Networks and Structural Themes

A Model of Biological Attacks on a Realistic Population

Kathleen M. Carley[1,2,a], Douglas Fridsma[4,b], Elizabeth Casman[3,a], Neal Altman[a], Li-Chiou Chen[1,2,a], Boris Kaminsky[1,2,a], Demian Nave[c], and Alex Yahja[1,2,a]

[1]Institute for Software Research International, [2]Center for the Analysis of Social and Organizational Systems, [3]Engineering and Public Policy, [4]Center for Biomedical Informatics

[a]Carnegie Mellon University, [b]University of Pittsburgh Medical Center, [c]Pittsburgh Supercomputing Center

Kathleen.Carley@cmu.edu, fridsma@cbmi.upmc.edu, casman@andrew.cmu.edu, na@cmu.edu, lichiou@andrew.cmu.edu, borisk@andrew.cmu.edu, dnave@psc.edu, ay@cmu.edu

1. Introduction

The capability to assess the impacts of large scale biological attacks and the efficacy of containment policies is critical and requires reasoning about social response and disease transmission within a complex social system. There is a close linkage among social networks, transportation networks, disease spread, and early detection. Spatial dimensions related to public gathering places such as hospitals, nursing homes, and restaurants, can play a major role in epidemics [Klovdahl et. al. 2001]. Like natural epidemics, bioterrorist attacks unfold within spatially defined, complex social systems, and the societal response can have profound effects on their outcome. This paper focuses on bioterrorist attacks, but the model has been applied to emergent and familiar diseases as well.

Physical "simulation" exercises such as Dark Winter can provide valuable insights on bioattacks. However, physical exercises are limited in number, size, scope, and scenarios due to cost, time, and cognitive constraints [Simon 1982]. It is faster, more cost-effective, and comprehensive to do computer simulations, which allow larger numbers of complicated, outside-the-box scenarios to be examined systematically. Computational models have been built to explore outbreak impacts for over 30 years. SIR models [Anderson and May 1991] typically assume homogeneously mixing

populations. Recent Geographic-Information-Systems-based models address the spatial dimension [Lawson 2001][Gimblett 2002]. Cellular-automata models, such as the individual-based model of smallpox epidemics [Epstein 2004] allow spatial operation and discontinuities. The geometry of cellular automata, however, could cause incorrect modeling of disease spread. System dynamics models can capture the general epidemic dynamics. This is similar to the discrete event simulation models such as the model of antibiotic distribution used to examine post-exposure prophylaxis [Hupert et. al. 2002]. The "genome"-based sensor-action models such as the Measured Response bioterrorism simulator based on Synthetic Environment for Analysis and Simulation (SEAS) [Chaturvedi and Mehta 1999] can model individual-based sensing and action pairs. Episims model of the Los Alamos National Lab uses transportation networks simulation to estimate contact graphs that are assumed to be social networks [Eubank 2002][Eubank et. al. 2004] and assumes viral load as the basis for disease spread. Most of existing bioattack models simulated a single disease or only bioattack diseases, made homogenous-mixing and perfect-information assumptions, discounted factors such as the urban geography and how people are embedded in social networks [Wasserman and Faust 1994], and are not able capture non-numerical knowledge such as knowledge about school districts, events, rules, etc.

BioWar is a model that combines state-of-the-art computational models of social networks, communication media, and disease transmission with demographically resolved agent models, urban spatial models, weather models, and a diagnostic error model to produce a single integrated model of the impact of a bioterrorist attack on a city. Unlike traditional models that look at hypothetical cities, BioWar is configured to represent real cities by loading census data, school district boundaries, and other publicly available information. Moreover, rather than just providing information on the number of infections, BioWar models the agents as they go about their lives – both the healthy and the infected. This enables on the observation of factors such as absenteeism, insurance claims, over-the-counter drug purchases, and hospital visit rates, among others. In this paper, we will only describe the cores of disease, agent, and social networks models. Comprehensive description of models including diagnosis, treatment, environmental, recreational, attack models can be found in the BioWar Technical Report [Carley et. al. 2004].

2. BioWar Model

BioWar is a city-scale spatial multi-agent network model capable of simulating the effects of weaponized biological and chemical attacks. Recent work has demonstrated that the failure to take the social network and physical locations into account leads to incorrect estimates of disease spread and of response policies [Morris 2001[[Lawson 2001][Morris et. al. 1996][Morris and Kretzschmar 1995][Morris and Kretzschmar 2000][Morris 1997][Morris 1993]. BioWar is socially (embedded in social, knowledge, and task networks) and spatiotemporally more realistic than previous bioterrorism attack models, as it enables;

1. Heterogeneous population mixing, defined by social networks.
2. Simultaneous modeling at multiple levels (pathogen dispersion, knowledge evolution, multiple layers of local, state, and federal authorities, etc.).
3. Detailed modeling and simulation of individuals and their social networks.

4.   Conformity of data: real world data can be used as input to simulation with minimal processing and simulation outputs match real world data formats.
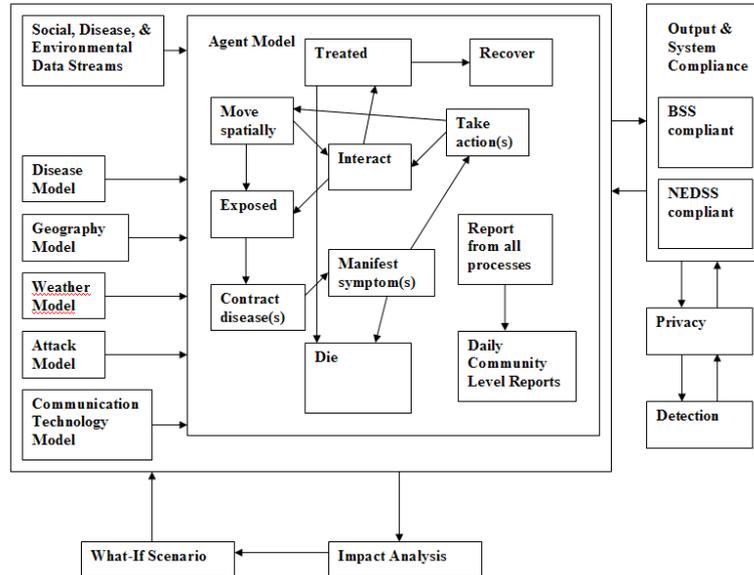
Figure 1 shows the components of BioWar.

Figure 1. The design of BioWar.

The agent-model takes as input the social, disease, and environmental data streams. It integrates various models of disease, geography, attack, weather, and communication technology. The agent-model itself is a finite state machine describing how agents interact, move, are exposed to bioagents, acquire disease, manifest symptoms, and seek and receive treatments. BioWar output conforms to the NEDSS (National Electronic Disease Surveillance System) and BSS (Behavioral Surveillance Surveys). Data used and produced by BioWar are screened for privacy and anonymity, including the data output for detection.

3. Agent-level Disease Model

The current version of BioWar simulates 62 diseases -- 4 weaponized diseases and 58 naturally-occurring diseases -- simultaneously in a population. We use a symptom-based general disease model. Each disease has its own set of symptoms, timing of disease phases, variability in presentation based on age, gender, and race, and contagiousness. Each symptom has its own severity and progression timing. Each instance of a disease infecting an agent is individually represented and progressed through time. Diseases propagation is probabilistically determined by agent risk factors influenced by demographic and occupation variables, the transmissibility of the disease, and the spatial and temporal proximity of uninfected agents to infected agents. Our disease model generates epidemic (or EPI) curves as output.

Each infectious disease progresses through up to five phases:

1. Incubation: the period of time before the agent begins presenting symptoms due to a bacterial or viral infection.
2. Early symptomatic (prodromal): the period of time during which an infected agent may experience mild or non-descriptive symptoms. Many diseases omit this phase, or have no known or identifiable early symptomatic period.
3. Late symptomatic (manifestation): the period of time during which an infected agent may experience severe and/or disease-specific symptoms. In many diseases, this phase may not be distinct from the early symptomatic phase.
4. Communicable: the period of time during which an infected agent may infect other agents. This phase may overlap with the above phases. Noncontagious diseases such as anthrax do not have this phase.
5. Recovery/death: a period of time during which an infection resolves or causes death.

In constructing our disease model, we used historical accounts of known anthrax releases [Inglesby 1999], documents from the October, 2001 bioterrorism attack, and disease knowledge bases [USAMRIID 2001][West 2001][Isada et. al. 2003]. We have also drawn on the experience of medical expert systems developed to assist in diagnosis [Miller et. al. 1982].

3.1. Symptoms

In BioWar, symptoms motivate behavior and determine the initial diagnosis of agents entering the medical system. This symptom-based disease model permits the representation of outliers and stochastic flux. The symptoms are assigned measures that influence which symptoms agents get and how that changes their behavior [Miller et. al. 1982]. The diagnosis is influenced by symptom severity and other factors such as other disease & symptom confounding, patient health history and demographics, and occupation. The true nature of the agent's health status is obscured to the agent so it only responds to sensed symptoms. Modulated by demographics, mild symptoms prompt a visit to a pharmacy, high ones to a doctor, and the severe ones to the emergency department.

3.2. Dose-Response Relationship for Weaponized Pathogens

Relationships of the infection probability, illness duration, and onset-of-illness versus inhaled dose were recently published for anthrax, botulism, pneumonic plague and Venezuelan equine encephalitis [Rickmeier 2001]. At this stage, BioWar takes into account dose and age-of-agent for infection probability for inhalational anthrax following formulas developed by Webb, Blazer, and Buckeridge [Webb and Blaser 2002][Buckeridge 2003] while for the other weaponized diseases the exponential model for infection probability is used.

4. Agent and Social Networks

The significance of social networks to contagious disease transmission is obvious. While non-contagious bioagents such as anthrax do not spread through social networks, social networks define the exposed subpopulation through co-location of agents at the time and place of an attack.

4.1. Representation of Agent

An agent is represented as a probabilistic finite state machine that has roles such as father, schoolmate, doctor, nurse, teacher, etc. Additionally, an agent has socio-demographic and economic status. An agent is located at specific spatio-temporal coordinates and exhibits behaviors. These behaviors include interaction (communicate, get infected, infect), recreation, going to school/work, seeking treatment, purchasing over-the-counter drugs, and moving to other places. Each agent has an ego net and natural biological rhythm (e.g., sleeping 8 hours a day). Moreover, an agent can exhibit symptoms and has mental model of diseases. The propensity of an agent to seek treatment is affected by socio-demographic position (age, race, gender, class, etc.), economic status, and severity of perceived symptoms. Note that even if an agent seeks treatment, treatment is not always available, such as when doctor office is closed. If an agent goes to pharmacy, then the agent's symptoms determine drug purchase choices influenced by symptom severity, symptom-drug relationships, and agent demographics.

4.2. Social Networks

In Biowar, each agent is linked to other agents in the simulated population by a set of links (relationships) modeling a social network. Each link is a descriptor of a pair of agents and their relationship type. Agents may be linked unidirectionally or bidirectionally. Relationship types currently implemented are family (spouse, parent, child, sibling, other family), proximity based (co-worker, schoolmate, group member, neighbor), and voluntary (friend, advisor, other). The relationship types were drawn from the General Social Services (GSS) survey data with the addition of "schoolmate" for younger agents, a population not covered by the GSS http://www.icpsr.umich.edu:8080/GSS/homepage.htm. The overall network size and distribution were drawn from Klovdahl's study along with some target numbers for individual relationship counts [Klovdahl 2002][Klovdahl et. al. 2001][Klovdahl 1997][Klovdahl 1995].

4.3. Agent Interaction

Agents interact with each other based on BioWar's CONSTRUCT model and on spatial and social network proximities. The core of CONSTRUCT contains the co-evolution and emergence of communication and knowledge networks, driven by homophily and expertise-seeking [Carley 1991]. The principle of homophily states that people are more likely to communicate with others who are similar to them, while the principle of expertise-seeking states that the information-poor are more likely to initiate communication with the information-rich to fulfill their information needs. Agents also interact with others nearby and also by random chance. As agents interact, they may exchange knowledge and diseases.

5. Model Validation

We have validated BioWar for anthrax and smallpox cases against existing epidemiological models (SIR – Susceptible-Infected-Recovered -- model and its variant) and empirical data [Chen et. al. 2003][Chen et. al. 2004] for various response

scenarios. We have also validated BioWar outputs against empirical school absenteeism, work absenteeism, pharmacy visits, drug purchases, doctor visits, and emergency department visits [Carley et. al. 2004][Yahja 2004]. Comprehensive validation of a model such as BioWar is however difficult to do manually or semi-manually due to model complexity, the significant number of input parameters, model parameters, output variables, and continuous model development, making automation necessary. An automated tool – named WIZER for **W**hat-**I**f Analy**ZER** – has been designed and partially implemented [Yahja and Carley 2004][Yahja 2004]. WIZER is an integrated inference and simulation engine to do validation and provide explanations. It extends the response surface methodology [Myers and Montgomery 2002] by performing knowledge-intensive data-driven search steps via an inference engine constrained by simulation. It checks the outputs of simulation and adjusts simulation parameters and meta-models to achieve validation of the simulator.

6. Implementation and Sample Runs

BioWar is designed to be modular and portable, and currently runs under Linux, Windows 2000 and XP, and Tru64 UNIX. Most experiments were performed using the Pittsburgh Supercomputing Center's TCS1 system, which comprises 64 4-way Alpha SMP nodes, each with 4GB of RAM and 4 667MHz Alpha 21264A (EV6.7) processors. The run time scales linearly with the number of agents.
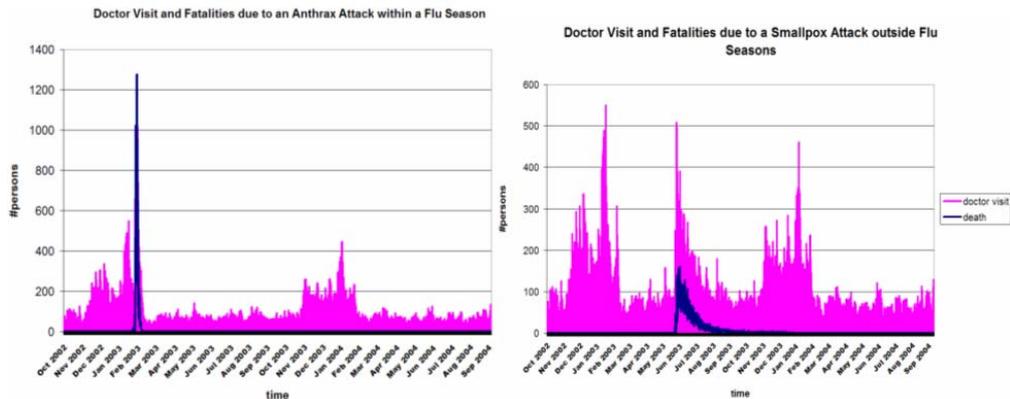

Figure 2. (Left) The number of deaths and doctor visits occurring with a 2.5 kg anthrax attack with an efficiency of 0.05 at 4pm of January 25, 2003, infecting 2122 people during a sports event at a stadium. **(Right)** The number of deaths and doctor visits occurring with a 2%-2.5% initial infection rate smallpox attack at 4pm of May 26, 2003.

We have simulated runs for several cities, including San Diego, Pittsburgh, Norfolk, San Francisco, DC, and Hampton city of Virginia. The scenarios simulated include no attack, anthrax attack, and smallpox attack. The runs for the city of Hampton cover 100% of its 145,665 inhabitants without scaling. The length of the runs is 2 years, with each tick of the simulation being 4 hours. There are 1 weaponized disease and 58 naturally-occurring diseases, including seasonal flu strains. The simulations start on September 15, 2001. Figure 2 above gives sample runs for the city of Hampton, Virginia. When an anthrax attack occurs during a flu

season, as shown in the left graph above, it increases the doctor visit rate significantly for a short period of time after the attack. The number of death by anthrax is also tightly clustered. The right graph shows a simulated aerosolized smallpox attack on Hampton, infecting 2%-2.5% of the agents in a locality. The smallpox attack caused a much longer footprint and deaths compared to the anthrax attack, due to the contagious nature of smallpox.

7. Conclusion

Several improvements for BioWar were planned, including adding infrastructures such as road & air transport networks and organizational structures & response, as well as parallelizing and distributing the code along spatial and social dimensions.

While more enhancements are in pipeline, current version of BioWar represents significant advancement over other numerical disease models. This includes BioWar's ability to model socially defined mixing and spatiotemporal effects, diverse outputs, emergent properties & unexpected outcomes based on local interactions, and ability to be configured with real world data to represent actual cities.

References

Anderson, R.M., and May, R.M., 1991, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, New York.

Buckeridge, D.L., 2003, private communication.

Carley, K.M., 1991, A Theory of Group Stability, *American Sociological Review*, **56(3)**: 331-354.

Carley, K.M., Altman, N., Kaminsky, B., Nave, D., and Yahja, A., 2004, *BioWar: A City-Scale Multi-Agent Network Model of Weaponized Biological Attacks*, Technical Report (CMU-ISRI-04-101). Pittsburgh, PA: CASOS, Carnegie Mellon University.

Chaturvedi, A.R., and Mehta, S.R., 1999, Simulations in Economics and Management: Using the SEAS Simulation Environment, *Communications of the ACM*, March 1999.

Chen, L-C, Carley, K.M., Fridsma, D., Kaminsky, B., and Yahja, A., 2003, Model Alignment of Anthrax Attack Simulations, accepted for publication in Decision Support Systems.

Chen, L-C, Kaminsky, B., Tummino, T., Carley, K.M., Casman, E., Fridsma, D., and Yahja, A., 2004, Aligning Simulation Models of Smallpox Outbreaks, 2004, *Proceedings of the Second Symposium on Intelligence and Security Informatics*, Tucson, Arizona, June 10-11, 2004 (forthcoming). Also to appear in *Springer-Verlag Lecture Notes in Computer Science* as volume **2665**.

Epstein, J.M., Cummings, D.A.T., Chakravarty, S., Singa, R.M., and Burke, D.S., 2004, *Toward a Containment Strategy for Smallpox Bioterror: An Individual-Based Computational Approach*, Brookings Institution Press, not yet published.

Eubank, S., 2002, *Scalable, Efficient Epidemiological Simulation*, SAC 2002.

Eubank, S., Guclu, H., Kumar, V.S.A., Marathe, M.V., Srinivasan, A., Toroczkal, Z., and Wang, N., 2004, "Modeling Disease Outbreaks in Realistic Urban Social Networks", in *Letters to Nature, Nature* Vol. 429.

Gimblett, H.R., 2002, *Integrating Geographic Information Systems and Agent-based Modeling Techniques*, Santa Fe Institute, Oxford University Press.

Hupert, N., Mushlin, A.I., and Callahan, M.A., 2002, Modeling the Public Health Response to Bioterrorism: Using Discrete Event Simulation to Design Antibiotic Distribution Centers, *Medical Decision Making*, **22**:1-9.

Inglesby, T.V., et.al., 1999, Anthrax as a Biological Weapon, *Journal of American Medical Association*, Vol. **281**, No. **18**, May 1999.

Isada, C.M., Kasten, B.L., Goldman, M.P., Gray, L.D., and Aberg, J.A., 2003, *Infectious Disease Handbook*, AACC.

Lawson, A.B., 2001, *Statistical Methods in Spatial Epidemiology*, John Wiley & Sons Publisher.

Klovdahl, A.S., 1995, Social Networks and Infectious Disease: HIV/AIDS. Guest Editor, Introduction, Special Issue, *Social Networks*, vol. **17**, no. **3-4**, pp. 163-165

Klovdahl, A.S., 1997, Social Network Analysis, in J.P. Keeves (ed.) *Educational research, methodology and measurement: an international handbook*, Pergamon.

Klovdahl, A.S., Graviss, E.A., Yaganehdoost, A., Ross, M.W., Wanger, A., Adams, G.J., and Musser, J.M., 2001, Networks and Tuberculosis: An Undetected Community Outbreak involving Public Places, in *Social Science and Medicine*, **52**:681-694.

Klovdahl, A.S., 2002, Social Networks in Contemporary Societies, *Dynamic Social Network Modeling and Analysis workshop*, Washington DC, November 7-9, 2002, *unpublished*.

Miller, R.A., Pople, H.E., and Myers, J.D., 1982, Interist-I, An Experimental Computer-based Diagnostic Consultant for General Internal Medicine, *N Engl J Med 1982*, **307**:468-76.

Morris, M., 1993, Networks and Epidemiology: Modeling Structured Diffusion, *Sociological Methods and Research*, **22**: 99-126.

Morris, M., 1997, Sexual Networks and HIV, *AIDS 97*: Year in Review, **11 (suppl A)**: S209-S216.

Morris, M., 2001, Concurrent Partnerships and Syphilis Persistence: New Thoughts on an Old Puzzle, invited editorial for Sexually Transmitted Diseases **28(9)**:504-7.

Morris, M., and Kretzschmar, M., 1995, Concurrent Partnerships and Transmission Dynamics in Networks, *Social Networks*, **17**: 299-318.

Morris, M. and Kretzschmar, M., 2000, A Micro-Simulation Study of the Effect of Concurrent Partnerships on HIV spread in Uganda, *Mathematical Population Studies* **8(2)**.

Morris, M., Podhisita, C., Wawer, M., and Handcock, M., 1996, Bridge Populations in the Spread of HIV/AIDS in Thailand, *AIDS* **10**:1265-71.

Myers, R. H. and D. C. Montgomery, 2002, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 2nd ed., John Wiley & Sons, Inc.

Rickmeier, G.L., McClellan, G.E., and Anno, G.A., 2001, Biological Warfare Human Response Modeling, *Military Operations Research*, **6(3)**:35-47.

Simon, H., 1982, *Models of Bounded Rationality*, MIT Press, Cambridge, MA.

USAMRIID, 2001, *USAMRIID's Medical Management of Biological Casualties Handbook*, US Army Medical Research Institute for Infectious Diseases.

Wasserman, S., and Faust, K., 1994, *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Webb, G.F., and Blaser, M.J., 2002, Mailborne Transmission of Anthrax: Modeling and Implications, *Proceedings of the National Academy of Sciences of the United States of America* **99(10)**:7027-32.

West, K.H., 2001, *Infectious Disease Handbook for Emergency Care Personnel*, ACGIH.

Yahja, A., 2004, WIZER: Automated Validation of Large-Scale Multi-Agent Systems, CASOS Technical Report, Carnegie Mellon University.

Yahja, A. and Carley, K.M., 2004, "The Design of Automated Validation and Explanation of Large-Scale Social Agent Systems", *2004 NAACSOS Conference*, accepted for publication.