# AN AGENT-BASED SIMULATION STUDY OF
# A COMPLEX ADAPTIVE COLLABORATION NETWORK

Özgür Özmen

Computational Sciences and Engineering Division
Computational Data Analytics Group
Oak Ridge National Laboratory
Oak Ridge, TN 37830, USA

Jeffrey Smith

Samuel Ginn College of Engineering
Industrial and Systems Engineering
Auburn University
Auburn, AL 36849, USA


Levent Yilmaz

Samuel Ginn College of Engineering
Computer Science and Software Engineering
Auburn University
Auburn, AL 36849, USA


## ABSTRACT

One of the most significant problems in organizational scholarship is to discern how social collectives govern, organize, and coordinate the actions of individuals to achieve collective outcomes. The collectives are usually interpreted as complex adaptive systems (CAS). The understanding of CAS is more likely to arise with the help of computer-based simulations. In this tutorial, using agent-based modeling approach, a complex adaptive social communication network model is introduced. The objective is to present the underlying dynamics of the system in a form of computer simulation that enables analyzing the impacts of various mechanisms on network topologies and emergent behaviors. The ultimate goal is to further our understanding of the dynamics in the system and facilitate developing informed policies for decision-makers.

## 1 INTRODUCTION

The collectives are becoming increasingly global and participatory due to online collaboration opportunities such as e-mailing, web-based social networking, and open-access collaboration platforms. Hence, individuals interact not only locally but also globally by constructing self-organizing collaboration networks. Specifically, scientific knowledge creation in such networks is called Global Participatory Science (GPS) (Zou and Yilmaz 2011).

A significant problem in organizational scholarship is to discern how social collectives govern, organize, and coordinate the actions of individuals to achieve collective outcomes (O'Mahony and Ferraro 2007). This tutorial summarizes the model development phase of the dissertation work that explores innovation potential in GPS (Ozmen 2013). In the following section, the problem and environment of interest are discussed. Subsequently, the first phase of the dissertation is summarized that explores micro-level (inter-scientist) socio-technical processes and mechanisms that explain emergent behaviors observed in scientific communities that collaborate over the cyber-infrastructure. Following the conceptual model section, the implementation of the model, verification and validation (V&V) efforts, and the analysis conditions are listed.

412

## 2 PROBLEM DESCRIPTION

The stakeholders of the problem are policy-makers of the Science of Science Policy (SoSP), which is "an emerging interdisciplinary field aiming to provide scientifically rigorous basis from which policy makers can assess the impacts of scientific enterprise, improve the understanding of its dynamics and assess the likely outcomes." ($http://scienceofsciencepolicy.net$ - As of 4.07.2013) The SoSP aims to evaluate various scenarios and components of science to develop informed decision-making, particularly to promote innovation and creativity.

After all parties of the problem domain are identified, the problem needs to be framed to motivate the simulation study. The objective in the research is to present the underlying dynamics of GPS in a form of computational model that enables analyzing the impacts of various behaviors of scientists on innovation potential of the collaboration network. Hence, the first question of interest is "Which interaction mechanisms in the literature explain operational behavior of GPS and its underlying socio-technical processes?"

To enable such an analysis, ethnographic study of the environment/system of interest should be delineated. Further understanding of the real-world helps us to identify the requirements of the model. At this stage, the modeler faces the decision of "How much complexity is sufficient to address the problems?" It is always easier to add new features to existing models, however a modeler should also consider the simplicity and V&V concerns. Therefore, the level of granularity is depended on the intended purpose of the model along with the previous studies of the fellow scientists. For example, while adding various aspects of the real-world to the existing models of physical environments is reasonable (for the sake of reality and prediction), when human entities are modeled, exploratory research on relatively simpler models might provide more effective perspective. Since we are still in the early stages of understanding human decision processes, development of various simplified models that incorporate different aspects and social theories will lead us to generate plausible explanations of the reality. Policy-makers encourage development of different models that bring different explanations to similar phenomenon. Eventually, with the availability of more data and cross-validation ability, the understanding developed by these models can be combined to create even better models.

The later focus of the research is on "How we can specify and implement these mechanisms in a form of a computational model to gain empirical insight and perform exploratory analysis?" For instance, there are three levels that science can be studied from: micro-level (inter-scientist interactions), meso-level (interactions among communities), and macro-level (communities of communities structures - ecosystem level. Figure 1 is introduced for better interpretation of the network topologies. The links between the nodes can be interpreted as any kind of relationship (i.e. collaboration, social, funding) and this interpretation is based on the intention and the purpose of the model developer. At early stages of model development, these specifications should be considered. The model explained in this tutorial is particularly interested in inter-scientist processes (micro-level).

## 3 COMPLEX ADAPTIVE SYSTEMS (CAS)

It is demonstrated by (Wagner 2008) that science is complex because researchers interact in both competitive and cooperative ways, with no imposed blueprint. Furthermore, she states that it is adaptive because scientists respond to environmental changes such as funding preferences and new discoveries. The CAS characteristics observed in GPS provide for the opportunity to study underlying micro-level (inter-scientist) behaviors in order to develop plausible explanations for this phenomenon.

CAS are formed of elements that have wide range in both form and capability (Holland 1996). Shlesinger (2007) describe CAS as "composed of interacting thoughtful (but perhaps not brilliant) agents." The phrase *not brilliant* raises concerns about *bounded rationality.* Axtell and Epstein (2006) discuss the empirical data, which demonstrate that all individuals should not necessarily be rational to produce efficiency in macro-level outcomes of a system in CAS, indeed human entities are not necessarily rational. In CAS, big

(a) Meso-Level
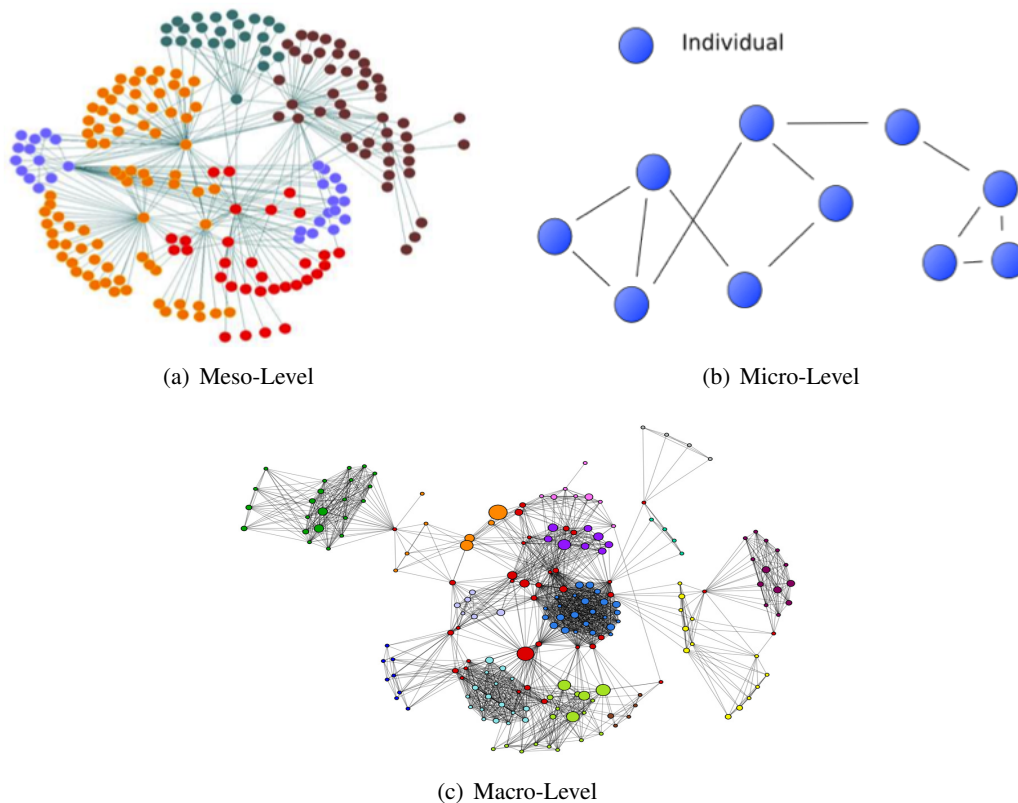


(b) Micro-Level



(c) Macro-Level

Figure 1: Network Visualizations of Micro-level, Meso-level, and Macro-level Science Studies

changes can generate small outcomes, while small perturbations can cause big emergent behaviors. (Yam 2005) defines emergence as the interdependence between details and the larger view of a system.

In addition to *bounded rationality*, Monge and Contractor (2003) describe the main elements of complex systems in terms of the network of agents, their attributes or traits, the rules of interaction, and the structures that emerge from these micro-level interactions. Authors list typical classes of agent traits as location, capabilities, and memory. Communication among the agents, as a main interaction mechanism, has an intense effect on the system level behavior of a complex adaptive system (Shlesinger 2007). In another seminal work *Hidden Order*, (Holland 1996) describes 7 basics (four properties and three mechanisms) that are common in CAS. Basic principles that are also adopted by this research are listed below:

- *Aggregation (a property):* In one sense, it means defining similar things in the same class. In another sense, it is emergent macro-level behaviors caused by aggregate micro-level relationships.
- *Tagging (a mechanism):* It results in selective interaction between the agents.
- *Nonlinearity (a property):* There are nonlinear interactions among the components of the system that means a change on a component non-linearly effects the state of another component.
- *Flows (a property):* There is a flow of components and information through the system.
- *Diversity (a property):* There are heterogenous agents and components in the system.
- *Internal Models (a mechanism):* It refers to anticipation among the agents.
- *Building Blocks (a mechanism):* There is a repetition of novel situations or structures that are emerged through building blocks (re-usable categorical parts).

The understanding of CAS is more likely to arise with the help of computer-based models (Holland 1996). Agent-based Modeling (ABM) gives us the opportunity to directly identify the entities along with the relationships and capabilities of them. ABM captures emergent phenomena because it has a holistic approach that perceives a system as more than the sum of its constituent parts. The system level behavior cannot be explained by the properties of the units in the system. Since ABM is used more with the behavioral entities, it provides an opportunity to model more realistically. Therefore in this research, Agent-based modeling mindset is adopted.

## 4 CONCEPTUAL MODEL DEVELOPMENT

Robinson, Brooks, Kotiadis, and Van Der Zee (2010) state that conceptual model starts from problem situation and moves through the questions about modeling such as, "What do we model," "What do we require to model?", and "How do we model?" The questions are iterative, and there is continuous feedback with revisions. The conceptual model is simplified. It is not the code or software model, and it considers client-side perspective as much as modeler's.

The model is grounded on theories and operating principles derived from observations on the system of interest. First, the model development is started with CAS principles in mind. In the model, interpretation of *collective action* theory and social interactions complement CAS principles (i.e., tagging, information-flows, diversity, non-linear interactions). Along with the theory base, the *information foraging* mechanism, which is inspired from food foraging in nature (Pirolli and Card 1999), is built and *preferential attachment* mechanism is designed as essential interaction process. *SEIR metaphor* and *population dynamics* are introduced which conclude the conceptual model development. Conceptual model development is followed by the implementation of it as a computer simulation. Table 1 summarizes the grounding theory base and the assumptions in the model.

Table 1: Conceptual Mechanisms and Assumptions

| Mechanism/Assumption | Grounding Theory/Study |
|---|---|
| *Artifact Selection* Mechanism | *Preferential Attachment* Process (Barabasi 2002) |
| *Bounded Rationality* Assumption | The model in (Axtell and Epstein 2006) |
| *Roulette Wheel* Algorithm | Inspired by *Predictably Irrational* (Ariely 2008) |
| *Information Foraging* Mechanisms | Metaphors in (Pirolli 2007) and *Marginal Value Theorem* from (Charnov 1976) |
| *Collective Action* assumption | *Collective Action* Theory (Olson 1974) |
| *Exposure* to Mutual-interest | The model in (Axtell and Epstein 2006)' s model |
| *Tension* within the Projects | Evolution of Project Life-cycles in OSSD Communities (Wynn 2003) |
| *Population Dynamics* | *Self-Organization* Principles in Camazine (2003) and Monge and Contractor (2003) |
| *Learning and Influence* Mechanisms | Information Flows Process in Page (2010) |
| *CAS* Principles (e.g. Tagging, Information Flows) | *Hidden Order* Book (Holland 1996), Assumptions in Yilmaz (2008) and Wagner (2008) |

### 4.1 Sample Mechanism - Artifact Selection

Let us imagine a web tool, in which motivated scientists, who believe in the collective, can browse the list of open artifacts (projects), select one of them, contribute to it, and thus learn from the artifact. Scientists browse the web-tool, which is metaphorically a grid in the model. But, not all scientists are equal in terms of time spent in browsing the online tool. Some scientists browse more titles while some browse fewer. That means the environment is heterogenous regarding the width of scopes among scientists. Each scientist has a scope that is bounded (they do not have the perfect information about the whole environment) and

scientists can only operate within that scope while searching for an artifact. The selection process is based on the calculation of three dimensions:

- *Popularity:* Scientists might select an artifact according to the artifact popularity; the more elaborated the artifact is, the more likely it is to be selected ($0 < pa < 1$).
- *Self-interest:* Scientists are more likely to select familiar artifacts ($0 < si < 1$).
- *Imitation:* Artifacts with greater number of active members are more likely to be selected ($0 < im < 1$).

Each dimension has a weight that signifies its importance in selection process. Initially, each weight is equal and $w_{pa} + w_{si} + w_{im} = 1$ . Each artifact $j$ has an incentive $P_j = w_{pa} \times pa + w_{si} \times si + w_{im} \times im$. In the case of being exposed to more than one artifact, a roulette wheel selection algorithm is used to assign probability $p_j$ to each artifact $j$ and select one of them based on the assigned probabilities.

$$p_j = \frac{P_j}{\sum_{i=1}^{N} P_i} \tag{1}$$

where $N$ is the total number of artifacts that are within the scope of a scientist. Principally, in the research, the roulette wheel algorithm is used in all selection processes since people in real life do not select the most likely decisions but they generally satisfice. That means, instead of selecting the choice that has the best value, they give their decision probabilistically relative to each decision element or criteria. Figure 2 illustrates the "artifact selection" process implemented. In the following section, the tool that is used and the implementation of the simulation model are summarized.
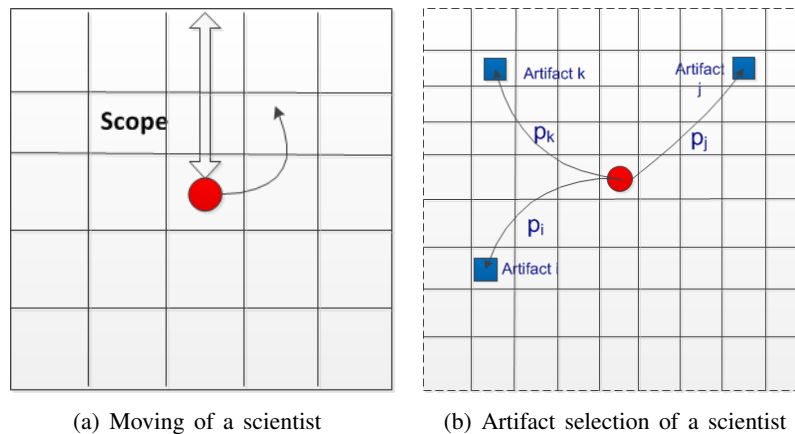


(a) Moving of a scientist  (b) Artifact selection of a scientist

Figure 2: Moving and Artifact Selection Processes in the Model

## 5 IMPLEMENTATION

Repast (Recursive Porous Agent Simulation Toolkit) is the toolkit used to create the simulation environment (North, Collier, Ozik, Tatara, Macal, Bragen, and Sydelko 2013). It is an open source simulation tool that allows for development of multi-agent simulation models in Java. Though it has a powerful framework that supports developers while building the context, it is open source and it has a shallow learning curve. Even though it is possible to encounter maintenance problems related to Repast, it has a highly active community that answers the submitted inquiries and helps users to fix the software bugs.

The main interaction context is a grid. Both scientists and artifacts are assigned to a cell. Cells are multi-occupancy, which means a cell can have more than one agent. Figure 3 represents a snapshot of RePast GUI. The right-top quadrant is the grid environment, and the right-bottom quadrant is the 3D social network representation. On the far left column, users can schedule the length of the simulations and

change their run speed. In the parameters column, users can explore the mechanism and parameter space by entering the values representing a desired scenario. On the grid, blue nodes are artifacts, red ones are the scientists.
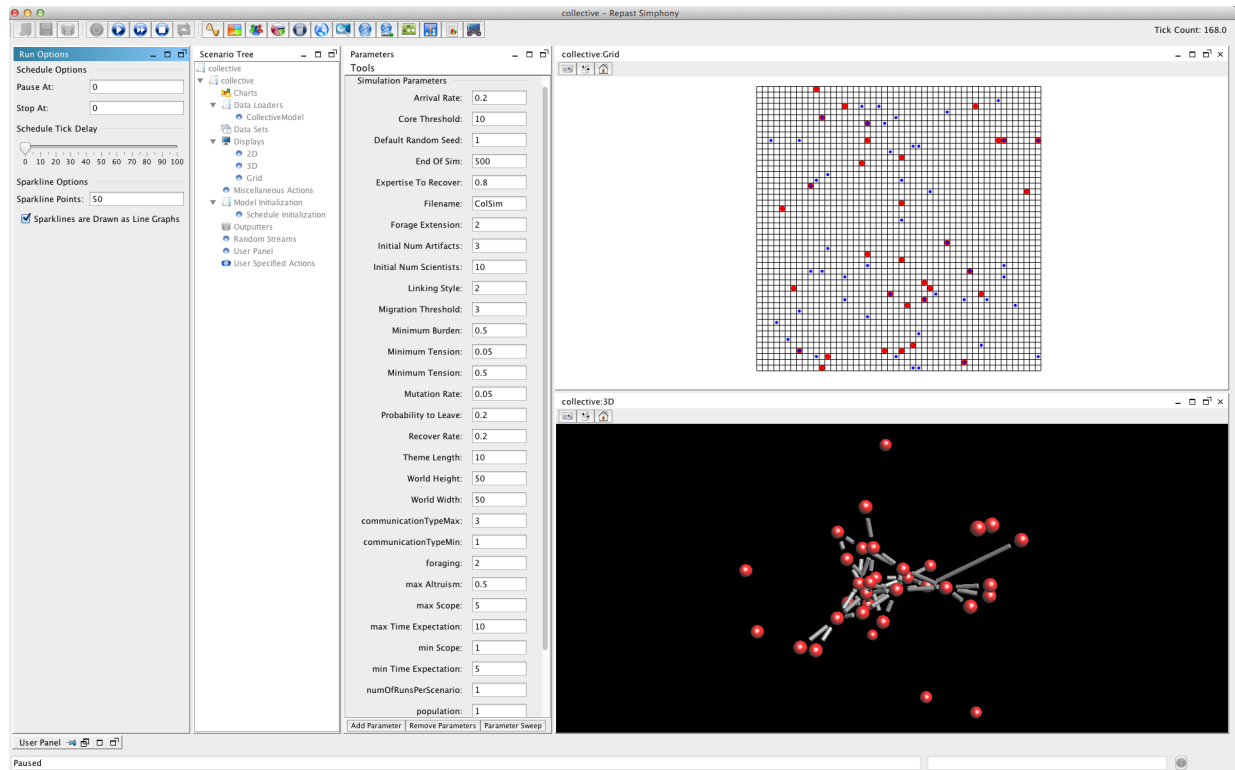


Figure 3: RePast GUI

## 5.1 Verification and Validation

Verification indicates how correctly the implemented model represents the conceptual model. When the conceptual model becomes more complex, the magnitude of complexity of the computational model increases significantly. The verification tests conducted in this research are listed below:

- Eclipse is used as an editor and a development environment. *Debugging* is used to detect anomalies at each implemented module. Step by step, each piece of code that updates variables is checked by the oversight in debugger through reasonabless analysis.
- *Unit-test* principles ensure that every algorithm and method are checked individually at extreme conditions (e.g. the behavior of certain outputs when there is only one scientist in the network) before implementation (Extreme Programming). Flow diagrams are used to verify the code.
- Performance metrics are manually calculated for small populations (typically a network of 5 scientists) and outputs of the simulations are compared with manual calculations.
- Input parameters are printed throughout the simulation to check any inconsistencies that might have been caused.
- The code is self-documented. Every piece of code that is assumed to be important has comments attached to it.
- Data interchange files (gdf) are generated for further verification in the format that records the node and edge data created by the simulations. Then, the network visualization tool *Scibrowser*, written

in Python and developed in the Auburn University Simulation and Systems Engineering Lab parses the *gdf* files and calculates the network metrics. Calculations of the implemented Java code are compared with *Scibrowser* outputs.

- Open Biomedical Ontologies (OBO) Foundry collaboration data is parsed in the Auburn University Simulation and Systems Engineering Lab. This data is explored to determine parameter and output ranges that can be encountered in GPS.

Sargent (2005) defines validation as "mean substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model." The approach of Bharathy and Silverman (2010) to validity assessment considers the life cycle of the entire simulation study and assesses the validity under the following four dimensions: "(1) methodological validity, (2) internal validity, (3) external validity, and (4) qualitative, causal and narrative validity." In methodological validity, authors consider modeling process and software process adequacy while obtaining inputs. Internal validity refers to the theoretical base of the behaviors in the model, and external validity examines how reasonable the output data is. Qualitative analysis consists of cross-validation techniques such as face validation, comparison of graphics, and visual analogies.

In this research, regarding methodological validity, the ethnographic analysis is conducted that occurs through the observation of the environment of interest. OBO data is analyzed to understand the possible outcomes and parameter ranges while determining initial conditions of the simulation runs. The simulation modeling process is followed and supported in order to avoid initialization bias, to support terminating state decisions, and to establish the number of replications. In the previous sections, the theoretical basis of the conceptual model was described for internal validity concerns. External validity is checked by observing the variability of outputs for different scenarios. Further, a single scientist is followed in two methods: debugging and visual tracking on Repast GUI to compare behavior against expected regularities. Qualitative analysis is performed representing different macro-level emergent patterns for validation purposes such as power law distributions, phase transition in the network, chaotic attractors.

## 5.2 Initial Conditions

Naturally, in complex adaptive models, the outputs are highly probable to have cyclic structures. In theory, these systems never stop running and usually do not reach equilibrium, because there are phase transitions in long run. The analysis of steady-state simulations is more difficult than terminating simulations and is not suitable for the model of this research. Also, no terminating state works perfectly for each scenario and each metric, following the stochastic nature of the study. While some scenarios converge to a desired network stage after 100 time ticks, some scenarios may need 400 time ticks to converge. Therefore, the terminating state is determined that is sufficient enough to observe the stage of the network that under various scenarios.

The next step is to decide on the number of replications (*n*). Sampling variability is a primary concern in the assignment of number *n*. In this research, Bonferroni Analysis is conducted to determine the number of replications because multiple performance measures are observed. The so-called problem of *multiple comparisons* should be mitigated. There may be a number sufficient for estimating an output metric with a given confidence interval. But for different scenarios and different metrics, the number of replications could vary. Regardless, Bonferroni inequality states that "all intervals should contain their performance measure simultaneously." Relative to this concern, the Bonferroni inequality addresses an overall probability of at least $1 - \alpha$ that the confidence interval of all $k$ metrics contain their own expected performance measures. If the confidence interval of metric *s* is $1\text{-}\alpha_s$, then Bonferroni inequality states:

$$P(\text{All intervals contain their respective performance measure}) \geq 1 - \sum_{s=1}^{k} \alpha_s \qquad (2)$$

## 5.3 Observables

Figure 4 represents the evolution of network topology over time. The network phases through four well-known stages that are also listed by Krebs and Holley (2002).



(a) Scattered
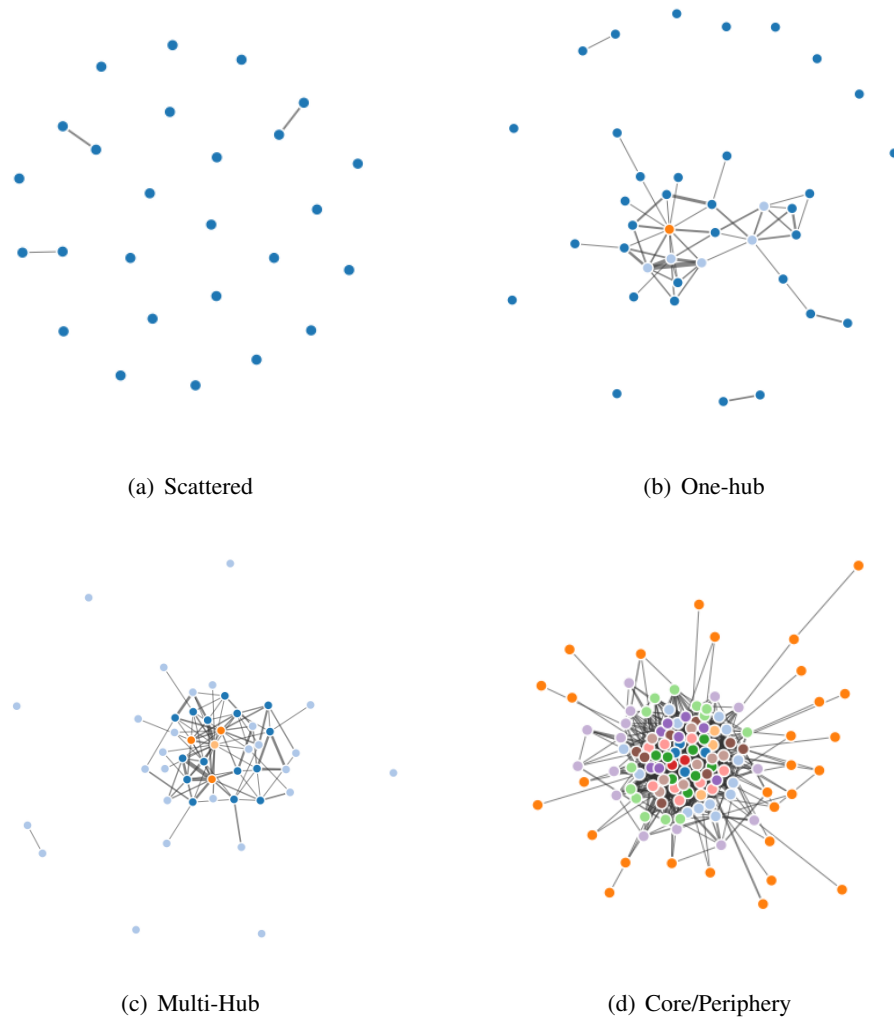
(b) One-hub

(c) Multi-Hub

(d) Core/Periphery

Figure 4: Emergent Network Patterns Over Time - Simulated

Relative to the interest of the stakeholders, innovation potential is the key feature to measure in emerging social networks and is likely to be captured by measuring various social network metrics. Some network structures are known to be more innovative than the others. The relative social network metrics that are identified as important are:

- *Network Density:* The calculation of the proportion of possible ties that exist in the network (Rowley 1997). *Density* is $\frac{2|E|}{N(N-1)}$, where $|E|$ is the total number of edges in a network and N is the total number of nodes. High density is an indicator of mobility, which means increased connectivity and transfer of ideas within the network.
- *Diversity in the network:* The number of different people connected in the network. Diversity can be measured through different skills, expertise, resources, and reputation.

- *Diversity in the population:* The perception of differences in others within the population over time. Difference can also be in terms of skills, expertise, resources, and reputation. How the diversity metrics are measured is defined in the following sections.
- *Degree:* An actor's total number of connections.
- *Degree centrality of an actor:* According to Wasserman (1994), people with the greatest number of ties are the most central actors in a network. It is the proportion of possible ties that exists for an actor. In this research, the *degree centrality* of scientist $i$ is $DC_i = \frac{Degree_i}{N-1}$, where $N$ is the total number of nodes in the network.
- Degree Centrality of a Network: The range of variability among degrees of the actors. *Degree centrality of a network* is $DC_{Network} = \frac{\Sigma_{i=0}^{N} DC_{max} - DC_i}{N-2}$, where $N$ is the total number of nodes and $DC_{max}$ is the maximum degree centrality a scientist has in the network.
- *Clustering Coefficient:* The number of edges in a neighborhood divided by the maximum possible number of edges that could exist in that neighborhood. The coefficient provides information about how actors in a network tend to cluster together. For each scientist $i$, a neighborhood is defined. The proportion of possible ties that exist between neighbor nodes is measured, assuming that the neighborhood is a network itself. The clustering coefficient of the whole network is the average of the clustering coefficients of individual scientists.
- *Average Path Length:* The average number of steps along the shortest paths for all possible pairs of network actors (Albert and Barabási 2002). It can also be stated as the average of the shortest paths from every scientist $i$ to scientist $j$. Lower values indicate higher cliquishness and fewer structural holes.
- *Core/Periphery Ratio:* The ratio of the number of core actors to the number of periphery actors. Boyd, Fitzgerald, and Beck (2006) state that "individuals in a group belong to either the core, which has a high density of ties, or to the periphery, which has a low density of ties." The *core-periphery ratio* calculation removes less central nodes recursively until only central nodes remain in the network. The remaining nodes are counted as core nodes and the removed nodes are counted as peripheral members (Borgatti and Everett 2000).

In addition to the social network metrics, the total maturity of the artifacts, the distribution of expertise levels of scientists, activeness in the population can also be discussed as collective creativity metrics. Also, diversity among the nodes and diversity among the links are also observed. To measure diversity in the environment, a *variation among a type* strategy is adopted (Page 2010). Basically, diversity can be defined as combination of three properties:

- *Variety:* It defines "how many types of things are there?" The number of different scientists, themes or interests can be perceived as variety depending on the metric. All else equal, the greater the variety, the greater the diversity.
- *Balance:* It defines "how much of each type of thing are there?" The greater equality of the balance, the better the diversity. For Balance formulation, *Shannon entropy* (Stirling 2007) is used in which $p_i$ indicates the proportion of members of given type in the total population:

$$-\sum_{i=1}^{N} p_i \times \ln p_i \tag{3}$$

- *Disparity:* It defines "how different from each other are the types of things?" Greater disparity is beneficial for diversity. The theme of artifacts or interests of scientists are used as attributes to calculate disparity in the population.

For example, using the metrics above, the high centrality and fewer structural holes argument can be adopted to measure innovativeness in a network (Burt 1995). This hypothesis promotes moderate level

average path length, higher degree centrality, and moderate level clustering coefficients in the network while it may decrease the density, which fosters mobility and diffusion of ideas. Moreover, diversity is also perceived as an attribute that generates interdisciplinary and spurs innovation.

## 5.4 Alternative Implementations

A developer can choose to calculate relative social network metrics by simply outputting the node and edge information and feeding the data to a commercial tool such as NetworkX. For visualization purposes, the library that Repast provides can be used or D3 libraries can be used as in Figure 4. These decisions are up to the model developer and depend on how the results are desired to be presented. For the sake of flexibility and performance, the metrics of the simulation are calculated by custom code. Since Repast is java-based and open-source, it provides the flexibility to overwrite or change the methods that are released in the framework. That enables fast evaluation of the networks if a co-evolutionary algorithm is implemented to explore the scenario-space.

Moreover, similar models can be developed using different tools. It would be beneficial to implement the same conceptual model with different libraries (i.e., StarLogo and NetLogo) and analyze whether built-in assumptions of the tools generate discrepancies between the outputs. Additionally, a developer is not limited to use a time-stepped approach that is provided by Repast time-execution. An event-driven methodology (i.e., DEVS) can also provide continuous time - continuous space implementations, which might provide better performance and realistic representations for some cases. In this research, time-stepped approach is adopted due to the easiness of implementation and suitability to cellular automata models (grid representation).

## ACKNOWLEDGMENTS

## REFERENCES

Albert, R., and A. Barabási. 2002. "Statistical mechanics of complex networks". *Reviews of modern physics* 74 (1): 47.

Ariely, D. 2008. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. 1 ed. HarperCollins.

Axtell, R., and J. Epstein. 2006. "Coordination in transient social networks: an agent-based computational model of the timing of retirement". In *Generative social science: Studies in agent-based computational modeling*. Princeton Univ Press.

Barabasi, A. 2002. *Linked: The New science of networks*. Perseus Books.

Bharathy, G. K., and B. Silverman. 2010. "Validating agent based social systems models". In *Proceedings of the Winter Simulation Conference*, 441–453. Winter Simulation Conference.

Borgatti, S., and M. Everett. 2000. "Models of core/periphery structures". *Social networks* 21 (4): 375–395.

Boyd, J., W. Fitzgerald, and R. Beck. 2006. "Computing core/periphery structures and permutation tests for social relations data". *Social networks* 28 (2): 165–178.

Burt, R. 1995. *Structural holes: The social structure of competition*. Harvard University Press.

Camazine, S. 2003. *Self-organization in biological systems*. Princeton Univ Pr.

Charnov, E. 1976. "Optimal foraging, the marginal value theorem". *Theoretical population biology* 9:129–136.

Holland, J. 1996. *Hidden order: How adaptation builds complexity*. Basic Books.

Krebs, V., and J. Holley. 2002. "Building sustainable communities through network building". Technical report.

Monge, P., and N. Contractor. 2003. *Theories of communication networks*. Oxford University Press, USA.

North, M. J., N. T. Collier, J. Ozik, E. R. Tatara, C. M. Macal, M. Bragen, and P. Sydelko. 2013, March. "Complex adaptive systems modeling with Repast Simphony". *Complex Adaptive Systems Modeling* 1 (3): 1–1.

Olson, M. 1974. *The logic of collective action: Public goods and the theory of groups*. Harvard University Press.

O'Mahony, S., and F. Ferraro. 2007. "The emergence of governance in an open source community". *Academy of Management Journal* 50 (5): 1079–1106.

Ozmen, O. 2013, February. *Exploring the Impact of Socio-technical Communication Styles on the Robustness and Innovation Potential of Global Participatory Science*. Ph. D. thesis, Auburn University.

Page, S. 2010. *Diversity and complexity*. Princeton Univ Pr.

Pirolli, P. 2007, April. *Information foraging theory: Adaptive interaction with information*. Oxford University Press, USA.

Pirolli, P., and S. Card. 1999. "Information foraging.". *Psychological review* 106 (4): 643.

Robinson, S., R. Brooks, K. Kotiadis, and D. Van Der Zee. 2010. *Conceptual modeling for discrete-event simulation*. CRC Press, Inc. Boca Raton, FL, USA.

Rowley, T. 1997. "Moving beyond dyadic ties: A network theory of stakeholder influences". *The academy of management review* 22 (4): 887–910.

Sargent, R. 2005. "Verification and validation of simulation models". In *Proceedings of the 37th conference on Winter simulation conference*, 130–143.

Shlesinger, M. F. 2007, September. "Complex Adaptive Systems: An Introduction to Computational Models of Social Life". *Journal of Statistical Physics* 129 (2): 409–410.

Stirling, A. 2007, February. "A general framework for analysing diversity in science, technology and society". *Journal of The Royal Society Interface* 4 (15): 707–719.

Wagner, C. 2008. *The new invisible college: Science for development*. Brookings Inst Pr.

Wasserman, S. 1994. *Social network analysis: Methods and applications*. Cambridge university press.

Wynn, D. 2003. "Organizational structure of open source projects: A life cycle approach". In *7th Annual Conference of the Southern Association for Information Systems, Georgia*.

Yam, B. 2005. *Making Things Work: Solving Complex Problems in a Complex World*. NECSI Knowledge Press.

Yilmaz, L. 2008. "Innovation systems are self-organizing complex adaptive systems". In *Association for the Advancement of Artificial Intelligence*.

Zou, G., and L. Yilmaz. 2011. "Dynamics of knowledge creation in global participatory science communities: open innovation communities from a network perspective". *Computational and Mathematical Organization Theory* 17 (1): 35–58.

## AUTHOR BIOGRAPHIES

**ÖZGÜR ÖZMEN** is a Postdoctoral Research Associate in Computational Sciences and Engineering Division at Oak Ridge National Laboratory. His research interests are Complex Adaptive Systems, Agent-based Simulation, and Verification and Validation of Multi-Agent Systems. He is a member of INFORMS, ACM, and SCS. He holds BIE degree from Yildiz Technical University in Istanbul, Master of Engineering Management degree from Galatasaray University in Istanbul, MISE and Ph.D. Degrees (both in Industrial and Systems Engineering) from Auburn University in Alabama. His email address is ozmeno@ornl.gov.

**JEFFREY SMITH** is a Professor of Industrial and Systems Engineering at Auburn University. He has served as the WSC Business Chair and General Chair and is currently on the WSC Board of Directors. He has a BIE from Auburn University and a MS and PhD (both in Industrial Engineering) from Penn State University. His email and web addresses are <jsmith@auburn.edu> and <http://jsmith.co>.

**LEVENT YILMAZ** is Associate Professor of Computer Science and Software Engineering and holds a joint appointment with the Industrial and Systems Engineering at Auburn University. He received M.S. and PhD. degrees from Virginia Tech. His research interests are in Modeling & Simulation, Agent-directed Simulation, and Complex Adaptive Systems with a focus in advancing the theory and methodology of modeling and simulation and developing computational models of autonomous, self-organizing, and self-adaptive systems. He is currently serving as the Editor-in-Chief of *Simulation: Transactions of the Society for Modeling & Simulation International*. He is member of ACM, IEEE Computer Society, Society for Computer Simulation International (SCS), and Upsilon Pi Epsilon. His email and web addresses are `<yilmaz@auburn.edu>` and `<http://www.eng.auburn.edu/users/yilmale>`.