Early Warning Analysis for Social Diffusion Events

Richard Colbaugh
Sandia National Laboratories and New Mexico Tech
Albuquerque, NM USA
colbaugh@nmt.edu

Kristin Glass
New Mexico Tech
Socorro, NM USA
kglass@icasa.nmt.edu

***Abstract*—There is considerable interest in developing predictive capabilities for social diffusion processes, for instance enabling early identification of contentious "triggering" incidents that are likely to grow into large, self-sustaining mobilization events. Recently we have shown, using theoretical analysis, that the dynamics of social diffusion may depend crucially upon the interactions of *social network communities*, that is, densely connected groupings of individuals which have only relatively few links to other groups. This paper presents an empirical investigation of two hypotheses which follow from this finding: 1.) the presence of even just a few inter-community links can make diffusion activity in one community a significant predictor of activity in otherwise disparate communities and 2.) very early dispersion of a diffusion process across network communities is a reliable early indicator that the diffusion will ultimately involve a substantial number of individuals. We explore these hypotheses with case studies involving emergence of the Swedish Social Democratic Party at the turn of the 20th century, the spread of SARS in 2002-2003, and blogging dynamics associated with potentially incendiary real world occurrences. These empirical studies demonstrate that network community-based diffusion metrics do indeed possess predictive power, and in fact can be significantly more predictive than standard measures.**

***Keywords*—**-social dynamics, predictive analysis, early warning, mobilization and protest.

I. Introduction

Understanding the way ideas, innovations, behaviors, and diseases propagate over social networks is of considerable importance to the national security community [e.g., 1-6]. Of particular interest are *predictive* capabilities for such social diffusion, for instance enabling early warning as to the emergence of violence or outbreak of an epidemic. As a consequence, vast resources are devoted to the task of predicting the outcomes of diffusion processes; however, the quality of these predictions is often poor. It is tempting to conclude that the problem is one of insufficient information. Clearly diffusion phenomena which "go viral" are qualitatively different from those that quickly fade into obscurity or they wouldn't be so dominant, the conventional wisdom goes, so in order to make good predictions we must collect enough data to allow these crucial differences to be identified.

Recent research in the social and behavioral sciences calls into question this conventional wisdom and, indeed, indicates there may be fundamental limits to what can be predicted about social diffusion. For example, the studies reported in [7-9] indicate that the intrinsic attributes normally considered to be essential when assessing the likelihood of diffusion success, such as the quality of the various options in a social choice situation, often do not possess much predictive power. This research provides evidence that when individuals are influenced by what others do, it is not possible to obtain reliable predictions using standard methods which focus on intrinsic characteristics of potential process outcomes.

We propose that, to be successful, social diffusion prediction must be based on both the intrinsics of the diffusion process *and* the way individuals influence one another, and we describe a theoretical examination of this idea in [10]. The modeling framework employed in this study is inspired by recent work in biology demonstrating that stochastic hybrid dynamical systems (S-HDS) are a useful mathematical formalism with which to represent multi-scale biological network dynamics [11,12]. Using S-HDS models and a novel reachability-based analytic methodology we show in [10] that, for a wide range of social diffusion phenomena, useful prediction requires consideration of the way the diffusion dynamics interacts with *social network communities*, that is, densely connected groupings of individuals that have only relatively few links to other groups. The concept of network community structure is illustrated in Figure 1 and is defined more carefully in the next section. This finding is important because it applies to social diffusion taking place on any network with community structure, and such structure is ubiquitous in real world social networks [13,14].

This paper presents an empirical investigation of two hypotheses which follow from the theoretical work reported in [10]: 1.) the presence of even just a few direct links between disparate network communities can make diffusion activity in one community a significant predictor of diffusion activity in other communities and 2.) very early dispersion of a diffusion process across network communities is a reliable early indicator that the diffusion will ultimately reach a substantial number of individuals. We explore the first hypothesis with case studies involving the emergence of the Swedish Social Democratic Party at the turn of the 20th century and the spread of SARS in 2002-2003. The second hypothesis is examined through an analysis of blogging dynamics associated with a collection of recent incidents which appeared at the outset to have the potential to trigger large protests. These empirical studies offer support for the two hypotheses, demonstrating that network community-based diffusion metrics possess significant predictive power; indeed, these metrics are found to be more useful than standard measures. In addition, the second of the hypotheses provides the basis for deriving a new, readily implementable early warning method for large mobilization events.

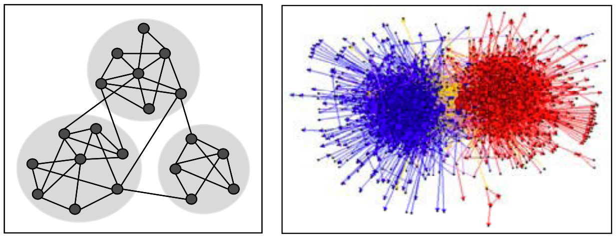

Figure 1. Network community structure. Cartoon at left depicts a network with three communities; graph at right is a network of political blogs in which communities of liberal (blue) and conservative (red) blogs are clearly visible [15].

II. SOCIAL DIFFUSION AND NETWORK COMMUNITIES

In social diffusion, individuals are affected by what others do. This is easy to visualize in the case of disease transmission, with infections being passed from person to person. Behaviors and innovations can also propagate through a population, as individuals pay attention to others in an attempt to obtain the benefits of coordinated action, infer otherwise inaccessible information, or manage complexity in decision-making. The dynamics of social diffusion can therefore depend upon the topological features of the underlying social network, for instance the degree distribution or presence of small world structure, and aspects of this dependence have been characterized (see [16] for a recent review). This dependence suggests that in order to derive useful prediction algorithms for diffusion phenomena, one should consider the topology of the underlying social network; however, standard prediction algorithms do not include such features [10].

While community structure is widely appreciated to be an important topological property in real world social networks, there is not a similar consensus regarding qualitative or quantitative definitions for this concept. Here we adopt the modularity-based definition proposed in [14], whereby a good partitioning of a network's vertices into communities is one for which the number of edges between putative communities is smaller than would be expected in a random partitioning. To be concrete, a modularity-based partitioning of a network into two communities maximizes the modularity Q, defined as

$$Q = s^T B\, s / 4m,$$

where m is the total number of edges in the network, the partition is specified with the elements of vector s by setting $s_i = 1$ if vertex i belongs to community 1 and $s_i = -1$ if it belongs to community 2, and matrix B has elements $B_{ij} = A_{ij} - k_i k_j / 2m$, with A_{ij} and k_i denoting the network adjacency matrix and degree of vertex i, respectively. Partitions of the network into more than two communities can be constructed recursively [14]. Note that modularity-based community partitions can be efficiently computed for large social networks and require only network topology data for their construction.

Despite the fact that community structure is ubiquitous in real social networks, little has been done to quantify the role of communities in diffusion dynamics or to incorporate such considerations into diffusion prediction methods. One reason for this gap in understanding could be that standard network analysis tools, such as those borrowed from statistical physics and econometrics, are not well-suited for investigating networks with community structure. In [10] we a present a theoretical analysis of social diffusion on networks with realistic topologies, including community structure. The analysis leverages S-HDS models for network dynamics. An S-HDS is a feedback interconnection of a discrete-state stochastic process, such as a Markov chain, with a family of continuous-state stochastic dynamical systems [11]. Combining discrete and continuous dynamics within a unified, computationally tractable framework provides an expressive, scalable modeling environment that is amenable to mathematically rigorous analysis. The analytic framework developed in [10] enables the derivation of provably-correct characterizations of social diffusion predictability on networks of real world scale and complexity. In particular, [10] presents predictability results for networks which possess realistic degree distribution, small world structure, modularity-based community structure, and exogenous inputs (e.g., from traditional media).

Using this theoretical framework in combination with empirically-grounded models for social diffusion [e.g., 16,17], we have demonstrated that the predictability of these models depends crucially upon network community structure. Indeed, one of the main conclusion of the study can be expressed in terms of two hypotheses concerning useful predictors of diffusion dynamics:

Hypothesis 1: The presence of even just a few (direct) links between network community A and a set of otherwise disparate network communities can make the diffusion activity at A's inter-community graph neighbors a significant predictor of the activity at A.

Hypothesis 2: Early dispersion of a diffusion process across network communities is a reliable early indicator that the ultimate reach of diffusion will be significant. For example, this measure is more predictive than the early magnitude of diffusion activity.

We now turn to an empirical assessment of these two hypotheses. In each case we begin by providing a more application-specific, and thus more precise, statement of the relevant hypothesis. Hypothesis 1 is explored via two case studies, the first involving the emergence and growth of the Swedish Social Democratic Party at the turn of the 20th century and the second focusing on the spread of the SARS virus during 2002-2003. Hypothesis 2 is investigated through an analysis of real world incidents that had the potential to trigger substantial mobilization activity and the associated diffusion of protest-related discussions through blog networks. The results of these case studies offer support for both Hypotheses 1 and 2, for instance showing that network community-based diffusion metrics possess significant predictive power, and in fact can be more useful than standard measures. The second case study also illustrates one way in which social network community structure can be leveraged to yield a simple, practically implementable early warning method for large diffusion events.

III. EMPIRICAL ASSESSMENT OF HYPOTHESIS 1

We begin assessment of Hypothesis 1 with a case study involving the Swedish Social Democratic Party (SDP). The SDP was founded in Stockholm in 1889 and grew to become one of the most successful political parties in the world. During the early years of the SDP, party activists traveled throughout Sweden attempting to generate interest and support. One consequence of this activity was the formation of a network of "long-range connections" linking previously disparate regions of the country [17]. Let us partition the set of all jurisdictional districts in Sweden into subsets of geographically-adjacent, regularly interacting districts (see [17] for a discussion); this partition will serve as a proxy for network community structure. The combination of intra-community links corresponding to geographic/demographic proximity and long-range inter-community links established by party activists enabled early diffusion of the SDP to take place via two main mechanisms: 1.) intra-community diffusion among geographically-adjacent districts *within* communities, and 2.) inter-community diffusion *between* geographically remote communities through activist-generated short-cut links (see Figure 2). These considerations lead to the following application-specific statement of Hypothesis 1: in the presence of direct inter-community links between a district A and a set of geographically remote districts, SDP membership growth in the remote districts is at least as predictive of A's growth as is membership growth in the districts geographically-adjacent to A. Because both membership dynamics and activist activities are well-documented [17,18], it is relatively convenient to gather the data required to empirically assess Hypothesis 1.

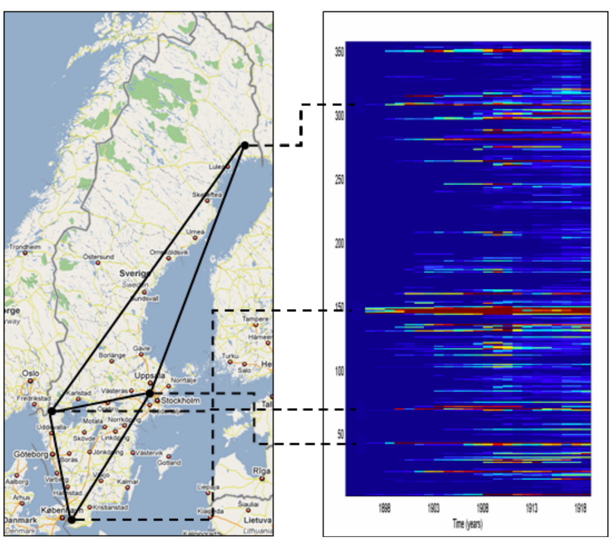

Figure 2. Illustration of SDP inter-community network. Map at left shows a few representative links connecting previously disparate network communities. Visualization at right depicts temporal evolution of concentration of SDP membership in 369 jurisdictional districts during the period 1889-1918.

Consider the evolution of SDP membership in the 369 jurisdictional districts of Sweden during 1889-1918 (the critical early years of party growth [17]). Figure 2 illustrates the basic features of this diffusion process. The visualization at the right of the figure depicts the temporal evolution of concentration of SDP members for all districts, with the vertical and horizontal axes corresponding to district index and time, respectively, and the colors indicating variation from minimum member concentration (dark blue) to maximum concentration (red); note that districts with similar indices are generally geographically proximate [18]. The map of Sweden on the left of Figure 2 illustrates the basic concept of an inter-community network by showing a few representative links in this graph. Time series analysis in [17] indicates that intrinsic characteristics of SDP diffusion, such as the appeal of certain activists and the demographics of local regions, are not predictive of membership growth. However, this analysis also suggests that membership levels in the K districts which form the vertices of the activist-induced inter-community network may possess predictive power, hinting at support for Hypothesis 1.

In order to assess Hypothesis 1 more directly, we examine the predictive power of the following simple model for party diffusion:

$$\Sigma_{SDP}: \quad m_i(t+1) = \alpha+\beta_1 m_i(t)+\beta_2 m_m(t)+\beta_3 m_{i,\ geo\text{-}m}(t)+\beta_4 m_{net\text{-}m}(t),$$

where t is time in years, $m_i(t)$ is party membership concentration (MC) in district i at time t, $m_m(t)$ is (country-wide) mean MC, $m_{i,\ geo\text{-}m}(t)$ is mean MC for the (geographically-based) network community in which i is located, $m_{net\text{-}m}(t)$ is mean MC for the K districts which define the inter-community network, and $\{\alpha, \beta_1, \beta_2, \beta_3, \beta_4\}$ are the model parameters. Obviously more sophisticated models could be employed to predict SDP MC growth. However, Σ_{SDP} is useful for our purposes because it provides a simple and natural representation of intra-community and inter-community influence, via the model terms $\beta_3 m_{i,\ geo\text{-}m}$ and $\beta_4 m_{net\text{-}m}$, respectively, and thereby enables quantification of the relative predictive power of these two components of network diffusion.

Estimating model Σ_{SDP} using data obtained from [18] for the K districts which possess both intra-community and inter-community interactions reveals that: 1.) all of the coefficients $\{\beta_1, \beta_2, \beta_3, \beta_4\}$ are statistically significant predictors of party membership growth ($p<0.05$), and 2.) $\beta_4 m_{net\text{-}m}$ is the most predictive term in the model, explaining nearly 60% of the variance in membership growth by itself. Observe that the latter result is somewhat surprising: inter-community interaction is more predictive of SDP membership for a given district for the next year than the current membership level for that district, and is much more predictive than membership levels in geographically-adjacent districts. Thus inter-community interaction, which is typically not considered in diffusion prediction algorithms, is more predictive than either process intrinsics or standard diffusion measures. The reader interested in a more complete description of the estimated model Σ_{SDP} is referred to the report [19].

The next case study illustrates the way Hypothesis 1 can be used to guide development of simple but still useful models for diffusion phenomena. Consider the outbreak and rapid propagation of severe acute respiratory syndrome (SARS) in 2002-2003. To obtain a very simple model for the global spread of SARS, we use individual countries as proxies for network communities and represent intra-community disease transmission with a standard stochastic susceptible-infected-removed model for "fully mixed" populations [e.g., 6]. Inter-community interactions are assumed to consist of individuals traveling between countries along commercial air travel routes (data on air travel was obtained from [20]). This characterization of inter-community dynamics results in a model which is easily and rapidly constructed from publicly available data. Moreover, despite its simplicity, we find that the model provides a useful description for the spread of SARS.

Figure 3 shows that simulations of the model are in good agreement with the actual spread of the SARS epidemic during 2002-2003. The utility of this simple model is a consequence of allocating modeling detail according to Hypothesis 1: because inter-community interaction is a key element of social network diffusion, this dynamics is modeled more carefully than intra-community interaction. Observe that network diffusion models of this form enable reasonably accurate assessments via rapid, inexpensive analysis, indicating that they can play an important role in national security applications. Additional details regarding the model are presented in [19].

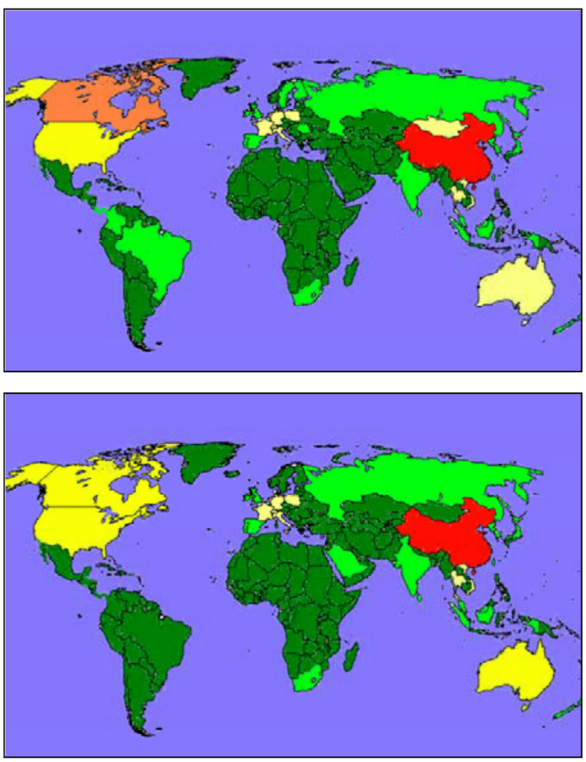

Figure 3. Geographic spread of SARS epidemic: actual (top) and simulated (bottom) cumulative SARS infection level by country (dark green is lowest, red is highest).

IV. Empirical Assessment of Hypothesis 2

There is considerable interest to develop methods for distinguishing successful mobilization and protest events, that is, mobilizations that become large and self-sustaining, from unsuccessful ones early in their lifecycle. This task is naturally cast as one of identifying reliable "indicators" which signal that a mobilization will be successful. Hypothesis 2 suggests that early dispersion of mobilization-related activity across disparate network communities may be such an indicator. In order to examine this possibility, we study Muslim reaction to nine recent incidents, each of which appeared at the outset to have the potential to trigger significant protests. This collection of events includes publication of cartoons depicting Mohammad in the Danish newspaper *Jyllands-Posten* in September 2005, the lecture given by Pope Benedict XVI in September 2006 quoting controversial material concerning Islam, and republication of the "Danish cartoons" in various newspapers in February 2008. Recall that the first Danish cartoons event ultimately led to substantial Muslim mobilization, including massive protests and considerable violence, while in contrast Muslim outrage triggered by the pope lecture and the second Danish cartoons event subsided quickly with essentially no violence. The set of nine events considered here contains four which became large and self-sustaining and five that quickly dissipated, so that taken together these events provide a useful setting for testing whether the presence of early diffusion across social communities can be used to distinguish nascent mobilization events which will become successful from those that will not.

More specifically, we consider the following application-relevant statement of Hypothesis 2: a reliable early indicator that a triggering event will lead to significant protest activity is the presence of "large" early dispersion of protest-related discussions across disparate social network communities. It is, of course, challenging to collect the data which would enable a quantitative assessment of this hypothesis. For example, such data should include social network information sufficient to allow identification and analysis of network communities as well as detect protest-related discussions among the individuals that make up the network. The present study addresses this challenge by using *online* social activity as a proxy for real world diffusion of protest-related discussions and information. More specifically, we use blog posts as our primary data set. The "blogosphere" is modeled as a graph in which the vertices are blogs and the edges represent links between blogs, with two blogs being linked if a post in one hyperlinks to a post in the other. Among other things, this blog graph model enables the identification of blog communities: we partition the blog graph into network communities by maximizing the modularity Q for the graph (see Section II and reference [14]). Recall that this computation yields groups of blogs with inter-group edge densities that are significantly smaller than expected; these blog groups serve as our proxy for social network communities.

We assess the veracity of Hypothesis 2 by first determining if the "early dispersion" indicator is present in blog data associated with the nine triggering events of interest, and then evaluating whether the presence or absence of this indicator is predictive of mobilization success. We check for the presence of the early dispersion indicator with the following procedure:

For each of the nine potential triggering events:

1. Perform a blog graph crawl to collect relevant blog posts and build the associated blog graph.
2. Partition the blog graph into network communities.
3. Assemble post volume time series and compute the post/community entropy (PCE) time series associated with the post volume dynamics.
4. Construct a synthetic ensemble of PCE time series from the post volume dynamics.
5. Compare actual PCE time series to the synthetic ensemble series to determine if the observed early diffusion of activity across communities is "large".

We now provide a few additional details concerning this procedure. Step 1 is by now standard, and various off-the-shelf tools exist which can perform this task [19]. In Step 2, blog network communities are identified with a modularity-based community extraction algorithm applied to the blog graph [14,19]. In Step 3, post volume for a given community i and sampling interval t is obtained by counting the number of relevant posts made in the blogs comprising community i during interval t. PCE for a given sampling interval t is defined as follows:

$$\mathrm{PCE}(t) = -\Sigma_i \, f_i(t) \log(f_i(t)),$$

where $f_i(t)$ is the fraction of total relevant posts made during interval t which occur in community i. Given the post volume time series obtained in Step 3, Step 4 involves the construction of an ensemble of PCE time series that would be expected under "normal circumstances", that is, if Muslim reaction to the triggering event diffused from a small seed set of initiators according to the S-HDS social dynamics model introduced in [10]. Observe that this step enables us to quantify normal dispersion in terms of PCE(t), so that we can recognize "large" dispersion. Finally, Step 5 is carried out by searching for time periods, if any, during which the actual PCE time series exceeds the mean of the synthetic PCE ensemble by more than two standard errors.

Illustrative time series plots associated with the proposed approach to mobilization/protest early warning are shown in Figure 4. Observe that in the case of the first Danish cartoons event the PCE of relevant discussions (blue curve) experiences a dramatic increase a few weeks before the corresponding increase in volume of blog discussions (red curve); this latter increase, in turn, takes place before any violence. In contrast, in the case of the pope event, PCE of blog discussions is small relative to the cartoons event, and any increase in this measure lags discussion volume. Similar time series plots are obtained for the other seven events, suggesting that early dispersion of discussions across blog network communities may be a useful indicator for large mobilization events.

To examine this possibility more carefully, we evaluate the predictive performance of two candidate early indicators of large protest events: 1.) the "early dispersion" PCE indicator summarized above, and 2.) a simple volume-based indicator, in which the presence or absence of significant post volume is used to predict which events will become large and which will dissipate. The performance of these warning schemes is compared with a (null hypothesis) "random" predictor, which assigns predictions of large/small protests at random (with the correct probability distribution for large/small events). We find that the PCE indicator, which is based on Hypothesis 2, provides statistically significant prediction accuracy ($p<0.005$) while the volume-based method does not (it is not significant at even the $p=0.1$ level). This case study therefore provides empirical support for Hypothesis 2, indicating that early diffusion of mobilization-related discussions across disparate social network communities is a reliable early signature of successful mobilizations. Moreover, this investigation yields a new, readily implementable early warning method for large diffusion events.

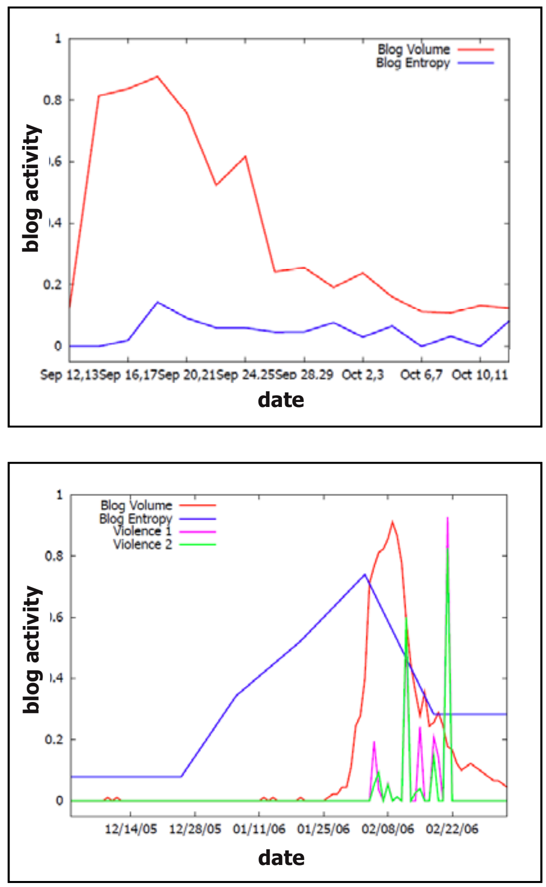

Figure 4. Sample results for the mobilization/protest case study. The illustrative time series plots shown correspond to the pope event (top) and first Danish cartoons event (bottom). In each plot, the red curve is blog volume and the blue curve is blog entropy; the Danish cartoon plot also shows two measures of violence (cyan and magenta curves). Note that while the volume and violence data are scaled to allow multiple data sets to be graphed on the same plot, the scale for entropy is consistent across plots to enable cross-event comparison.

Acknowledgements

This research was supported by the U.S. Department of Homeland Security, the U.S. Department of Defense, and the Laboratory Directed Research and Development program at Sandia National Laboratories.

References

[1] Moghadam, A., *The Globalization of Martyrdom: Al Qaeda, Salafi Jihad, and the Diffusion of Suicide Attacks*, Johns Hopkins University Press, Baltimore, 2008.

[2] Myers, D. and P. Oliver, "The Opposing Forces Diffusion Model: The Initiation and Repression of Collective Violence", *Dynamics of Asymmetric Conflict*, Vol. 1, pp. 164-188, 2008.

[3] Chen, H., C. Yang, M. Chau, and S. Li (Editors), *Intelligence and Security Informatics*, Lecture Notes in Computer Science, Springer, Berlin, 2009.

[4] Krueger, A. and J. Maleckova, "Attitudes and Action: Public Opinion and the Occurrence of International Terrorism", *Science*, Vol. 325, pp. 1535-1536, 2009.

[5] Bergin, A., S. Osman, C. Ungerer, and N. Yasin, "Countering Internet Radicalization in Southeast Asia", ASPI Special Report, March 2009.

[6] Anderson, R. and R. May, *Infectious Diseases of Humans*, Oxford University Press, 1992.

[7] Walls, W., "Modeling Movie Success When 'Nobody Knows Anything': Conditional Stable-Distribution Analysis of Film Returns", *Journal of Cultural Economics*, Vol. 29, pp. 177-190, 2005.

[8] Salganik, M., P. Dodds, and D. Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market", *Science*, Vol. 311, pp. 854-856, 2006.

[9] Colbaugh, R., K. Glass, and P. Ormerod, "Predictability of 'Unpredictable' Cultural Markets", *Proceedings of International Conference on Social Computing, Behavioral Modeling, and Prediction*, Washington DC, March 2010.

[10] Colbaugh, R. and K. Glass, "Predictive Analysis for Social Processes I: Multi-Scale Hybrid System Modeling, and II: Predictability and Warning Analysis", *Proceedings of the 2009 IEEE Multi-Conference on Systems and Control*, Saint Petersburg, Russia, July 2009.

[11] El-Samad, H., S. Prajna, A. Papachristodoulou, J. Doyle, and M. Khammash, "Advanced Methods and Algorithms for Biological Networks Analysis", *Proceedings of the IEEE*, Vol. 94, pp. 832-853, 2006.

[12] Yuan, C., X. Mao, and J. Lygeros, "Stochastic Hybrid Delay Population Dynamics: Well-posed Models and Extinction", *Journal of Biological Dynamics*, Vol. 3, pp. 1-21, 2009.

[13] Newman, M., "The Structure and Function of Complex Networks", *SIAM Review*, Vol. 45, pp. 167-256, 2003.

[14] Newman, M., "Modularity and Community Structure in Networks", *Proceedings of the National Academy of Sciences USA*, Vol. 103, pp. 8577-8582, 2006.

[15] Adamic, L. and N. Glance, "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog", *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, Chicago, IL, 2005.

[16] Easley, D. and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, 2010.

[17] Hedstrom, P., R. Sandell, and C. Stern, "Mesolevel Networks and the Diffusion of Social Movements: The Case of the Swedish Social Democratic Party", *American Journal of Sociology*, Vol. 106, pp. 145-172, 2000.

[18] Swedish National Data Service, Popular Movement Archive, 1881-1950 Social-Democratic Labor Party of Sweden, Data Set 062, accessed 2007.

[19] Colbaugh, R. and K. Glass, "Predictive Analysis for Social Processes", Sandia National Laboratories SAND Report 2009-0584, January 2009.

[20] http://www.iata.org/index.htm, accessed 2007.