Annual Report for Period: 10/2010 - 09/2011

Principal Investigator: Sayama, Hiroki

Organization: SUNY Binghamton

Submitted By:
Sayama, Hiroki - Principal Investigator

Title:
CDI-Type I: Modeling and Predicting State-Topology Coevolution of Complex Adaptive Networks

---

**Project Participants**

**Senior Personnel**

Name: Sayama, Hiroki

**Worked for more than 160 Hours:** Yes

**Contribution to Project:**

---

**Post-doc**

**Graduate Student**

Name: Schmidt, Jeffrey

**Worked for more than 160 Hours:** Yes

**Contribution to Project:**

Jeff worked as a primary developer of the software system this project aims to produce. He also worked with Ben on the implementation of computer simulation models of adaptive networks. He presented one abstract at ICCS 2011, and participated in the overseas research collaboration with Thilo Gross at the Max Planck Institute for the Physics of Complex Systems in Dresden, Germany, over the summer of 2011.

Name: Bush, Benjamin

**Worked for more than 160 Hours:** Yes

**Contribution to Project:**

Ben worked on computer simulation models of evolving social networks as an application of the adaptive network framework that this project is proposing. He also worked with Jeff on the development of adaptive networks modeling and analysis software. He presented one abstract at ICCS 2011, and participated in the overseas research collaboration with Thilo Gross at the Max Planck Institute for the Physics of Complex Systems in Dresden, Germany, over the summer of 2011.

Ben actively participated in the NetSci High high school research competition program as a co-supervisor of the two student teams at Maine-Endwell High School, Endwell, NY. He also played the leading role in developing those students' posters.

**Undergraduate Student**

Name: Krell, Steven

**Worked for more than 160 Hours:** No

**Contribution to Project:**

Steve helped the PI organize the STCAN 2011 workshop and the ICCS 2011 conference as a student administrative assistant. He was able to attend the workshop/conference and learned a lot about complex systems and network sciences.

---

**Technician, Programmer**

---

**Other Participant**

---

**Research Experience for Undergraduates**
Max Planck Inst. Phys. Complex Systems

Thilo Gross at the Max Planck Institute for the Physics of Complex Systems (MPI-PKS) has been collaborating with the PI. In Summer 2011, Gross arranged a visiting scholar program at MPI-PKS for the PI and the two graduate students to stay at the Institute and conduct collaborative research. Accommodation and local allowance costs were covered by the Visitors Program of the Institute. The PI and the students were given access to various facilities, including the library and computational resources.

Maine-Endwell High School

The PI, Bush and Akaishi worked with Mrs. Julie Gallagher, Assistant Principal of the Maine-Endwell High School, Mr. John Endress, the School's IT specialist, and two student teams (seven female high school students) on their research projects. This was part of the NetSci-High high school research competition program, which was created and run as part of this grant. The research culminated in two poster presentations that were presented at the 2011 International Conference on Network Science (NetSci 2011) in Budapest, Hungary, as well as at the Eighth International Conference on Complex Systems (ICCS 2011) in Boston, MA.

New York Hall of Science

The PI collaborated with Stephen Uzzo and Tara Chudoba at the New York Hall of Science (NYSci) to organize and run the NetSci-High high school research competition program, which was part of this grant. The PI visited NYSci a few times for meetings and hosting an initial teachers' workshop there. The PI's institution provided funding to NYSci as a participant's cost to fund the NetSci-High program. Chudoba and the PI traveled to Budapest, Hungary, to run the high school students' poster session collaboratively.

University of Connecticut

The PI collaborated with Junichi Yamanoi, a Ph.D. student in the University of Connecticut Business School, to develop a computer simulation model of corporate merger processes represented as an adaptive network. The result was presented at the 2011 International Conference on Network Science (NetSci 2011) in Budapest, Hungary, and also at the Eighth International Conference on Complex Systems (ICCS 2011) in Boston, MA. Yamanoi and the PI are currently working on writing a journal article on this collaborative research.

DRDC CORA

We have recently made an agreement with DRDC CORA (Defence Research and Development Canada Centre for Operational Research and Analysis, Ottawa, ON, Canada) that we will apply our GNA framework to the modeling and analysis of their operational networks for Arctic Search and Rescue (SAR) incidents.

Other Collaborators or Contacts

Activities and Findings

Research and Education Activities:
======================================================================
[Developing and implementing core algorithms for automatic discovery of network rewriting rules]

In Year 2010-2011, we have designed and developed a preliminary version of the Generative Network Automata (GNA) framework in Python using the NetworkX module. The development is being done using SourceForge.net, a widely used open-source software development website, where all source code is publicly available for free.
The preliminary version of the GNA framework has the following capabilities:

(a) Input/output interface of network evolution data. Our framework can read network evolution data from a file in a graph ML format. Graph ML is an XML-based graph mark-up language that can represent multiple graphs in a single file, which we adopted as a general data format to represent time series of networks. Our framework can also output network time series into a graph ML file.

(b) Identification of network rewriting events and compression of network evolution data. Our framework can scan a given graph ML file and sequentially detect differences between two consecutive snapshots of networks, and then utilize this information to compress the network evolution data. When the network differences are identified, several statistical measurements associated with the differences are also recorded, which will be used for estimation of extraction mechanisms in (c) described below.

(c) Automatic discovery of subgraph extraction rules. When multiple candidate models of subgraph extraction mechanisms are given, our framework can calculate the likelihood of each mechanism based on the compressed network evolution data. This likelihood calculation algorithm was newly developed over the past year (i.e., not described in our original proposal), whose details will be disclosed in a conference or journal publication shortly. The calculated likelihood values can be presented back to the user, then either the user can choose the plausible extraction mechanism by assessing those values, or the software can automatically select the mechanism with the largest likelihood value.

(d) Automatic discovery of subgraph rewriting rules and reconstruction of network evolution. Our framework can use the collection of network changes generated in (b) above as a set of subgraph rewriting rules and perform network rewritings so as to reconstruct network evolutionary processes from an initial configuration (i.e., simulation of network evolution).

All of the above capabilities have already been implemented and are working in our software. There is still much room to make each capability more reliable and computationally more efficient. We will continue to work on the improvement of our software.

[Testing algorithms with abstract network models]

We have tested the implemented algorithms for automatic rule discovery using artificially generated network evolution data. Specifically, we have used several different network growth models, including (a) degree-based preferential attachment, (b) state-based preferential attachment, (c) degree-based preferential attachment applied to nodes with a specific state only, and (d) pure random attachment, among others. We have applied our software to these data and confirmed that in most cases the correct model was identified with the largest likelihood values. There are still some cases when our software gives wrong answers. Possible causes of such wrong estimations are being investigated and will be fixed.
Our software can also reconstruct (i.e., simulate) network evolution using the automatically discovered extraction/rewriting rules. We are currently evaluating the quality of the reconstruction results by comparing them with the original network evolution data.

[Collecting real-world social network data]

Real-world network evolution data are necessary for future evaluation of our framework. Through the process of searching, we realized a fundamental lack of such data of network evolution over time. To address this problem, we have developed an original web search engine based method for collecting approximated historical data of temporally changing social adaptive networks. In our method, a search query string is combined with additional keywords that specify the inclusion/exclusion of specific years to limit the search results to a particular time point. Using the proposed method, we reconstructed the temporal evolution of a social network from 2005 to 2009 of 93 individuals who are important in the US economy as a test case.

We also used this method to reconstruct a network that represents the relatedness between scientists and their peers as well as various research topics. We characterized each individual scientist's most notable research topics (expertise) by measuring a 'visibility boost', defined as the increase of the scientist's visibility resulting from focusing on a particular research topic. We analyzed the correlation between scientists' expertise and their interdisciplinary nature. The interdisciplinarity of a scientist was characterized in two different ways: (a) diversity in their expertise (i.e., how many research topics they are associated with), and (b) their betweenness centrality in the scientists' network (i.e., how much they connect multiple scientific communities). While this research did not consider the dynamical nature of the network (yet), it did demonstrate the effectiveness of our data collection method.

We plan to apply this method to collect more data of real-world social network evolution, which will be used in the evaluation of our modeling framework at a later stage of the project.

[Expanding domains of application]

Several applications of adaptive networks for biological and social systems modeling have been initiated, including:

(a) Adaptive network models of social network evolution. We have developed an adaptive network model of idea-exchanging social networks to investigate the relationship between individual behavior and network position. Our individual agents divide their time into three distinct actions: thinking about and reconciling one's own ideas, disseminating ideas to others, and listening to ideas from one's peers. Each individual is host to a local population of ideas, some of which are fixed, while others can be replaced as a result of social interaction. Individuals evaluate ideas based on their local idea population, and these evaluations are used to update the (directed) link weights whenever ideas are exchanged. Using this model, we
calculated the distribution of network centrality and other measurements over the space of all possible individual behaviors.

(b) Adaptive network models of corporate merger. In collaboration with Junichi Yamanoi at the University of Connecticut Business School, we have developed an adaptive network model of cultural integration in merging firms, represented as a network initially made of two communities. Nodes receive information from their neighbors and update their cultural states, while the link weights also change so that links between nodes with similar/different cultural states are encouraged/discouraged. We explored two experimental parameters that determine initial link distributions: how concentrated the information sources of within-firm social ties are on a small number of key individuals, and how concentrated the between-firm links are on information sources of each firm. Model implementation and Monte Carlo simulations were conducted in Python using the NetworkX module.

(c) Adaptive network based models of collective behavior of swarms. In collaboration with Thilo Gross at the Max Planck Institute for the Physics of Complex Systems, we have started using adaptive networks to represent interaction networks of marching locust swarms that may switch the direction of motion between left and right. Gross's recent work considered phase transitions in a homogeneous swarm. We have extended their model to heterogeneous swarms made of multiple types. To facilitate analytical work, we have developed a computer program that automatically generates all the network rewriting rules from a smaller set of model assumptions, and also a set of dynamical equations that describe dynamical change of network motif densities using a technique called 'moment closure'. This program may eventually be integrated into our Python-based GNA framework in the future.

(d) Incorporation of GNA into network models of joint systems. We have recently made a research contract with Defence Research and Development Canada Centre for Operational Research and Analysis (DRDC-CORA) to develop GNA-based computer simulation models and analysis tools of the behavior of dynamical operational networks made of heterogeneous specialized agents, each working in distinct environmental domains. A target problem is the SAR (Search and Rescue) operational network for the Canadian Arctic. This project will be the first instance of real-world applications of our GNA framework used in non-academic domains.

======================================================================
[
[Dissemination]]

[Publications]

During Year 2010-2011, we have published seven conference proceedings publications listed below (two full papers and five short abstracts) on general overviews of GNA framework and/or technical details of automatic rule discovery and data collection methods. Of particular note is that the PI's presentation at the IPSJ SIG Mathematical Modelling and Problem Solving (2010-MPS-81) won the Best Presentation Award. See the attached supplementary documents for details.


[Developing project websites]

We have developed the project website at http://coco.binghamton.edu/NSF-CDI.html. Our GNA framework source
codes are also available from http://gnaframework.sourceforge.net/.

[Organizing academic meetings]

We have organized the following meetings on adaptive network research during Year 2010-2011:

* STCAN 2010: Special Track on State-Topology Coevolution in Adaptive Networks. This track was held as part of the Fifth International ICST Conference on Bio-Inspired Models of Network, Information, and Computing Systems (BIONETICS 2010), Boston, MA, on December 1-3, 2010. The track consisted of seven peer-reviewed full papers on various topics relevant to adaptive networks. Thilo Gross (collaborator for this project) helped organizing this special track as a co-chair, and planned to give a keynote talk. Unfortunately his travel was canceled because of the inclement weather in Europe.

* STCAN 2011: Workshop on State-Topology Coevolution in Adaptive Networks. This workshop was held as part of the Eighth International Conference on Complex Systems (ICCS 2011), Boston, MA, on June 27, 2011. The workshop hosted seven presentations selected based on abstract reviews. Compared to STCAN 2010, a broader range of application research areas was represented in this workshop, such as management and organizational sciences.

* NetSci High: International High School Student Poster Competition on Network Science. This new educational outreach program was developed in collaboration with the New York Hall of Science. Two poster sessions were organized: one in the 2011 International School and Conference on Network Science (NetSci 2011), on June 6-10, 2011, Budapest, Hungary, and the other in the Eighth International Conference on Complex Systems (ICCS 2011), Boston, MA, on June 26-July 1, 2011. More details will be provided in the Educational Outreach section of this annual report.

[Other dissemination activities]

We also made several oral and poster presentations about the research outcomes of this project, as listed below:


* Hiroki Sayama and Junichi Yamanoi, An adaptive network model of
cultural integration in corporate merger, presented as a talk at
NetSci 2011: International School and Conference on Network Science,
June 6-10, 2011, Budapest, Hungary.

* Cara Boothroyd, Brianna Benson, Deanna Blansky, Christina Kavanaugh,
Julie Gallagher, John Endress, Benjamin James Bush, and Hiroki
Sayama, Academic achievement and personal satisfaction in high
school social networks, presented as a poster at the NetSci High:
International High School Student Poster Competition on Network
Science, in NetSci 2011: International School and Conference on

* Jessica Calderone, Emma Valentine, Josie Trichka, Julie Gallagher,
Benjamin James Bush, Jin Akaishi, and Hiroki Sayama, A comparative
study on the social networks of fictional characters, presented as a
poster at the NetSci High: International High School Student Poster
Competition on Network Science, in NetSci 2011: International School
and Conference on Network Science, June 6-10, 2011, Budapest,
Hungary.

* Hiroki Sayama, Computational approaches to adaptive network
modeling, invited talk at the Adaptive Network Dynamics
minisymposium in SIAM Conference on Applications of Dynamical

======================================================================
[[[Education]]]

[Course development]

The PI proposed a new graduate-level course 'BME-523X: Dynamics of
Complex Networks', which has been approved by the Binghamton
University Graduate Studies Committee and will officially run in
Spring 2012. This course will provide students with concepts and
mathematical/computational tools of network science, for modeling,
analyzing and simulating the dynamics of various complex adaptive
networks. Specific topics to be introduced in this course will
include: Graphs, complex networks (random networks, small-world
networks, scale-free networks), network analysis techniques (diameter
and density measures, degree distribution, centrality measures,
community identification, modularity, motifs), scaling and power laws,
dynamical networks, adaptive networks, robustness and vulnerability of
networks, and applications to biological, ecological, social and
engineered systems. Python and NetworkX will be used for modeling and
analysis of complex networks. The GNA framework will be integrated in
the course materials.

[Graduate student training]

Two graduate research assistants (Benjamin James Bush and Jeffrey
Schmidt) were hired and trained in this project. Bush joined in Fall
2010 and has been working on the application of adaptive networks to
social network modeling and analysis. Schmidt joined in Spring 2011
and has been working on the implementation of the GNA framework in Python/NetworkX. They gave oral presentations on their respective research results at the Eighth International Conference on Complex Systems (ICCS 2011) in June 2011. Both of them traveled to the Max Planck Institute for the Physics of Complex Systems in Dresden, Germany, in July/August 2011, and participated in research collaboration with Thilo Gross and his staff/students.

Findings:
[Testing algorithms with abstract network models]

Testing the implemented algorithms for automatic rule discovery using artificially generated network evolution data produced the following outcomes:

* Experiments with data generated by (1) degree-based preferential attachment ('degree only') and (2) degree-based preferential attachment applied to a specific state only ('degree-state') produced correct results.

* Experiments with data generated by (3) state-based preferential attachment ('state only') gave wrong results. Specifically, our algorithms estimated that the data should be generated using 'degree-state' mechanisms, instead of what was actually used ('state only').

We are currently investigating into potential causes of the problem above. One possibility is that even if the true mechanism of data generation used information about node states only, the resulting network growth process may actually have shown some degree correlation as an emergent property. If this is the case, that would mean our algorithms did correctly identify observed correlations in the data.

Another issue that was realized through the software development process is how to handle subgraph rewriting cases that did not occur in the given network evolution data. This problem is particularly crucial when one tries to simulate network evolution based on a relatively small set of empirical data. To address this issue, we are currently implementing an option to simulate network evolution stochastically by measuring the distances between the extracted subgraph and the actual subgraph rewriting events, and by using the measured distances as probabilities of rule selection.

[Collecting real-world social network data]

Using our web search engine-based data collection method, we reconstructed the temporal evolution of a social network from 2005 to 2009 of 93 individuals who are important in the US economy. We measured centralities of those individuals for every year and found several illustrative cases where the temporal change of centrality of an individual correctly captured the actual events that are related to him/her over this time period. These results indicate the effectiveness of the proposed method.

We also used our data collection method to reconstruct a network that
represents the relatedness between scientists and their peers as well as various research topics. Analysis of the results revealed several distinct classes of research topics with regards to the interdisciplinarity metrics. Topics in 'classic' science tend to correlate positively with diversity of the scientist's expertise, but negatively with his/her betweenness centrality, while topics that are more common terms tend to correlate in an opposite way. Interestingly, topics such as 'complex network' and 'complex system' tend to sit between these two classes. These results demonstrate a novel way of characterizing the interdisciplinarity of scientists and their research topics. We are currently conducting more detailed analysis of the obtained data, with one journal article manuscript in preparation.

[Expanding domains of application]

(a) Adaptive network models of social network evolution. Preliminary results of computer simulation showed that the centrality of individuals were affected by their behavioral propensity, though the implications of results differed significantly for different experimental conditions. We are currently checking the simulation code and also revising the data analysis code to calculate the probability for each individual to be the origin of the ideas being discussed in the society.

(b) Adaptive network models of corporate merger. Using an adaptive network model, we investigated the impact of network structures within and between two merging firms on cultural integration and organizational dysfunctions that are derived from cultural distance. The simulation results demonstrated that the highest cultural integration was achieved when the network structure is more centralized within the merging firms while connections between the merging firms are less concentrated to central individuals.

The following application subprojects have just started recently, so there is no substantial finding to be reported at this point.

(c) Adaptive network based models of collective behavior of swarms
(d) Incorporation of GNA into network models of joint systems

Training and Development:
Through collaboration with Thilo Gross at the Max Planck Institute for the Physics of Complex Systems, the PI has acquired mathematical knowledge and skills for analytical study of adaptive network dynamics using moment closure techniques.

We have hired two graduate research assistants for this project: Benjamin James Bush and Jeffrey Schmidt. Bush is from an underrepresented group (Hispanic). They both enrolled in the Ph.D. program in Systems Science at Binghamton University, and the PI supervises their work. Both students took a graduate course on computational modeling taught by the PI. They gave oral presentations on their work on adaptive networks at the Eighth International Conference on Complex Systems (ICCS 2011) in June 2011. They also went to the Max Planck Institute in July/August 2011, to accompany the PI
for the collaboration with Gross. Their reflections on this international research experience are attached to the annual report.

In addition, the PI, Bush and Jin Akaishi (collaborator) participated in the supervision of local high school student teams during the Fall 2010/Spring 2011 semesters. They had weekly meetings with two student teams on a regular basis to discuss research progresses, tasks and directions. They all gained valuable educational experiences through these outreach activities.

Moreover, we hired one undergraduate administrative assistant, Steven Krell (Bioengineering junior), to assist the PI in organizing the STCAN 2011 workshop and the ICCS 2011 conference itself (for which the PI served as a Program Chair). While he did not conduct any research, Krell was able to attend the workshop/conference as a staff member and learned a lot about complex systems and network sciences.

Outreach Activities:
We have organized a pilot run of NetSci High: International High School Student Poster Competition on Network Science. This was done in close collaboration with Stephen Uzzo and Tara Chudoba at the New York Hall of Science, as well as several network science research labs at top-ranked research universities.

In early Fall 2010, we organized initial teachers' workshops at three locations (New York City, Boston, and Binghamton) to recruit high school science/math teachers and students to this program. The PI attended the New York City workshop and also hosted the Binghamton workshop. As a result, seven student teams and their teachers participated in the program during 2010-2011 and worked with their local research laboratories. In Binghamton, the PI, Bush and Akaishi collaboratively supervised two student teams (seven female students) in Maine-Endwell high school in Endwell, NY, in close collaboration with Mrs. Julie Gallagher, Assistant Principal of the school.

Each of the seven participating teams worked on its own research project in network science and submitted a poster electronically by mid-April, 2011. The submissions were then reviewed by a scientific committee made of leading network scientists, who selected two winning posters. As a result, four high school students and their teachers (three of them, each representing a different school) traveled to Hungary, Budapest, in June 2011, to participate in the 2011 International Conference on Network Science (NetSci 2011) and present their work there in person. The travel costs were covered by this grant. All of the seven posters were printed and put up at NetSci 2011. Unfortunately none of the Binghamton student teams were selected as winners.

The PI later arranged another poster session at the Eighth International Conference on Complex Systems (ICCS 2011) in Boston, MA, for which he served as a Program Chair. Some of the students who could not go to NetSci 2011 were able to come and present their work at ICCS 2011 on their own.

At each poster session, the audience was asked to fill in a comment sheet to give encouragement and feedback to the students and teachers.
who participated in this program.

Selected comments from participating students, as well as some photos taken at these conferences, are attached to this annual report.

**Journal Publications**

**Books or Other One-time Publications**

Editor(s): Hiroki Sayama, Ali A. Minai, Dan Braha, and Yaneer Bar-Yam

Editor(s): Hiroki Sayama, Ali A. Minai, Dan Braha, and Yaneer Bar-Yam

Editor(s): Hiroki Sayama, Ali A. Minai, Dan Braha, and Yaneer Bar-Yam

Editor(s): Hiroki Sayama, Ali A. Minai, Dan Braha, and Yaneer Bar-Yam

Bibliography: IPSJ SIG Mathematical Modelling and Problem Solving Technical Report 2010-MPS-081, no. 28, Fukuoka, Japan, December 16-17 (Won the Best Presentation Award)


Web/Internet Site

URL(s):
http://coco.binghamton.edu/NSF-CDI.html
http://gnaframework.sourceforge.net/

Description:
(1) Project website
(2) Generative Network Automata framework software on SourceForge.net

Other Specific Products

Product Type:
Software (or netware)

Product Description:
We are developing the Generative Network Automata (GNA) framework, free software implemented as a Python module. This framework allows a user to construct adaptive network models based on GNA, either based on the user's own network rewriting rules or rules that are automatically reconstructed from given network evolution data. The framework also allows dynamic simulations of network behavior as well as their visualizations.

Sharing Information:
The software is currently under development, yet its source codes and other information are already publicly available from our SourceForge.net site.

Contributions

Contributions within Discipline:
This project will produce a novel theoretical framework and a computational toolkit that will transform the ways of studying the dynamics on and of complex networks and thereby achieve significant advances in the modeling and prediction of their self-organization.

Contributions to Other Disciplines:
The outcomes of this project will be useful in many cutting-edge fields, including social network science, organizational research, network ecology and epidemiology, systems biology, bioinformatics, and many others. Possible areas of application include social network analysis, modeling and analysis of organizational behavior, and modeling and analysis of biological network formation. The developed framework will also serve as a generalized conceptual/mathematical 'language' for modeling, analyzing and discussing the dynamics of various complex systems, which will galvanize interdisciplinary discussion and collaboration across many different areas of applications. To facilitate cross-disciplinary discourses, international workshops/special sessions have been organized at a couple of conferences.

Contributions to Human Resource Development:
This project involves two graduate students (Jeff Schmidt and Benjamin Bush). They are receiving multidisciplinary research training under financial support from this NSF award, including international collaboration with Thilo Gross at the Max Planck Institute for the Physics of Complex Systems (MPI-PKS) in Dresden, Germany. The PI stayed at MPI-PKS for two weeks, and the two students for three weeks, in late July and early August of 2011. The PI learned from this collaboration experience new analytical techniques, including
automatic equation construction with moment closure. The students' reflections on their learning at MPI-PKS are attached to the annual report 2010-2011.

**Contributions to Resources for Research and Education:**
This project is producing a free software package of the GNA framework for modeling and analysis of complex adaptive networks, which will be a useful information resource for broader research communities.

The new graduate course on complex adaptive networks produced in this project will serve as an educational resource for faculty and students at Binghamton University.

Moreover, the educational outreach program developed as part of this project, NetSci High: International High School Student Poster Competition on Network Science, will be a continuing educational program that will bridge secondary education with state-of-the-art scientific research conducted by leading scientists.

**Contributions Beyond Science and Engineering:**
The proposed GNA framework may help model and understand many real-world social/organizational/operational networks, which may help better inform policy makers. As a starting point, we have recently started collaboration with DRDC CORA on GNA-based modeling of their operational networks. The outcome of this project will inform Canadian Defence authorities about what kind of organizational improvements will help make their SAR response joint systems more efficient.

### Conference Proceedings

### Special Requirements

**Special reporting requirements:** None

**Change in Objectives or Scope:** None

**Animal, Human Subjects, Biohazards:** None

### Categories for which nothing is reported:

Any Journal
Any Conference