

## **Epidemiologic Considerations in Network Modeling of Theoretical Disease Events**

**Marcus Lem, MD, MHSc, FRCP(C)**

Communicable Disease Control Division  
Primary Health Care and Public Health Directorate  
First Nations and Inuit Health Branch, Health Canada  
A.L. 1915C, Tunney's Pasture  
Ottawa, Ontario K1A 0K9  
Canada

[marcus\\_lem@hc-sc.gc.ca](mailto:marcus_lem@hc-sc.gc.ca)

### **1.0 INTRODUCTION**

Social network modeling is a relatively new addition to the armament of public health and epidemiology. Epidemiologists and communicable disease control researchers have been turning to network analysis to address and understand gaps in traditional outbreak management techniques such as contact tracing. Network analysis has shown utility in the study of a range of communicable disease outbreaks affecting both health and commerce, including SARS, tuberculosis, syphilis and foot-and-mouth-disease, and may have applications in automated disease surveillance systems.

Visualization of these communicable disease networks is an integral component of such analysis. However, visualization of more complex relationships will require consideration of a variety of epidemiologic factors which affect these relationships, and the development of techniques to display them. Any analysis of case level health data has the potential for compromising privacy, and network visualization is no exception. Like other analysis tools such as data mining, or Geographic Information Systems (GIS), network visualization will need to incorporate techniques to ensure confidentiality.

In this paper we shall discuss the role of network analysis in communicable disease outbreak control, epidemiologic considerations in visualizing networks, and the emerging issue of confidentiality.

### **2.0 NETWORK ANALYSIS AND OUTBREAK CONTROL**

Scale-free networks have certain characteristics which are pertinent to the control of communicable diseases.

Firstly, disease outbreaks occurring in scale-free networks have no epidemic threshold [1], making them resistant to random deletions. Even one infected individual can result in an epidemic. Compared to random networks, the initial spread of disease in scale-free networks is much faster and disease is detected at a later stage. Hence traditional contact tracing for disease control requires a much higher effort and accuracy to control an epidemic for a scale-free network than for a corresponding random network [2].

Secondly, while scale-free networks may be resistant to “shotgun” interventions, they are vulnerable targeted attacks which can disrupt and fragment the network [3]. In a retrospective analysis of the 2001 British foot-and-mouth disease epidemic, Shirley & Rushton examined the network of livestock markets and farms involved. The authors concluded that removal of any one of only three key vertices would have

Lem, M. (2006) Epidemiologic Considerations in Network Modeling of Theoretical Disease Events. In *Visualising Network Information* (pp. 11-1 – 11-6). Meeting Proceedings RTO-MP-IST-063, Paper 11. Neuilly-sur-Seine, France: RTO. Available from: <http://www.rto.nato.int/abstracts.asp>.

prevented over 80% of the infected premises and stopped the epidemic [4]. The authors also concluded that it is necessary to “understand the contact pattern of susceptible populations before embarking on any strategy for disease control, which means that populations at risk from disease need to be characterized topologically before an outbreak occurs.” This has implications for public health planning. Prospective characterization of a variety of networks could help inform the design of potential public health interventions before outbreaks occur. Such networks may include commercial air travel routes, hospital and extended care facility separations and transfers, and food supply chains.

Network analysis of disease transmission need not be restricted to known pathogens. The emergence of new organisms, like the SARS-associated corona virus, demonstrates that future outbreaks and epidemics may take the public health community completely unaware. Greater utility may be found in prospectively analyzing potential and theoretical disease events with a range of transmission characteristics and developing intervention options *a priori*.

Also, understanding of communicable disease network topology will be of great use in the implementation of future public health interventions which can mimic the spread of the diseases they are designed to prevent. Examples of this type of intervention include the oral polio vaccine, which like wild-type polio virus is spread by fecal-oral route, and, potentially, commensal bacteria bioengineered to produce HIV-microbicidal peptides or proteins, which are currently under development [5].

### **3.0 EPIDEMIOLOGIC CONSIDERATIONS IN NETWORK VISUALIZATION OF COMMUNICABLE DISEASE EVENTS**

One of the greatest strengths of network models is the ability to visually represent complex and heterogeneous distributions of interactions in an intuitively comprehensible and interactive fashion. However, some epidemiologic characteristics related to the host, the disease agent, and the environment are less easy to represent, but should still be considered in theoretical network models. This is especially important when modelling new or emerging diseases with unknown or changing characteristics. Considerations may be divided into characteristics of the host, the disease agent, and the environment. Listed below are descriptions of such epidemiologic factors and possible ways to address their visualization:

- ◆ The host factor of susceptibility to infectious disease exists on a spectrum which may or may not have definable thresholds. Disease susceptibility may be non-modifiable (e.g. genetic, co-morbidities), modifiable (e.g. by immunization or natural sub-clinical infection, lifestyle, such as smoking), or even intermittent in cases of waning immunity in the absence of wild-type disease challenge. Co-factors for disease, where a single infective or toxic agent or genetic predisposition is necessary, but not sufficient to cause disease may also exist.  
Visualization technique: Node colour gradient may be used to show susceptibility. Alternately, numeric probability of susceptibility could also be displayed by each node.
- ◆ The pathogenicity of a disease-causing organism also may not remain constant either, independent of changes in host immunity. Evolutionary pressures favour increasing infectiousness and decreasing pathogenicity of an organism with successive passes through a population. “Successful” organisms tend not to kill their hosts. Conversely, genetic shifts and interspecies jumps, such as can occur with influenza, may lead to dramatically more virulent organisms.  
Visualization technique: Edge pattern may be used to denote virulence of transmissible organism. Transmissibility may be reflected by line intensity or pattern density.

- ◆ Environmental factors may profoundly affect the spread of a disease, regardless of whether the spread is person-to-person, food-borne, water-borne, or by fomites (contaminated objects). Diseases such as malaria or lyme disease are spread by arthropod vectors which have their own habitat and dispersion characteristics. Social network modelling of tuberculosis and sexually transmitted illnesses have shown that a geographic location, rather than a specific individual, may be the critical vertex between infected persons.

Visualization technique: Electron orbital models for the representation of chemical interactions may be useful for environmentally-related disease transmission. Environmental spread is not best represented by lines/edges, but by two-dimensional or 3-dimensional shapes such as circles or spheres (“clouds”) around nodes, or ellipses (“orbitals”) or cones where directionality is required. In cases where disease is the result of multiple factors, including both infectious and toxic/environmental elements, visualization may take the form of orbital interactions.

Another way to potentially improve the visualization of communicable disease networks may be to structure the model to reflect the public health response, as well as the disease transmission. For example, response to outbreaks of infectious diseases such as tuberculosis typically occur in an expanding “bullseye” fashion: contact tracing and screening starts with the small group of close and family contacts of the index case, and moves to progressively larger, more distant connections until the limits of the outbreak are defined. Circles have been widely used in the literature to illustrate networks since they allow naïve observers to recognize differences between a specific solution and the idealized pattern from which it deviates [6]. Placement of nodes with similar risk for infection in concentric circles around the index case in a network visualization makes intuitive sense, and may assist in identifying previously unknown or improperly characterized relationships.

Furthermore, network analyses of different aspects of the same issue may yield insight into more appropriate interventions. Thus analyses of both disease spread and utilization of health care or other services may identify optimal intervention points or locations. Another public health area where network visualization may have utility is the analysis of organizational structure for response to communicable disease events. Such analysis has been used to examine command and control for emergency services such as police, fire and hazardous materials [7]. Three-dimensional animation is an effective way to display hierarchical information [8] and may be useful in the analysis and planning of integrated federal, provincial/territorial and local public health responses to communicable disease emergencies. In recent months there has been an unprecedented international effort to prepare for the threats of avian and pandemic influenza (PI). In Canada there has also been an associated proliferation of committees, secretariats and organizational structures to address the issue at all levels of government and public health, although the number of public health officials has not substantially changed and remains small. The Communicable Disease Control Division of the First Nations and Inuit Health Branch of Health Canada has undertaken a preliminary analysis of the PI committee structures and memberships using a custom built visualization application (VITA 6.0b, Exocortex Technologies 2006). This exercise has been instructive in identifying gaps in organizational connections, as well as overlaps and redundancies which could be streamlined.

A final technique which may improve visualization of communicable disease networks is the simultaneous combination of animation with graphing. Existing software has the capability for generating epidemic curves from animated models. Presentation of both the network animation and the epidemic curve on the same screen could be integrated with a “drop-out” animation in which infected nodes fall out into the graph, generating the curve.

#### **4.0 PRIVACY CONSIDERATIONS IN THE VISUALIZATION OF COMMUNICABLE DISEASE EVENTS**

One method of increasing the amount of information presented in a network diagram is to superimpose network information on a map. This technique is intuitive and has been used in analyses such as sexual networks in the one Canadian province [9] and networks of livestock movements in Great Britain [10], [11]. GIS and the common use of geocodes, such as address or postal code, for billing or case follow-up in clinical and public health administrative databases make imaging of geographic distributions fairly easy.

Although geographic information may be instructive in understanding communicable disease networks, its use presents privacy issues. Depending on the population density of the area under study, the precision of the geocode, the background detail on the map, and the visualization algorithm, the information conveyed may be sufficient to personally identify an individual. Even identification of individual communities or parts of communities may be considered sensitive. For example, in Canadian First Nations and Inuit (FN/I) communities, sufficient stigma is still attached to communicable diseases such as tuberculosis (TB), that communities with outbreaks may not wish to be identified.

Like geographic location, graphic representation of human relationships may have the potential for inadvertent violations of confidentiality, even when the nature of only a single relationship link in the network is publicly known. Examples of this may include disease spread within an extended family, religious group, or close-knit community. If one individual in the network can be personally identified, then the other members of the network may be deduced.

Spatial epidemiologists have developed a variety of methodologies for masking data to ensure confidentiality, including transformation, random perturbation and aggregation [12]. Aggregation in particular may be applicable to network visualization of disease transmission. Cases or nodes with the same or similar characteristics and relations may be aggregated, simplifying the visualization but destroying much relational information in the process.

#### **5.0 CONCLUSIONS**

Combining network analysis with other public health tools such as traditional epidemiologic statistics and geographic information systems may lead to better characterization of disease events. Some possible implications for network analysis research are as follows:

1. Network analyses of the flow of humans, information, and commerce need to be performed before crises occur and the effects of vertex deletion assessed. Network visualization may be useful in the analysis and improvement of public health command and control structures, organization and incident responses and may help model “unforeseen” events such as unintentional/accidental breakdowns in infrastructure, criminal and terrorist acts, and outbreaks of new or emerging communicable diseases.
2. Network analyses of disease events should consider the possibility of situations where vertices have different natures (e.g. person, place, time), may appear, disappear or otherwise change (i.e. changes in immunity or susceptibility, presence of disease co-factors), or have intermediaries (e.g. vectors or fomites). Continued consideration should be given by the research community to interventions and investigations which mimic the behaviour of the disease, especially in cases where network analysis cannot identify appropriate control points. Such interventions are only warranted when the danger of the disease is greater than the risk associated with the intervention, as is the case for oral polio vaccine in endemic countries.

3. As visualization techniques improve and increase the amount of information one is able to present, the potential for breeches in privacy also increase. Attention needs to be paid to developing methodologies which preserve the information critical for both analysis and confidentiality.
4. Innovative ways to visually represent aspects of disease epidemiology in network models need to be explored. Example of this may include use of 3D animations, edge patterns, and “electron orbital-like” models for disease transmission relationships.

## **6.0 REFERENCES**

- [1] Pastor-Satorras R, Vespignani A. Epidemic spread in scale-free networks. *Phys Rev Lett* 2001; 86:3200-3203.
- [2] Kiss IZ, Green DM, Kao RK. Infectious disease control using contact tracing in random and scale-free networks. *J R Soc Interface* (2006) 3, 55-62.
- [3] Shirley MDF, Rushton SP. Where diseases and networks collide: lessons to be learnt from a study of the 2001 foot-and-mouth disease epidemic. *Epidemiol Infect* (2005), 133, 1023-1032.
- [4] Ibid.
- [5] Klasse PJ, Shattock RJ, Moore JP. Which topical microbicide for blocking HIV-1 transmission will work in the real world? *PLoS Med* 2006; 3(9):e351.DOI:10.1371/journal.pmed.0030351.
- [6] Krempel L. Simple representations of complex networks: strategies for visualizing network structure. 3rd European Conference for Network Analysis, Muenchen, 1993.
- [7] Houghton RJ, Baber C, McMaster R, Stanton NA, Salmon P, Stewart R, Walker G. Command and control in emergency service operations: a social network analysis. *Ergonomics* 2006 Oct. 15; 49(12):1204-1225.
- [8] Robertson GG, Mackinlay JD, Card SK. Cone trees: animated 3D visualizations of hierarchical information. *Human factors in computing systems conference proceedings on Reaching through technology*, 1991, pp.189-194.
- [9] Jolly AM, Wylie JL. Gonorrhoea and chlamydia core groups and sexual networks in Manitoba. *Sex Transm Infect* 2002;78(Suppl 1):i145-i151.
- [10] Green DM, Kiss IZ, Kao RR. Modelling the initial spread of foot-and-mouth disease through animal movements. *Proc R Soc B* (2006) Aug 1; 273, 2729-2735.
- [11] Kao RR, Danon L, Green DM, Kiss IZ. Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. *Proc R Soc B* (2006) Aug 22; 273, 1999-2007.
- [12] Armstrong MP, Rushton G, Zimmerman DL. Geographically masking health data to preserve confidentiality. *Stats. Med.* 18, 497-525 (1999).

