Social Network Modeling and Agent-Based Simulation in Support of Crisis De-escalation

Michael J. Lanham, *Member, IEEE*, Geoffrey P. Morgan, Kathleen M. Carley, *Senior Member, IEEE*

***Abstract*— Decision makers need capabilities to quickly model and effectively assess consequences of actions and reactions in crisis de-escalation environments. The creation and what-if exercising of such models has traditionally had onerous resource requirements. This research demonstrates fast and viable ways to build such models in operational environments. Through social network extraction from texts, network analytics to identify key actors, and then simulation to assess alternative interventions, advisors can support practicing and execution of crisis de-escalation activities. We describe how we used this approach as part of a scenario-driven modeling effort. We demonstrate the strength of moving from data to models and the advantages of data-driven simulation, which allow for iterative refinement. We conclude with a discussion of the limitations of this approach and anticipated future work.**

***Index Terms*—Computer simulation, Social Network Analysis, information diffusion, text mining**

I. Introduction

Effective crisis response requires thinking through the implications and interactions of complex sets of events, an error-prone process for humans with high-stakes in the deterrence domain. War games, and modeling and simulations (M&S) in support of war games, can mitigate lack of experience and support forward thinking by providing safe venues for assessing alternatives. Simulation efforts intended to support forward thinking, however, often have long time-cycles to develop; the need for the tool is overtaken by events before the tool is ready. This paper presents a rapid meta-network (multi-mode, multi-link) modeling approach using concept extraction techniques to develop models for examining scenarios within a useful time-span.

We present and discuss our process for rapidly developing useful meta-network and information diffusion models through semi-automated analyses of text corpora; how we applied the approach to deterrence and crisis de-escalation scenarios, and our lessons learned. We discuss how our method's outcomes were triangulated using a multi-modeling approach and offer caveats and potential future work.

Manuscript received March 17, 2012; revised September 21, 2012; accepted November 13, 2012. This work was supported in part by the Office of Naval Research - ONR -N000140811223 and by the Air Force Office of Sponsored Research –MURI, 600322. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research, the Air Force Office of Sponsored Research or the U.S. government.

All the authors are with the Center for Computational Analysis of Social and Organizational Systems (CASOS), Institute of Software Research, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213. (e-mail: mlanham | gmorgan | kathleen.carley@cs.cmu.edu).

II. Related Works

Conflict reduction is an often-researched area of human knowledge. Indeed there are entire journals dedicated to the study of international conflict: *The Journal of Conflict Resolution; The Journal of Conflict & Security Law; Peace, Conflict, and Development; The International Journal of Conflict Management* among others. Assessment of conflict reduction and de-escalation efforts can be as simple as "is there no longer a shooting war" to much more nuanced sets of measures of effectiveness and measures of performance. Use of computer-aided M&S has ranged from human-based experiments [1] to efforts to include environmental and cultural framing to contextualize information [2] as well as correlation models [3]. Richardson introduced a purely mathematical set of models in [4] while Ruloff used system dynamics to model international relations scenarios in [5]. Yilmaz presents a summary of modeling efforts of the past 30 years from game theory to early social phenomena models in [6].

Building models from unstructured data also has a wide-ranging application history, from cell tower data for social network inference [7] to developing emergent ontologies [8]. Indeed, a grand vision of the "Semantic Web" was to bring structure and computable meaning to the World Wide Web [9, 10].

Social network construction from web-based sources has been done from the mid-90's to the current day through the use of web search tools' APIs [11, 12]. We differ from these methods because the constructed networks are both analyzed and serve as inputs into M&S environments. Like [13, 14], we need to rapidly build M&S capable models, but we use socio-linguistics theory and machine-learning-based topic modeling as introduced in [15] to provide a model construction mechanism.

Belief modeling plays a key role in the effort though we are not using a belief-desire-intention (BDI) modeling paradigm [16]. Modelers have used beliefs to support individual goal-oriented behaviors [17] as well as simulating threshold-based behaviors [18]. Beliefs, by which we mean attitudes, convictions, and opinions, do not require exposure to or awareness of knowledge and facts for people to sustain them.

Friedkin developed a belief modeling method with his social influence theory in [19] that brought four innovative concepts to social theory and modeling. One, he relaxed previous assumptions where previously agents had to either conform or deviate from a fixed consensus (the public choice model). Two, his method did not have to lead to consensus, and could support stable patterns of disagreement. Three, he provided a multi-level theory where micro-level cognitive processes could influence and constrain macro-level changes. Fourth, his methods supported quantitative analysis of the systematic consequences of social structures.

The agent-based modeling performed in this effort makes direct use of the *Construct* information diffusion simulation developed at CMU's Center for Computational Analysis of Social and Organizational Systems (CASOS). *Construct* has its roots in constructuralism and combines structuralism with social influence theory [20, 21]. Constructuralism, in brief, asserts that agents' actions, perceptions of selves and others, learning and forgetting knowledge and beliefs are all constantly influenced by the agents' surrounding environment, particularly their surrounding social environment.

We use *Construct* to demonstrate a fast method for building simulation models. These models enable practicing crisis de-escalation and deterrence. They also allow policy analysts to evaluate multiple counter-factual scenarios. We use *Construct* as a belief diffusion model, arguing that if policy makers believe they "should" go to war, then the deterrence calculus has failed. *Construct* is a validated model of belief and knowledge diffusion [22-24] shown to fit a wider range of data than reinforcement theory and information processing theory [25].

The multi-modeling component of this article is limited to a discussion of the simultaneous development and use of three different modeling tools, each with very different origins and theories of function to arrive at congruent results.

III. The Data to Model Process

The data to model process (D2M) we use is a systematic, computer-assisted, repeatable approach with these steps [26, 27]:

1. Collect data
2. Clean the text corpus
3. Ontological cross classification
4. Generate static data for analysis

Collecting data is the first step. The D2M process focuses on the challenges associated with unstructured data, although other forms of data can contribute to later analysis. We convert large amounts of unstructured texts into rich multi-mode, multiplex and multi-level relational networks (i.e., meta-networks) for use in dynamic simulations. The second step in our process is *cleaning the text corpus*. Text data, like all language, is rife with ambiguity. Data cleaning removes and/or clarifies redundant or ambiguous references, removes noise words, performs pronoun resolution, and acronym disambiguation. Step three is *ontological cross classification*; this step classifies phrases, for example "President" is classified as an agent, and also resolves ambiguities when words have two meanings such as "battery", which can be either a resource or an agent. Illustrative classes are Agents, Knowledge and Tasks [28, 29]. Typical semantics of the networks between them are shown in Table I. An analyst or planner iterates through steps two to four as many times as is appropriate to the demands of their leadership. Ideally, she would maintain up-to-date models through periodic additions to her corpora with new information and sources. The data-to-model process creates intermediate artifacts, allowing the process to be run without modification on new data or tweaked to improve the resulting model(s). Improvements can be subjective in the eyes of subject matter experts (SMEs), objective with respect to leader-specified network analytics and metrics or a combination of the two.

TABLE I SIMPLE META-NETWORK COMPRISED OF SIX (6) NETWORKS FROM THREE (3) TYPES OF NODES

	Agents	*Knowledge*	*Tasks*
Agents	Social	Knowledge	Assignment
Knowledge		Information	Needs
Tasks			Precedence

The final step in the process, *generate static data for analysis*, identifies linkages among the nodes through windowing, i.e., through proximity of the cleaned nodes in the text. These linkages are across multiple modes, creating a meta-network such as that seen in Fig. 1. The analyst can then use this meta-network for point-in-time analysis as well as input to simulations—in our case, a diffusion simulation. An important difference between these networks, and traditional network science's focus on agent-by-agent interactions, is the inclusion of the non-agent node classes in the networks and in the analysis [30].

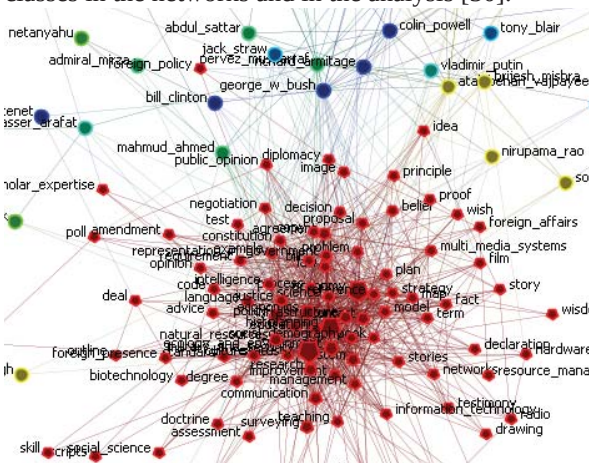


Fig. 1 A sample Multi-Mode network of Agents (circles, multi-colored by country) and Knowledge (hexagons, red), sized by Eigenvector Centrality

These four steps are not, themselves, unique. What is unique, from searches of related literature, is the use of these four steps in an operational-environment-friendly, low barrier-to-entry and low maintenance cost process of text mining to build multi-mode social networks that become the inputs into agent based simulations. Fig. 2 depicts this larger process in an abbreviated flow chart.

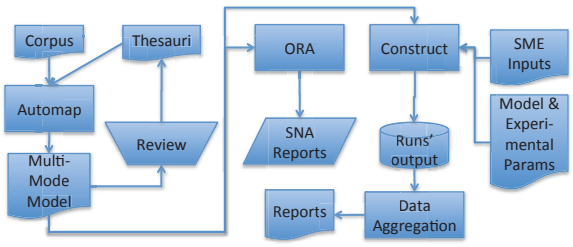


Fig. 2 A flowchart of this paper's process

IV. BELIEF FORMATION IN CONSTRUCT

Construct is a widely validated, agent-based model, with a focus on information diffusion and belief change [22-24, 31, 32]. Agents interact with those with whom they are similar (e.g., homophily) [33] which is a proven cross-cultural phenomenon [34]. Agents also interact with those from whom they seek information they do not have (e.g., expertise seeking) [32, 35]. Agents exchange and learn correct and incorrect information (implemented as vectors of 0/1 bits that we refer to as knowledge bits) as well as exchange information about ego's and alters' beliefs [21, 25].

In *Construct*, agents' beliefs may be anchored to knowledge—sets of knowledge bits can contribute positive or negative valence for a belief and each agents' belief values range [-1.0, 1.0]). For this effort, we used knowledge-anchored beliefs only. Belief formation, on a per-agent per-turn basis, is a summative function between an agent's prior beliefs mitigated by their ability to be influenced by their alters (extended from Friedkin in [19]) and their similarity to their alters magnified by their ability to be influenced by their alters. We will build to this formal equation (12) in the following paragraphs as *Construct* implemented it in [36, 37].

Agents in *Construct* also have error-prone perception of who-knows-what and who-believes-what. We use the term transactive memory for this perception. *Construct* implements transactive memory as a three dimensional binary matrix denoted Knowledge Transactive Memory (KTM) with indices *i*, *j*, and *k* where agent *i* (the ego) perceives agent *j* (the alter) is in possession of knowledge bit *k*. The same convention applies to the Belief Transactive Memory (BTM) for each belief, *b*, in the simulation. The expressions are shown below.

$$\forall i,j \in Agents\ (A),\ \forall k \in Knowledge\ Bits\ (K),\ \forall b \in Beliefs\ (B),\ KTM_{ijk}\ and\ BTM_{ijb} \quad (1)$$

As alluded to above, homophily preference is driven by a measure of knowledge similarity (SK) and belief similarity (SB). As shown below in (2) and (3), using the expressions in (1), it is the sum of self-perception per-bit multiplied by the perception of each connected agent's per-bit knowledge or belief.

$$SK_{ij} = \sum_k (KTM_{iik} \times KTM_{ijk}) \quad (2)$$

$$SB_{ij} = \sum_b \left((BTM_{iib} \times BTM_{ijb}) \right) \quad (3)$$

The ability of an agent (ego) to affect its connected neighbors (alters) is called social influence. Social Influence (RS) is a function of connectedness

between an ego *i* and its *n* alters as well as Knowledge and Belief Similarities. It is shown below in (4) and incorporating (2) and (3). α below is an exogenous parameter that is the weight an agent places on Knowledge Similarity versus Belief Similarity. In this experiment, we set α to 0.50 for equal weighting of the two factors of Social Influence.

$$RS_{ij} = \left[\frac{\alpha(SK_{ij})}{\sum_{j\neq i}^{n} SK_{ij}}\right] + \left[\frac{(1-\alpha)SB_{ij}}{\sum_{j\neq i}^{n} SB_{ij}}\right] \quad (4)$$

Still building to the belief calculation, we require a value to quantify the expertise seeking (EXP) discussed at the beginning of this section. It is a pair-wise function of the knowledge not shared between agents and their *n* alters and calculated using (5) below.

$$EXP_{ij} = \frac{\sum_{k,i=j}(\overline{KTM_{iik}}\times KTM_{ijk}}{|K_i|-\sum_k KTM_{iik}} \quad (5)$$

We've noted that this model uses fact-based beliefs so we had to provide the model a belief by fact weighting matrix, *V*, that provides valence weights for each belief to each fact from [-1.0,1.0]. This matrix allows facts to impact more than one belief as well as have no impact at all with a weight of zero (0). Intermediate outputs of Pythia—a different component of the multi-modeling effort that is a timed-influence Bayesian network tool [38-40]—provided these weights and are shown in Fig. 2 as "Model & Experimental Params." We'll denote the valence weight *V* for belief *b* using fact *k* as shown below in (6).

$$V_{bk} \quad (6)$$

To account for the ability for an agent to have perceptions of its beliefs, as well as to generalize self-perception to three (3) states (strongly agree, strongly disagree, no opinion), we use (7) for self-perception of belief *b* for later use.

$$B'_{iib} = \begin{Bmatrix} 1, \sum_k \frac{V_{bk}\times KTM_{iik}}{|K|} > 0.2 \\ -1, \sum_k \frac{V_{bk}\times KTM_{iik}}{|K|} < 0.2 \\ 0, otherwise \end{Bmatrix} \quad (7)$$

We also need to calculate the expected influence of self-perceived knowledge for each belief (EI), or how strongly the agent holds fact-based-belief, *b*, and we use (8) for that purpose.

$$EI_{iib} = \frac{\sum_k (V_{bk}\times KTM_{iik})}{\sum_k V_{bk}} \quad (8)$$

We need three more values to calculate the per-agent per-belief value per-turn. They are the influentialness of alters on the ego (INF_{ji}) in (9); the resistance to being influenced by alters (INF_{ii}) in (10); and the total influentialness ($TotalInf_{ji}$) in (11). The value of $influentialness_j$ in (9) is an exogenous parameter set by the experimenter and represents the ability of alters to influence the ego—the *ji* notation in (9) and (11) is intentional and not a typographic error. The value of BI_i in (10) is also exogenous and represents the propensity of an ego to be influenced by alters, also called belief influenceability. In this model, these values were kept constant and used successful default settings from prior validated work with *Construct*.

$$\forall j \text{ (alters) connected to } i\text{(ego)}:$$

$$INF_{ji} = influentialness_j \times \frac{RS_{ji}\times EXP_{ji}}{2} \quad (9)$$

$$INF_{ii} = (1 - BI_i)\times \frac{RS_{ii}+EI_{iib}}{2} \quad (10)$$

Summing (9) for each *j* connected to *i* helps us generate (11).

$$\forall j \text{ (alters) connected to } i\text{(ego)}:$$

$$TotalInf_i = \sum_j INF_{ji} \quad (11)$$

In compressed form, we can now represent agent *i*'s self-perception of belief *b* at time *t* in (12).

$$BTM_{iib_t} = BTM_{iib_{t-1}}\times\left(1 - BI_{i_t}\right) + BI_{i_t}\times \left[\left(\sum_j \left(\frac{INF_{ij_t}}{TotalInf_{i_t}}\times BTM_{ijb_t}\right)\right) + \left(\frac{INF_{ij_t}}{TotalInf_{i_t}}\times B'_{iib_t}\right)\right] \quad (12)$$

V. The India-Pakistan Crisis Scenario

We used this data-to-model approach, as operationalized in the software application *AutoMap [28, 41]*, as part of a scenario-driven exercise. The intent of the exercise was to illustrate the value of two organizations coordinating to assess the impact of different courses of action (COAs). The organizations were two US Regional Combatant Commands (COCOM): US Pacific Command (USPACOM) and US Central Command (USCENTCOM). The scenario for this fictional situation used a mixture of fictional scenario events and real-world events from a specific time-period, from 2 June 2002 to 5 August 2002. It also used fictional and real-world interactions among agents along with real names for people and places. The interactions were generated through SME elicitation and use five of the seven categories from [42]: public appeals; communication facilitation; mediation; fact-finding; and humanitarian aid. The scenario location is along the disputed territorial border regions of Jammu and Kashmir between India, Pakistan and China. The scenario begins with a fictitious raid into the parliament building of Srinagar, India by gunmen on 2 June 2002. The scenario continues to 5 August 2002 with a number of actions by Pakistan, India, the United States, and select other countries of interest.

A. Data-to-model Process Applied to the Scenario

We used 3,000 LexisNexis®-provided text files that met the search criteria of the scenario's dates and the terms: "India" and "Pakistan." These newspaper

articles provided background and supporting cultural and contextual data to the scenario-provided information. We also scraped each nation's national security apparatus' official web sites (circa 2010) as well as the official web sites of USCENTCOM and USPACOM. By national security apparatus we mean the functional equivalents of the US National Security Council (NSC), Department of Defense (DoD), and Department of State (DoS). After these web scrapes, there were approximately 27,000 files in our corpus. We built the synonym and classification thesauri as well as the delete list from scratch: there was no COCOM planning staff from which we could borrow thesauri or delete lists. The development of these lists took approximately 160 man-hours, though subsequent improvements to *Automap's* "Data to Model Wizard" have demonstrated significant speedup [26].

Identification of specific persons relevant to a border-crisis scenario was an iterative process of identifying a term or sets of terms (e.g. "Prime Minister of India," "Vajpayee") then using web-based searches to determine the nature of the term and resolve uncertainties. This allowed us to remove the multitude of cricket players and Bollywood stars within the corpus. We used social network measures such as degree centrality, betweenness centrality, and eigenvector centrality to estimate different aspects of a node's criticality in our resultant networks as well as consultation with our multi-modeling partners. Table II describes the end-state of the network model.

TABLE II NODE COUNTS, PER COCOM, VIGNETTE A & B

	Vignette A		Vignette B	
Node Type	**PACOM Count**	**CENTCOM Count**	**PACOM Count**	**CENTCOM Count**
Agents	42	47	42	47
Knowledge	145	145	148	148
Belief	21	21	32	32
Other node sets not depicted here: Event, Location, Organization, Resource, Role, Task. USPACOM was principally focused on India while USCENTCOM was principally focused on Pakistan.				

B. Network Analysis Applied to Scenario Models

Following the scenario outline, we divided the data set into three vignettes: 1) initial crisis incident plus eight days; 2) mid-crisis when the COCOMs were using independent analysis and actions; and 3) a final period when the two COCOMs would, in the scenario, collaborate and merge their respective models and COAs to present to US national leadership. For each vignette, we used the Organizational Risk Analyzer (ORA) network analysis software [43] to calculate numerous static node and static network measures and to visualize the interconnections of strategic decision-makers (labeled "SNA Reports" in Fig. 2). A more comprehensive discussion of which network measures we used is available in Chapter 14 of the efforts final technical report [44]. ORA includes over 157 different network measures applicable to two-mode and multi-mode networks [45].

Using this methodology, we discerned shifts in relative rankings of the top ten agents across the vignettes. Fig. 3 is an example graphic from the "Key Agent" report, which is a component of the "Key Entities" report, for Vignette B from the USCENTCOM perspective. The graphic identifies the agents that are most commonly in the top ten (10) rankings across twenty-two (22) different social network analytic measures relevant to agents. In this report, President Musharraf is in the top ten agents 90% of the time, or twenty (20) of twenty-two (22) measures. The numbers after the agents' title in the arrows reflect the change from Vignette A to Vignette B, with three (3) new agents appearing in Vignette B. The Secretary of State's (SecState) involvement suggests that the diplomacy instrument of national power is increasing its level of effort.

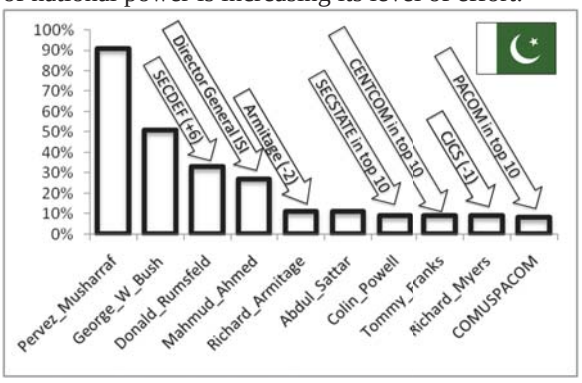


Fig. 3 The change in actor relevance indicates that the scenario is shifting from a diplomatic to a military situation.

The appearance of PACOM and CENTCOM, meanwhile, is consistent with an interpretation that the scenario is rapidly moving from a diplomacy-centric situation to one involving the US military. This finding is in accord with the tenor of the scenario and the impressions of our SMEs. The Chairman of the Joint Chiefs' (CJCS) drop in relative ranking is consistent with the increasing presence of both COCOM commanders in direct discussions and interactions with the President. Their direct involvement with the President is consistent with the DoD moving from planning for action with the CJCS as the principal military advisor to executing action

through the COCOMs.

C. Dynamic analysis through diffusion simulations

The *Automap*-extracted meta-network built by the D2M process described in Sections A and B above, became the primary input to the *Construct* simulation, as shown in Fig. 2. For this scenario, the primary output measure of interest was the number of strategic decision-makers who possessed a "pro-war" belief as calculated using (12). We harvested the strategic decision-makers and their relationships directly from the text-mined multi-mode data. The remainder of the multi-mode data, drawn primarily from LexisNexis®, was not pertinent to the scenario and questions of interest.

Table III Experimental Conditions

Variable	Possible Values	Combinations
Variables:		
Response Timing	None Early Middle Late Scenario	5
Constants:		
Provocation Timing	Scenario	1
Social Network	Automap	1
Knowledge Network	Stylized	1
K x Belief Network	Stylized	1
TM False Negative	0	1
TM False Positive	0.5	1
Total Combinations		5
Runs per Condition		25
Total Runs		125

We still needed some form of knowledge in the simulation so we implemented stylized representations of knowledge; these representations were created to reflect national identity, culture, as well as knowledge both for and against aggression. The size of these pools and their distribution to agents was based on the text-mined outputs of *AutoMap*. We also seeded the simulation with the majority of agents to have more "anti-war" knowledge than "pro-war"—representing the status quo that there are far more people disinclined to go to war than there are those inclined to start a war. The error-prone transactive memory was instantiated with a population-wide false-negative rate of 0.5 (that is an ego wrongly perceives an alter does not have a particular piece of knowledge).

We implemented the scenario within *Construct* as a set of forty-one exogenous "provocations," twenty-nine "responses," and twenty-four general events that added general facts to agents' knowledge pools. All these events had a magnitude, a start-time, and an end-time. Within *Construct*, we used a special-purpose agent for each individual provocation and response. The knowledge bits these agents transferred were bit strings exogenously tied to the beliefs of 'pro-war' for provocations, and 'anti-war' for responses. We modeled the impact of these events by constraining the duration the special agents were active within the simulation—inactive agents do not interact with others and thus do not share knowledge with other agents. The creation of links to decision maker agents, as well as durations of activity, were drawn from the multi-modeling team, subject matter experts, and intermediate outputs of Pythia—all exogenous to the social network from AutoMap. The duration of activity of the special agents increased the probability that the knowledge it was communicating would be learned by its interaction partners, thereby impacting the knowledge-anchored beliefs of those interaction partners. The complete collection (~120KB) of provocations and responses, durations, names, and start times, as well as the complete input files (~400KB) for Construct are available from the authors on request.

As shown in Table III, the principal question we explored was one of timing the interventions, with three relevant sub-questions, specifically: 1) how many strategic decision-makers will possess the pro-war belief if the United States does not intervene (the "None" case); 2) given the scenario and all of its deterrence and de-escalation actions, how many Indian and Pakistani decision-makers will possess pro-war beliefs (the "Scenario" case); and 3) given a stable set of deterrence actions, how does changing the timing of this action set change the number of decision-makers with pro-war belief (the "Early", "Middle", and "Late" cases)?

These virtual experiments showed that without US or others' work to tamp down tensions, within thirty days more than 60% of the Pakistani and Indian strategic decision makers believe that war is the right choice. Fig. 4 also indicates that the conventional studied diplomatic US response set in the scenario document was insufficient to avoid war despite producing a shift toward anti-war beliefs in the minds of decision-makers. Early interventions produced the most significant impact (see also Fig. 5) – as agents then chose to pass along knowledge with negative valence towards the "pro-war" belief.

There are a number of reasons for this outcome. Agents in the model suffer the same "echo chamber" effect as people in the real world—their interactions with agents like themselves reinforce their beliefs and existing knowledge, forming a feedback loop that

gains ever larger portions of the population [46]. Interupting that feedback loop early by exposing agents to additional or alternate information is critical. Fig. 5 was very sensitive to additional provocations (e.g., troop deployments, missile launches, riots and media coverage thereof)—reinforcing the perception that actions will frequently overwhelm talks.

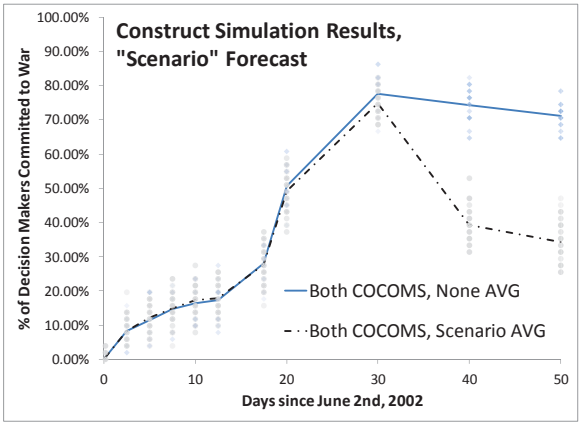


Fig. 4 Construct forecasts that the majority of strategic decision makers in both India and Pakistan will possess the pro-war belief within thirty days. Deterrence actions from the scenario have only minimal impacts until more than a month after the crisis begins.

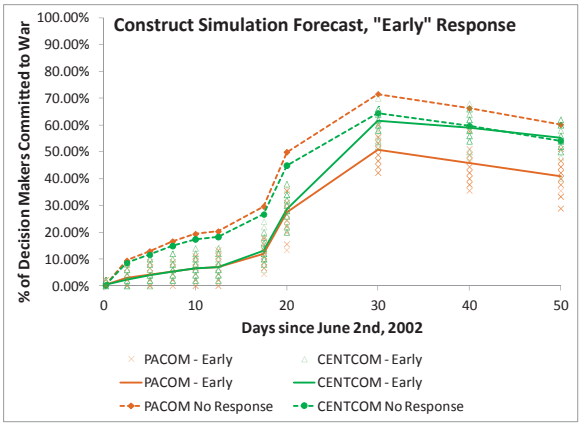


Fig. 5 Early interventions allow more time for comprehensive response.

D. Model implications for policy

These results suggest that: US and others must use levers of deterrence quickly; levers of deterrence may need repeated use to have an effect; continued provocations will rapidly overwhelm US instruments of national power; early and fast action may not, by itself, lead to de-escalation though it may buy time to bring additional resources to bear.

E. Validation

Since the basis of the model is a fictitious scenario, there is no mechanism available to do empirical or historical validation of this specific model. As such, we turn to other forms of confirming face validity, plausibility and usefulness of the model. Of most important note, the conclusions from this model were consistent across three diverse sets of assumptions, paradigms, modeling languages and tools: Construct, CAESAR III [47, 48] and Pythia. CAESAR III is a colored petri-net tool for assessing decision-making organizations largely omitted from this discussion. The congruence in outcomes, despite the distinctly different operating assumptions and paradigms of each tool, provides increased confidence in the plausibility of the constructed models, the techniques to build them, and the assessments that derived from the master scenario event list. The larger multi-modeling effort would also support incorporation of other models that incorporate other motivations for interactions (e.g., social capital, exchange theory, balance theory).

More important than the particular validation of this specific instance is the confirmation that this semi-automatable and repeatable approach of moving from large quantities of unstructured text to a well-developed meta-network is worthwhile. The static analysis using social network analysis tools and techniques generated reasonable results in the context of the scenario. The approach was further shown effective in generating the basis for an agent-based simulation model that might otherwise have taken significantly longer to build. Together, the presented approach shown in Fig. 2 provides techniques for decision makers to assess a wide variety of COAs in safe and controlled environments.

As for the scenario itself, the effort tells decision makers that there is little decision space within which they can maneuver—not a ground breaking result, but one based on more than intuition and the personal experience of individuals. This outcome was also accepted by the SMEs of the project sponsor when reviewing the effort.

VI. Discussion

We demonstrated a rapid model development approach that allows integration of multiple data-sources to produce a meta-network which forms the basis for the simulation. We showed the feasibility of how multiple organizations can take these meta-networks and examine possible futures—during long-range deliberate planning and execution as well as crisis and time-sensitive environments. Finally, we established the ability of diffusion simulations and network science to provide estimations of action-

reaction cycles in situations that require the coordinated use of multiple instruments of change—in this case the elements of national power. These estimations were valid in the eyes of the SMEs and the results were congruent with the output of models from very different backgrounds.

This first iteration of this approach and its application to deterrence was a useful demonstration; however, there are challenges. Data drawn from LexisNexis® is not an accurate reflection of the information each COCOM staff could or should maintain. We believe that as the data improves in quality and topicality, the utility and explanatory power of such models will improve. Further, each COCOM's information assets are likely to have distinct and important differences that we did not reflect – and these differences may lead to diverse but more useful final results. Follow-on efforts will need to incorporate a more sustained collection of text and other unstructured data. Fictional scenarios will require additional synthetic data.

The diffusion simulation made some additional simplifying assumptions. We did not conduct SME elicitation and profile the strategic actor set to learn and program their starting inclinations towards the pro-war belief. The simulation is able to use such information, with a few minor changes. We relied primarily on the agent-by-agent networks due to the paucity of the agent-to-knowledge links in the collected data. We estimate that COCOM staff's would have richer data sets that would support use of additional networks within the *Automap*-extracted meta-network and avoid or reduce the use of stylized knowledge sets within *Construct*. Another simplifying assumption was the deliberate exclusion of India's Cold Start doctrine [49-51], their 'no first use against non-nuclear states' policy [52], as well as Pakistan's published responses to the Indian doctrine. We did not incorporate meta-cognition reasoning into the simulation—agents being aware that others are attempting to influence them. *Construct* is very robust to trends and population/group level analysis. It does not predict the precise actions of individuals at specific times nor should decision-makers use *Construct* for per-agent analysis or predictions.

VII. Conclusion

This work demonstrates that the data-to-model approach enables rapid model development and supports model reuse, merging, and extension when a network analytic approach is taken. This approach meets the needs of decision makers to quickly model, simulate, and assess consequences of actions and reactions in crisis de-escalation environments.

VIII. Acknowledgment

The authors wish to acknowledge and thank their collaborators at George Mason University's System Architectures Lab. Without the combined efforts of Alex Levis, PhD, Sayed Abbas K. Zaidi, PhD, Lee Wagenhals, PhD, Robert Elder, PhD, Tod Levitt, PhD, Ahmed Jbara Y Abu, and Syed Hasan Ali Rizvi, this project would not have met the successes it did. This paper represents only a small part of the work this project entailed. We would also like to thank Mike Kowalchuck for his expert knowledge.

References

[1] J. Wilkenfeld, K. Young, V. Asal, and D. Quinn, "Mediating International Crises," *Journal of Conflict Resolution*, vol. 47, pp. 279-301, June 1, 2003 2003.

[2] D. A. Sylvan and S. J. Thorson, "Ontologies, Problem Representation, and the Cuban Missile Crisis," *Journal of Conflict Resolution*, vol. 36, pp. 709-732, December 1, 1992 1992.

[3] P. A. Schrodt and D. J. Gerner, "An Event Data Analysis of Third-Party Mediation in the Middle East and Balkans," *Journal of Conflict Resolution*, vol. 48, pp. 310-330, June 1, 2004 2004.

[4] L. F. Richardson, *Arms and insecurity; a mathematical study of the causes and origins of war*. Pittsburgh: Boxwood Press, 1960.

[5] D. Ruloff, "The Dynamics of Conflict and Cooperation between Nations. A Computer Simulation and Some Results," *Journal of Peace Research*, vol. 12, pp. 109-121, 1975.

[6] L. Yilmaz, T. I. Oren, and N. Ghasem-Aghaee, "Simulation-based problem-solving environments for conflict studies," *Simulation & Gaming*, vol. 37, pp. 534-556, December 1, 2006 2006.

[7] N. Eagle and A. S. Pentland, "Reality mining: sensing complex social systems," *Personal and Ubiquitous Computing*, vol. 10, pp. 255-268, 2006.

[8] M. Peter, "Ontologies are us: A unified model of social networks and semantics," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, pp. 5-15, 2007.

[9] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, May 17 2001.

[10] N. Shadbolt, T. Berners-Lee, and W. Hall, "The Semantic Web Revisited," *IEEE Intelligent Systems*, vol. 21, pp. 96-101, May/June 2006.

[11] H. Kautz, B. Selman, and M. Shah. (1997, Summer) The Hidden Web. *The AI Magazine*. 27-36. Available: http://www.cs.cornell.edu/selman/papers/pdf/97.aimag.refweb.pdf

[12] Y. Matsuo, J. Mori, M. Hamasaki, T. Nishimura, H. Takeda, K. Hasida, *et al.*, "POLYPHONET: An advanced social network extraction system from the Web," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, pp. 262-278, 2007.

[13] C. Weinstein, W. Campbell, B. Delaney, and G. O'Leary, "Modeling and Detection Techniques for Counter-Terror Social Network Analysis and Intent Recognitio," 2009.
[14] A. McCallum, X. Wang, and N. Mohanty, "Joint Group and Topic Discovery from Relations and Text," in *Statistical Network Analysis: Models, Issues and New Directions, Lecture Notes in Computer Science*. vol. 4503, ed New York: Springer, 2007, pp. 28-44.
[15] J. Diesner and K. M. Carley, "A methodology for integrating network theory and topic modeling and its application to innovation diffusion," in *IEEE International Conference on Social Computing (SocComp), Workshop on Finding Synergies Between Texts and Networks*, Minneapolis, MN, 2010.
[16] A. S. Rao and M. P. Georgeff, "Modeling rational agents within a BDI-architecture," *Readings in agents*, pp. 317-328, 1997.
[17] Z. Lin and K. M. Carley, "Organizational decision making and error in a dynamic task environment," *Journal of Mathematical Sociology*, vol. 22, pp. 125-149, 1997.
[18] M. Granovetter, & Soong, R.,, "Threshold models of diffusion and collective behavior," *Journal of Mathematical Sociology*, vol. 9, pp. 165-179, 1983.
[19] N. Friedkin, *A Structural Theory of Social Influence*. New York: Cambridge University Press, 1998.
[20] K. M. Carley, "An Approach For Relating Social Structure To Cognitive Structure," *Journal of Mathematical Sociology*, vol. 12, pp. 137-189, 1985.
[21] K. M. Carley, M. K. Martin, and B. R. Hirshman, "The Etiology of Social Change," *Topics in Cognitive Science*, vol. 1, pp. 621-650, 26 June 2009.
[22] Y. Ren, K. M. Carley, and L. Argote, "The Contingent Effects of Transactive Memory: When Is It More Beneficial to Know What Others Know?," *Management Science*, vol. 52, pp. 671-682, May 2006.
[23] K. M. Carley, "Group Stability: A Socio-Cognitive Approach," in *Advances in Group Processes: Theory and Research*. vol. Vol III, E. Lawler, B. Markovsky, C. Ridgeway, and H. Walker, Ed., ed Greenwhich, CT: JAI Press, 1990, pp. 1-44.
[24] C. Schreiber and K. M. Carley, "The Impact of Databases on Knowledge Transfer: Simulation Providing Theory," in *North American Association for Computational Social and Organizational Science (NAACSOS)*, Pittsburgh, PA, 2003.
[25] K. M. Carley, "On the Persistence of Beliefs," *Sociological Abstracts*, 1991.
[26] K. M. Carley, M. Bigrigg, C. Papageorgiou, J. Johnson, F. Kunkel, M. Lanham, *et al.*, "Rapid Ethnographic Assessment: Data-To-Model," in *HSCB Focus 2011: Integrating Social Science Theory and Analytic Methods for Operational Use*, Chantilly, VA, 2011.
[27] K. M. Carley, M. Bigrigg, D. Garlan, M. J. Lanham, Y. Lu, G. P. Morgan, *et al.*, "Experimentation Testbeds: Using SORASCS to Run and Process HSCB Virtual Experiments," in *HSCB Focus 2011: Integrating Social Science Theory and Analytic Methods for Operational Use*, Chantilly, VA, 2011.
[28] K. M. Carley, D. Columbus, M. W. Bigrigg, and F. Kunkel. (2011, 4 November 2011). *AutoMap User's Guide 2011*. Available: http://www.casos.cs.cmu.edu/publications/papers/CMU-ISR-11-108.pdf
[29] J. Diesner and K. M. Carley, "Revealing social structure from texts: Meta-Matrix text analysis as a novel method for network text analysis," in *Causal Mapping for Information Systems and Technology Research: Approaches, Advances, and Illustrations*, V. K. Narayanan and D. J. Armstrong, Eds., ed Harrisburg, PA: Idea Group Publishing, 2005, pp. 81-108.
[30] K. M. Carley, "Dynamic Network Analysis," in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, Committee on Human Factors, National Research Council*, R. Breiger, K. M. Carley, and P. Pattison, Eds., ed Washington, DC: National Research Council, 2003, pp. 133-145.
[31] C. Schreiber and K. M. Carley, "Agent Interactions in Construct: An Empirical Validation using Calibrated Grounding," in *Behavior Representation in Modeling & Simulation (BRIMS)*, Norfolk, VA, 2007.
[32] C. Schreiber, S. Singh, and K. M. Carley, "Construct - A Multi-agent Network Model for the Co-evolution of Agents and Socio-cultural Enviroments," Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Pttsburgh PA 15213, Pittsburgh, Pa.7 January 2011 2004.
[33] M. McPherson, L. Lovin, and J. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, vol. 27, pp. 415-444, 2001.
[34] C. E. Osgood, W. H. May, and M. S. Miron, *Cross-cultural universals of affective meaning*. Urbana-Champaign, IL: University of Illinois Press, 1975.
[35] D. Krackhardt, "Social Networks," in *Encyclopedia of group processes and intergroup relations*. vol. 2, S. Otten, C. Cohrs, J. M. Levine, and M. A. Hogg, Eds., ed Los Angeles, CA: Sage Publishing, 2010, pp. 817-821.
[36] B. R. Hirshman, J. St. Charles, and K. M. Carley, "Leaving Us in Tiers: Can Homophily be used to Generate Tiering Effects," *Computational and Mathematical Organization Theory*, vol. 17, pp. 318-343, 2011.
[37] M. Kowalchuck, "Construct Belief Model," M. J. Lanham, Ed., ed. Pittsburgh, PA, 2012.
[38] P. W. Pachowicz, L. W. Wagenhals, J. Pham, and A. H. Levis, "Building and Analyzing Timed Influence Net Models with Internet-Enabled Pythia," in *SPIE Defense and Security Symposium*, Orlando, FL, 2007, p. 292.
[39] L. W. Wagenhals, "Course of Action Development and Evaluation Using Discrete Event System Models of Influence Nets," George Mason University, Fairfax, Ph.D. Dissertation2000 2000.

[40] L. W. Wagenhals, & Levis, Alexander H., "Course of Action Development and Evaluation," in *The 2000 Command and Control Research and Technology Symposium*, Fairfax, 2002.
[41] J. Diesner and K. M. Carley, "AutoMap 1.2 : extract, analyze, represent, and compare mental models from texts," Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Pttsburgh PA 15213, Pittsburgh, Technical Report CMU-ISRI-04-100, 23 December 2004.
[42] W. J. Dixon, "Third-Party Techniques for Preventing Conflict Escalation and Promoting Peaceful Settlement," *International Organization*, vol. 50, pp. 653-681, 1996.
[43] K. Carley, J. Pfeffer, J. Reminga, J. Storrick, and D. Columbus. (2012, 5 March 2011). *ORA User's Guide 2012*. Available: http://www.casos.cs.cmu.edu/publications/papers/CMU-ISR-12-105.pdf
[44] K. M. Carley, I.-C. Moon, G. P. Morgan, and M. J. Lanham, "Adversary Modeling –Applications of Dynamic Network Analysis," in *Computational Modeling of Cultural Dimensions in Adversary Organizations*, K. M. Carley and A. Levis, Eds., ed. Fairfax, VA: The Volgenau School of Engineering Dept. of Electrical and Computer Engineering System Architectures Laboratory, George Mason University, 2010, pp. 172-203.
[45] K. M. Carley and J. Pfeffer, "Title," unpublished|.
[46] K. Wallsten, "Political Blogs and the Bloggers Who Blog Them: Is the Political Blogosphere and Echo Chamber?," presented at the American Political Science Association Annual Meeting, Washington, D.C., 2005.
[47] A. H. Levis and D. M. Perdu, "CAESAR II: A System for the Design and Evaluation of Command and Control Organizations," in *Proc. 2nd International Command and Control Research and Technology Symposium*, Market Bosworth, England, 1996.
[48] S. K. Kansal, A. M. AbuSharekh, and A. H. Levis, "Computationally Derived Models of Adversary Organizations," in *Computational Intelligence in Security and Defense Applications, 2007. CISDA 2007. IEEE Symposium on*, 2007, pp. 92-99.
[49] A. Ahmed. (2010, 4 January 2011). The 'Cold Start and Stop' strategy. *Insitute for Defence Studies and Analyses - Comment* [Electronic OpEd]. Available: http://www.idsa.in/idsacomments/TheColdStartandStopstrategy_aahmed_280910
[50] G. B. R. Kanwal. (2010, 4 January 2011). India's Cold Start Doctrine and Strategic Stability. *Insitute for Defence Studies and Analyses - Comment* [Electronic OpEd]. Available: http://www.idsa.in/node/5442/372
[51] H. V. Pant. (2010, 4 January 2011). India's quick-strike doctrine causes flutter. *The Japan Times* [Electronic Newspaper]. Available: http://search.japantimes.co.jp/cgi-bin/eo20100202a1.html
[52] S. S. Menon, "Speech by [Indian National Security Advisor] NSA Shri Shivshankar Menon at NDC on "The Role of Force in Strategic Affairs"," N. S. Council, Ed., ed. New Delhi, India: Government of India, 2010.



Michael J. Lanham (M'11) LTC Michael J Lanham earned his MS in Computer Science (CS) from University of Florida in 2002.

He is a PhD candidate at Carnegie Mellon University (CMU). He is a Lieutenant Colonel in the US Army and an Information Systems Management (FA53) officer and former Infantry officer. He has served as US Military Academy CS faculty, US Strategic Command/J65 Program Manager, Joint Functional Component Command-Integrated Missile Defense/J35, J63 and Deputy CIO, and US Army Central Information Assurance Program Manager. His current research interests are resilient command and control in contested cyber environments.



Geoffrey P. Morgan earned a BS degree in Information Sciences and Technology from the Pennsylvania State University in 2005.

He is a PhD candidate in Carnegie Mellon's Computation, Organization, and Society program. Previously, Morgan worked in industry, developing high fidelity models of human processes and prototypes of mixed-initiative robotic command and control system. Previous job titles include Artificial Intelligence Engineer, Technical Lead, and Advanced Technologies Lead. He has experience in agent-based modeling, cognitive modeling, and network-centric simulation. Morgan is interested in organizations, individuals, and how the two influence each other. He is an active member of INSNA and the Cognitive Science Society.



Kathleen M. Carley (M '06–SM '11) received her Ph.D. in mathematical sociology from Harvard in Boston, MA, USA and an S.B. in political science and an S.B. in economics from Massachusetts Institute of Technology, Boston, MA, USA.

Dr. Carley is currently a Professor of Computation, Organizations and Society in the department Institute for Software Research in the School of Computer Science at Carnegie Mellon University (CMU). She is also the director of the center for Computational Analysis of Social and Organizational Systems (CASOS). She holds affiliated positions and courtesy appointments in the departments of Social and Decision Sciences, Engineering and Public Policy, and Electrical and Computer Engineering at CMU. Her research centers on dynamic network analysis and agent-based modeling applied to issues such as security, counter-terrorism, information diffusion and social change.

She is a member of INSNA, ASA, Informs, and AAAS. She was the 2011 winner of the Simmel Award for outstanding contributions to network science and has served on numerous NRC committees, and IEEE workshop committees.