APPLYING 3D PRINTING AND GENETIC ALGORITHM-GENERATED ANTICIPATORY SYSTEM DYNAMICS MODELS TO A HOMELAND SECURITY CHALLENGE

Michael J. North
Pam Sydelko
Ignacio Martinez-Moyano

Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL, USA 60439

ABSTRACT

In this paper we apply 3D printing and genetic algorithm-generated anticipatory system dynamics models to a homeland security challenge, namely understanding the interface between transnational organized criminal networks and local gangs. We apply 3D printing to visualize the complex criminal networks involved. This allows better communication of the network structures and clearer understanding of possible interventions. We are applying genetic programming to automatically generate anticipatory system dynamics models. This will allow both the structure and the parameters of system dynamics models to evolve. This paper reports the status of work in progress. This paper builds on previous work that introduced the use of genetic programs to automatically generate system dynamics models. This paper’s contributions are that it introduces the use of 3D printing techniques to visualize complex networks and that it presents in more detail our emerging approach to automatically generating anticipatory system dynamics in weakly constrained, data-sparse domains.

1 INTRODUCTION

In this paper we apply 3D printing and genetic algorithm-generated anticipatory system dynamics models to the homeland security challenge of understanding the interface between transnational organized criminal networks and local gangs. The organization of this paper is as follows. The second section discusses the interface between transnational organized criminal networks and local gangs. The third section reviews the related work. The fourth section introduces the use of 3D printing to visualize complex criminal networks. The fifth section discusses the use of genetic programming to automatically generate anticipatory system dynamics models. The sixth section presents conclusions and next steps.

This paper reports the status of work in progress. This paper’s contributions are that it introduces the use of 3D printing techniques to visualize complex networks and that it presents in more detail our emerging approach to automatically generating anticipatory system dynamics in weakly constrained, data-sparse domains. Please note that this paper builds on a paper under accepted for the 2015 International Conference of the Systems Dynamics Society (North, Sydelko, and Martinez-Moyano 2015). The new content in this paper includes the discussion of 3D printing and a more detailed discussion of system dynamics model generation.
2 TRANSNATIONAL ORGANIZED CRIME

Governments are increasingly faced with challenges that present themselves as “wicked problems” (Rittel and Webber 1973). These problems are complex systems that have many interdependent elements. They are typically not owned by one organization, but instead have a myriad of stakeholders with different and sometimes conflicting perspectives on the system. Finally, these problems become especially challenging for areas related to security, where the complex systems being addressed are highly adaptive and covert.

The U.S. government typically addresses these types of complex wicked problems by dissecting them and parsing out the pieces to individual agencies and organizations. It is unreasonable to expect congressionally mandated agencies to reorganize themselves for every wicked problem. Therefore, interagency coordination is often accomplished through committees and task forces. Unfortunately, these approaches have not always been effective against wicked problems.

Collaborative interagency groups need to act like a meta-organization, with a trans-agency structure. To tackle complex adaptive systems, the ‘trans-agency’ itself needs to become a complex adaptive system. A meta-organization is an organization composed of organizations. McChrystal et al.’s “team of teams” (2015) offers a compelling example. Methods must be developed that can design these trans-agency meta-organizations to be systemically aligned to the wicked problem they are charged with tackling.

Our case study looks at U.S. strategies for addressing the convergence of transnational organized crime with domestic local gang crime. Transnational organized crime (TOC) groups engage in many kinds of trafficking including that of people, drugs, arms, dangerous chemicals, biological materials, nuclear materials, and funds. Even the illicit transfer of information over the Internet can be categorized as transnational crime. Today, TOC has shocking scale that creates enormous costs for both the global economy and the human community (United Nations 2004). For example, one estimate sets the financial cost of TOC at $870 billion annually (United Nations Office on Drugs and Crime 2012). An example TOC influence diagram developed from interviews with subject matter experts is shown in Figure 1.

![Figure 1: An influence diagram for the CUE Project. Nodes represent focal points such as actors, places, or resources. Links represent interactions or connections between focal points.](image)

Criminal networks are not only expanding their operations, but they are also diversifying their activities, resulting in a convergence of transnational threats that have evolved to become increasingly complex, volatile, and destabilizing. These networks threaten U.S. interests in many ways including the formation and feeding of TOC alliances with corrupt elements of governments worldwide.

Several global trends—including dramatically increased trade volumes and velocity, the growth of cyberspace, and population growth—have facilitated an explosion of violent non-state actors,
strengthened TOC, supported the emergence of a new set of transcontinental supply chains, and driven the expansion of existing illicit markets. The resourcefulness, adaptability, innovativeness, and ability of illicit networks to circumvent countermeasures make them formidable foes for national governments and international organizations alike (Miklaucic and Brewer 2013).

The complexity of the challenge requires attention to all levels of the illicit trafficking supply chain. The U.S. government has traditionally sought to address these challenges vertically, with agencies acting largely in isolation from one another. The Drug Enforcement Agency focuses on controlling narcotics; the Food and Drug Administration concentrates on stopping counterfeit pharmaceuticals; the Department of Energy targets dual-use components of weapons of mass destruction; and the State Department limits conventional weapons flows to name but a few. These organizations are all short on resources, are highly overworked, are evaluated using divergent metrics by Congress, and are unable to develop the interagency responses necessary to disrupt the increasingly interconnected illicit enterprises they are charged with fighting.

Studies aimed at anticipating the evolution of TOC emphasize the need to understand the cultures and subcultures that yield and shield organized crime (Miklaucic and Brewer 2013). In addition, enhanced awareness of the political and economic incubators for criminal enterprises are needed. Facilitating the development of appropriate precautions or countermeasures by law enforcement agencies requires anticipating the long-term risk management strategies of criminal enterprises.

Recent empirical research on drug trafficking networks in Central and South America confirms that illicit networks are not only composed of the expected unlawful social agents but also include critical “gray agents” (Salcedo-Albaran and Salamaca 2012). Gray agents are defined as social agents with conflicting organizational and functional roles. Examples are public servants, political actors, or security specialists who also promote criminal interests. Such agents are said to engage in preference falsification driven by their divergent, and mutually incompatible, commitments. As a result, interactions with gray agents produce different social relationships than those seen amidst the typical confrontation between bright (i.e., lawful) and dark (i.e., unlawful) social agents (Salcedo-Albaran and Salamaca 2012). These unexpected interactions contribute to the already high level of complexity found in TOC-infused systems. This naturally leads to a discussion of complex adaptive systems.

The complex adaptive system used in our case study is that of worldwide TOC network connections to local gangs. We are working to apply the system dynamics model generation approaches discussed later to the TOC control problem.

3 RELATED WORK

There are two threads of related work. The first is the use of various technologies for visualizing data in three dimensions. The second is the application of genetic algorithms to system dynamics modeling. Both threads are individually reviewed in this section.

3.1 Visualizing Data in Three Dimensions

A variety of approaches have been used to visualize data in three dimensions including static holograms (Benton 1982), dynamic holographs (Chen, Su, and Jhun 2014), augmented reality systems (Carmigniani et al. 2011), static stereoscopic displays (Holmes 1859), dynamic stereoscopic displays (Benzie et al. 2007), direct volume devices (Clifton and Wefer 1993), 3D printing (Hull 1984), and handmade 3D artifacts (Batson 2014). Some of these options can be combined with haptic (i.e., touch) feedback technologies. The main disadvantages of 3D printing compared to most of the other options are lower resolution and limited display extent. This constrains the size and complexity of networks that can be printed. The main disadvantage of 3D printing compared to the dynamic options is that it is static. Nonetheless, 3D printing is unique in that it fully integrates visual and haptic channels into a single media that is supported by inexpensive devices with low operating costs that are rapidly becoming ubiquitous.
North, Sydelko, and Martinez-Moyano

(Moilanen and Vadén 2013). Even if one or more of the other 3D data visualization options become widely used, it is likely that 3D printing can be used as a complementary technology.

3D printing produces three dimensional shapes using additive manufacturing processes (Freedman 2011). 3D printing began with Hideo Kodama’s and Chris Hull’s pioneering work in the early 1980’s (Hull 1984) and is rapidly growing in popularity (Moilanen and Vadén 2013). 3D printing has been used to generate a range of visualizations for subject areas including mathematical equations (Séquin 2005, Batson 2014), cartography (Gruen and Muraib 2002; Rase 2011), and medicine (Rengier et al. 2010; Ebert, Thali, and Ross 2011).

The use of 3D printers to visualize complex abstractions builds on previous work with handmade 3D visualizations. As early as 1893, German mathematician Felix Klein displayed a host of handmade 3D mathematical models at the World’s Columbian Exposition in Chicago (Batson 2014). This led to an explosion of interest in handmade 3D mathematical models in the late 1890’s (Batson 2014). The excitement waned after the turn of the century due to the rise of the Bourbaki group’s intentionally visualization-free approach to mathematics (Batson 2014). Later in this paper we build on both the previous handmade and 3D printed visualization work by using 3D printing for the first time to produce visualizations of complex networks.

3.2 Genetic Algorithms

John Holland’s (1992) genetic algorithms are a category of biologically inspired search methods that implement some of the central features of natural selection. Genetic algorithms (Goldberg 1989; Mitchel 1996) evolve a population of individuals, each of which represents a candidate solution to a user-selected problem. Genetic algorithms use a fitness function to rank each individual’s effectiveness as a solution to the chosen problem. Genetic algorithms modify their populations over a series of generations using events patterned after natural selection. The events include the deaths of uncompetitive members of the population, the reproduction of competitive individuals, and random mutations among survivors. Reproduction, in particular, can include crossover events where children gain a mixture of the traits of their two parents.

Genetic algorithms have been widely used for a wide range of search and optimization tasks with substantial success (Goldberg 1989; Mitchel 1996). In the case of search, the fitness function quantitatively estimates how well each individual candidate meets the search criteria. For optimization, the fitness function is usually the objective function to be approximately minimized or maximized. It should be noted that for nontrivial problems and achievable run times, genetic algorithms are usually heuristics that do not guarantee optimal results. In practice, genetic algorithms often do quite well despite the caveats. Here we use genetic algorithms for a special kind of optimization to be detailed later.

Genetic algorithms have been used in three main ways in the published system dynamics literature. First, genetic algorithms have been used in constrained domains to construct system dynamics models that match selected times series data. Second, genetic algorithms have been used to optimize the parameters of existing system dynamics models relative to an objective function. Third, genetic algorithms have been used to calibrate existing system dynamics models. We will now review each of these uses.

3.2.1 System Dynamics Model Creation

Creation of models in constrained domains takes advantage of the natural structure in a subject area to narrow the range of possible system dynamics models to a large, but manageable, size. A variant of genetic algorithms called genetic programming (Koza 1990) is often used. Genetic programming can be understood as a special class of genetic algorithm where each individual is a small computer program. These programs are often represented as branching trees of appropriate instructions. Koza et al. (2001) and Pawlas and Zall (2012) discuss applications of genetic programming to system dynamics.
Koza et al. (2001) “reverse engineered both the topology and sizing...of a network of chemical reactions.” Pawlas and Zall (2012) used genetic programming to determine the equations and parameters in selected nodes of a system dynamics model of economic activity designed by subject matter experts. Koza et al. (2001)’s pioneering work is the closest to that presented in this paper, particularly the molecular approach. Nonetheless, there are significant differences that build on both Koza et al. (2001)’s and Pawlas and Zall (2012) contributions. First, we are modeling a less constrained domain so the search space of potential models is much larger. Second, we have less data and therefore must use techniques appropriate to a data-sparse domain. Third, we are not necessarily attempting to reproduce a known network but rather anticipate networks that might emerge in the future. Fourth, we use a nested two-stage optimization that considers dimensional analysis then system performance rather than a single stage.

3.2.2 System Dynamics Model Optimization

Optimization of existing models involves using a genetic algorithm to find system dynamics model parameters that approximately minimize or maximize an objective function. The objective function in turn is either the cumulative or final value of a chosen model output as generated by executing the system dynamics model for a fixed period of time using the selected candidate parameters. Examples papers that use this kind of approach include those by Linard (2000), Alborazi (2008), and Eksin (2008). Our work preforms optimization on both the model’s input parameters and the model’s structure.

3.2.3 System Dynamics Model Calibration

Calibration of an existing system dynamics model using a genetic algorithm usually involves evolutionary tuning of the model’s parameters. Typically, each individual in the population represents one candidate set of model parameters. The fitness function result for each candidate is found by running the system dynamics model for a fixed period of time using the associated input parameters and recording either the cumulative or final value of a chosen model output. The calibration may be run once or it may be executed repeatedly to match varying entries in a target time series. Example papers that apply variations of this method are those by Jeng, Chen, and Liang (2006); Shuhong (2008); and Yu and Wei (2012). The work presented in this paper includes calibration, but goes beyond this by also evolving model structure.

This approach to calibration can be understood as a special kind of optimization with minimizing the difference between the model outputs and the calibration data being the objective function. The two approaches are a considered separately here since they are usually treated distinctly in the literature.

3.2.4 System Dynamics Model Calibration and Optimization Debate

There is some debate as to how effective genetic algorithms are for calibration and optimization of system dynamics model parameters. Ventana Systems Inc., the makers of Vensim, have stated that “we have experimented extensively with genetic algorithm optimization and found that the results are very poor” (2015). Other system dynamics tools are reported to have included genetic algorithms for optimization (Linard 2000), at least at one time.

4 3D NETWORK PRINTING

Understanding and communicating the structure of complex abstractions is known to be an extremely difficult problem (Séquin 2005). One answer pioneered by Felix Klein (Batson 2014) is to use people’s natural ability to manually work with three dimensional structures in the physical world. 3D printing makes this answer much more practical than was possible with previous handcrafted approaches.

We use a multistep, free and open source software-focused process to produce 3D printouts of complex networks such as that shown in Figure 1. The workflow is shown in Figure 2. A custom script
was created to help automate the workflow. First, the network data is created using the free and open source Cytoscape computational biology environment (Cytoscape 2015). Second, the network is rendered in 3D using a special 3D version of Cytoscape. Third, the network is stored in an Extensible Markup Language (XML) version of Graph Modeling Language (GML) format called XGML. Fourth, the XGML is compiled into the solid computer aided design (SCAD) file format. The SCAD file is then converted into the stereolithography or standard tessellation language (STL) file format via the free and open source OpenSCAD application (OpenSCAD 2015). The STL file can then be loaded into common 3D printer drivers to determine the required temporary overhang supports and then print. Temporary overhang supports are needed for long horizontal spans without intermediate support. They are manually removed once printing is complete.

Figures 3 and 4 show the resulting 3D printouts of the 2D influence diagram in Figure 1. Figure 3 shows two views of a raw 3D printout including the temporary overhang support structures. The left side of Figure 4 shows the same 3D printout after cleaning by removal of the temporary overhang supports. The right side of Figure 4 shows the same cleaned 3D printout atop a properly aligned 2D printout of the network.

Early discussions with a variety of stakeholders indicate that the use of 3D printouts dramatically improves their ability to understand the networks being studied, at least for networks small enough to print. The use of 3D network printouts also seems to substantially increase their level of engagement.
Two fundamental approaches to generating system dynamics models are being pursued for this paper. Both approaches use the genetic algorithm technique known as genetic programming to dynamically create a series of candidate system dynamics models to be evaluated by subject matter experts. Both approaches are being implemented in Repast Simphony (2015).
Each individual in the genetic program under development consists of a set of assignments to a fixed list of system-level output variables. Each output variable includes a time series of values and an associated unit. The domain-specific output variables are chosen by subject matter experts to represent the system’s critical measures of interest. The genetic program is being designed to determine the assignments to these output variables using the approaches discussed later.

The fitness function uses rising values to represent increasingly preferred candidates. We use a nested two-stage fitness function that considers dimensional analysis then system performance. The dimensional analysis evaluation adds up any dimensional arithmetic or assignment errors. The system then gives an undesirable value to the fitness function that is inversely proportional to the total number of errors.

System performance evaluation only occurs when there are no dimensional analysis errors. System performance evaluation involves running the model for a specific period of time and then collecting the output results. Then, the system performance is assigned a desirable function of the output variables selected by the subject matter experts (e.g., maximizing profit). A candidate with the minimum system performance is arranged to have a higher fitness value than un-executable candidates with even a single dimensional analysis error.

Either the molecular or the atomic approaches discussed next will tend to produce generations of models with increasing fitness levels. Once the fitness levels become high enough, the resulting models will be considered as possible future networks. Of course, having a high fitness level alone does not necessarily mean that a given TOC network is a realistic possibility now or in the future. It is possible for networks with high fitness levels to contain subtle problems or to simply be unreachable by transformations from the current state. Nonetheless, novel networks with high fitness values may offer interesting anticipatory windows into possible futures. These candidate networks will be shown to subject matter experts to determine if the networks contain useful insights. ‘Delta’ techniques are expected to be used to highlight the differences between the expert’s suggested networks and the new candidate networks.

5.1 The Molecular Approach
The molecular approach uses Eberlein and Hines’ (1996) concept of system dynamics “molecules” to build candidate models. Molecules are small to medium-sized sets of system dynamic components that represent common themes or motifs in a domain of interest. For example, Eberlein and Hines (1996) identify the simple stock and flow structure called a “decay process” as a common example molecule.

The molecular approach to dynamically generating system dynamics models involves subject matter experts identifying common behavioral patterns in their domain that are candidates to become molecules. These molecules are then given standardized interfaces that minimize the chances of dimensional analysis errors and maximize the opportunities for interoperability with other molecules. For example, common units are chosen when possible (e.g., all currency values in are U.S. dollars). It is intended that this will reduce the time spent in finding dimensionally consistent candidate models.

5.2 The Atomic Approach
The atomic approach allows equations to be built up from raw terms rather than molecules. This approach allows greater flexibility than the molecular approach, but also substantially increases the range of possible models. In particular, it is expected that this will increase the optimization time spent on dimensionally inconsistent models.

5.3 Current Implementation
The current implementation is designed to support both the molecular and atomic approaches using a single code base. The engine has been implemented except for the mutation operator. The core code has
been unit tested with 97.7% test coverage as shown in Figure 5. System dynamics models are implemented using arrays of terms and functions to be discussed next.

5.3.1 Terms

Terms can have optional units that are defined using the JScience library (JScience 2015). Terms are subdivided into constants and variables.

Constants are conversion factors appropriate for the domain being modeled. Variables are further subdivided into primitive variables and indexed variables. Primitive variables are variables without indices. Indexed variables are variables with bounded integer indices in the range \{1..N_v\} for \(N_v\), the maximum index value for variable \(v\). Units are standardized for each optimization run to maximize interoperability of terms. Terms without defined units are considered to be dimensionless.

Each variable is assigned an equation. An equation is a tree of functions and terms. Equations are limited in depth. The default is four layers. Variables are assigned to equations as follows:

1. Primitive variables are assigned combinations of primitive variables and indexed variables with constant indices for variable \(v\) indexed by \(i\) in \{1..\(N_v\)\}.
2. Indexed variables (variable \(v\) indexed by \(i\)) are assigned combinations of primitive variables, indexed variables with index \(i\), or indexed variables with index \(\text{MOD}(i + j, N_v) + 1\) with \(j\) in \{1..\(N_v\)\}.

5.3.2 Functions

Functions implement mathematical operations that combine terms. Functions have a defined set of input units, consistency rules, conversion rules, and output units. Inputs are either dimensionless or in fixed units (e.g., kilograms). Consistency rules allow for no constraints (e.g., for multiplication) on units or the restriction that all units must be the same (e.g., for addition). The conversion rules allow for an optional constant with optional units times fixed units (e.g., dimensionless results or dollars), the units of the \(i\)th term of the function, or the product of all of the input units for the function. The currently implemented functions are numerical integration (i.e., system dynamics stocks), addition, subtraction, multiplication, and division.
6 CONCLUSIONS AND NEXT STEPS

In this paper we discussed the automatic generation of system dynamics models using a kind of genetic algorithm known as a genetic program. This approach allows both the structure and the parameters of the system dynamics models under study to be evolved. The paper provides an outline of how this technique is being applied to the example domain of TOC. This paper also showed how 3D printing techniques can be used to visualize complex networks such as those for transnational organized crime.

This paper reports the status of work in progress. The next step is to complete the work introduced in the paper. Later steps include generating system dynamics models for other domains, applying 3D printing to visualize other complex networks, and studying in depth the impact that 3D printing technologies can have on the visualization of complex networks.

ACKNOWLEDGMENTS

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. Argonne National Laboratory's work was supported under U.S. Department of Energy contract DE-AC02-06CH11357. Argonne National Laboratory's work was supported under U.S. Department of Energy contract DE-AC02-06CH11357.

REFERENCES


North, Sydelko, and Martinez-Moyano


AUTHOR BIOGRAPHIES

MICHAEL J. NORTH MBA, Ph.D. is the Group Leader of the Integrated Analytics Group and a Principal Computer Scientist within the Global Security Sciences Division of Argonne National Laboratory. He is also a Senior Fellow in the joint Computation Institute of The University of Chicago and Argonne. Dr. North has over 20 years of experience developing and applying analytical methods for industry, government, and academia. Dr. North has published two books, five conference proceedings, one journal special issue, eight book chapters, four invited encyclopedia entries, more than 20 journal articles, and over 70 conference papers. Dr. North is also the lead developer of free and open source Repast agent-based modeling suite. His email address is north@anl.gov.

PAM SYDELKO is the Director of the Systems Science Center, which is part of the Global Security Sciences Division at Argonne National Laboratory. She is also a doctoral research student at the Centre for Systems Studies at University of Hull, United Kingdom. She has 28 years of experience in systems science and analysis and has led the development of numerous innovative modeling, analysis and decision support technologies, focusing on integrated multi-component software systems. Key modeling/analysis domains include national security, sustainability, and critical infrastructures. Pamela Sydelko earned her MBA from the University of Chicago, her M.S. in Soil Science from the University of Illinois at Urbana-Champaign, and her B.S. in Botany/Ecology from North Dakota State University. Her email address is psydelko@anl.gov.

IGNACIO MARTINEZ-MOYANO is the Deputy Director of Center for Integrated Resiliency Analyses, which is part of the Global Security Sciences Division at Argonne National Laboratory. He is also a Computational Social Scientist and an Institute Fellow at the Computation Institute of The University of Chicago. Dr. Martinez-Moyano’s research focuses on the application and theoretical development of system dynamics modeling to advancing the theories and understanding of human judgment, decision making, and behavior in complex systems, particularly in high uncertainty security systems. Dr. Martinez-Moyano’s current research projects include work related to understanding resiliency in complex systems, aviation security, energy efficiency in buildings, modeling crime in cities, and understanding the cognitive processing of accumulation and dynamics. His email address is imartinez@anl.gov.