# Towards the Realization of a Public Health System for Shared Secure Cyber-Space

Jeff Rowe
UC Davis
One Shields Ave,
Davis, CA 95616
rowe@cs.ucdavis.edu

Karl Levitt
UC Davis
One Shields Ave
Davis, CA 95616
levitt@cs.ucdavis.edu

Mike Hogarth
UCD Medical Center
2315 Stockton Blvd
Sacramento, CA 95817
mahogarth@ucdavis.edu

## ABSTRACT
It has been a longstanding goal of the cyber-security community to improve the collective security of the general computing population by reducing the attack incidents and the overall susceptibility to attack; we refer to this as improving the *public* cyber-security. Traditionally, computer security techniques have tried to accomplish this by focusing upon securing specific computing systems and networks. This approach is akin to the *practice of medicine* in the health care industry to treat illness and disease in individuals. In the field of health care, the collective health of the population is treated by the *practice of public health*. Currently, there is no analogous public cyber-security system to treat the collective cyber-health of the computing population. We assert that a public cyber-security system, based upon well founded, sound principles currently used in the public health care discipline is needed in order to satisfy our longstanding goal of improving public cyber-security. We outline some of the technical features of such a system and how it might operate.

## Categories and Subject Descriptors
K.4.1 [**Computers and Society**]: Public Policy Issues – *Abuse and crime involving computers, Regulation, Computer-related health issues*

## Keywords
Public Health, Governance, Monitoring, Incident Response, Conformance

## 1. INTRODUCTION
There has been a longstanding interest among cyber-security researchers in using the public health system model as a new paradigm for cyber-security governance. As a complement to traditional security techniques applied to specific computer systems, a public health inspired approach would deal with the state of cyber-health for the population of computing systems in general. We use cyber-health to mean the susceptibility of computers to and incidents of various forms of attack. A very large body of work investigates the high level implications of this model. Mulligan and Schneider[1] approach the question from a legal standpoint. Charney[2], from Microsoft Research, brings the prospective of a large commercial software vendor. Rowe, et.al.[3] investigate the economic consequences of such a model and [4] perform a preliminary evaluation using social science techniques. While the public health model has obvious attractive features for questions of cyber-security governance and policy, computer scientists have yet to investigate the technical implications that such a public health inspired cyber-security policy might have if it were implemented. For example, what types of monitoring architectures are required, which algorithms are suitable and what novel new data sources could be brought to bear to estimate population risk and effective treatments? Conversely, rarely has treatment of the high level public-health policy and governance models addressed the significant unique technical requirements of cyber-security that differ from public health, and how the governance model must differ accordingly. The goal of fostering the public cyber-security will require a new approach to the traditional security methods of prevention, monitoring and recovery. Specifically:

- What new technical methods would be suitable for the realization of a public cyber-security model?
- What specific prevention and intervention measures are effective at improving the public cyber-security?
- What architecture and algorithms are needed to realize a public cyber-security incident response system?
- What unique technical issues in cyber-security require significant modification of the cyber-health model?

This paper describes how techniques borrowed directly from the public health medical discipline could be adapted and applied to a public cyber-security architecture. This is a collaboration between computer scientists working in cyber-security and physicians specializing in public health in the School of Medicine.

## 2. THE PUBLIC HEALTH MODEL
When people think of health care, they usually think of services provided by doctors, hospitals and pharmacies in the service of specific patients. These services are all part of medical practice, which has the goal of treating illness in individuals. Another aspect of health care, that only becomes apparent in special circumstances, is the practice of public health. Rather than dealing with individual medical cases, the goal of public health is to estimate the health of the overall population, to enact large scale


*NSPW'13*, September 9–12, 2013, Banff, AB, Canada.

http://dx.doi.org/10.1145/2535813.2535815

health remediation measures, and to respond to outbreaks and incidents affecting the greater public. As such, public health tasks are multifold: 1) to assess the health status of the aggregate population, 2) to diagnose widespread health problems, 3) to determine the causes of these problems and to 4) develop solutions to these problems. It is important to note that remediation can fall both within standard medical practice, such as cancer screening, and beyond the field of medicine, such as seatbelt laws and anti-smoking outreach. In the broadest of terms, computer scientists have traditionally focused upon the "medicine" aspects of cyber-security; how individual computing systems are attacked and techniques for their protection. The primary question we wish to address is, given the obvious attractive features of a public cyber-security policy model, what novel monitoring, diagnosis, attack prediction and remediation techniques would be required to realize such a system, borrowing directly from the sound methods used currently by our public health system.

To realize the public health system goals, the medical community has implemented a system based upon five major categories: education, monitoring, epidemiology, immunization, and incident response. Given that education and epidemiology have been previously addressed by the computer science community, we address only the monitoring, prevention and incident response policies and methods that would compose a public cyber-security system.

# 3. PUBLIC CYBER-SECURITY MONITORING

In the arena of public health medicine, monitoring is primarily conducted by the CDC and operates over a wide variety of data sampled from a subset of the larger population. The main goal for public health monitoring is not to identify sick individuals, or even to diagnose the nature of a new illness, but to detect changes in the overall health of the population, to identify risk categories, and to serve as an outbreak early-warning system. Currently, no such function is performed to detect changes in the overall public cyber-security of computer systems. Cyber-security monitoring techniques have traditionally focused upon detecting attack instances against specific computer systems. This is analogous to the practice of medicine in identifying and treating disease in an individual, or a group of individuals. Missing in current cyber-security approaches are systems and methods that can detect *changes* in the security of large segments of the computer system population over a wide variety of users and environments. The question we wish to address here is, could such a system be architected based upon our current public health infrastructure and prove to be effective in detecting changes in the overall security of the nation's computer systems, to identify high risk cyber-security behaviors and configurations, and as an early warning system for new, widespread attacks?

## 3.1 Cyber Syndrome Surveillance

A key function of population health monitoring is *syndrome surveillance*[5]. Syndrome surveillance refers to methods to detect population (and individual) health indicators before any diagnosis is made. Prior to any laboratory confirmation of disease, individuals exhibit patterns of behavior, symptoms and signs that can be tracked to detect undiagnosed changes in the health of the overall population. Syndrome surveillance of the public's cyber-health to detect early stages of new attacks and compromise isn't currently performed and would be a useful function of a public cyber-security system.

As a basis for the design of a public cyber-security system we follow the syndrome surveillance guidelines used by the medical community[6]. The system functions are broken down into six categories:

1) Selecting the population and the data.
2) Acquisition and organization of the data.
3) Data integration across multiple sources.
4) Privacy protection.
5) Outbreak detection.
6) Integration with public response.

Syndrome surveillance in our current medical public health system relies primarily upon patient data supplied by medical practitioners: doctors, pharmacies and hospitals. There are also a variety of other data sources outside of standard medical practice that have proved useful. The sales of over-the-counter medicines, such as pain killers and flu remedies, school absentee records, and Internet searches for symptom keywords are all used to detect outbreaks of previously unknown maladies and bioterror incidents.

One of the key questions in designing a public cyber-security system is which data can be used as indicator for widespread changes in the security of computing systems given the lack of a cyber-health reporting infrastructure. As a first step, we address this by looking at what data is publically available and how it can be used in cyber-security syndrome surveillance.

### *3.1.1 Twitter hashtag tracking*

With the rapid adoption of online social networks comes a wealth of data regarding social communications and interactions. These communities reflect both real world social relationships as well as purely online social communities. Twitter is a well-known social networking service whose interactions are open to the public. The community of Twitter users spontaneously created a hashtag feature for labeling individual posts. If the hash symbol, #, precedes a word, that word represents a tag that is publically searchable. By tracking cyber-security relevant hashtags, relative frequency of label occurrence in Twitter communications could serve as syndrome surveillance data. As an example, we tracked the Twitter hashtag *"#hacked"* over the course of several weeks. The idea here is that individuals reporting malicious cyber-activity, either as victims or as perpetrators, might include this as a label in their posts. Figure 1 shows the frequency of tweets containing the *#hacked* hashtag over a two week time period. Notice excess in tweet counts in the April 6-7 time period.

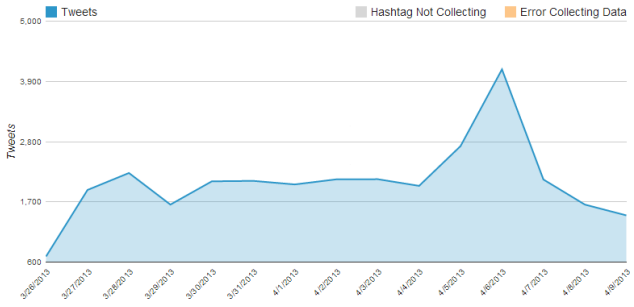


**Figure 1: Counts of tweets with the *#hacked* hashtag over a two week period.**

Had this been a data element in a cyber-syndrome surveillance system, this period with excess *#hacked* tweets represents a symptom of a yet unknown public cyber-security event. We track the additional *#spam* hashtag as a potential public cyber-security symptom. The idea here is that the goal of Twitter account compromise is to obtain a platform for sending spam tweets to all followers. An increase in tweets with this hashtag indicates a response by multiple spam tweet victims. Figure 2 shows a distribution of the *#spam* hashtag in tweets covering the same time period as the *#hacked* hashtag in figure 1. There is a clear excess of tweet events in the April 3-4, 2013 time period.

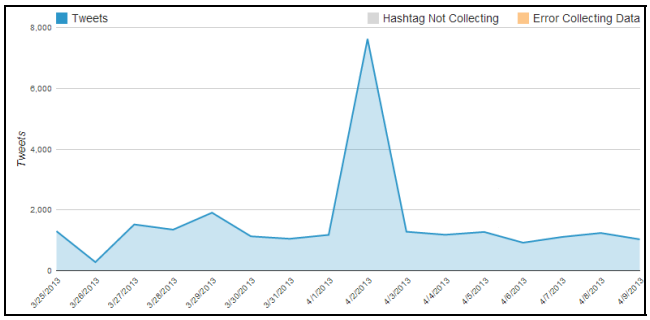


**Figure 2: Counts of tweets with the *#spam* hashtag over a 2 week period.**

We investigated further into the specific Tweets in these two time periods. For the "#hacked" hashtag, most tweets were related to hacking attempts against Israel based companies evidently, however most of the tweets were not by English speakers so it is very difficult to determine the source. Similarly, tweets with the "#spam" hashtag were generated primarily by Turkish users. Most of these tweets pointed to a URL which describes Turkish anti-sedition laws that include immoral Internet use and criticism of the government. Further investigation showed that a variety of accounts from Germany, Holland and Belgium had been used to send child pornography spam in Turkish to a large number of Turkish Twitter users.

Just as in syndrome surveillance for public health issues, monitoring this data simply provides an indication for changes in symptoms for a large number of users. No diagnosis of the cause of these events is provided by this type undirected Twitter hashtag monitoring.

### *3.1.2 Google Trends*

Another potential publically available data source for cyber-syndrome surveillance monitoring is the Google Trends service. Google Trends makes available statistics and correlated information regarding keyword searches performed by Google search users. The nature of this data should be quite different from the Twitter hashtags. Search keywords represent the type of information that individuals are trying to access, whereas Twitter hashtags represent information communicated publically or to other individuals. Consider searches on the keywords used previously in the Twitter hashtags: spam and hacked. Figure 3 and 4 show the normalized Google search volume for both of these terms.

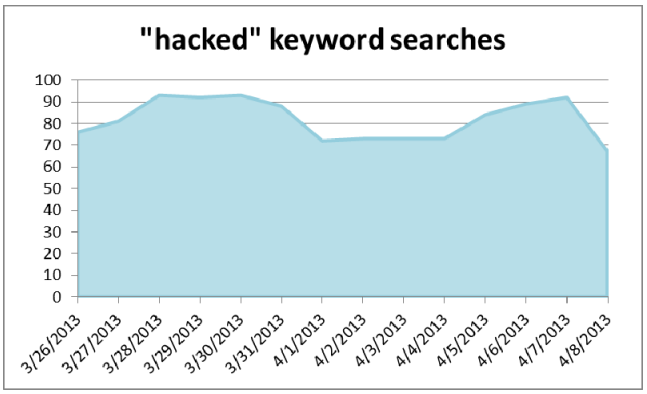


**Figure 3: Normalized Google search volume for the "hacked" keyword over a 2 week period.**

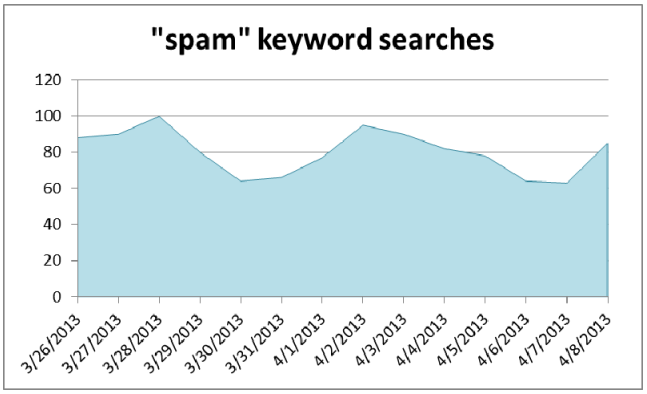


**Figure 4: Normalized Google search volume for the "spam" keyword over a 2 week period.**

Notice that the relative volume of Google keyword searches show no apparent excess in either of these terms. Next, consider the normalized search volume for the keyword "anonymous" shown in figure 5.

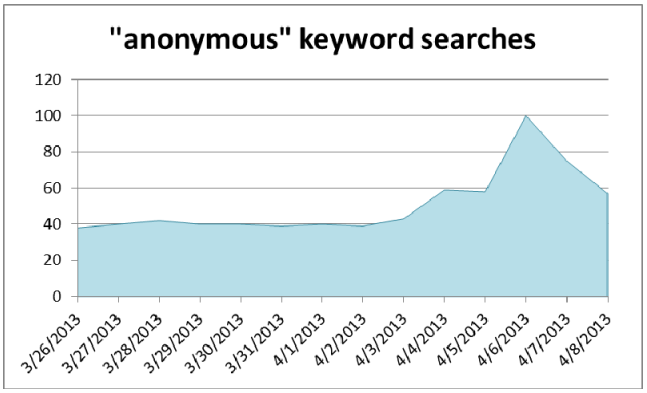


**Figure 5: Normalized Google search volume for the "anonymous" keyword over a 2 week period.**

Notice the excess of queries that roughly corresponds to the excess in Twitter "#hacked" hastag postings. If one searches for news articles after April 6, 2013 using both the "hacked" and the

"anonymous" keywords, articles describing a coordinated attack called OpIsrael against Israeli targets and accounts. The "hacked" keyword in Google searches shows no excess. Only the Twitter "#hacked" hashtag posts gives an indication in this case. We present this very crude analysis, not to prove the efficacy of detecting this specific event, but to motivate the design of a public cyber-security syndrome surveillance system. This very simple example has many features of the syndrome surveillance approach employed by the medical community in our public health system. The different modes of usage across different data sources provide very different signals that must be aggregated without any prior diagnosis of the specific incident underway. One doesn't know in advance which signals will be correlated. Multiple signals across a wide variety of independent sources, however, provides a fairly robust method for event recognition which can be used to trigger a more detailed diagnosis.

#### *3.1.3 Voluntary, incentive-based public data collection*

The previous sections describe data sources that are publically available as by-products of user's interaction with existing services on the Internet. It would be more advantagous to have public cyber-security surveillance data sources that indicated when user's were seeking assistance for recognized problems, analogous to public health monitoring of visits to the hospital and doctor offices. Obviously, individuals are motivated to report symptoms to their medical care providers because it is a key action in maintaining personal health and wellbeing. An analogous service could be implemented for a public cyber-security surveillance system by collecting data from user's attempts to diagnose suspected security problems using automated software system.

The first approach would be to collect reports from existing anti-virus and personal firewall software. Currently, anti-virus vendors strive for a near zero false positive rate in their products, which is understandable given the technical expertise of their market – the general public. For public cyber-security data collection, an additonal set of signatures would be useful that monitor for malware, OS modification or unusual file changes. For this signature set, matches would not trigger an emergency alert on the user's screen, but would be logged for the user's later examination and sent to the public cyber-security syndrome surveillance system for emerging threat recognition. There are obvious privacy concerns with this type of monitoring that we will address in later sections.

A second approach is motivated by web based, on-demand security and configuration checks currently available. Several online services exist which allow users to upload files they suspect are malicious or infected with a virus[7]. These services, upon receiving an upload, runs all available A/V software on the file for a comprehensive signature scan. Another type of service currently used to collect network configuration is NetAlyzr[8], which runs a java application on the user's local machine to send a variety of packets to the server. By controlling both ends of the network connection, the NetAlyzr service is able to diagnose potential network problems and network routing vulnerabilities, like system-in-the-middle transport paths. As a by-product of it's execution, NetAlyzr collects data about Internet configurations on paths from all users that access the service. Users have a considerable incentive to use the service because of its value in diagnosing networking problems. A similar system might be deployed in a public cyber-security surveillance system that, instead of network diagnosis, performs a local security check instead. Symptoms found, whether diagnosed as a serious known problem or not, would be aggreated and provided to the syndrome surveillance system. Again, we are well aware of the serious privacy concerns here that must be addressed.

## 3.2 Population change detection

The primary statistical tool used by the medical community for tracking public health indicators is the cumulative sum (Cusum) control chart, a sequential analysis technique for change detection developed for process control[9]. For sequential data with a large variance, subtle changes to the mean are masked for short time intervals after the change. The Cusum algorithm tracks a moving summation of excursions in one direction, positive or negative, from the mean. New date points accrue by assigning likelihood weights for the measurement. Intuitively, the goal is to maintain a running sum of "surprise" events, and apply a threshold to this sum as an alarm mechanism.

The Cusum method operates over sequential time-series data that might be continually sampled, where: 1) the random deviation from mean follows a normal distribution, and 2) there is no serial time correlation between data points under normal conditions.

Briefly, for two competing hypothesis, $H_0$ – normal operations, and $H_1$ – process fault, the Cusum method plots the series,

$$X_t = \max(0, X_{t-1} + W_t)$$

Where,

$$W_t = \ln\left(\frac{L(H_0)}{L(H_1)}\right)$$

the log-likelihood of normal versus faulty operations for sample $t$. In practice, $X_0=0$ in nearly every case. The Cusum series signals an alarm when $X_t > h$, where $h$ is some upper boundary that has been set according to the tolerable false alarm rate. On an alarm, a remedial action is taken whose effectiveness can be monitored using the same Cusum sequence. Whereas other hypothesis tests work once on a particular data set, the Cusum algorithm provides a sequential check for ongoing processes and most significantly *can never accept the $H_0$ hypothesis* which would terminate the procedure, only eventually accepting $H_1$ when the system is faulty.

To our knowledge this technique, common to the public health community, has never been applied to track the cyber-security state of a large population of heterogeneous computers. More than simply an analogy, an implement of Cusum tracking borrowed directly from the medical public health community, but operating on multiple public cyber-security indicators as discussed above, and others such as number of suspected viruses, firewall access log counts, spam emails, etc. could detect how widespread a particular vulnerability may be and when a new cyber-attack is imminent.

## 3.3 BioSense inspired cyber-health monitoring

The medical community's primary public health monitoring program is the BioSense 2.0 system. BioSense was instituted shortly after Sept. 11, 2001 as a national bio-surveillance tool to

track bio-terror attacks. Subsequently, its usefulness turned out to be much more effective in general public health incidents, successfully predicting the outbreak of H1N1, and monitoring the health of communities following natural disasters and during flu season. Given this proven effectiveness, BioSense was expanded to cover non-infectious diseases, like chronic illness, injury and substance abuse trends.

BioSense collects and aggregates large amounts of data from doctors, hospitals, pharmacies and government health services to detect major health hazards and outbreaks. The notional top-level architecture of BioSense is shown in figure 6. Currently, only the data sources and user analysts are distributed. All data stores and access control is centralized.

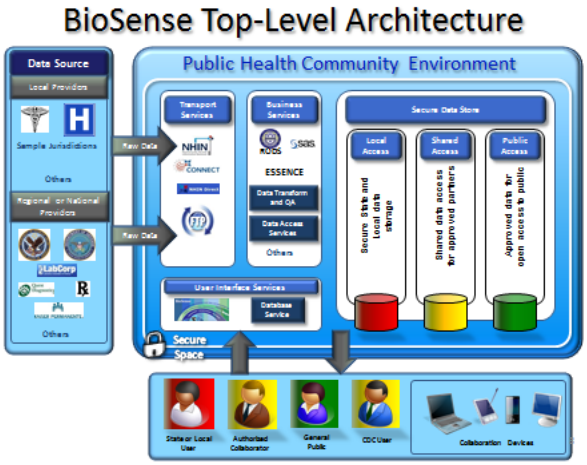


**Figure 6: Top level architecture of the BioSense 2.0 public health surveillance system.**

One difficulty with this architecture is that data collection is performed by local government health departments and gets stored and controlled in a central system at the CDC, which is a national level organization. However, the majority of health outbreaks are localized and recognized by the same local health departments that provided the data, but who don't necessarily have the permission or skills to access the relevant data.

A public cyber-security surveillance architecture that can support similar wide-spread monitoring of the public cyber-security could be similarly architected. For cyber-security monitoring it is easier to imagine novel, cloud-based systems to support distributed data collection, analysis and collaborative defense that might also address scalability issues with BioSense 2.0. Analysis for detection and diagnosis could be performed in the cloud rather than in an existing medical infrastructure, which could allow for a useful incremental adoption strategy. The localized nature of medical health outbreaks drives the architecture for future distributed BioSense systems. One question for a public cyber-security built upon this model, is, how are public cyber-security incidents grouped in cyber-space? For example, malware most likely attacks computers with similar vulnerabilities rather than located in the same networks, whereas propagating worms might be spread to LANs first before attempting WAN attacks.

Finally, the medical community has invested considerable work in markup formats for public health incident reporting. Fortunately, the design and requirements of a language for cyber-incident reports is similar to previous work on CiDF and IDMEF[10] which could be leveraged.

# 4. CYBER-PREVENTION POLICY AND STRATEGY

Significant work has been performed in cyber-immunization techniques to prevent attacks on individual computers. The relevant public cyber-security question that has received less attention is, if a cyber-immunization is available, what countermeasure implementation strategy would best treat the vulnerable segment of the population? Medical immunization policies depend upon risk of contracting an illness from a vaccine versus well understood benefits in preventing an illness. How are these policies to be modified and implemented in a cyber-security environment when the corresponding risks are ill-defined?

## 4.1 Pervasiveness of voluntary prevention measures

A critical question for public health practitioners is to determine how many individuals have taken or are willing to take measures to improve their health voluntarily. People will willingly get a booster vaccine every 5 years, for example, yet be unwilling to voluntarily change their daily eating habits to prevent diabetes. The research question we wish to address is, how willing are individuals to install and manage cyber-health prevention tools voluntarily, and is there a correlation between the use of voluntary prevention and the risk of malicious compromise? This isn't a traditional cyber-security problem addressed using techniques from computer science. This would be an opportunity to enlist social scientists to conduct public surveys coupled with voluntary lightweight monitoring to classify test populations into risk groups, and for testing specific treatments.

## 4.2 Mandated immunizations

Nearly all public health immunization programs are voluntary, yet the vast majority of the population willingly accepts vaccination by their health care providers because it is so obviously beneficial. There are, however, cases of mandated immunization, such as vaccination requirements for children attending public schools. To realize a public cyber-security system, what similar mandated cyber-health prevention measures are reasonable? Social scientists could also contribute here to perform public surveys and questionnaires, as is currently practiced by the medical public health community, to judge what is considered acceptable to the most vulnerable populations. For example, recently run virus scanning software is oftentimes a requirement for connecting to private networks. Might a similar requirement be in place for accessing government services online, such as online tax filing, vehicle registration and other e-government functions?

## 4.3 Vaccine for the immunocompromised

Nearly all vaccine programs managed by the public health system are target at all individuals. Oftentimes, due to risks associated with certain vaccine programs, only a limited portion of the population is targeted, such as only individuals with compromised immune systems. In a public cyber-security system, how are the most vulnerable members of the population to be identified and provided with a more rigorous set of prevention measures that may be too costly for the bulk of the computing population.

# 5. CYBER-HEALTH INCIDENT RESPONSE

Incident response in traditional cyber-security involves blocking an attack and recovering the affected machines. In public health, incident response can take the form of identifying the source of an outbreak, tracking down individuals exposed to a pathogen, isolation of the source, and in extreme cases, quarantine of an affected population. These functions are currently absent for widespread cyber-security incidents. Determining attack sources, potential victims and remediation measures would be critical functions of a cyber-health system. Cyber-quarantine has been the focus of extensive prior work in automated dynamic quarantine against spreading malicious attacks such as worms and botnets. It is unrealistic to expect a public cyber-security system to implement such an automated quarantine on privately owned networks which are unlikely to accept that type of direct control. Instead, we attempt to borrow directly from the public health system of the medical community to determine the tasks necessary to implement a less rigid cyber-security incident response.

## 5.1 Determining an outbreak source

One critical function of the US Center for Disease Control is investigation of outbreaks to determine a source, such as a tainted component of the food supply, or a patient zero in a viral outbreak. In the analogous public cyber-security system, the corresponding function would be identifying the root cause of common widespread attacks. Unlike the public health system, geographic proximity would be of little use in this analysis. Outbreak policy in cyber-space would need new models of security "distance" that would be outbreak and vulnerability dependent. Using the same ISP, reception of the same spam email, or installing the same software would be components of such distance measure. Finding the few systems hosting malware that has caused an observed outbreak is also an important cyber-security problem with features significantly different than finding patient zero in a health outbreak.

## 5.2 Identification of outbreak victims

The second CDC function, once an outbreak is suspected, is to discover all other victims outside of the monitored sample. If one farm has been identified as the source of an O157 bacteria contamination, for example, after observing cases of illness, the critical problem becomes tracing the currently unknown distribution of the farm's products throughout the food supply system. In an analogous cyber-security incident, data sources such as customer contracts, shipping records and geographic limits wouldn't be readily available or even relevant in many cases. A major component of such a public cyber-security system would involve policies for data collection that would be useful for identification of victims as well as outbreak detection.

# 6. Limitations of a cyber-security public health doctrine

When deciding whether public cyber-security system built along the lines of our medical public health system, one must consider also the fundamental limitations? Our current public health system relies almost exclusively upon the extensive existing medical infrastructure. Monitoring is performed using data from doctors, pharmacies and hospitals. Such an infrastructure would be difficult and costly for a public cyber-security system, so data would most likely be collected directly from the public. We see two potential avenues for a voluntary public cyber-security infrastructure. First, instead of using dedicated professionals like physicians, diagnosing security problems in computer systems and software are best done with programs. There is a very large body of work on using static and dynamic analysis to detect the presence of malware. Individual users or administrators that suspect a compromise have a strong incentive to voluntarily run diagnosis programs on their systems to find problems. These programs are analogous to medical doctors in the public health system. Even if the specific cause of the suspected problem isn't found, telltale symptoms might be collected as a byproduct of the analysis that could be reported for use in syndrome surveillance.

Collecting data from individuals, however, introduces obvious privacy concerns that aren't present in aggregated patient statistics from medical professionals. There are a variety of methods that are currently in use to provide anonymity for both publishing and accessing information. Tor is a well-known system using onion routing to create a privacy preserving overlay network that supports both anonymous publishing and anonymous access[11]. For a privacy preserving public cyber-security system, onion routing distribution mechanisms could be built into the data collection infrastructure. This would serve to hide the location of remote users who are submitting suspicious files for inspection, for example. Allowing diagnostic programs to execute on a user's local computer, however, is more problematic. Ensuring that executable code doesn't leak data is an unsolved problem in computer security. We turn to the medical community for their analogous problem. Patient health information is perhaps even more sensitive than cyber security information might be in many cases. The medical community has a long-standing tradition of doctor/patient confidentiality backed by the legal system, and through legislation such as HIPAA. This is seen to be essential to health care since it encourages individuals to enlist the help of professionals regardless of the nature of the health issue they face. If the public and the government are serious about fostering cyber-security, similar laws limiting what can be disclosed regarding an individual's computer configuration.

Perhaps most difficult of all is the recognized problem with the current public health high-risk prevention strategy. Health risk groups are defined from aggregate population studies; however prevention is implemented by medical practice on individuals. Risk factors highly correlated to the health of populations aren't accurate predictors of risk for any given individual. High-risk prevention measures (like yearly mammograms for women over 40) clearly shift the risk to the population but aren't accurate predictors for survival rates of individuals. A critical research question is, how can a public cyber-security doctrine provide useful prevention measures for attacks based upon aggregate population statistics?

# 7. Conclusion

Prior work has explored the implications of using the public health system as a model for cyber-security governance. We, on the other hand, have attempted to describe what technical features and engineering requirements are needed to implement a public cyber-security system, just as the medical community has implemented a public health monitoring system. We believe this represents a new security paradigm for monitoring, diagnosis and incident response that complements current cyber-security approaches. Rather than focus upon securing individual computers and system, this paradigm would monitor the entire

compute population at large, finding widespread emerging threats rather than identification of attacks, diagnosing the cause of widespread incidents rather than individual break-ins and attempting a population level incident response.

# 8. REFERENCES


[1] D. K. Mulligan and F. B. Schneider, "Doctrine for Cybersecurity," *Daedalus*, vol. 140, no. 4, pp. 70–92, Sep. 2011.

[2] Scott Charney, "Collective Defense: Applying the Public-Health Model to the Internet," vol. 10, no. 2, pp. 54–59, 17-Oct-2012.

[3] Rowe, Brent, (second) Halpern, Michael, and (third) Lentz, Tony, "Is a Public Health Framework the Cure for Cyber Security?," *Crosstalk*, no. Nov/Dec 2012, pp. 30–38, 2012.

[4] Rowe, Brent, (second) Halpern, Michael, (third) Lentz, Tony, and (fourth) Wood, Dallas, "Understanding Cyber Security Risk Preferences: A Case Study Analysis Inspired by Public Health Research," Institute for Homeland Security Solutions, Nov. 2012.

[5] R. Heffernan, F. Mostashari, D. Das, A. Karpati, M. Kulldorff, and D. Weiss, "Syndromic surveillance in public health practice, New York City," *Emerg. Infect. Dis.*, vol. 10, no. 5, pp. 858–864, May 2004.

[6] K. D. Mandl, J. M. Overhage, M. M. Wagner, W. B. Lober, P. Sebastiani, F. Mostashari, J. A. Pavlin, P. H. Gesteland, T. Treadwell, E. Koski, L. Hutwagner, D. L. Buckeridge, R. D. Aller, and S. Grannis, "Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience," *J. Am. Med. Inform. Assoc.*, vol. 11, no. 2, pp. 141–150, Mar. 2004.

[7] "VirusTotal," 11-Apr-2012. [Online]. Available: http://www.virustotal.com. [Accessed: 11-Apr-2012].

[8] "NetAlyzr." [Online]. Available: netalyzr.icsi.berkeley.edu. [Accessed: 11-Apr-2013].

[9] W. H. Woodall, J. Brooke Marshall, M. D. Joner Jr, S. E. Fraker, and A.-S. G. Abdel-Salam, "On the use and evaluation of prospective scan methods for health-related surveillance," *J. R. Stat. Soc. Ser. A Stat. Soc.*, vol. 171, no. 1, pp. 223–237, 2008.

[10] Debar, H., Curry, D., and Feinstein, B., "The Intrusion Detection Message Exchange Format (IDMEF): RFC 4765." IETF, 2007.

[11] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: the second-generation onion router," in *Proceedings of the 13th conference on USENIX Security Symposium - Volume 13*, Berkeley, CA, USA, 2004, pp. 21–21.