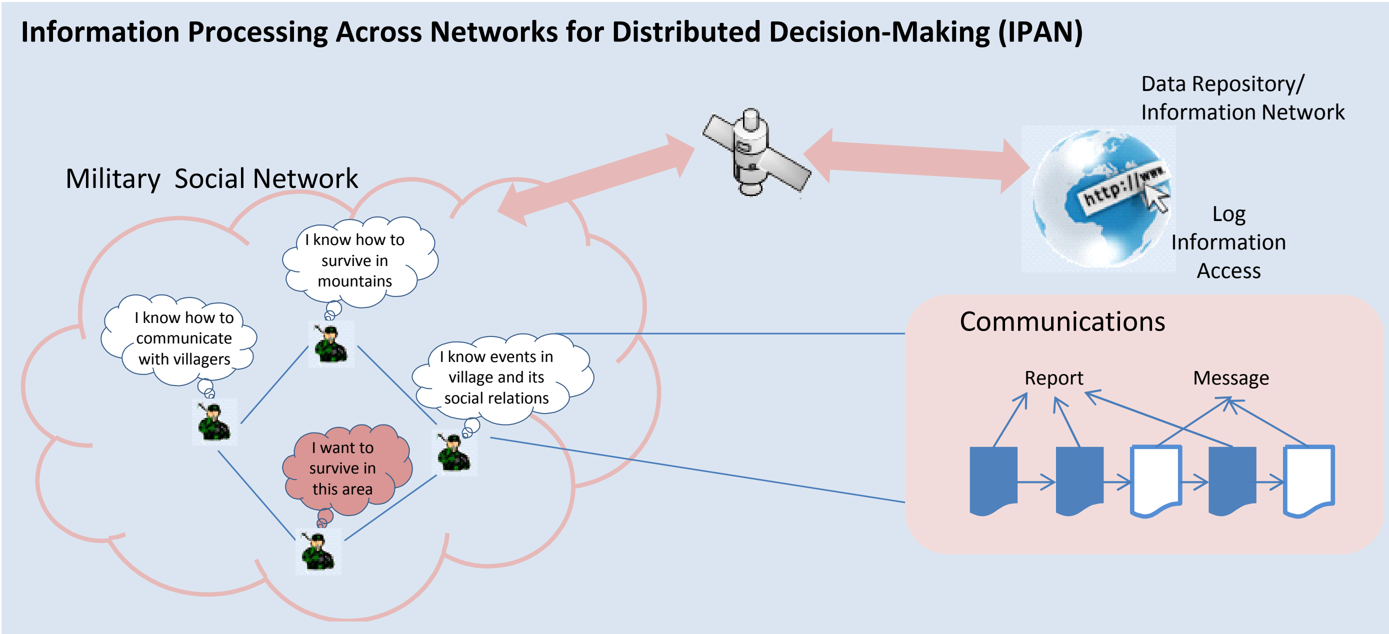
Year 6 Prospects

In Year 6, the Co-EDIN focus shifts to studying dynamic phenomena in multi-genre networks in Army and civilian settings that can potentially co-evolve with each other. The structure of a network usually affects the behavior of the participating nodes as network processes that are executing on them. Similarly, behavioral dynamics may alter the network structure. Such co-evolution can happen when one network is exerting stress on the other as in competitive settings or even under normal modes of operation. Co-evolution can result in births, deaths, and changes to nodes and the processes running over them. The modeling techniques in Co-EDIN should have adequate expressive power to capture such phenomena and should lend themselves to tractable mathematical analysis.

To model multi-genre network interactions and their co-evolving dynamics, new tools and mathematical frameworks may be necessary to achieve succinct yet mathematically tractable representations. These tools should be suitable for modeling interactions between nodes in multiple genres of networks and also for better understanding how groups of nodes and their properties evolve over time as they interact with other groups. The task "Modeling and analysis of groups in multi-genre networks," is addressing this issue.

The task "Discovering network processes in time-evolving networks" aims to understand the structure and behavior of large inter-genre networks, which are not always globally exposed to an observer. The related task, "Advanced techniques for learning and prediction over inter-genre, time-evolving networks," will develop techniques for efficiently discovering certain substructures (desirable or undesirable) by following direct or indirect methods.

The final focus of Co-EDIN is to design and control improved networks. New methods are being developed for designing optimal composite networks over time in the task named "Designing and controlling composite networks." In "Controllability of complex networks," mathematical advances will enable the control of collections of temporal, layered, co-evolving, and nonlinear networks spanning multiple genres. We will attempt to relax the assumptions of linear continuous time control, time-invariant networks, and centralized control for deriving practical techniques for designing and controlling real-world networks.



Information Processing Across Networks for Distributed Decision-Making (IPAN)

The IPAN research thrust addresses the fundamental understanding of how to transform data and observations from multi-genre networks into relevant, situational understanding for users in a highly constrained environment (e.g., at the tactical edge) to enable them to make effective decisions relative to their mission space. In this thrust, research has been conducted on multiple frontiers, partitioned into three major topics: (1) context-aware knowledge synthesis and analytics over multi-genre networks, (2) uncertainty management across multi-genre networks, and (3) distributed processing and user interactions for situational understanding.

Topic 1, Context-aware knowledge and analytics over multi-genre networks, is based on the concept that many network indicators in multi-genre networks are inherently unstructured, e.g., social *synthesis* media, situation reports, news and human communications. Knowledge about human behavior and other structures provides an untapped opportunity to infer more knowledge by extracting and exploiting the latent semantic structures. To realize this potential, this task will investigate a two-layer construction process integrated with a multi-dimensional analysis of truth and text. It will also validate and integrate social cognitive behavioral models into the multi-genre network construction.

Our premise is that a deep understanding of how information processing extracts knowledge and presents it to users in light of uncertainty can lead to ground breaking improvements in human decision making. Research in Year 5 on network construction and reliable knowledge synthesis with noisy and untrustworthy text data has generated multiple effective methods for (i) construction of high-quality, semi-structured, multi-genre networks from unstructured, noisy text data, (ii) wikification of high-quality knowledge synthesis, and (ii) multi-dimensional summarization and knowledge discovery within heterogeneous socio-informational networks.

In Year 6, ongoing work is *constructing unified, structured knowledge networks by integrating text, social media and multi-genre networks* (Task I1). Progress includes a new framework for extraction of high-quality phrases and typing of these phrases. To enhance the robustness of entity linking without requiring training data, we adopted a novel hypergraph-based abstract meaning representation to select high quality sets of entity "collaborators" in collective inference. Finally, the task is developing new models from the Theory of Planned Behavior to improve entity linking and build predictive models of humans/communities.

Topic 2, Uncertainty management across multi-genre networks, studies the idea that, in the real world, the data and observations are inherently noisy, incomplete, and sporadic, which leads to uncertainty in the synthesized knowledge resulting from processed information. Moreover, many observations originate with human-generated reports or network proxies such as web clicks or phone logs. The noise is the result of semantic ambiguity, innocent or maliciously-intended errors, and/or deletion, which is more complicated than noise from sensor readings that can be modeled with well-known statistical distributions, e.g., Gaussian, Poisson, etc.

Research in Year 5, generated new principles and effective methods for fact-finding, i.e., extracting reliable information from sources whose reliability is generally unknown. This problem was based on analytic foundations that allow the fact-finder to rigorously assess the reliability of its conclusions. The new analytic framework was demonstrated by extracting reliable information from Twitter and extending the original fact-finding algorithm and system, Apollo, in several major ways. Apollo was extended to include analysis of non-independent claims and non-independent sources, accommodation of time-varying state, fact-finding with polarized sources, conflict resolution, and handling uncertain links.

In Year 6, the work focusses on *taming uncertainty in social channels: algorithms, bounds, and team performance* (Task I2). Three areas of emphasis are: (i) investigating how human data sources are influenced by their social networks, and quantifying fundamental error bounds and information recovery rates achievable by fact-finding algorithms that collect their inputs from social channels; (ii) investigating whether and how algorithms can achieve these bounds; and (iii) validating the new algorithms in multiple human team experiments designed to assess the impact of reliable uncertainty characterization on team decision making performance.

Topic 3, distributed processing and user interactions for situational understanding, addresses the need to understand how network manipulations can enhance information processing over multi-genre networks. It studies how synthesized knowledge can be formed over distributed computing resources and displayed to the users based on their requests and need to enhance situational understanding. This deep understanding also needs to encompass social/cognitive models for understanding social networks under surveillance as well as understanding the social/cognitive state of the users receiving the information. In this topic, the work is partitioned into three major tasks:

- **Task I3, Resource aware, multi-modal content fusion.** Task I3 is a new task that focusses on information fusion at low resource computing nodes connected in the communications network via low bandwidth and sporadic links. In Year 6, we are investigating meta-data representations and novel network resource management methods to determine what should be fused and where, within the network, should it be fused. The goal is to optimize information throughput through a deep study of the information fusion process for objects extracted from disparate video sources. The task will determine theoretical network capacity with fusion (CWF) that is possible when distributing the workload optimally over the communications network. We are especially interested in developing a theoretical framework around the concept of correlated quality of information, based on a newly proposed *correlated quality of information (CQoI)* for a specific query/event. CQoI weighs how close the source is to the query and how much benefit one sensing can provide to answer the query. Each sensing datum has its own amount of CQoI, which is also the upper bound for how it contributes to the answer.
- **Task I4, Distributed, user-oriented multi-scale network summarization and online analytical processing.** This task builds on the Year 5 work on multi-dimension text information extraction, aggregation, and mining in EventCube. It incorporates structured and unstructured text data along with cognitive principles to better adjust to the needs of the user. In Year 6, the work is developing theory and technology for user-oriented, multi-dimensional, multi-scale network summarization and online analytical processing (OLAP) methodologies for flexible and efficient information processing, situation analysis, and decision support. It has a special focus on extraction of inter-related dimensions and concept hierarchies from unstructured and semi-structured data for multi-dimensional and multi-level aggregation and knowledge discovery. Recently, new approaches have been proposed for using meta-graph and ego-network structures to explore multi-dimensional network discovery, summarization, and exploration across multi-genre networks.
- **Task I5, Problem solving in socio-information networks.** Task I5 is built on Year 5 work on the exploration of search and recommendation over socio-information networks. It targets collaborative networks that are composed of information repositories and human experts who cooperate with each other and use the repositories to complete specific tasks, such as resolving problems reported by customers. In Year 6, the work will enable a human-in-the-loop study to validate and update the multi-theoretical multi-level (MTML) social model for explaining how humans search and query multi-genre networks. The MTML model and other socio-information models will be incorporated in a new framework to determine how human users interact with human experts and data repositories in socio-information networks. The framework will also help determine how querying systems can better interact with humans to refine queries for decision making.

IPAN has also been collaborating with other thrusts on various research frontiers. Collaboration efforts include work with QoI-SAN on bridging data to decisions for delivery of semantically meaningful information over communication networks; with TIME to understand group cognition over the social network; and with Co-EDIN to exploit information sources that belong to latent groups so we can better understand the evolution of the latent networks. IPAN is also actively looking for transition opportunities that support military applications, including the Apollo fact-finding tool, the querying-answering technology via user interaction, as well as the network-based OLAP technology that is planned to emerge from the research.

Publications

K. Avrachenkov, P. Basu, G. Neglia, B. Ribeiro, and D. Towsley, "Pay Few, Influence Most: Online Myopic Network Covering," in *Proc. Sixth IEEE International Workshop on Network Science for Communication Networks* (NetSciCom '14), Toronto, Canada, May, 2014. [Best paper award]

Miguel Araujo, Spiros Papadimitriou, Stephan Gunnemann, Christos Faloutsos, Prithwish Basu, Ananthram Swami, Vagelis Papalexakis, and Danai Koutra "Com2: Fast Automatic Discovery of Temporal ('Comet')Communities," in *Proc. PAKDD '14*, Taiwan, May 2014. [Runner Up to Best Student paper]

The work of Prof. Brian Uzzi of Northwestern University was featured in these media articles:

- "The No. 1 Predictor of Career Success According To Network Science," *Forbes*, Jan., 2015.
- "Entrepreneurship through Networking," *LinkedIn Pulse*, Jan. 24, 2015)
- "Big Data Drawing Big Student Enrollments," *Poets & Quants*, Dec., 2014.
- "The One Thing You Should Know About Deeply Connecting With Anyone," *Forbes*, Nov., 2014.
- "Fixing a Work Relationship Gone Sour," *HBR Blog Network*, Aug., 2014.

A. Pilny, M.S. Poole, A. Yahja, & M. Dobosh, "Dynamic communication patterns and performance in small groups," received a top paper award from the Group Communication Division at the National Communication Association Annual Conference, Chicago, IL, Nov., 2014.

About NS CTA Research

Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.