Routledge
Taylor & Francis Group

Journal of Applied Security Research

ISSN: 1936-1610 (Print) 1936-1629 (Online) Journal homepage: https://www.tandfonline.com/loi/wasr20

Modeling Hacktivism Using System Dynamics

Jakub Drmola, Martin Bastl & Miroslav Mares

To cite this article: Jakub Drmola, Martin Bastl & Miroslav Mares (2015) Modeling Hacktivism Using System Dynamics, Journal of Applied Security Research, 10:2, 238-248, DOI: 10.1080/19361610.2015.1004514

To link to this article: https://doi.org/10.1080/19361610.2015.1004514

Journal of Applied Security Research, 10:238–248, 2015

ISSN: 1936-1610 print / 1936-1629 online
DOI: 10.1080/19361610.2015.1004514

Routledge
Taylor & Francis Group

Modeling Hacktivism Using System Dynamics

JAKUB DRMOLA, MARTIN BASTL, and MIROSLAV MARES
Faculty of Social Studies, Masaryk University, Brno, Czech Republic

This article deals with modeling hacktivism. The methodological approach—system dynamics—is described. Concept of hacktivism is analyzed in relation to various dimensions of the use of this term. The model of hacktivism is explained. The core of this model is formed by an adapted risk equation and then expanded upon using causal feedback loops. The authors come to conclusion that system dynamics model is useful in the field of cybersecurity research, however, further work on this scientific approach is necessary.

KEYWORDS *Hactivism, system dynamics, cyberattacks*

INTRODUCTION

In the field of research of and attempts to shed light on cyberthreats, system dynamics enables us to complement the technical perspective in a very comprehensive way. This approach not only takes into consideration the social, political, cultural, or economic and military dimensions, but is also free of shortcomings inherently present when we work with analogies.

Generally speaking, we come across at least two types of analogies that might mislead us when we try to understand the nature of cyberthreats. The first type is represented by analogies of deeds and activities performed in the real, physical world. Crime versus cybercrime in the framework of criminality, and naval or airborne combat in the framework of warfare (Ventre, 2011, pp. 184–202) are typical examples. The second type consists in the analogy of environment that has led to the elaboration of the disputable concept of the so-called "fifth domain" and introduction of this concept into research and doctrinal materials.

Both analogies as well as attempts to examine cyberthreats as isolated phenomena (case studies), show a significant extent of reductionism.

Address correspondence to Dr. Miroslav Mares, Faculty of Social Studies, Masaryk University, Jostova 10, 602 00, Brno, Czech Republic. E-mail: mmares@fss.muni.cz

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/wasr.

System approach, on the contrary, enables us to study and understand the mechanism of individual cyberthreats in their complexity or at least with an acceptable extent of reduction while taking into account the most important factors and ties between them. It is flexible and universal enough to be applied on a whole spectrum of threats without being tied to specific, narrowly delimitated cases.

While it can be doubted that in the case of a large military conflict between states, cyberwarfare should have roles other than that of a subsidiary means or multiplier of force, in cases of low intensity conflicts and conflicts on the social level (Arquilla & Ronfeld, 1996; Arquilla & Ronfeld, 2001) cyberattacks, as means of asserting interests, have been used for a long time. This is represented by hacktivism, which will be dealt with in the following text. Understanding the mechanism that influences courses of conflicts and ways they happen could have a very significant practical impact, and system dynamics, as one of possible approaches, offers a powerful tool for its effective analysis.

An analysis of hacktivist movements seems to be an important object of research from the point of view of various other subfields. The cybersecurity (as a field integrating technological and societal threats emerging from cyberspace) is the most significant one. Because of engagement of extremist oriented persons among hacktivists the regime security is threatened. It means that the research on extremism and terrorism is involved in this research too. The misuse of these people by organized crime structures can be also mentioned from the point of view of criminology, especially from the point of view of research on organized crime. Hacktivists can be—deliberately or nondeliberately—used by foreign enemies of a country as a specific "fifth column" (manipulation of domestic hacktivist movements by foreign secret services is a possible way of such event). This is a research issue for strategic studies. Last but not least, the human and social aspect of hacktivism demands attention from psychology and sociology as well. The model and methods used in this article seem to be useful in all the previously mentioned fields.

SYSTEM DYNAMICS: THE METHOD

The method used here is still relatively uncommon, despite being quite old. Therefore, it is probably appropriate to briefly introduce it. Its roots go all the way back to the beginnings of the Cold War and to the first electronic guidance and radar systems. From there, mainly through the work of Jay Forrester, it had spread into the domain of economy and business management in the early 1960s (see Forrester, 1961, 1968). Underlying ideas of system dynamics also influenced the entire concept of natural ecosystems and ecology.

But it was only very recently that the potential of system dynamics was recognized in social sciences and, more importantly, security studies. The breakthrough came in 2000s, probably under the influence of increased interest in international security and asymmetrical conflicts during the so-called *War on Terror*. Terrorism itself became a phenomenon to be modeled using this approach. There was a model of ethnic terrorism (Akcam & Asal, 2005), model of terrorist groups in the Middle East and their interaction with the U.S. military (Gil et al., 2005), and even models of internal structure, management, and coordination of terrorist groups (see Madnick & Siegel, 2007 or Schoenenberger, Schenker- Wicki & Beck, 2012). Systemic approach was also used to model insurgencies (Choucri et al., 2006; Leweling & Sieber, 2007; Smith, 2002; Wakeland & Medina, 2010) and information security (Behara, Huang, & Hu, 2007; Foroughi, 2008). These few examples show the growing role system modeling plays in security research.

Its basic premise, as the name suggests, is that every object or phenomenon can be treated as a system whose properties and behavior emerge from its underlying structure. This structure can be represented as a collection of interconnected nodes or, more specifically, stocks and flows which contain and channel material, money, people, energy or information throughout the system. The crucial advantage of system dynamics is the inclusion of feedback loops which transform otherwise linear path of progression into a nonlinear, circular network of causes and effects without a clear beginning or an end. As a result of these internal interactions and causalities (positive or negative), the entire system can exhibit varying overall behavior on observed variables, such as oscillation, run-away effect, balanced state, or their combinations (see Cavana & Mares, 2004 or Radzicki & Taylor, 1997 for more details).

So when a new group, conflict, or event is being analyzed, the goal is to essentially crack open the proverbial black box, shine some light into it, explore its innards, and build a corresponding abstract model which allows one to understand its behavior and, under ideal circumstances, even predict and influence its future states. This naturally has great potential utility but it is also very difficult to fully achieve, especially when modeling complex social systems such as terrorism or hacktivism (often due to lacking data). While technical or economical models are usually fully numerical (quantitative), security researches sometimes do away with exact mathematical expressions of causalities (see previous examples) and create what might be termed qualitative systems models, which accentuate conceptual coherence and intelligibility. That is also the case of this model of hacktivism, which is a very recent phenomenon (see Coleman, 2013 or Norton, 2012 for more). Another issue common to complex social systems is the difficulty of deciding which variables and links ought to be included in their models. These systems rarely have any clear-cut natural boundaries so it often comes down to the purpose of the model itself.

HACKTIVISM

Hacktivism itself is a relatively new concept and there is not a broad consensus on what exactly it actually means, therefore, it needs to be delineated before trying to build a model of it. The optimal starting point seems to be Denning's spectrum of political action in cyberspace, where hacktivism lies between online activism and cyberterrorism (Denning, 2001). Its distinguishing feature is its disruptive nature. Whereas activism aims to promote its point of view, lecture, persuade, or even disseminate propaganda, hacktivism is directed at suppressing or subverting opposing messages, blocking communication channels, and misappropriating their digital assets and information. Tools of activism are Web sites, blogs, podcasts, mailing lists, forums, and social networks. Tools of hacktivism are Distributed Denial of Service attacks (DDoS), Web site defacements, SQL database injections, social engineering or peer-to-peer file transfer networks for distributing their spoils. So while activists will strive to be more visible and compelling than their opponents, hacktivists would directly undermine, displace, and ridicule them by disrupting their information flows in cyberspace.

Cyberterrorism sits on the other end of the spectrum. It differs from hacktivism by its destructive nature that goes far beyond mere disruption. There is of course no clear break-off point or boundary where hacktivism would end and cyberterrorism begin, but some form of serious fear-generating violence would probably need to take place to make its outcome effectively comparable with the conventional, noncyber terrorism. To achieve that, the attackers would most likely need to break through the so-called kinetic barrier and produce real-world physical effects leading either to death or significant material damage. Cyberattacks targeting dams, chemical plants, or transportation infrastructure are often proposed as possible scenarios leading to such dramatic effects. But as of now, achieving these physical results remains exceedingly difficult and rare and so the cyberterrorism itself remains a hypothetical concept.

As much as it is important to determine approximate boundaries of hacktivism regarding its tools and produced effects, it is also crucial to establish how its motives and goals set it apart from other forms of aggression in cyberspace. Probably the biggest overlap in tools exists with cybercrime. But telling these two activities apart is rather straightforward. Cybercrime is overwhelmingly driven by monetary profit, usually through confidence tricks, exploitation of stolen credentials and extortion. Whether it is large scale spam operations or cyberattacks directed at individuals, the goal is virtually always to extract money from the victims. Hacktivism, on the other hand, is a thoroughly political and ideological endeavor motivated by gaining influence or diminishing someone else's influence by shutting down their message, making them look silly or attacking their reputation. While some

individuals can regularly engage in both activities, the disparity between the two kinds of attacks is quite glaring.

But there exists one another class of cyberattacks, mainly in the form of sabotage (such as Stuxnet) and espionage (such as Flame or Red October). These pursue somewhat different kinds of political advantage and are generally closer to intelligence and military operations rather than ideological ones. They visibly differ from hacktivism by their covert or even clandestine nature, which is a characteristic usually shared by cybercrime as well. They shun attention whereas hacktivism covets it. Hacktivists actively seek publicity and sometimes even go as far as to announce their forthcoming targets to the media days in advance. So this basically sums up the notion of hacktivism being modeled here: disruptive and subversive (but not destructive) attention-seeking cyberattacks executed by nonstate actors for political or ideological gains.

THE MODEL

The model itself (see Figure 1) was built gradually, from the ground up and by progressively expanding causal connections. However, it is not possible to include the entire process here as it would be significantly beyond the scope of this article. A more advanced version of this qualitative model is presented here instead, together with commentary pertaining to the variables, causalities, and their interactions.

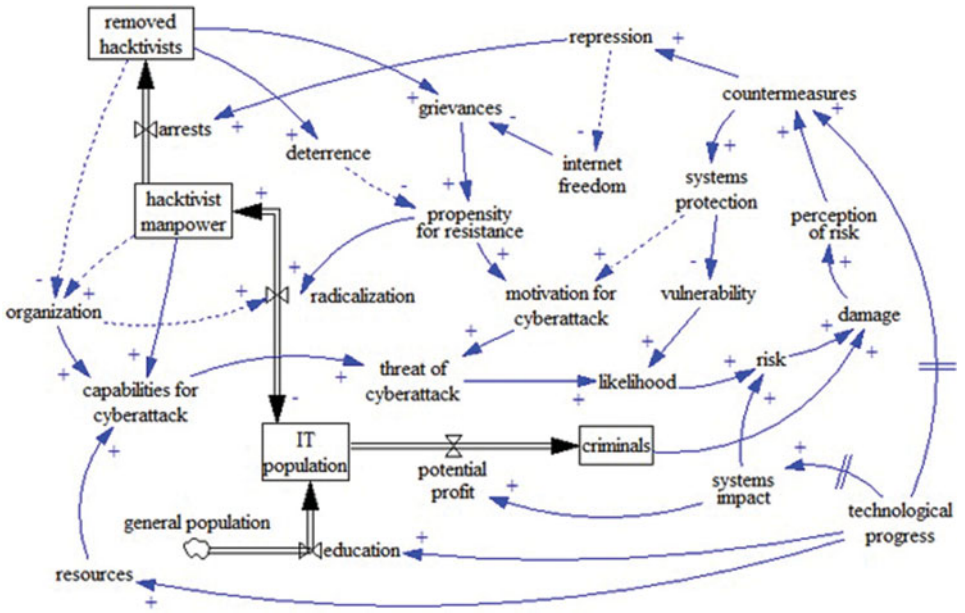

FIGURE 1 Dynamic model of hacktivism.

The core of this model is formed by an adapted risk equation, which can be seen in the central horizontal plane. Motivation and capabilities for cyberattack come together to form a threat which combined with vulnerabilities increases (hence the plus sign) the likelihood or probability of hacktivist cyberattack occurring. Risk is then produced by the addition "systems impact" variable that describes the value of compromised systems. Hacking into a police database will generally have a higher impact than compromising someone's twitter account, for example.

When this risk produces some appreciable damage, it generates increased perception of risk. Because the risk itself is abstract and cannot be perceived directly by organizations, institutions, or the society, the observed and reported damage facilitates this transfer of information from one set of actors to the other. As the perception of risk feeds into countermeasures, the path of causes and effects starts to loop back on itself and produce feedbacks.

The first and most straightforward one of these is the reduction (minus sign) of vulnerabilities. As risks increase and generate countermeasures via the rising perception of risks, incentives to better protect vulnerable systems are produced. Consequently, a balancing feedback loop is created as risks are being reduced because of decreasing vulnerabilities.

An example of reinforcing or escalating feedback loop branches off at the systems protection variable. Growing efforts to protect targeted systems can paradoxically entice some hacktivists to pursue the challenge and attempt to defeat any new protection put in place. This behavior can emerge from a belief that rewards are bound to be worth the stepped up effort or even something as simple as bragging rights can play a role. However, this link is comparatively weak and of much less significance, which is why it is denoted by the dashed line.

Further to the left are feedback loops emerging from repressive countermeasures. While the aforementioned protection of vulnerable systems can be seen as a passive defense or precautionary steps taken to prevent the attacks from being successfully executed in the first place, repressive measures usually take place after attacks and are meant to punish the hacktivists for their actions. This can also have some deterrent effect by dissuading potential attackers from these endeavors and it can additionally decrease the threat of future cyberattacks by weakening existing hacktivist organizations and pruning their ranks of their most active members. These effects would mostly manifest themselves in decreased capabilities for cyberattacks.

Yet these are not the only effects. This repression often induces considerable grievances among existing hacktivists, their supporters and sympathizers, as well as wider population. Arrests and indictments can be perceived as unjust or unreasonably harsh. Limiting of Internet freedoms can be seen as excessive. Such grievances can in turn attract new supporters or further

radicalize existing ones, thus countering intended decline of capabilities and even leading to potential increase of overall threat by boosting hacktivist numbers and their dedication to the cause.

Therefore, any active and repressive measures should be carefully considered in advance and analyzed for their expected net effect on risk. Removal of individuals will not always have a very large impact on the performance of hacktivist groups due to their horizontal, autonomous, and free-for-all structure. The degree of anonymity provided by cyberspace also limits effectiveness of deterrence. For that reason, repression should be highly targeted and employed only against individuals who clearly pose higher risk than any radicalization their removal might bring about, lest these countermeasures become counterproductive.

These transformations and movements of people within the model are shown as stocks and flows. The most important stock is the hacktivist manpower, which directly influences their capabilities. The outflow from this stock leads to the removed hacktivists stock and the rate of this flow is determined by the arrests variable. The second connected stock is, for the lack of better concise name, called information technology (IT) population. This denotes segment of general population which has considerable information and communication technology skills and is basically a recruitment pool for hacktivists. Crucially, this flow can run both ways depending on the current value of the radicalization variable. So when it is high, which would be mostly consequence of existing grievances, people flow towards the manpower stock. When it is low, due to absent grievances or somehow successful deterrence, they can deradicalize and flow back into the original stock.

The other outflow from that stock leads towards the stock of cybercriminals. This part of the model is very simplified and does not describe an actual system of cybercrime. But it serves the purpose of showing both the alternative outflow and also an alternative source of damage and risk perception. The rate of flow here is an overall potential profitability of this kind of crime.

The source of IT population is the general population which becomes transformed by some type of formal schooling or even self-education in this field. In other words, hacktivists are those members of the population who possess both the necessary skills and the ideological zeal.

Interestingly, technological progress, located at the bottom right, seems to cause growth of many variables across the model. Not only it increases hacktivist capabilities and the population conversion rate (thus expanding the recruitment pool), but it also feeds into countermeasures. Another consequence of continuing technological progression is growing pervasiveness of these technologies which in turn makes our society more dependent upon them, thus increasing value of these information systems and impact of

possible attacks on them. In a sense, the entire dynamic model of hacktivism is driven forward by this ever-growing variable which serves as a main external input into the system.

THE CYCLE

To demonstrate the applicability of the model (Figure 1) on the reality we can look to the recent history to see if it fits. Due to the nature of system dynamics it makes no sense to look for any sort of beginning or an end—there are none. But a notable and well-documented event can be selected and sequence of its preceding causes and subsequent effects explored. For the purposes of this article that notable event is attack on HBGary Federal in 2011 (see Bright, 2011).

So what were the causes? The attack itself was a realization of a risk growing out of surprising vulnerability on part of the company itself and a considerable threat of attack from the Anonymous movement. While the capabilities of this movement did not change much preliminary to the attack, their motivation for it grew significantly in the space of just a few days. Right before the attack the company announced that it managed to deanonymize members of the hacktivist movement and it planned to sell its know-how. The hacktivists perceived this as an attack on the freedom of the Internet and decided to hack into HBGary's servers.

And why did HBGary try to deanonymize these hacktivists? Because just 2 months earlier, in December 2010, Anonymous DDoSed several financial institutions (such as Visa and MasterCard) in a retaliation for their refusal to process donations for WikiLeaks (Bryan- Low & Grundberg, 2010). Unresponsive servers of these corporations were a visible damage which prompted new countermeasures. The effort to infiltrate Anonymous and uncover their identities was one of them.

While it might be possible to continue this regression, let us look as the sequence of effects which followed after this attack instead. Empowered by their success, several of the hacktivists involved in the HBGary attack formed a highly active group known as LulzSec. Their activities soon provoked significant countermeasures leading to several arrests, convictions, and their removal from the stock of active hacktivists (Bright, 2012). In turn, these arrests induced some new grievances, but due to the rather low popularity of LulzSec among other hacktivists (mostly because of their recklessness) the net effect was an overall decrease of threat resulting from decreased capabilities and possibly some successful deterrence. It should be noted that this would probably not be the case if the targeted hacktivists exercised more restraint in their behavior. This causal sequence can be shown as a table (see Table 1).

TABLE 1 Causality Table

Event	Variable	Timeframe
Renouncing WikiLeaks	Decrease of Internet freedom	Winter 2010
DDoS wave	Manifested risk	From December 6, 2010
Unresponsive servers	Apparent damage	December 8–10, 2010
Containing Anonymous	Countermeasures	From December 2010
Deanonymization effort	Decrease of Internet freedom	January 2011
HBGary going public	Motivation for the attack	February 5, 2011
HBGary attack	Manifested risk	February 5–6, 2011
Leaking documents	Apparent damage	February 2011
LulzSec investigation	Repression	Spring– Autumn 2011
Arrests and convictions	Removed hacktivists	From Spring 2011
Declaration of reprisals	Motivation for future attacks	From Summer 2011

CONCLUSION

This article can be seen as another manifestation of the expanding number of disciplines where system dynamics can provide new perspectives and new insights into existing complex problems our society currently faces. Given the rapidly rising prominence of cybersecurity these new perspectives cannot come too soon and we should use all the tools at our disposal to broaden our understanding. Moreover, even as system dynamics modeling proves its mettle in less technical and mathematical disciplines all current security models still need to be refined and enhanced. The ultimate goal would be building a fully mathematical working model of hacktivism with some predictive capabilities regarding leverage of specific policies. Secondly, connecting existing models to the wider "supersystem" of cybersecurity is also a goal worth pursuing.

It also brings us one step closer towards understanding hacktivism as a complex system made up of larger number of interacting nodes, a dynamic clockwork of sorts, instead of singular and insulated phenomenon composed solely of disruptive attacks and their perpetrators. This level of understanding is crucial if we intend to try influence the system because it can potentially react in a counterintuitive way. Our possibly erroneous understanding combined with the nonlinear structure of the system might, if we are not careful, lead to counterproductive measures that exacerbate issues instead of fixing them. More research is necessary in order to avoid that and system dynamics modeling just might be one of the appropriate tools to combine social, political, cyber, and security elements required to make headway.

FUNDING

This contribution was prepared as part of the research project "Methods of Predicting Long-term Geopolitical Development in Central

Europe–VF20102015005," funded by the Ministry of Interior of the Czech Republic.

REFERENCES

Akcam, B. K., & Asal, V. (2005). *The dynamics of ethnic terrorism*. System Dynamics Research Colloquium. University at Albany, State University of New York, Albany, NY. Retrieved from http://www.systemdynamics.org/conferences/2005/proceed/papers/AKCAM225.pdf

Arquilla, J., & Ronfeldt, D. (1996). *The Advent of Netwar*. Santa Monica, CA: RAND.

Arquilla, J., & Ronfeldt D. (2001). *Networks and netwars: The future of terror, crime, and militancy*. Santa Monica, CA: RAND.

Behara, R., Huang, C. D., & Hu, Q. (2007). A system dynamics model of information security investments. *ECIS 2007 Proceedings* (Paper 177). Florida Atlantic University. Retrieved from http://is2.lse.ac.uk/asp/aspecis/20070016.pdf

Bright, P. (2011, February 16). Anonymous speaks: The inside story of the HBGary hack. *ArsTechnica*. Retrieved from http://arstechnica.com/tech-policy/2011/02/anonymous-speaks-the-inside-story-of-the-hbgary-hack/

Bright, P. (2012, March 10). With arrests, HBGary hack saga finally ends. *ArsTechnica*. Retrieved from http://arstechnica.com/tech-policy/2012/03/the-hbgary-saga-nears-its-end/

Bryan-Low, C., & Grundberg, S. (2010, December 8). Hackers rise for WikiLeaks. *The Wall Street Journal*. Retrieved from http://online.wsj.com/article/SB10001424052748703493504576007182352309942.html

Cavana, R. Y., & Mares, E. D. (2004, Fall). Integrating critical thinking and systems thinking: From premises to causal loops. *System Dynamics Review*, *20* (3), 223–235. doi:10.1002/sdr.294

Choucri, N., Electris, C., Goldsmith, D., Mistree, D., Madnick, S. E., Morrison, J. B., … Sweitzer-Hamilton, M. (2006, March 5–12). Understanding & modeling state stability: Exploiting system dynamics. Proceedings of the 2006 IEEE Aerospace Conference, Big Sky, Montana. Retrieved from http://web.mit.edu/smadnick/www/wp/2006-02.pdf

Coleman, G. (2013). *Anonymous in context*. CIGI, Internet Governance (Paper No. 3). Retrieved from http://www.cigionline.org/sites/default/files/no3_8.pdf

Denning, D. E. (2001). *Activism, hacktivism, and cyberterrorism: The Internet as a tool for influencing foreign policy*. Nautilus Institute. Retrieved from http://faculty.nps.edu/dedennin/publications/Activism-Hacktivism-Cyberterrorism.pdf

Foroughi, F. (2008). The application of system dynamics for managing information security insider-threats of IT organization. Proceedings of the World Congress on Engineering 2008, Vol I, WCE 2008, London, UK. Retrieved from http://www.iaeng.org/publication/WCE2008/WCE2008_pp528-531.pdf

Forrester, J. W. (1961). *Industrial dynamics*. Waltham, MA: Pegasus Communications.

Forrester, J. W. (1968). *Principles of systems*. Cambridge, MA: Wright-Allen Press.

Gil, B. R. A., Matsuura, M., Monzon, C. M., & Samothrakis, I. (2005, June). *The use of system dynamics analysis and modeling techniques to explore policy levers in the*

fight against Middle Eastern terrorist groups.Naval Postgraduate School, Monterey, CA. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a435682.pdf

Leweling, T., & Sieber, O. (2007). Using systems dynamics to explore effects of counterterrorism policy. Proceedings of the 40th Hawaii International Conference on System Sciences–2007. Naval Postgraduate School, Monterey, CA. Retrieved from http://www.computer.org/csdl/proceedings/hicss/2007/2755/00/27550198.pdf

Madnick, S., & Siegel, M. (2007). A System Dynamics (SD) approach to modeling and understanding terrorist networks. *Proactive Intelligence (PAINT): Model Development*.Massachusetts Institute of Technology, Sloan School of Management, Political Science Department, Engineering Systems Division. Retrieved from http://web.mit.edu/smadnick/www/Projects/PAINT/PAINT%20Proposal.pdf

Norton, Q. (2012, March). How anonymous picks targets, launches attacks, and takes powerful organizations down. *Wired*, Threat Level. Retrieved from http://www.wired.com/threatlevel/2012/07/ff_anonymous/all/

Radzicki, M. J., & Taylor, R. A. (1997). *Introduction to System Dynamics*. U.S. Department of Energy, Office of Policy and International Affairs. Retrieved from http://www.systemdynamics.org/DL-IntroSysDyn/inside.htm

Schoenenberger, L., Schenker-Wicki, A., & Beck, M. (2012). *Analysis of a terror network from a system dynamics perspective*. University of Zurich, Department of Business Administration, UZH Business (Working Paper No. 322). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2171240

Smith, R. (2002). *Counter terrorism simulation: A new breed of federation*. Simulation Interoperability Workshop–Spring 2002. Titan Systems Corporation, Florida. Retrieved from http://www.modelbenders.com/papers/02S-SIW-004.pdf

Ventre, D. (2011). Cyberconflict: Stakes of power. In D. Ventre (Ed.), *Cyberwar and information warfare (*pp. 113–244*)*. London, UK: ISTE Ltd.

Wakeland, W. W., & Medina, U. E. (2010). *Comparing discrete simulation and system dynamics: Modeling an anti-insurgency influence operation*.28th International Conference of the System Dynamics Society, Korea. Retrieved from http://www.systemdynamics.org/conferences/2010/proceed/papers/P1276.pdf