Modeling the Structure and Effectiveness of Intelligence Organizations: Dynamic Information Flow Simulation

Modeling and Simulation Track
Student Paper

Robert Behrman
Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA 15213
412-268-1876
rbehrman@andrew.cmu.edu

Kathleen Carley
Center for the Computational Analysis of Social and Organizational Systems
Carnegie Mellon University
Pittsburgh, PA 15213
412-268-6016
kathleen.carley@cmu.edu

Abstract

This paper describes the Dynamic Information Flow Simulation (DIFS), an abstract model for analyzing the structure and function of intelligence support organizations and the activities of entities within them. In order to do so, DIFS simulates the flow of tasks and reports between various units (decision makers, collectors, processors (analysts), databases, etc.) and agencies within an intelligence organization. DIFS is a dynamic, discrete, multi-agent, networked simulation. The structure of the simulation – i.e., the types and properties of entities, communication ties, agencies, and operating procedures - is described. The performance metrics used in, equations used in, and assumptions behind its design are discussed. Methods of conducting a virtual experiment using the simulation, output generated, and a strategy for validating the results is given. DIFS is intended to provide a method of analysis of intelligence support effectiveness abstracted from sources, methods, and content. Using this analysis, marginal performance change caused by change in organization structure or policy can be quantitatively modeled. Over- and under-loaded units, units that are not re-tasked when intelligence requirements change, and changes in information flow can be identified and modeled. Finally, the weaknesses, benefits, and additional applications of DIFS and areas where further research is desired are discussed.

1. Introduction

A quotation of Lord Mountbatten that was once paraphrased to me reads (as my memory serves): "Horatio Nelson once said that under the best of circumstances while at sea he could send a message to the Admiralty, by courier, and get a response back in about six weeks … Now, I can dictate a message to my orderly, and have it sent via radio to the Admiralty in minutes … and get a response back in about six weeks." Technological advancement has decreased the amount of time required to communicate with even the most distant parts of the globe. But this technological progress – advancements in communication and computers – is of limited value if not accompanied by corresponding progress in techniques of command, control, and intelligence (Alberts, Garstka, and Stein, 1999).

Decisions facing planners on how to facilitate needed progress in command, control and intelligence are complicated by the inherent difficulty to evaluate, predict, and compare intelligence support organizations (Berkowitz and Goodman, 1989, 2000). This difficulty is rooted in the context and consumer dependent (subjective) nature of intelligence (Kent 1951), the need to protect sources and methods, and the uncertainty of validating information. This paper will describe the Dynamic Information Flow Simulation (DIFS), an abstract, computational, discrete, dynamic, multi-agent, networked simulation of intelligence organizations (Law and Kelton 1991, Carley 2002). DIFS attempts to address the problems described above by abstracting content- or context-dependent information from the consideration of the intelligence organization and instead focusing on communication, cooperation, and coordination between units (nodes) in the organization. By doing so, DIFS can model the marginal effect of changes in organizational design and operating procedures upon information flow within the organization as a whole and through individual nodes in particular. This helps to identify and rank trade-offs between various factors of information flow, which is of particular use to those concerned with organizational design (Carley, 2002). Additionally, simulations designed in DIFS using parameters modeled after actual organizations can help identify potential bottlenecks in information flow – from overused nodes (Law and Kelton, 1994), underutilized resources, and communication inefficiency – and criteria or threshold values that could lead to these bottlenecks.

This model was informed by a combination of ideas from the organization theory and military theory literatures. The concept of abstracted information flow is not new to either theoretical organization theory literature or computational organization simulations. Concepts such as the meta-matrix (Carley 2000, 2001; Krackhardt and Carley, 1998) describe information flow and cognitive load in organizations in terms of agents, knowledge, tasks, and organizations. Social network methods can be used to identify positions with similar importance based on correlation between meta-matrix values. In computer simulation, the garbage can model (Cohen, March, and Olsen; 1972) simulate organizational operation as a black box (or garbage can…) that serves to combine input problems and solutions (knowledge) into organization decisions. OrgAhead, by Kathleen

Carley and Ju-Sung Lee, models organization structure as a self-organizing neural network that attempts to maximize prediction accuracy (organization output) based on an abstract binary block of perceived data (organization input). In military theory literature, abstracted military decision making models such as the O-O-D-A loop model concepts such as initiative and optempo in terms of organizational perception and response to random change and initiated change in the environment.

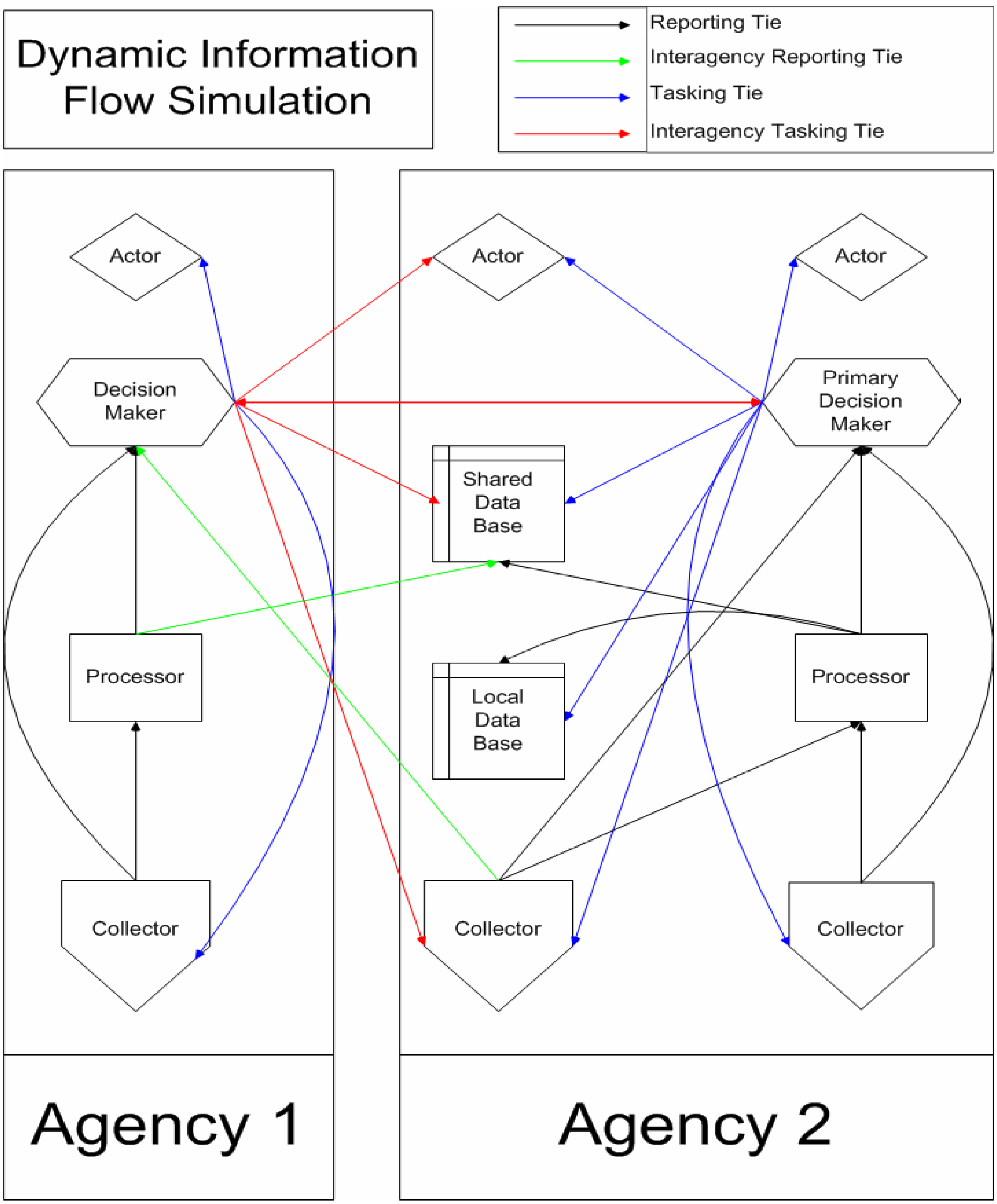

Figure 1: A depiction of a simple intelligence organization modeled in DIFS.

2. Dynamic Information Flow Simulation

The DIFS model defines intelligence organizations as a collection of entities (nodes) and communication links (ties), collected within agencies, that support a number of designated decision makers (which are in themselves entities within the organization). Figure 1, above, depicts a very simple intelligence organization modeled using DIFS. The nodes are depicted as 5 types of geometric shapes, indicating actors, decision makers, databases, processors, and collectors. DIFS is a multi-agent model insofar as all of the separate nodes have individual rule behavior and aggregate rule behavior (agency rules), and simulation output is a product of the report and task handling of the various individual units (Schelling, 1978). It describes networked agents in terms of various, finite, communication ties. All nodes in the model are connected to some other node, which allows the model to be of use in describing intelligence organization assets in relation to each other (Carley, 2002).

DIFS is primarily concerned with the flow of tasks and reports through the organization. Tasks represent collection or information requirements sent to units in the intelligence organization. Reports indicate intelligence or information collected by certain units in the organization and sent between units. Reports have a generic, randomly determined 'quality' value that determines the amount of information contained therein. Simulation rules would require decision makers to receive a certain amount of information of various types in order to make decisions. Communication ties represent tasking or reporting authority, communicated along certain channels. All nodes, ties, and phenomena (tasks and reports) are coded with additional attributes – criteria, sensitivity, priority, etc. These attributes affect phenomenon routing (a decision maker who needs a certain criteria of report will not send the task to a different criteria collector), queuing, and certain performance metrics. Reports are routed according to rule sets, which are defined in agency objects. Agencies determine how phenomena are handled within the agency and between other agencies, which allows the simulation to model organizational command and control differently for various subunits of the organization.

Decision makers are the most important node type in the model. 'Decision maker' nodes are responsible for tasking intelligence organization assets to collect needed information. In the figure above, each agency has a decision maker responsible for tasking actors, processors, databases, and collectors controlled by its agency or another agency. Decision makers can also be intelligence consumers, insofar as they 'read' intelligence reports generated by the intelligence organization and plan tasking requirements accordingly. This intelligence consumption function can be used as a simple metric for intelligence organization effectiveness. As designated in Figure 1, the decision maker in agency 2 is the 'primary' decision maker. If we are concerned with intelligence organization effectiveness in supporting the primary decision maker, we can model this in terms of the number of reports he receives.

Collectors represent the input stream of data into the organization. Collectors can respond to specific tasks or generate random reports (to represent organization perception of environmental events) corresponding to their criteria parameter. Information reports generated by collectors are routed according to agency rules to either processors, databases, or directly to decision makers. Processors represent the analytic functions of intelligence organizations. In the simulation, processors condense various reports into single reports that contain more information (and are thus more valuable for decision makers to read). Databases represent organizational memory – processors or decision makers can task databases to send stored or outdated reports should organizational parameters, routing rules, or interagency cooperation rules require. Actor nodes exist simply as a method of modeling decision maker effectiveness, and are not necessary to the simulation itself. Decision makers nodes can generate and send tasks to actors if they get enough information, and organization effectiveness can be modeled in terms of the number of tasks completed by actors in the organization.

When the simulation is first run at time 0 all of the nodes, communication ties, outputs, etc. are initialized. Decision maker nodes begin making tasks, and other nodes are queried for responses to idle behavior (primarily in order to determine random report generation by collectors). Random report generation is normally distributed with a mean and variance determined by input parameters and a uniformly distributed inter-arrival time with user-input mean. As the simulation progresses, the timer advances itself between each subsequent event generated by some unit in the organization, determined by timing values for node functions and communication links. Reports and tasks move along communication ties between nodes according to routing rules in and between agencies (for example, reports shared between agencies may be incremented or decremented in priority determined by rules existing in the agency.). The simulation ends at a user-defined stopping condition – for example, after a predetermined amount of time, a predetermined number of reports, or a set number of actions by actor nodes.

3. Virtual Experimentation of Intelligence Organizations Using DIFS

Virtual experiments in DIFS can be used to answer a number of questions regarding information flow and efficiency within the organization. For example, one could ask what the effect would be on organization performance if the number of randomly generated reports was varied; if interagency rules governing the priority of extra-agency taskings were varied, or if intra-agency rules governing the handling of high priority reports were varied. Organization performance would be evaluated in terms of three different output measures – intelligence flow to the primary decision maker, overloaded resources, and underused resources.

Intelligence flow to the primary decision maker can be modeled as percent of total information generated sent to decision maker:

$$IQflow = \frac{\sum_{\Omega} x^{\inf o}}{\sum_{i=0}^{n} x_i^{\inf o}}$$

Where n is the total number of reports generated, x^{info} is the amount of information contained in the particular report, and Ω is the set of all reports received by the primary decision maker.

Unused resources are trivially easy to identify – simply compute the percentage of time nodes spend working on tasks over the total simulation time. Unused or underutilized resources will have low values. Overloaded resources are more difficult to identify (as opposed to fully utilized resources, which is desired). Possible metrics for overloaded resources include average queue length:

$$MeanQueue = \frac{\int (Q_0 + Q_1 + ... + Q_k)dt}{t_{total}}$$

Where $Q_0 - Q_k$ equals time spent with queue of length $0 - k$.

Some possible simulation configurations:

	High	Medium	Low
Random report inter-arrival rate	Inter-arrival rate less than processing time	Inter-arrival rate greater than processing time	Infrequent or no random report generation
Interagency task priority	Extra-agency task priority increased	No change in task priority between agencies	Extra-agency task priority decreased
Intra-agency priority report handling	High and medium priority reports are sent directly to decision makers	High priority reports are sent directly to decision maker	No reports are sent directly to decision maker (all go through processor)

The various simulation configurations described above could lead to interesting conclusions for organization design and management. Using standard report routing procedures, as the rate of random report generation increases, past a certain threshold intelligence flow to the primary decision maker could decrease or stay constant as processors get backlogged. However, greater marginal gains in intelligence flow at high random report inter-arrival rates would be experienced for increases in the amounts of reports sent directly to decision makers as cognitive load is shifted from collectors and processors to the decision makers themselves. Similarly, for high inter-arrival rate conditions decreasing interagency task priority might lead to greater intelligence flow within agency 2 by forcing decision makers in agency 1 to query database assets instead of getting reports directly. Multiple simulations can be compared holding different simulation variables constant to model marginal tradeoffs and multicollinearity between various simulation parameters.

4. Validation Strategy, Strengths, Weaknesses, and Conclusion

The DIFS model provides a powerful framework for the analysis of tradeoffs in simulated intelligence flows; however the level of abstraction used in making the model makes some direct application to extant intelligence organizations difficult to model with certainty. DIFS uses many assumptions about report handling, bounds of processor and decision maker rational capability, and organization/environment interface (Lawrence and Lorsch, 1969). Though all of these assumptions lend themselves to sensitivity analysis over multiple simulation runs (DIFS saves random number seeds to be saved with other state/simulation constructor variables), actually modeling corresponding attributes of real-world organizations in terms of the DIFS state variables within the sensitive ranges requires detailed case study and analysis.

In order to provide an initial framework to validate DIFS simulations, case studies of relatively small (or automated) organizations with formalized report contents would allow for convincing practical validation of DIFS. Simulations composed of linked, small, elements validated above could then be compared to historical logs and records of larger real-world agencies to validate the model for larger intelligence organizations. In the end, some difficulty will remain in validating DIFS for certain organizations merely because of the unavailability of validating information or the need to protect methods.

The fact that DIFS does not concretely model decision quality is its primary weakness. DIFS cannot be used to predict whether intelligence organizations will allow decision makers to make correct decisions; nor whether intelligence organizations will provide complete and reliable information. On the other hand, it is not intended to – DIFS instead provides a model for comparison between different intelligence organizations based on their ability to efficiently handle the information with them; not the information itself. Modeling information or decision accuracy would require another wide range of assumptions, would be very difficult to validate, highly uncertain, and possibly even impossible to quantify.

Keeping this in mind, however, one of the main strengths this allows for DIFS is its ability to model information flow through a wide variety of organizations. Careful determination of time, node, and agency parameters can allow DIFS simulations to model information flows between organizations as large as the UK Royal Navy to as small as ad-hoc computer or radio networks. A perfect example is field artillery fire direction – local commanders (subordinate decision makers) task forward observers (collectors) to send information to fire direction centers (processors), which in turn send information to field artillery commanders (decision makers) to task fire missions to gun batteries (actors). Combining this with the ability to compute trade-offs between information policies, DIFS provides a potentially powerful tool for the analysis and understanding of many different types of intelligence support organization, and useful metrics of information flow in any organization.

References

[Alberts, Garstka, and Stein, 1999] Alberts, David S.; Garstka, John J.; and Stein, Frederick P. (1999) *Network Centric Warfare: Developing And Leveraging Information Superiority.* National Defense University Press, Washington, DC.

[Berkowitz and Goodman, 1989] Berkowitz, Bruce D. and Goodman, Allan E. (1989) *Strategic Intelligence for American National Security*. Princeton University Press, Princeton, New Jersey.

[Berkowitz and Goodman, 2000] Berkowitz, Bruce D. and Goodman, Allan E. (2000) *Best Truth: Intelligence In The Information Age*. Yale University Press, New haven, Connecticut.

[Carley, 2000] Kathleen M. Carley (2000) "Smart Agents and Organizations of the Future." Unpublished Draft.

[Carley, 2001] Kathleen M. Carley (2001) "Computational Organizational Science and Organizational Engineering." *Companion To Organizations*, Baum, Joel C. ed. Blackwell, Oxford, UK

[Carley, 2002] Kathleen M. Carley, forthcoming, "Dynamic Network Analysis" in the Summary of the NRC workshop on Social Network Modeling and Analysis, Ron Breiger and Kathleen M. Carley (Eds.), National Research Council.

[Carley, 2002] Kathleen M. Carley, (2002) "Information Technology and Knowledge Distribution in C^3I Teams" in *Proceedings of the 2002 Command and Control Research and Technology Symposium.* Conference held in Naval Postgraduate School, Monterey, CA. Evidence Based Research, Vienna, VA.

[Cohen, March, and Olsen, 1972] Cohen, Michael D.; March, James G.; and Olsen, Johan P. (1972) "A Garbage Can Model of Organizational Choice" in *Administrative Science Quarterly*

[Comfort, 2002] Comfort, Louise K. (2002) "Institutional Re-Orientation and Change: Security As A Learning Strategy." *The Forum*, Volume 1, Issue 2. Berkeley Electronic Press http://www.bepress.com/forum.

[Connolly, Conlon, and Deutsch, 1980] Connolly, Terry; Conlon, Edward J.; Deutsch, Stuart Jay; "Organizational Effectiveness: A Multiple-Constituency Approach." *Academy of Management Review*, Volume 5. Academy of Management.

[Feldman and March, 1981] Feldman, Martha S.; and March, James G. (1981) "Information in Organizations as Signal and Symbol" *Administrative Science Quarterly*, Volume 26, Issue 2.

[Kent, 1951] Kent, Sherman. (1951) *Strategic Intelligence for American World Policy.* Princeton University Press, Princeton, New Jersey.

[Lawrence and Lorsch, 1969] Lawrence, Paul R.; and Lorsch, Jay W. (1969) "Organization-Environment Interface" in *Classics of Organization Theory*, third edition, eds. Shafritz and Ott. Wadsworth Publishing Company, Belmont, California.

[Krackhardt and Carley, 1998] Krackhardt, David and Kathleen M. Carley (1998) "A PCANS Model of Structure in Organizations" Pp. 113-119 in *Proceedings of the 1998 International Symposium on Command and Control Research and Technology*. Conference held in June. Monterey, CA. Evidence Based Research, Vienna, VA.

[Law and Kelton, 1991] Law, Averill M. and W. David Kelton (1991) *Simulation Modeling and Analysis*. McGraw-Hill, New York, NY.

[Liang and Xiangsui, 1999] Qiao Liang and Wang Xiangsui (1999) *Unrestricted Warfare*. PLA Literature and Arts Publishing House, Beijing.

[March and Simon, 1958] March, James G. and Simon, Herbert A. (1958) *Organizations*. John Wiley & Sons, Inc.

[Mizruchi and Galaskiewicz, 1994] Mizruchi, Mark S.; and Galaskiewicz, Joseph. (1994) "Networks of Interorganizational Relations" in *Advances In Social Network Analysis,* eds. Wasserman and Galaskiewicz. Sage Publications, London.

[Perrow, 1994] Perrow, Charles. (1992) "Small-Firm Networks" in *Networks and Organizations: Structure, Form, and Action*, eds. Nohria and Eccles. Harvard Business School Press, Boston, Massachusetts.

[Powell, 1990] Powell, Walter W. (1990) "Neither Market Nor Hierarchy: Network Forms of Organization" *Research In Organizational Behavior,* Volume 12. JAI Press, Inc.

[Simon, 1956] Simon, H., (1956). "Rational choice & the structure of the environment." *Psychological Review,* 63: 129-138.

[Simon, 1955] Simon, H., (1955), "A behavioral model of rational choice." *Quarterly Journal of Economics,* 69: 99-118.

[Schelling, 1978] Schelling, Thomas, (1978), *Micromotives and Macro Behavior*.

[Scott, 2000] Scott, John. (2000) *Social Network Analysis: A Handbook*. Sage Publications, London, England

[Wasserman and Faust, 1994] Wasserman, Stanley and Katherine Faust. (1994) *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.