Dynamic Network Analysis for Understanding Complex Systems and Processes

Irene Pestov
Central OR Team

DRDC CORA TM 2009–020
July 2009

Defence R&D Canada
Centre for Operational Research and Analysis

National Defence Défense nationale

Canada

Dynamic Network Analysis for Understanding Complex Systems and Processes

Irene Pestov
Central OR Team

Defence R&D Canada – CORA

Technical Memorandum

DRDC CORA TM 2009-020

July 2009

Principal Author

Original signed by Irene Pestov

Irene Pestov

Approved by

Original signed by Paul Massel

Paul Massel
Central Operational Research Team Leader

Approved for release by

Original signed by Dale Reding

Dale Reding
CORA Chief Scientist

Abstract

Many complex systems and processes of interest to defence are characterized by varying topologies that can be modeled as Complex Dynamical Networks with nodes representing entities and links representing relationships among those entities. Complex Dynamical Networks give a flexible, powerful and useful representation of structural changes in such systems. Unfortunately, the theoretical aspects of these network classes remain largely unexplored. Dynamic Network Analysis provides an approach and a tool to deal with such cases of practical significance in spite of the lack of theoretical knowledge regarding structural dynamics. The presented technical memorandum conducts a detailed analysis of this approach, its major concept - the metanetwork, and a software tool that implements it. The significance of Dynamic Network Analysis for modelling multi-relational data, visualizing networks, and identifying hidden links is demonstrated on two examples. The first example shows how the metanetwork representation can be used to gain additional insight into the national survey data on complexity research in Canada. The second example illustrates the application of Dynamic Network Analysis to modelling the capability engineering process. A new modelling framework for the identification of capability gaps in terms of people's expertise is suggested in this example. The proposed modelling framework will be useful in large multi-disciplinary S&T projects.

Résumé

Parmi les systèmes et processus complexes qui intéressent la défense, un grand nombre sont caractérisés par des topologies diverses qu'on peut modéliser en tant que réseaux dynamiques complexes dotés de núds qui représentent des entités, et de liens qui représentent les relations entre entités. Les réseaux dynamiques complexes offrent une représentation souple, puissante et utile des changements structurels dans de tels systèmes. Malheureusement, les aspects théoriques de ces classes de réseau sont en grande partie inexplorés. L'analyse des réseaux dynamiques offre une méthode et un outil permettant de gérer de tels cas, qui ont une importance pratique en dépit du manque de connaissances théoriques au sujet de la dynamique structurale. Le document technique présenté contient une analyse détaillée de cette méthode, de son principal concept – le métaréseau – et d'un outil logiciel qui le met en úuvre. On utilise deux exemples pour illustrer l'importance de l'analyse des réseaux dynamiques pour la modélisation des données multirelationnelles, la visualisation de réseaux et la détection des liens cachés. Le premier exemple montre comment la représentation sous forme de métaréseau peut servir à mieux connaître les données d'enquêtes nationales sur la recherche sur la complexité au Canada. Le deuxième exemple illustre l'application de l'analyse des réseaux dynamiques à la modélisation du processus d'ingénierie des capacités. Dans cet exemple, on propose un nouveau cadre de modélisation servant à identifier les lacunes en matière de savoir-faire du personnel. Le cadre de modélisation proposé sera utile dans les grands projets multidisciplinaires de science et technologie.

This page intentionally left blank.

Executive summary

Dynamic Network Analysis for Understanding Complex Systems and Processes

Irene Pestov; DRDC CORA TM 2009-020; Defence R&D Canada – CORA; July 2009.

Background: The presence of many interacting entities is a key feature of complex systems and processes, and networks are well-suited to represent this feature. The nodes in the network can represent system entities, and the links can model relationships among those entities. The variety of network models matches the variety of complex systems. In some network models the node functions may change with time, whereas in the others the network structure may also vary. As far as structural dynamics is concerned, Complex Dynamical Networks give a flexible, powerful and useful representation of many systems and processes of interest to defence. Unfortunately, the theoretical aspects of the majority of Complex Dynamical Networks remain largely unexplored.

The lack of theoretical knowledge regarding underlying network dynamics is a major impediment to understanding the evolution of complex systems. There is clearly a need for tools and models capable to unravel underlying rules of structural dynamics. In the defence settings, where many research tasks are centered at studying adversarial systems, this need becomes even more pressing.

Adversarial systems are characterized by cellular and distributed structures, hidden relationships, and the rules of structural dynamics that are difficult to infer. Dynamic Network Analysis provides a tool and an approach to tackle such systems, in spite of existing theoretical gaps concerning structural dynamics.

Principal results: The aim of this technical memorandum is to investigate the potential of Dynamic Network Analysis (DNA) for defence applications. This work presents a detailed analysis of the DNA approach, its major concept – the metanetwork, and the ORA software tool that implements it. The findings of this work are listed as follows.

DNA is an innovative approach that goes beyond the limitations of classical social network analysis and brings network techniques to the level of the complex systems thinking. The metanetwork concept, which is central to DNA and upon which DNA is based, is a truly complex systems concept. A metanetwork is a set of inter-linked networks each of which consist of multiple types of nodes connected by probabilistic links. The multi-mode multi-plex metanetwork concept is a significant extension of social network techniques.

Structural dynamics is captured in DNA through representing complex systems as a time series of metanetworks and through incorporating social, cognitive, and political processes that govern the system evolution. The ORA software tool operates on metanetworks and on a time-series of metanetworks. It allows users to visualize, analyze and explore metanetworks, and to develop new networks, as we demonstrate on two examples.

The first example shows how the metanetwork representation can help to gain insight into the Canadian complexity research scene. The ORA capabilities are used in this example to determine concentrations of complexity researchers with respect to different locations across Canada and the areas of research strength; to identify the major contributors to the field of Complexity and the researchers with unique skill sets of interest to defence; and to create and analyze collaboration networks which engage different modes of collaboration.

The second example illustrates the application of DNA and ORA to modelling the capability engineering process. A new modelling framework for the identification of capability gaps in terms of people's expertise and tools, which are needed to develop and evolve a capability during the course of the capability engineering process, is suggested in this example. The proposed framework can be easily extended to accommodate the full scope of the capability engineering process through addition of such node types as stakeholders, resources and events. The resulting metanetwork will provide a comprehensive picture of what comprises a capability. A time series of such metanetworks will represent the evolution of a capability from its current 'as-is' state to desired 'to-be' state. The task of detecting capability gaps in available expertise, tools or resources will be reduced to automated matrix algebra operations on networks.

Significance of results: The analysis results reveal the significance of DNA and ORA for defence and security applications, especially for visualization and analysis of multi-relational data and for identification of hidden or missing links in massive data sets. The proposed modelling framework will facilitate planning and execution of capability engineering projects in the S&T project settings.

Future work: This technical memorandum is the first in a series that reports on the results of the feasibility study into computational and modelling approaches pertaining to complex defence systems and operations. DNA is one of three closely-related groups of quantitative approaches identified by the study. Approaches from two other areas of quantitative expertise, statistical machine learning and multi-agent technologies, will be discussed in the works that follow.

Sommaire

Dynamic Network Analysis for Understanding Complex Systems and Processes

Irene Pestov ; DRDC CORA TM 2009-020 ; R & D pour la défense Canada – CARO ; juillet 2009.

Contexte : La présence de bon nombre d'entités en interaction constitue une des caractéristiques clés des systèmes et processus complexes, et les réseaux conviennent bien à la représentation de cette caractéristique. Les nœuds du réseau peuvent représenter des entités de système, et les liens peuvent modéliser les relations entre entités. La diversité des modèles de réseaux correspond à la diversité des systèmes complexes. Dans certains modèles de réseau, les fonctions du nœud peuvent évoluer au fil du temps, tandis que dans d'autres, la structure même du réseau peut aussi être modifiée. Sur le plan de la dynamique structurale, les réseaux dynamiques complexes offrent une représentation souple, puissante et utile des changements structurels et la représentation utile de bon nombre de systèmes et de processus qui intéressent la défense. Malheureusement, les aspects théoriques de la majorité des réseaux dynamiques complexes sont en grande partie inexplorés.

Le manque de connaissances théoriques au sujet de la dynamique des réseaux constitue une barrière considérable, qui empêche de comprendre l'évolution des systèmes complexes. On a clairement besoin d'outils et de modèles permettant de dégager les règles qui sous-tendent la dynamique structurale. Dans le contexte de la Défense, qui concentre bon nombre de ses activités de recherche à l'étude des systèmes de l'adversaire, ces besoins prennent encore plus d'importance.

Les systèmes de l'adversaire comprennent des structures cellulaires et réparties, ainsi que des liens cachés, et les règles de la dynamique structurale sont difficiles à deviner. L'analyse des réseaux dynamiques présente un outil et une méthode permettant de s'attaquer à de tels systèmes en dépit des lacunes théoriques qui existent au sujet de la dynamique structurale.

Principaux résultats : Le but de cette note technique est d'étudier la possibilité d'utiliser l'analyse de réseaux dynamiques (ARD) dans le contexte de la Défense. Ce document présente une analyse détaillée de la méthode d'ARD, de son principal concept – le métaréseau – et de l'outil logiciel de l'ORA qui l'utilise. Voici les conclusions de ces travaux.

L'ARD est une méthode novatrice qui dépasse les limites habituelles de l'analyse traditionnelle des réseaux sociaux et qui amène les techniques réseau sur le plan de la conception de systèmes complexes. Le concept de métaréseau, qui est au cœur et à la base de l'ARD, est un concept de systèmes vraiment complexe. Le métaréseau est un ensemble de réseaux interreliés dont chacun consiste en de multiples types de nœuds connectés par des liens probabilistes. Le concept du métaréseau multimode et multiplexe constitue un dérivé important des techniques de réseaux sociaux.

Dans l'ARD, on représente la dynamique structurale en considérant les systèmes complexes comme une série chronologique de métaréseaux et en y incorporant les processus sociaux, cognitifs et politiques qui régissent l'évolution des systèmes. L'outil logiciel de l'ORA s'utilise avec les méta-

réseaux et avec des séries chronologiques de métaréseaux. Il permet aux utilisateurs de visualiser, d'analyser et d'explorer les métaréseaux et de développer de nouveaux réseaux, comme l'illustrent les deux exemples.

Le premier exemple montre comment la représentation sous forme de métaréseau peut aider à cerner le domaine de la recherche de complexité au Canada. Dans cet exemple, on utilise les capacités de l'ORA pour connaître combien il y a de chercheurs sur la complexité à divers endroits au Canada et savoir quels sont les points forts de cette recherche ; de connaître le nom des personnes qui contribuent de manière importante au domaine de la complexité et les chercheurs ayant des compétences particulièrement prisées par la défense ; et, enfin, de créer et d'analyser les réseaux de collaboration qui font appel à divers modes de travail collaboratif.

Le deuxième exemple illustre l'application de l'ARD et de l'ORA à la modélisation du processus d'ingénierie des capacités. Dans cet exemple, on propose un nouveau cadre de modélisation servant à déceler le savoir-faire et les outils qui manquent au personnel et qui sont essentiels à la mise en place et à l'avancement des capacités dans le cadre du processus d'ingénierie des capacités. Le cadre proposé peut être élargi facilement. On peut y intégrer la pleine portée du processus d'ingénierie des capacités grâce à l'ajout de types de nœuds, par exemple les intervenants, les ressources et les événements. Le métaréseau ainsi créé offrira un portrait complet des éléments qui constituent une capacité. La série chronologique de tels métaréseaux illustrera l'évolution d'une capacité, de son état actuel à l'état souhaité. La tâche qui consiste à déceler les lacunes au niveau du savoir-faire, des outils ou des ressources sera réduite grâce à l'exécution automatique d'opérations d'algèbre matricielle dans les réseaux.

Importance des résultats : L'analyse des résultats révèle l'importance que représentent l'ARD et l'ORA pour les applications de défense et de sécurité, surtout pour la visualisation et l'analyse des données multirelationnelles et le repérage de liens cachés ou manquants dans d'énormes ensembles de données. Le cadre de modélisation proposé facilitera la planification et l'exécution des projets d'ingénierie des capacités dans le contexte des projets de science et technologie.

Recherches futures : Cette note technique est la première d'une série de rapports sur les résultats de l'étude de faisabilité des méthodes de traitement informatique et de modélisation qui ont trait aux systèmes et opérations complexes de défense. L'ARD est l'un parmi trois groupes de méthodes quantitatives étroitement liés révélés par l'étude. Dans les recherches à venir, on abordera les méthodes issues de deux autres domaines du savoir-faire quantitatif, de l'apprentissage machine par statistiques et des technologies multi-agent.

Table of contents

List of figures

List of tables

Acknowledgements

The author thanks Paul Massel, Central OR Team Leader, for his continuing support to this research effort. The guidance on defence-related CAS applications from Paul Comeau, DRDC Director S&T Integrated Capabilities, is gratefully acknowledged.

This page intentionally left blank.

1 Introduction

The current interest in networks is part of a broader movement towards research on complex systems.

Steven H. Strogatz "Exploring Complex Networks"

Almost every definition of a complex system that can be found in the scientific literature contains a statement about interactions among system components, entities, or players [1]. Complex social systems, in particular, can be defined as those characterized by dynamical interactions between diverse individual actors [2]. Individuals interact with each other through formal and informal networks, join new networks, and establish new relationships within those networks. As a result of these structural alterations, their social systems change and evolve.

The presence of many interacting components is indeed a prominent feature of complex systems and processes. If we look for a model to reflect this feature, then we should look no further than networks. Undoubtedly, networks are well-suited models for representing complex systems and processes: the nodes in the network can have functions assigned to them to model behaviour of various entities; and the links can represent relationships among those entities. The variety of models that networks offer matches the variety of complex systems: from gene regulation to human brain to social and organizational systems.

A use of networks as models of complex systems and processes brings a number of advantages. Firstly, networks have a convenient graphical representation: network nodes and links can be visualized as graph vertices and edges respectively. Visualization of networks as graphs has value in itself for understanding intricate relationships within complex systems. Secondly, graph theory [3], a discipline that studies graphs, is well-developed and can provide a solid theoretical foundation to such new areas as complex social systems. Thirdly and most importantly, graph measures and computational tools are readily available, thanks to generations of mathematicians and computer scientists who contributed to graph theory and its applications.

Graph measures, such as for example centrality measures, provide important information about the properties and function of an entity in a network based on the number of edges connected to a vertex which represents that entity. The nature of connections is explored through quantitative measures representing the number of connections to influential nodes, connections between cliques, boundary spanners, and so on. There are also graph measures available which provide performance indicators for the network as a whole (e.g. shortest path length as an indicator of the average communication speed within a network). The representation of networks by graphs allows to better understand the interdependencies within the network structure and the global effects which may arise from them. (Examples of graph measures can be found in Annex A.)

It is important to realize, though, that networks are more than their graphical representations, as a graph is only a snapshot of a network at a particular moment in time. Unlike static graphs, networks are dynamic creatures. In some networks the node functions vary with time, whereas in the others the structure may also change.

For instance, the dynamical network models of Random Boolean Networks (RBN) and Artificial Neural Networks (ANN) have been broadly used for modelling complex systems which show rich dynamical behaviour but the wiring of which can be assumed fixed. The RBN and ANN assume a fixed structure consisting of prescribed numbers of nodes and links. The dynamics of these networks is expressed through functional interdependencies between the network nodes. In the RBN, each node has a randomly generated Boolean function assigned to it, which depends on inputs from other randomly chosen nodes [4]. In the ANN, interconnected artificial neurons receive inputs from a prescribed number of other neurons in the network [5]. Some limitations of the RBN are lifted here: the neural network architecture is not randomly assigned, and the function choice is not limited to the Boolean functions. As a result of a greater flexibility, the ANN can learn and adapt. While in both cases the topology of a network remains fixed, its state and the functions of its nodes change with time.

There were recent attempts to incorporate structural changes into the above models. In [6], a new network class called Generative Network Automata (GNA) was proposed for modelling complex systems which change their states and topologies according to a set of generative update rules. In this context, ‘generative’ means that structural changes are driven by local network dynamics, not by external perturbations. In the GNA, states and topology are updated through a rewiring algorithm using binary dynamics similar to the RBN and random extraction mechanism. Computer simulations of all possible replacement mechanisms determined by the Boolean dynamics reveal a large diversity of complex systems that may be modeled using the GNA framework [6].

In real social networks, structural changes may occur due to internal dynamical update rules, as in the case of the GNA [6], and also in response to external perturbations (e.g. attacks or interventions). Both internal and external factors influence the structural dynamics of real networks: links appear and disappear, and nodes become or cease to be members of networks. In response to structural alterations, the node functions also change. Such phenomena as shifting hubs have been observed in real social networks. As shown in [7], roles in large social networks change dramatically from day to day. This suggests that interventions which target hubs may become less effective. The underlying structural dynamics has been now recognized as one of the most important factors determining the evolution of social networks.

The networks with varying topologies are often referred to in literature as Complex Networks or Complex Dynamical Networks [8, 6]. (The latter term is used in the presented work.) Complex Dynamical Networks (CDN) combine structural and dynamical complexity, as they dynamically change their architecture (i.e. structural diagrams as represented by graphs) and hence their state due to both external and internal factors. Structural changes in the CDN may occur in the form of dynamical addition and removal of components (nodes) and their interactions (links). Some classes of the CDN are well studied and understood, whereas the majority of the CDN still present a challenge to mathematicians [8].

Two relatively well-studied classes of the CDN are small-world and scale-free networks. Both classes allow dynamical alterations of structural components, but the rules of attachment are different. In a small-world network, new links are added to *randomly* chosen pairs of nodes. Whereas in a scale-free network, new links appear based on the *preferential* attachment rule when links are added to the nodes which already have more links than the other nodes.

A small-world network is constructed as follows. Let us take a regular lattice in which each node is connected to its nearest neighbours within a specified radius. Such a network is characterized by long characteristic pathways (i.e. a typical distance between any two nodes) and high clustering coefficients (i.e. clustering of a typical neighbourhood). Then we add a small number of links to randomly chosen pairs of these locally clustered nodes with some assigned probability. As observed in [9], for a broad interval of probability values the characteristic pathway falls dramatically, whereas the clustering coefficient remains relatively high. The resulting effect is known as the small-world phenomenon. The network of film actors (where the nodes are actors and the links are appearances in the same movie) is an example of the small-world network [9]. According to [10], short-term memory may use the small-world network between neurons.

A scale-free network is constructed by progressively adding nodes to an existing network and introducing links to existing nodes with preferential attachment, so that probability of linking to a given node is proportional to the number of existing links the node already has. The scale-free topology has been introduced by Barabási and Albert in [11]. Examples of scale-free networks include the Internet (where the hubs are popular sites such as Google), peer-reviewed scientific publications (where the hubs are prolific authors with many publications), and Hollywood actors (where the hubs are stars with many film appearances) [8].

There are many spectacular examples of small-world and scale-free networks [8], but many more can be found falling outside the scope of these two classes of the CDN. As pointed out in [12], many real networks are neither small-world nor scale-free but something in between. In other words, the underlying rules that determine structural dynamics in such networks are neither random nor preferential but something yet to be unraveled by network theory. As of today, the majority of the CDN classes remain largely unexplored. The lack of theoretical knowledge regarding structural dynamics of the CDN is a major impediment to understanding the evolution of complex systems. There is clearly a need for tools and models capable to unravel hidden rules/patterns by which systems evolve. In the defence settings, where many research tasks are centered at studying adversarial organizations [13], this need becomes even more pressing.

Adversarial organizations have network structures that are cellular and distributed; their members can use multiple aliases; and the information that is available about their activities may be intentionally misleading [13]. Nonetheless, there should be a way to exploit their weaknesses, to develop isolation strategies, and to destabilize their networks. What we need is a tool that will allow us to handle systems and processes characterized by a multitude of intricate relationships, hidden links, and complex structural dynamics. What we also need is an approach that will help us to overcome existing theoretical gaps concerning the CDN classes and their underlying structural dynamics. Dynamic Network Analysis (DNA) promises to provide such a tool and an approach.

The aim of this study is to evaluate the potential of DNA for defence and security applications. In Section 2, we review the DNA approach, its major concept – the metanetwork, and the ORA tool which implements it. Two examples of interest to defence are presented in Section 3. The first example illustrates the application of DNA and ORA to the survey data on complexity research in Canada (Section 3.1). The second examples shows how the capability engineering process can be modeled using the metanetwork representation. A new modelling framework for the identification of capability gaps in terms of people's expertise is outlined in Section 3.2.

2 Dynamic Network Analysis

Dynamic Network Analysis (DNA) has been proposed in the works of Kathleen Carley [14, 15, 16] for analyzing complex social networks with varying levels of uncertainty. Although the original purpose of DNA was to tackle terrorist and other covert networks, the approach is generally applicable to a broad range of complex systems and processes, including social and organizational networks, alliances, financial systems, supply chains, semantic networks, and many more. DNA has already been tested on many practical problems and the results are encouraging [16].

The following sections review the DNA approach and give pointers to its practical implementations. Section 2.1 begins with a discussion of the concept of the metanetwork, which is central to DNA and upon which DNA is based. The discussion proceeds to Section 2.2, where the similarities and differences between DNA and classical Social Network Analysis are examined. Section 2.3 outlines the ORA tool that implements DNA techniques.

2.1 The Concept of a Metanetwork

The concept of a metanetwork has been introduced by Kathleen Carley in [14, 15]. Carley has combined techniques from cognitive science, operations research and social theory to derive a concept that employs multiple types of nodes and links of varying levels of uncertainty. In the metanetwork, the nodes could be people, knowledge, resources, locations, etc., and the links that connect them could be of different strength/weight. In fact, the metanetwork is a system of inter-linked networks, each of which represents a system of interacting entities.

The advantages of the metanetwork representation are clearly identifiable. In the metanetwork, 'hidden' relationships between entities can be identified through analyzing their connections to nodes of different kinds. For example, attendance at the same event, as reflected by links between the node representing that event and the nodes representing individuals who attended the event, implies that there could be a relationship between those individuals as well. Even more significantly, changes in one network cascade through the entire metanetwork. When one node is isolated or some other changes occur in one network, it is immediately reflected in other networks. A broader range of isolation strategies and destabilization methods can be identified and effectively tested using the metanetwork concept.

Extended graph measures, which the metanetwork brings into play, are valuable additions to the analysts' toolkit, as they use data drawn from multiple networks simultaneously. Cognitive demand is an example of an extended graph measure. It is the measure of the effort an individual has to employ to hold their role within the group. Cognitive demand takes into account all information regarding that individual across the entire metanetwork. According to [16], cognitive demand could be a good predictor of emergent leadership. Other extended graph measures include exclusivity measures which determine uniqueness of an entity in terms of access to knowledge or resources. For instance, an entity with unique expertise in weapons (as indicated by an agent-versus-knowledge link) and with access to explosive materials (an agent-versus-resource link) could be a far more valuable target than an entity with high degree centrality in the social network (which could be just a secretary and hence easy to replace).

The probabilistic treatment of relationships between entities is another advantage of the metanetwork representation. Whereas in social networks the links are often deterministic (0 or 1, depending on whether or not the relationships exist), in the metanetworks the links are assigned different weights to represent varying levels of uncertainty in our knowledge about the relationships. (For example, 0.5 means 50% chance that the link exists.) In this way, the metanetwork representation addresses the problem of incomplete, inaccurate, or intentionally-misleading information. There is no doubt that many analysts who work with terrorist and other covert networks will find this feature useful. As pointed out in [13], the nature of relationships in the adversarial networks is such that the reliability of apparent links and the probability that hidden links may exist need to be addressed.

2.2 Dynamic Network Analysis versus Social Network Analysis

DNA is an extension and generalization of classical Social Network Analysis (SNA) [17]. Both methodologies are well-founded in graph theory, as both employ graph-theoretical measures for determining key players and network vulnerabilities. But DNA goes beyond that. Unlike classical SNA which deals with one node type (usually people) and deterministic ties, DNA employs metanetworks, inter-linked multiple network representations of complex systems in which nodes of different types are linked by ties of varying probability (see Section 2.1). The utilization of metanetworks in DNA is a significant generalization of classical SNA. On the basis of this generalization new extended graph measures have been developed to provide better insight into functional characteristics of complex systems [16]. The probabilistic treatment of relationships addresses problems of missing information and hidden structural dynamics.

As Carley warns in [16], isolating a key actor today does not mean that the network will be destabilized tomorrow. According to Carley [16], what we have today is only a sample of the real network. Classical SNA is limited to representing networks by static graphs which reflect the states of social systems at one point in time. To overcome this limitation, DNA represents social systems/ groups/networks as a time series of metanetworks. Each metanetwork in a time series represents entities and probabilistic connections between them at one particular moment in time. In this way DNA handles dynamical networks that evolve, change and adapt [16].

In addition to the time-series representation, the network dynamics is captured in DNA through incorporating a whole range of basic governing processes. As Carley puts it in [16], there are several basic processes that influence structural dynamics in social networks. These processes are social, cognitive, and political in nature. The concept of the metanetwork helps to specify these processes, and then to set up agent-based simulations to generate structural changes. When combined with agent-based simulation technologies, the nodes in the metanetwork become active adaptive agents. They can learn and acquire new knowledge, take part in events, do tasks, gain access to resources and through their actions alter their networks. The social and other networks included in the metanetwork co-evolve. In this way, DNA informs ABM by providing essential information for model development. The fact that DNA can provide a framework for modelling the evolution of complex systems is the key distinction of DNA. The distinctions between DNA and SNA are summarized in Table 1.

DNA	SNA
Based on graph theory	Based on graph theory
Multiple types of node	One type of node
Probabilistic links (weighted between 0 and 1)	Deterministic links (weighted 0 or 1)
Extended graph measures	Graph measures
Time-series of metanetworks representation	Single network representation
Incorporates social, cognitive and political processes	Incorporates social processes
Linked with agent-based models	—

Table 1: *A comparison of DNA and classical SNA*

2.3 ORA – the Dynamic Network Analysis and Visualization Tool

The Organizational Risk Analyzer (ORA) has been developed in the Center for Computational Analysis of Social and Organizational Systems (CASOS) by Kathleen Carley, who is the center director, and by her team of researchers and graduate students. It was originally conceived as a risk assessment tool for identifying individuals of potential risks to groups or organizations. ORA implements the DNA techniques to identify those entities (individuals, knowledge, tasks or resources) that are critical from a performance and security perspective [18].

ORA operates with metanetworks and takes advantage of extended graph measures. It combines techniques, methods and algorithms from graph theory, machine learning, operations research, social psychology, probability theory, and matrix algebra. With its advanced GUI, it is possible to visualize, analyze, and explore metanetworks, develop new metanetworks, and test isolation strategies. Free for research purposes, the complete ORA suite can be downloaded from [19].

The graph-drawing algorithms, which are employed by ORA, allow the display of multi-colour multi-linked metanetworks. Different node types are depicted by different colors. Different node sizes and colour schemes can be used to visually represent node characteristics and the roles they play in the network. The links can be directed or undirected, and the link colour can match the source or target node color. One can zoom and rotate graphs, group and isolate nodes, create metanodes, add and remove links, and publish figures. The ORA visualization capabilities are illustrated in Sections 3.1 and 3.2.

ORA can run and publish reports on various aspects of network performance in a convenient HTML format. The ORA reports include Key Entity, All Measures, Sphere of Influence, Missing Links, Immediate Impact, and many more. ORA reports provide quantitative information about key players, emergent leaders, impacts of interventions, and network performance (e.g. average communication speed and network density). An example of the ORA Key Entity Report is given in Annex A.

The ORA data management tool can be used to perform transformations of networks and to develop new networks from the networks already included in the metanetwork. New networks are obtained using matrix algebra operations such as multiplication, subtraction, addition, and transposition. The features of the ORA matrix algebra tool are described in Annex B.

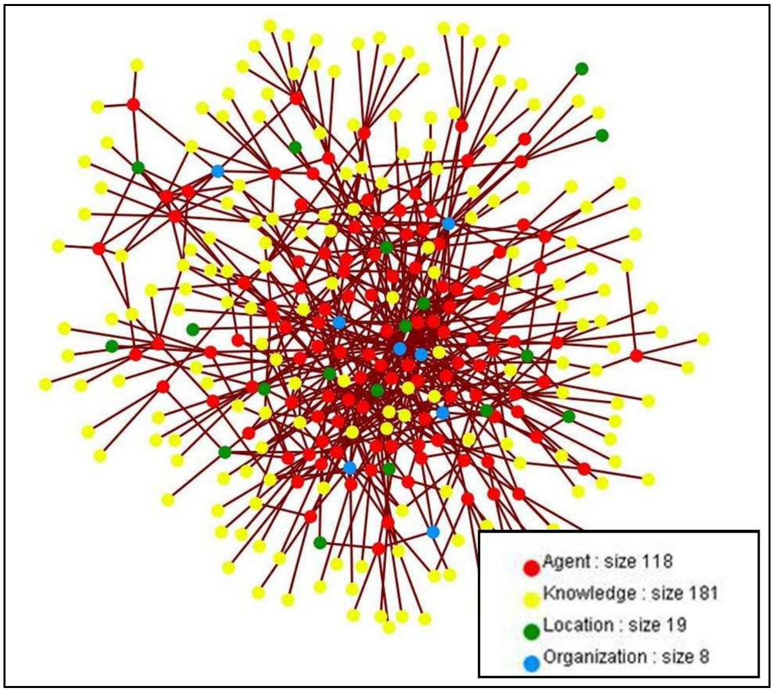


Figure 1: *Complexity research metanetwork*

3 Modelling Systems and Processes with Metanetworks

Practically any complex system can be represented by metanetworks, and many illustrations of their use can be found in the literature. The following examples demonstrate the application of DNA and ORA. The first example deals with the survey data on the complexity research in Canada [20]. The analysis and visualization results are presented in Section 3.1. Section 3.2 illustrates the application of DNA and ORA to modelling the capability engineering process. A new modelling framework for the identification of capability gaps in the S&T project settings is developed in this example.

The application of DNA and ORA to analyzing terrorist networks can be found in [21], [22], and [23]. In [21], the DNA approach to assessing destabilization strategies is outlined. The work [22] discusses challenges arising from a cellular distributed structure of covert networks. In [23], the ORA Network Generator tool is used to identify the C2 structure of an adversarial group. A demonstration of the ORA capabilities using an example of a real terrorist network is given in [18].

3.1 Complexity Research in Canada

This example illustrates the application of DNA and ORA to data collected during a survey of complexity research in Canada. The survey data contains names, contact details and the information about research interests of 118 Canadian researchers [20]. It has been processed using Excel and loaded into ORA. The resulting Canadian complexity research metanetwork is shown in Figure 1.

	Researchers (Agent)	**Research Interests** (Knowledge)	**Cities** (Location)	**Groups/Societies** (Organization)
Researchers	Collaboration net Who collaborates with whom	Knowledge net Who has interest in what	Location net Who resides where	Membership net Who is a member of what

Table 2: *Complexity research network data*

There are four types of nodes and three inter-linked networks in the complexity research metanet shown in Figure 1. The node types are: individual researchers (ORA class 'agent'), areas of research interests (ORA class 'knowledge'), Canadian cities (ORA class 'location'), and scientific groups and societies (ORA class 'organization'). The included networks are: knowledge network which links researchers and research areas, membership network which shows affiliations of researchers with groups and societies, and location network which gives a geographical layout of the Canadian complexity research scene. The network data that is required for the development of the complexity research metanet is summarized in Table 2.

The metanetwork shown in Figure 1 gives a multi-mode multi-plex representation of the Canadian complexity research scene. As stipulated in [18], 'multi-mode multi-plex' networks are those that connect entities of different types. For instance, if a researcher lists machine learning as her research interest and if she resides in Ottawa, then there is a link between a red node, which represents that researcher, and a yellow node, which represents machine learning in the knowledge net. There is also a link between the red node and a location node (shown in green in Figure 1), which represents Ottawa in the location net. The researcher in question might be also linked to one or several blue nodes which represent scientific groups in the membership net, if she is a member of those groups. The multi-mode multi-plex representation brings additional insights into the complexity research data, as we shall see further.

According to Table 2, it is also possible to link researchers in a 'collaboration' network to represent collaboration potential. In fact, multiple collaboration networks are possible, depending on the criteria for possible collaboration. Collaboration networks do not have to be provided with input data, but can be developed using the knowledge and membership nets shown separately in Figure 2 and Figure 3. Before proceeding further with constructing collaboration nets, let us have a close look at these two networks.

There are 181 knowledge nodes in the knowledge network shown in Figure 2. Knowledge nodes represent areas of research interests, as described by researchers in personal websites. (Original descriptions are preserved in this analysis.) The largest proportion of knowledge nodes is related to network topics such as bio-molecular networks, communication networks, neural networks, and network analysis. Machine learning topics are the second largest. They include statistical machine learning, pattern recognition, complexity of large sets of data, information retrieval, and cluster analysis. A significant proportion of research topics is related to multi-agent technologies. These topics include artificial life, cellular automata, automata theory, and combat modelling. (For complete listing of research areas see [20].) In terms of individual disciplines, mathematical biology ranks highest as dominant knowledge, while neural networks and machine learning rank second highest. Ranking of research areas is derived using the ORA Key Entity report (Key Entity - How) and the details can be found in Annex A.

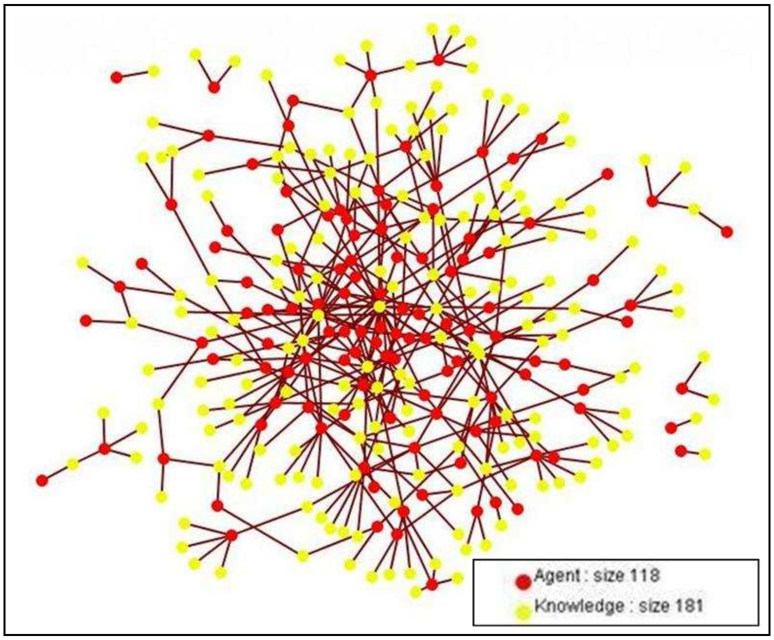


Figure 2: *The knowledge network: researchers versus research interests*

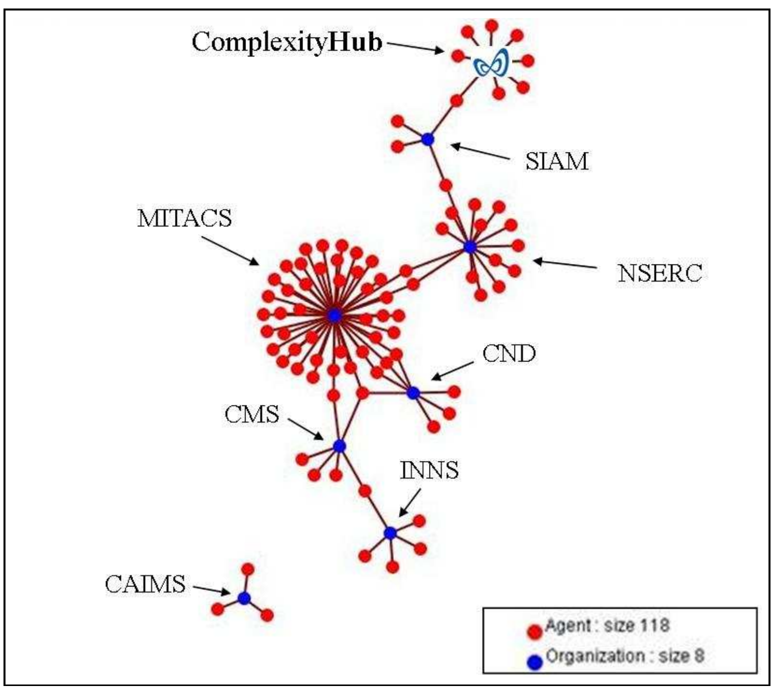


Figure 3: *The membership network: researchers versus research groups*

In Figure 2, hubs correspond to researchers with interests in many areas of Complexity Science, whereas dyads identify individuals specializing in a unique area. The ORA Key Entity report provides ranking of researchers with respect to a number of connections to knowledge nodes (row degree centrality). Two researchers rank highest in the row degree centrality measure with 8 connections each (Annex A). The Key Entity report also identifies researchers with connections to the knowledge nodes that are unique in the network, and ranks them with respect to the knowledge exclusivity measure (Annex A). (Knowledge exclusivity is proportional to the number of connections between a particular researcher and knowledge nodes which no other researcher in the knowledge network has [18].) This information is important for identifying potential S&T providers with unique skill sets.

Major complexity research groups and societies are the hubs in the membership network shown in Figure 3. The Mathematics of Information Technology and Complex Systems centers of excellence (MITACS) is the main scientific body for complexity research activities in Canada. The NSERC hub represents an intersection of the NSERC Canada Research Chairs network with the complexity research metanetwork. Its members are Canada Research Chairs who work in the areas of science related to Complexity. It is the second largest organizational node in the membership network. *ComplexityHub* [24] is an emerging on-line collaboration group for the defence complexity research community and the third largest organizational node. The international society nodes in the membership network are the intersections of the Canadian complexity research metanetwork and corresponding society memberships.

In this example, two collaboration networks have been developed from the knowledge and membership networks respectively using the ORA Data Management tool [18]. The first potential collaboration network has been obtained from the knowledge network, based on the assumption that researchers are likely to collaborate if they have at least one common research interest. The second potential collaboration network has been derived from the membership network based on the assumption that researchers collaborate through scientific groups they belong to. In both cases, potential collaboration networks have been obtained from the existing networks by 'folding' them. The folding operation is automatically performed by ORA, and the process is described in Annex B. In the first case, a link between two individuals has been added if they are connected to the same knowledge node in the knowledge network. In the second case, two individuals have become connected if they are linked to the same organizational node in the membership network. The resulting potential collaboration networks are shown in Figure 4 and Figure 5 respectively.

In Figure 4, node sizes represent the total degree centrality of individual researchers. (Total degree centrality is the number of direct connections, which a particular researcher has in the network, divided by the number of all possible connections [18].) Researchers with broader research interests in many areas of Complexity Science are shown by larger nodes, and researchers with interests in fewer areas are shown by smaller nodes. Isolates in Figure 4 represent researchers with unique research interests. In this context, isolates are important and should be preserved in data sets. Although the total degree centrality for the isolates is zero, they rank high in extended graph measures of knowledge exclusivity (Annex A). Total degree centrality values for this network vary between 0 and 0.453. The expected value for a random network of the same size and density is 0.109 or almost four times smaller, which means that random network models are not applicable to this system.

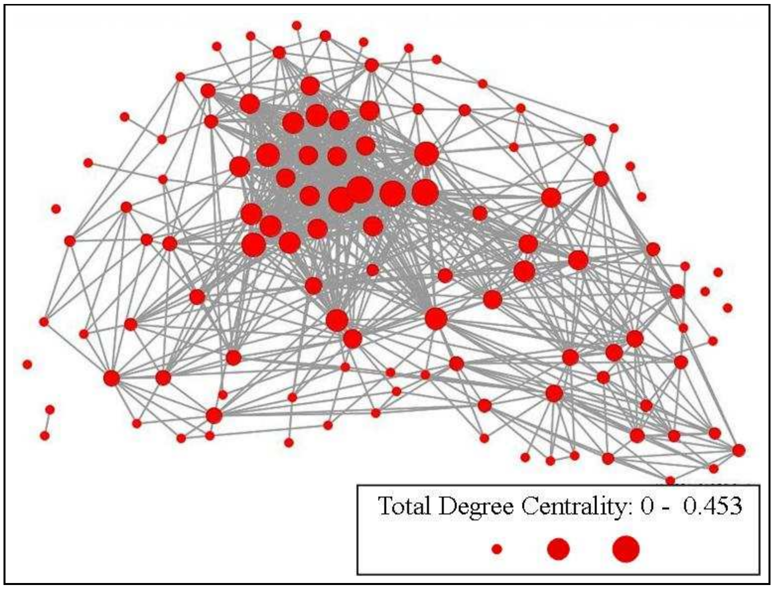


Figure 4: *Potential collaboration by research interest*

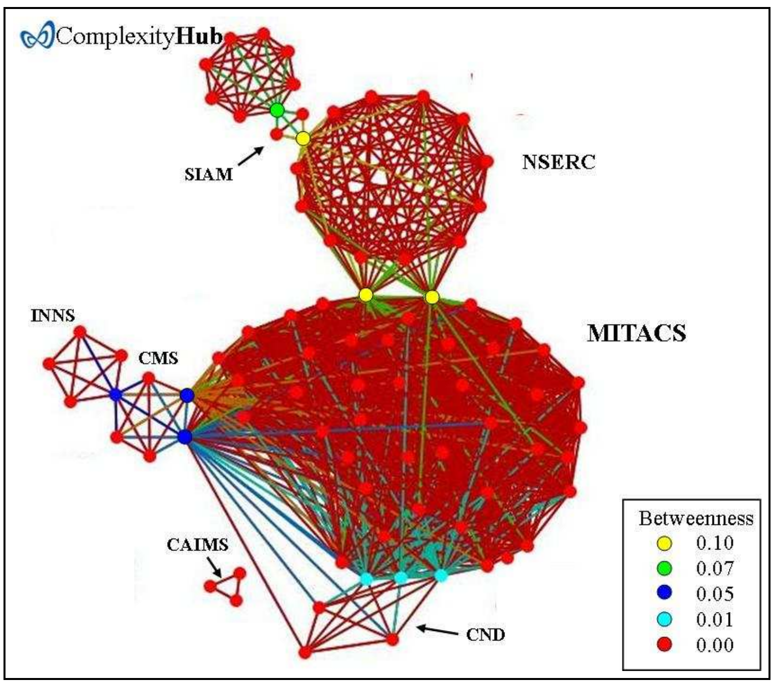


Figure 5: *Potential collaboration by membership*

In Figure 5, nodes are coloured with respect to their betweenness centrality values. (Betweenness centrality is proportional to the number of connections of the shortest length across all node pairs in a network that pass through a single individual node [18].) High betweenness centrality values indicate which nodes are the most connected to other parts of the network. There are three large and five small cliques in this network. All cliques are tightly connected internally but have a few outside links. Nodes with high betweenness values connect different cliques. They are gatekeepers of information flow, influential nodes that hold the network together. The maximum betweenness value in the collaboration-by-membership network is around 0.10 or ten times higher than the expected value of 0.01 for a random network of the same size and density. This is another indication that random network models are not applicable to the system considered in this example.

In both collaboration networks, the same researchers rank highest as emergent leaders in the extended graph measure of cognitive demand (Annex A). (Cognitive demand measures the total amount of cognitive effort expended by each agent to do its tasks including the number of other agents with whom this agent interacts and the number of tasks / resources / knowledge areas managed by this agent [18].) The emergent leader in the collaboration-by-interest net coincides with the person who has the highest degree centrality in that network. However, these roles are held by two different individuals in the collaboration-by-membership net, which the fact stresses the distinction in the nature of collaboration.

A comparison of network performance shows that the collaboration-by-membership net outperforms the collaboration-by-interest net in terms of communication speed, although the fragmentation of the former is about 4 times higher than that of the latter (see Annex A, Performance Indicators). (Communication speed and fragmentation as measures of network performance are defined in Annex A.) There are more entities that are disconnected in the collaboration-by-membership net. However, the average communication speed in this network is 16% higher: 0.474 as compared to 0.409. The following explanation of the above fact can be suggested. The nodes with high betweenness centrality values that connect tightly wired cliques in the collaboration-by-membership net serve as shortcuts for information flow. In the collaboration-by-interest net, such nodes are absent and information has to follow longer paths to reach distant parts of the network. One may conclude that the wiring diagrams similar to the collaboration-by-membership net will be more efficient in transmitting information across large networks.

In addition to the above quantitative measures and performance indicators, the ORA Key Entity report computes group sizes (Agency Size), determines concentrations of complexity researchers with respect to different locations across Canada (Key Entity - Where), identifies areas of research strength (Key Entity - How), and detects leaders of strong cliques who have high eigenvector centrality and boundary spanners who have high betweenness and low degree centrality (see Annex A). This information provides valuable insights into the Canadian Complexity Research scene and helps to answer questions such as: who are the major contributors to the field of Complexity in Canada, what are the major research groups of interests, are there emerging scientific leaders or groups, and how might the collaboration networks look like when different modes of collaboration are engaged. The summary of the Key Entity report is given in Annex A.

3.2 A Network Perspective on Capability Engineering

Capability Engineering (CE) is a complex process that applies the system-of-systems approach to support planning and acquisition within the Department of National Defence [25]. According to [25], the CE goals are to facilitate strategic readiness and responsiveness in capability-based planning and to transform defence acquisition into the agile evolutionary process. A reduced timescale for strategic planning necessitates the ability to respond to a rapidly changing environment, which needs to be supported through a robust CE process. The requirement for the agility in defence acquisition stems out from the need to deliver and evolve inherently *joint*[1] capabilities with acceptable performance over extended time frames [25]

From the CE perspective, a capability is viewed and managed as a complex system [25] comprising people (members of integrated project teams), processes, tasks, and materiel facilities (real property, utilities and equipment). The goal of the CE process is to evolve the current 'as-is' capability towards a desired 'to-be' end state. The process begins with a rigorous assessment and validation of the 'as-is' capability through application of modelling tools and also through bringing stakeholders in the process. Getting stakeholder validation is important for projecting future shortfalls that may occur as a result of evolutionary changes. Next, the gap between the 'as-is' and 'to-be' architectures is identified, and the decision-maker is presented with options for augmenting the capability. The CE process has been explored in detail during the course of the DRDC Valcartier CapDEM project tasked by the Joint Capability Review Board in January 2003 [26].

An integrated project team is an important element of the CE process [25]. The people who comprise an integrated project team build models and architectural frameworks, develop a suitable set of capability metrics, and use system engineering tools to identify capability gaps 'turning capacity into capability' [25]. The effective team work is achieved via a collaborative engineering environment, which consists of system engineering tools (e.g. CADD, modelling and risk analysis tools) and tools supporting knowledge sharing and information exchange (e.g. web technologies) [26]. Given the above features, the metanetwork representation will be very suitable for modelling the CE process. In what follows, we address the issue of the identification of the capability gap in terms of people's expertise and tools which are needed to develop and evolve a capability during the course of the CE process.

The metanetwork models, which represent people, expertise, and tools comprising multi - disciplinary projects, could be rather large and intricate. Not surprisingly, the identification of the capability gap, even in this seemingly simplified perspective, could be a daunting task. The following example demonstrates how DNA and ORA can facilitate this task. The purpose of the presented example is to provide a modelling framework which could be used during the project planning stage for the identification of the current gap in people's expertise and tools. With this purpose in mind, a hypothetical project set-up with limited numbers of nodes and node classes has been developed in this example. For illustrative purposes, the chosen project tasks resemble tasks routinely performed in a typical OR project. Note that a limited number of nodes does not impose serious restrictions on the applicability of the proposed modelling framework, as it can be easily increased to accommodate large projects.

[1] In the defence context, jointness means interoperability across several environmental domains: Air, Land, Maritime, Space, and Cyber.

	Experts	Knowledge	Tasks
Experts	Team net Who works with whom	Expertise net Who knows what	Assignment net Who does what
Knowledge		Information net What informs what	Needs net What is needed to do what
Tasks			Workflow net What precedes what

Table 3: *Capability engineering network data*

Basic network data which is required for constructing a CE metanetwork is show in Table 3. There are three types of nodes and six inter-linked networks in a basic CE metanet. The node types are experts, knowledge, and tasks. Experts are people with expertise in selected S&T areas. They form a team network (who works with whom). Knowledge nodes are methods and techniques that can be used to perform various project tasks. Knowledge nodes are linked in the information net (what informs what), where links indicate the direction of flow of information between the knowledge nodes. The expertise network links experts with corresponding knowledge nodes (who knows what). Task nodes are project tasks, as specified in a project plan. The workflow net represents a sequence of tasks according to the project plan (what precedes what). The needs network determines what knowledge is needed to perform the task (what is needed to do what), and the assignment net assigns personnel to the tasks (who does what).

Some of the networks listed in Table 3 should be provided with input data, whereas others can be developed using the ORA tools. For instance, the assignment net does not have to be provided, but can be obtained by multiplying the expertise and needs nets using the ORA Matrix Algebra tool [19]. One of the advantages of the metanetwork representation is that networks are mathematical objects, and various mathematical operations can be performed on networks. Thus, with the metanetwork representation the task of personnel assignment is reduced to the multiplication of two matrices: the matrix representing the expertise network (Expert x Knowledge in Table 3) and the matrix representing the needs network (Knowledge x Task). (The details of matrix multiplication, as applied to operations on networks, are explained in Annex B.) This feature becomes very useful when dealing with large metanets which are comprised of multiple node classes and in which many inter-linked networks can be created.

To accommodate the full scope of the CE process, such node classes as stakeholder organizations, resources and events should be added to Table 3. Addition of new node classes will result in new networks. For instance, the ownership network will link stakeholder organizations with facilities which they own; and the attendance network will connect the event nodes (e.g. meetings with stakeholders) and people who attended those events. The resulting metanet will provide a comprehensive picture of what comprises a capability with both the product architecture and embedded processes represented. A time series of such metanets will represent the evolution of a capability from a current ‘as-is’ state to a desired ‘to-be’ state. As many intermittent metanets as needed can be included in the time series to represent different stages of the CE process. The full representation of the CE process is beyond the purpose of this example, as it deserves a detailed investigation.

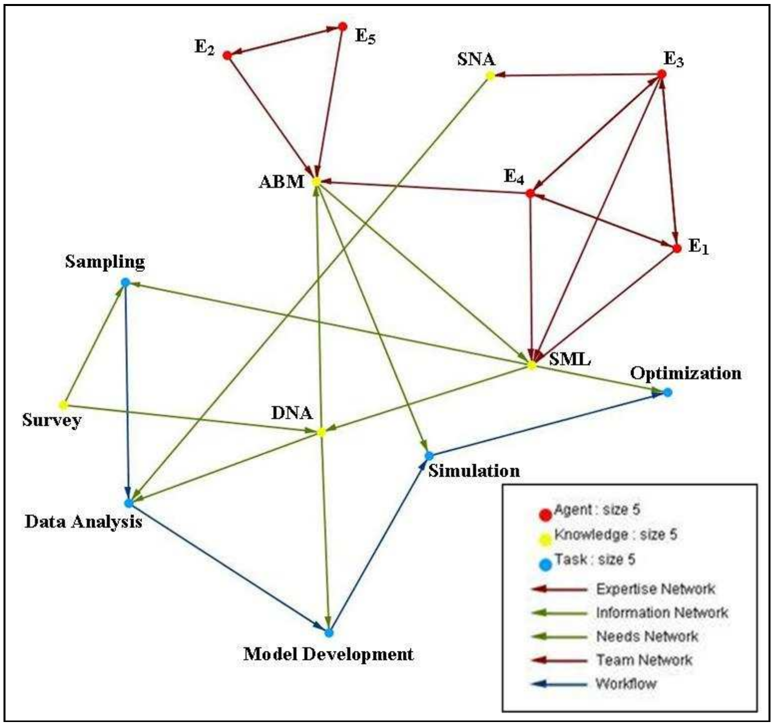


Figure 6: *A hypothetical capability engineering metanetwork*

Figure 6 shows a basic CE metanet consisting of three node classes and five inter-linked networks, as specified in Table 3. In Figure 6 experts, knowledge and tasks are shown by red, yellow and blue dots respectively. The link colour matches the colour of the source node. In this simplified example, the workflow net consists of a sequel of 5 tasks: sampling, data analysis, model development, simulations, and optimization. Blue arrows indicate the task precedence. For instance, data analysis normally precedes model development, and model development precedes simulations. The information net is represented by green arrows connecting knowledge nodes according to flow of information. For instance, DNA informs ABM, and hence there is a directed link from the DNA node to the ABM node. The needs network is shown by directed green arrows leading from knowledge nodes to tasks nodes. There are two teams in the CE metanet, which are shown by bi-directed arrows between expert nodes.

There are two different ways of how the gaps can be detected. One way is inward looking, as it focuses on identifying deficiencies in available expertise within an organization through analyzing the assignment net. Another way is outward looking, as it benchmarks organizational practices against the current state-of-the-art through comparing the following two networks:

1. Needs network that represents methods routinely used in an organization to perform certain tasks;

2. Best-practice network that reflects best applicable methodologies and procedures derived from a survey of the state-of-the-art.

In both cases, the ORA Matrix Algebra tool is used to perform the tasks (see Annex B).

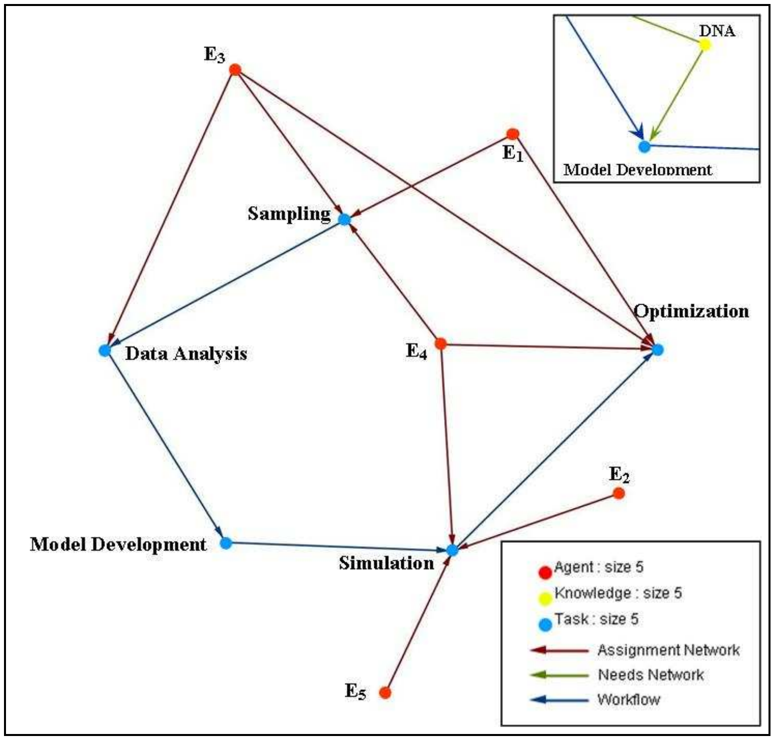


Figure 7: *Identification of capability gaps using the assignment network*

The assignment net, a product of network multiplication (see Annex B), is shown in Figure 7. An isolated task node in the assignment net detects a potential capability gap for the organization in question. In Figure 7, the Model Development node is not linked to any expert nodes, which means that there is a lack of expertise in the knowledge area required to perform this task. As one can deduce from the displayed part of the needs network in the upper left corner of Figure 7, the knowledge area in question is DNA. Thus, the visualization of the assignment and needs nets provides a way for detecting a potential deficiency in available expertise within an organization.

Again, the ORA Matrix Algebra tool can be used to compare the needs and best-practice networks (see Annex B). First the dot product of the above two networks can be obtained. The dot product is the intersection or common part that two networks share. Then the dot-product net can be subtracted from the best-practice net. The difference network represents the capability gap:

$$[\text{Difference net}]=[\text{Best-Practice}]-[\text{Best-Practice}].*\ [\text{Needs}]$$

This idea is illustrated in Figure 8, where the difference net (which represents the capability gap) is displayed together with the workflow net. The difference net identifies DNA as a potential capability gap for a hypothetical organization of this example.

As mentioned earlier, there could be more node types, and hence more networks in the CE metanet to represent the full scope of the CE process.

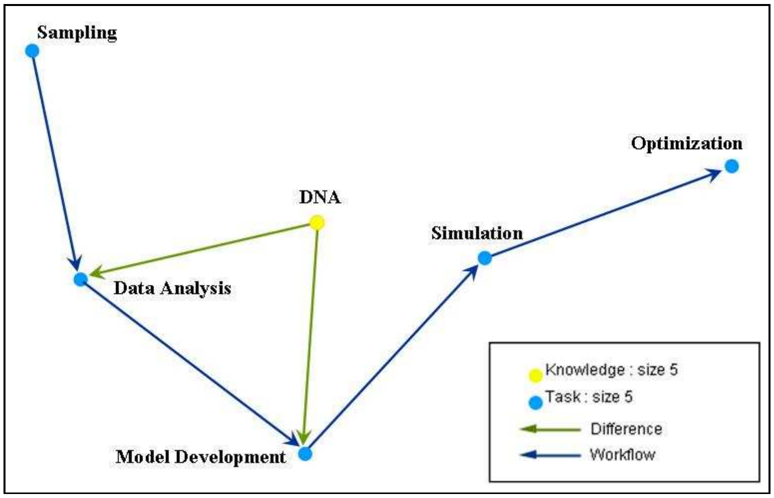


Figure 8: *Identification of capability gaps using the best-practice network*

The above techniques will work most effectively in multi-disciplinary S&T projects represented by large inter-linked networks. For such projects, text analysis tools can be used for automated extraction of information from texts. For example, the needs net can be derived from internal technical reports, and the best-practice network can be obtained from the state-of-the-art survey. After capability gaps have been identified, multi-agent simulations can be used to develop and test policies for improvement of organizational practices.

4 Discussion and Conclusions

The presence of many interacting components is a key feature of complex systems and processes, and networks are well-suited to represent this feature. There is a variety of network models to choose from, but the Complex Dynamical Networks (CDN) stand out. The CDN combine structural and dynamical complexity, as they can dynamically change their structure due to both external and internal factors. From this view point, the CDN-based models give a rich and flexible representation of many real systems and processes of interest to defence and security. Unfortunately, the theoretical aspects of the majority of the CDN remain largely unexplored. There is clearly a need for an approach and a tool capable of handling cases of practical significance. Dynamic Network Analysis (DNA) and the ORA software tool, which implements it, attempt to address this need.

DNA is an innovative approach that goes beyond the limitations of classical social network analysis and brings network techniques to the level of the complex systems thinking. The metanetwork concept, which is central to DNA and upon which DNA is based, is a truly complex systems concept. Extended graph measures, which this concept provides, are valuable additions to the analysts' toolkit, especially when dealing with adversarial systems and networks. Structural dynamics is captured in DNA through representing complex systems as a time series of metanetworks, and also through incorporating a range of social, cognitive, and political processes that govern the system evolution. In this way, DNA provides a framework for modelling the system evolution and informs the agent-based model development.

ORA, the dynamic network analysis and visualization tool, operates on metanetworks and on time-series of metanetworks. With its advanced GUI, it is possible to visualize, analyze and explore metanetworks, develop new metanetworks, and test various intervention strategies. The ORA analytical and visualization capabilities can be used to gain insight into multi-relational data, to visualize relationships, and to identify hidden or missing links. Even with incomplete information, ORA is still able to provide useful insights into essential features of complex systems, and at a relatively low cost in terms of data requirements. The tool has been built to gracefully degrade, so that it calculates only those metrics for which there is information [18].

As any model, DNA and ORA are not free from limitations. There are many challenges yet to be addressed. Some of the challenges, such as for instance the treatment of uncertainty, are clearly articulated in [18] and in other CASOS publications which can be found on the CASOS website [27]. Both DNA and ORA are under active development: new algorithms and intelligent tools are being implemented, tested and added to overcome the tool limitations.

The current tool limitations are mainly related to the node classes and how those are specified in the metanet models. It is easy to imaging an entity that could fit into two different categories, and the one that does not fit in either. Emergency service providers, such as police units for example, could be classed as organizations and agents in the critical infrastructure models. What node class would be appropriate to model beliefs, ideas and emotions? Some node types can be complex systems in themselves (e.g. a weapon platform as a resource node), which the fact raises the question about multiple scales within the metanet models.

Also, the idea of separating functions from nodes may prove to be limiting in some applications, especially in relation to the representation of causality and feedback loops. Another important area which could be affected by this limitation is the occurrence of emergence, that is patterns of behaviour arising out of relatively simple interactions.

In spite of the above limitations, DNA and ORA have significant potential for Defence R&D and can be used in a variety of defence and security applications. Potential applications include adversarial systems and networks, counter-IED, critical infrastructures, capability engineering, and capability-based planning. The advantages of using DNA and ORA have been discussed in Section 2.

Two illustrative examples of potential applications of interest to defence have been considered in Section 3. The first example shows how the metanetwork representation can be used to gain insight into the Canadian complexity research community (Section 3.1). The second example illustrates the application of DNA and ORA to modelling the capability engineering process (Section 3.2). A new modelling framework for the identification of capability gaps in terms of people's expertise has been suggested in Section 3.2. It has also been shown how the proposed framework can be extended to accommodate the full scope of the capability engineering process.

This technical memorandum reports on the results of the feasibility study, entitled "Complexity in Defence Systems and Operations: a Survey of Computational and Modelling Approaches." DNA is one of three closely-related groups of approaches, which have been identified by the study as important for quantitative support of military decision making. The other two areas, statistical machine learning and multi-agent technologies, will be discussed in the works that follow.

References

[1] Bar-Yam, Y. (2004), Making Things Work: Solving Complex Problems in a Complex World, NECSI Knowledge Press.

[2] Miller, J. H. and Page, S. E. (2007), Complex Adaptive Systems: An Introduction to Computational Models of Social Life, Prinston University Press.

[3] Bollobás, B. (1998), Modern Graph Theory, Springer-Verlag, New York.

[4] Aldana, M., Coppersmith, S., and Kadanoff, L. P. (2002), Boolean dynamics with random couplings, *arXiv:nlin.AO/0204062*, 2, 1–71.

[5] Arbib, M. A. (2003), The Handbook of Brain Theory and Neural Networks, The MIT Press, Cambridge (Massachusetts).

[6] Sayama, H. (2007), Generative network automata: a generalized framework for modeling complex dynamical systems with autonomously varying topologies, *Proc. IEEE Synposium on Artificial Life (CI-ALife 2007)*, pp. 214–221.

[7] Braha, D. and Bar-Yam, Y. (2006), From centrality to temporary fame: dynamic centrality in complex networks, *The Bell System Technical Journal*, 27, 5 pp.

[8] Strogatz, S. H. (2001), Exploring complex networks, *Nature*, 410, 268–276.

[9] Watts, D. J. and Strogatz, S. H. (1998), Collective dynamics of small-world networks, *Nature*, 393, 440–442.

[10] Cohen, P. (2004), It's a small world inside your head, *New Scientist*, 182(2448), 12.

[11] Barabási, A. L. and Alberts, R. (1999), Emergence of scaling in random networks, *Science*, 286, 509–512.

[12] Carley, K. M. (2007), Modeling social organizations, *International Conference on Complex Systems 2007 (ICCS2007)*, October 28 – November 2, Boston, MA (Keynote Lecture).

[13] Badia, A. and Skillicorn, D. (2008), The challenges of knowledge discovery in adversarial settings, *SIAM News*, 41(7), 7.

[14] Carley, K. M. (2002), Inhibiting adaptation, *Proc. 7th ICCRT Symposium*, Naval Postgraduate School, Monterey, CA (electronic publication), 10 pp.

[15] Carley, K. M. (2002), The Handbook of New Media (L. Lievrouw and S. Livingstone, Eds.), Chapter 12: Smart Agents and Organizations of the Future, pp. 206–220, Sage, Thousand Oaks, CA.

[16] Carley, K. M. (forthcoming), Dynamic network analysis, *The NRC Workshop on Social Network Modeling and Analysis, National Research Council*, (R. Breiger and K. M.Carley, Eds.), 13 pp.

[17] Wasserman, S. and Faust, K. (1994), Social Network Analysis: Methods and Applications, Cambridge University Press, Cambridge.

[18] Carley, K. M., Columbus, D., DeReno, M., Reminga, J., and Moon, I. C. (2008), ORA User's Guide 2008, The CASOS Technical Report CMU-ISR-08-125, Carnegie Mellon University, Pittsburgh, PA.

[19] ORA (online), CASOS, `www.casos.cs.cmu.edu/projects/ora/software.php` (Access Date: February 2009).

[20] Pestov, I. (2008), Complexity Research in Canada: A National Survey, (DRDC CORA TN 2008-020) DRDC CORA.

[21] Carley, K. M. (2003), Destabilizing terrorist networks, *Proc. 8th ICCRT Symposium*, National Defense War College, Washington, DC (electronic publication), 6 pp.

[22] Frantz, T. L. and Carley, K. M. (2005), A Formal Characterization of Cellular Networks, The CASOS Technical Report CMU-ISRI-05-109, Carnegie Mellon University, Pittsburgh, PA.

[23] Moon, I. L., Carley, K. M., and Levis, A. H. (2008), Vulnerability assessment on adversarial organization: unifying command and control structure analysis and social network analysis, *SIAM International Conference on Data Mining, Workshop on Link Analysis, Counterterrorism and Security*, Atlanta, Georgia (electronic publication), 10 pp.

[24] ComplexityHub (online), DRDC, `complexityhub.org` (Access Date: July 2008).

[25] Pagotto, J. and Walker, R. S. (Aug 2004), Capability Engineering: Transforming Defence Acquisition in Canada, (DRDC OTTAWA SL 2004-074) DRDC OTTAWA.

[26] CapDEM TDP: The Collaborative Capability Definition, Engineering and Management Technology Demonstration Program (online), Department of National Defence, `www.capdem.forces.gc.ca` (Access Date: January 2009).

[27] CASOS Home (online), CMU, `www.casos.cs.cmu.edu` (Access Date: February 2009).

Annex A: ORA Key Entity Report

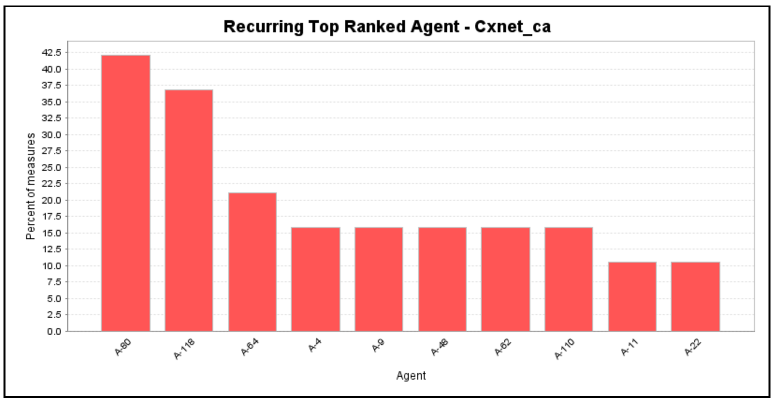
The following measures have been generated by ORA as part of the Key Entity report for the Canadian complexity research metanetwork discussed in Section 3.1. The output data format has been slightly modified for presentation purposes.

Key Entity - Who

Input data: Canadian complexity research metanetwork (Figure 1)

Top Ranked Agents

This chart shows ten agents that repeatedly rank in the top three in the measures. The value shown is the percentage of measures for which the Agent was ranked in the top three.



Emergent Leader (cognitive demand)

Measures the total amount of cognitive effort expended by each agent to do its tasks [18].

Input network(s): Collaboration by research interest (Figure 4)

Rank	Value		Agent	
1	0.194541	(0.452991)*	A-80	(A-80)*
2	0.184779	(0.452991)*	A-64	(A-118)*
3	0.184779	(0.435897)*	A-118	(A-64)*

*Values in brackets show total degree centrality measures and corresponding top three agents in that measure

Input network(s): Collaboration by membership (Figure 5)

Rank	Value		Agent	
1	0.194541	(0.504274)*	A-80	(A-48)*
2	0.184779	(0.504274)*	A-64	(A-106)*
3	0.184779	(0.478632)*	A-118	(A-4)*
*Values in brackets show total degree centrality measures and corresponding top three agents in that measure				

In-the-Know (total degree centrality)

The Total Degree Centrality of a node is the normalized sum of its row and column degrees [17, 18].

Input network(s): Collaboration by research interest (Figure 4)

Rank	Value	Unscaled	Agent	Context*
1	0.452991	106	A-80	12.0319
2	0.452991	106	A-118	12.0319
3	0.435897	102	A-64	11.4348
4	0.427350	100	A-9	11.1363
5	0.376068	88	A-101	9.3452
*Number of standard deviations from the mean if links were distributed randomly				

Potentially Influential (betweenness centrality)

Betweenness Centrality of node v is the percentage of the shortest length connections across all node pairs that pass through v [17, 18].

Input network(s): Collaboration by membership (Figure 5)

Rank	Value	Unscaled	Agent	Context*
1	0.10168	1380	A-62	13.7578
2	0.09320	1265	A-48	12.4834
3	0.09320	1265	A-106	12.4834
4	0.07427	1008	A-81	9.6353
5	0.05021	682	A-4	6.0170
*Number of standard deviations from the mean if links were distributed randomly				

Most Knowledge (row degree centrality)

The Row Degree Centrality of a node is its normalized out-degree [17, 18].

Input network(s): Knowledge (Figure 2)

Rank	Value	Unscaled	Agent
1	0.044199	8	A-29
2	0.044199	8	A-110
3	0.038674	7	A-17

Complete Exclusivity - knowledge

Detects entities that have ties that no other entity has.

Input network(s): Knowledge (Figure 2)

Rank	Value	Unscaled	Agent	Speciality
1	0.028	5	A-110	real-time and embedded systems, distributed simulation, web-service oriented simulation, defence and emergency response, crowd and evacuation simulation
2	0.028	5	A-26	self-organized criticality, stability of complex systems, phase transition, critical phenomena, combat modelling
3	0.022	4	A-61	modelling cognitive processes, decision support systems, radar and wireless communications, multi-sensor surveillance

Leader of Strong Clique (eigenvector centrality)

Calculates the principal eigenvector of the network. A node is central to the extent that its neighbors are central [17, 18].

Input network(s): Collaboration by membership (Figure 5)

Rank	Value	Agent	Context*
1	0.0240475	A-4	-2.3077
2	0.0229689	A-80	-2.3118
3	0.0229689	A-99	-2.3118
*Number of standard deviations from the mean if links were distributed randomly			

Connects Groups (high betweenness and low degree)

The ratio of betweenness to degree centrality; higher scores mean that a node is a potential boundary spanner [17, 18].

Input network(s): Collaboration by membership (Figure 5)

Rank	Value	Agent
1	0.314853	A-81
2	0.253558	A-62
3	0.208236	A-82
4	0.066971	A-48
5	0.066971	A-106

Agency Size (column degree centrality)

The column Degree Centrality of a node is its normalized in-degree [17, 18].

Input network(s): Membership (Figure 3)

Rank	Value	Unscaled	Organization
1	0.38983	46	MITACS
2	0.12712	15	NSERC
3	0.06780	8	ComplexityHub

Key Entity – How

Dominant Knowledge (total degree centrality)

The Total Degree Centrality of a node is the normalized sum of its row and column degrees [17, 18].

Input: all graphs based on the node class(es) Knowledge

Rank	Value	Unscaled	Knowledge
1	0.220339	26	mathematical biology
2	0.110169	13	neural networks
3	0.110169	13	machine learning
4	0.093220	11	dynamical systems
5	0.084746	10	nonlinear dynamics
6	0.076271	9	computational complexity
7	0.050848	6	communication networks
8	0.050848	6	complex adaptive systems
9	0.050848	6	cellular automata
10	0.042373	5	disease modelling

Key Entity – Where

Highest concentration of actors (row degree centrality)

The Row Degree Centrality of a node is its normalized out-degree [17, 18].

Input network(s): Location

Rank	Value	Unscaled	Location
1	0.169492	20	Montreal QC
2	0.152542	18	Waterloo ON
3	0.152542	18	Ottawa ON
4	0.152542	18	Toronto ON
5	0.084746	10	Vancouver BC
6	0.059322	7	Calgary AB
7	0.042373	5	Quebec QC
8	0.042373	5	Kingston ON
9	0.025424	3	Winnipeg MB
10	0.025424	3	London ON

Performance Indicators

Node class	Size
Agent	118
Knowledge	181
Location	19
Organization	8

Overall Complexity (density of the metanetwork as a whole)	0.086510
Social Density (density of Agent x Agent networks)	
Collaboration by research interest	0.108504
Collaboration by membership	0.176010
Social Fragmentation (fragmentation of Agent x Agent networks)	
Collaboration by research interest	0.147038
Collaboration by membership	0.541793
Average Communication Speed (based on the inverse of the shortest path lengths between node pairs)	
Collaboration by research interest	0.409116
Collaboration by membership	0.473716

Average Communication Speed is the average inverse shortest communication path (number of links) between any pairs of nodes in a network [17, 18].

Fragmentation is a measure of network performance that shows the proportion of disconnected entities in a network [17, 18].

This page intentionally left blank.

Annex B: ORA Matrix Algebra Tool

The ORA Matrix Algebra tool allows users to perform various operations on networks [18]. Such operations as transposition, addition, subtraction, multiplication and folding of networks are performed by ORA in accordance to the matrix algebra rules. The matrix operations are included under the Matrix Algebra tool in the Data Management menu.

Network Multiplication

In Section 3.2, network multiplication was used to obtain the assignment network (Expert x Task) which assigns experts to tasks according to their expertise areas. The product of two networks was obtained by multiplying the corresponding matrices of the expertise net (Expert x Knowledge) and the needs net (Knowledge x Task), i.e.:

$$\begin{pmatrix} E_1K_1 & \cdots & E_1K_m \\ E_2K_1 & \cdots & E_2K_m \\ \cdots & \cdots & \cdots \\ E_nK_1 & \cdots & E_nK_m \end{pmatrix} \times \begin{pmatrix} K_1T_1 & \cdots & K_1T_p \\ K_2T_1 & \cdots & K_2T_p \\ \cdots & \cdots & \cdots \\ K_mT_1 & \cdots & K_mT_p \end{pmatrix} = \begin{pmatrix} E_1T_1 & \cdots & E_1T_p \\ E_2T_1 & \cdots & E_2T_p \\ \cdots & \cdots & \cdots \\ E_nT_1 & \cdots & E_nT_p \end{pmatrix} \tag{B.1}$$

In Equation B.1, E_i, K_J, T_k are experts, knowledge and tasks respectively $(i = \overline{1,n},\ j = \overline{1,m},\ k = \overline{1,p})$. Matrix elements E_iK_j, K_jT_1 and E_iT_k are equal to 0, if there is no link between corresponding nodes, and are more than 0, if the corresponding nodes are linked. For instance, if there is a link between an expert node and a knowledge node in the expertise net and if there is a link between this knowledge node and a task node in the needs net, then there will be a link between this particular expert and that particular task in the assignment net. Note that the number of columns in the first matrix must be equal to the number of rows in the second matrix, meaning that only networks with matching node types can be multiplied. That is, the network, which links a node of type A with a node of type B, can only be multiplied by the network, in which a node of type B is the source.

Dot Product Multiplication

The dot product of two networks is obtained by multiplying each element of the binary matrix representing the first network by the corresponding elements of the binary matrix of the second network. The resulting binary matrix will represent the common part of two networks. The dot product multiplication was used in Section 3.2 to obtain the intersection of the best practice and needs networks.

Folding Networks

Folding of networks is performed by ORA outside the Matrix Algebra tool. But the matrix multiplication rules are followed, as this operation involves matrix transposition and multiplication. Two nodes in the folded network are linked, if they are linked to the same node in the original network. The network-folding operation was used in Section 3.1 to obtain two collaboration networks shown in Figures 4 and 5 from knowledge and membership networks respectively.

List of symbols/abbreviations/acronyms/initialisms

ABM	Agent-Based Models
ANN	Artificial Neural Networks
C2	Command and Control
CADD	Computer Aided Design and Drafting
CAIMS	Canadian Applied and Industrial Mathematics Society
CAS	Complex Adaptive Systems
CASOS	[Center for] Computational Analysis of Social and Organizational Systems
CE	Capability Engineering
CMS	Canadian Mathematical Society
CDN	Complex Dynamical Networks
CND	Centre for Non-linear Dynamics
CORA	Centre for Operational Research and Analysis
DNA	Dynamic Network Analysis
DND	Department of National Defence
DRDC	Defence Research and Development Canada
GNA	Generative Network Automata
GUI	Graphical User Interface
INNS	International Neural Network Society
MITACS	Mathematics of Information Technology and Complex Systems
OR	Operational Research
ORA	Organizational Risk Analyzer
RBN	Random Boolean Network
S&T	Science and Technology
SIAM	Society for Industrial and Applied Mathematics
SML	Statistical Machine Learning
SNA	Social Network Analysis

Distribution list

DRDC CORA TM 2009-020

Internal distribution

1 DRDC CORA DG

1 DRDC CORA Chief Scientist

1 DRDC CORA DDG

1 DRDC CORA MORT (Mr. Paul Massel)

3 DRDC CORA CSSORT (Mr. Paul Chouinard)

2 DRDC CORA Library

3 DRDC Valcartier, System of Systems (Dr. M. Ducharme)

2 DRDC Ottawa, FSSE (Dr. G. Fusina)

1 DRDC CSS DDG (Dr Mark Williamson)

1 DRDC DSTIC (Mr. Paul Comeau)

10 Author

Total internal copies: 26

External distribution

International recipients

6 Dr. Anne-Marie Grisogono
TTCP JSA AG14 (CAS) Chair
Land Operations Division
Defence Science and Technology Organisation PO Box 1500 Edinburgh 5111 SA
Australia

Total external copies: 6

Total copies: 32

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence R&D Canada – CORA Dept. of National Defence, MGen G. R. Pearkes Bldg., 101 Colonel By Drive, Ottawa ON K1A 0K2, Canada	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Dynamic Network Analysis for Understanding Complex Systems and Processes

4. AUTHORS (Last name, followed by initials – ranks, titles, etc. not to be used.)

Pestov, I.

5. DATE OF PUBLICATION (Month and year of publication of document.) July 2009	6a. NO. OF PAGES (Total containing information. Include Annexes, Appendices, etc.) 44	6b. NO. OF REFS (Total cited in document.) 27

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Memorandum

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – CORA
Dept. of National Defence, MGen G. R. Pearkes Bldg., 101 Colonel By Drive, Ottawa ON K1A 0K2, Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) PG0 Project 10bc	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC CORA TM 2009-020	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
( ) Defence departments and defence contractors; further distribution only as approved
( ) Defence departments and Canadian defence contractors; further distribution only as approved
( ) Government departments and agencies; further distribution only as approved
( ) Defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11)) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Many complex systems and processes of interest to defence are characterized by varying topologies that can be modeled as Complex Dynamical Networks with nodes representing entities and links representing relationships among those entities. Complex Dynamical Networks give a flexible, powerful and useful representation of structural changes in such systems. Unfortunately, the theoretical aspects of these network classes remain largely unexplored. Dynamic Network Analysis provides an approach and a tool to deal with such cases of practical significance in spite of the lack of theoretical knowledge regarding structural dynamics. The presented technical memorandum conducts a detailed analysis of this approach, its major concept - the metanetwork, and a software tool that implements it. The significance of Dynamic Network Analysis for modelling multi-relational data, visualizing networks, and identifying hidden links is demonstrated on two examples. The first example shows how the metanetwork representation can be used to gain additional insight into the national survey data on complexity research in Canada. The second example illustrates the application of Dynamic Network Analysis to modelling the capability engineering process. A new modelling framework for the identification of capability gaps in terms of people's expertise is suggested in this example. The proposed modelling framework will be useful in large multi-disciplinary S&T projects.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

complex systems
structural dynamics
dynamic network analysis
system visualization
system analysis
Canadian complexity research network
capability engineering

DEFENCE DÉFENSE

DRDC CORA

www.drdc-rddc.gc.ca