Movement and spatial specificity support scaling in ant colonies and immune systems: Application to national biosurveillance

Tatiana Flanagan¹ ² Walt Beyeler² Drew Levin² Patrick Finley² Melanie Moses¹

1. Department of Computer Science, University of New Mexico, Albuquerque, New Mexico
2. Sandia National Laboratories, Albuquerque, New Mexico

Corresponding author: Tatiana Flanagan, tpflana@sandia.gov

Data obtained from biosurveillance can be used by public health systems to detect and respond to disease outbreaks, and save lives. However, existing data is distributed across large geographic areas, and both the quality and type of data vary in space and time. We discuss a framework for analyzing biosurveillance information to minimize detection time and maximize detection accuracy while scaling the analysis over large regions. We propose that strategies used by canonical biological complex systems, which are adapted to diverse environments, provide good models for the design of a robust, adaptive and scalable biosurveillance system. Drawing from knowledge of the adaptive immune system, and ant colonies, we examine strategies that support the scaling of detection in order to search and respond in large areas with dynamic distributions of data. Based on this research, we discuss a bio-inspired approach for a distributed, adaptive, and scalable biosurveillance system.

Introduction

Analyzing and responding to biosurveillance information is fundamental for the health of modern societies. Strategies for the detection of disease outbreaks have received considerable attention from researchers and policymakers, but making the development of a detection and actionable analysis approach scalable over large areas, and applicable to diverse populations, environments and social contexts, remains a formidable challenge. Disease outbreaks are inevitable, and early detection is necessary for adequate containment. However, more data exist than can effectively be analyzed, and those data are distributed across large geographic areas. Further, data sources are diverse (Gajewski et. al 2014, Althouse et al. 2015), noisy, variable in space and time, and have locally distinct contexts that can affect their interpretation. Therefore, intelligent, distributed and adaptive sampling, detection and response is required. Sampling and analyzing data with the goals of reducing detection time and maximizing accuracy is challenging. Additionally, scaling results to large and diverse areas is not feasible without a distributed approach. Current biosurveillance architectures that enable both local and global data analysis could also benefit from a scalable approach with decentralized authority to detect and respond.
How effectively strategies used by biological systems perceive information depends on how it is distributed in time and space. Biological systems respond to distributions of information across time by bet-hedging on uncertain information (Donaldson-Matasci 2001), and evolve rules (aka algorithms) that produce behaviors, or structures, shaped in response to the dynamics of the environment (Gordon 2016). Distributed and parallel solutions adapted to the distribution of resources or information in a system’s environment provide efficient strategies and structures adapted to the distribution of information. Marzen and Dedeo (2017) provide a theoretical framework for achieving optimal information perception under competing constraints on accuracy and misclassification. This approach involves the evolution of a perceptual mapping operating on environmental information where the environment is defined by a stationary probability distribution as well as by a penalty function that imposes costs for misclassification by the perceptual mapping. Mapping accuracy is constrained by a cost proportional to the sophistication of the mapping as measured by the mutual information between the perceived state and the environment that induces the perception. This work highlights tradeoffs in the coupling between perception and environment, but focuses on a unitary rather than a distributed perceptual system.

Complex systems in nature have evolved solutions adapted to distributed search and response. Immune systems and ant colonies scale search and automated response in dynamic environments using a distributed approach (Banerjee and Moses 2011) in which many agents can sample information from different locations. Ant colonies and immune systems have evolved solutions that rely on distributed local sensing to perceive their environment, establish an appropriate response, and dynamically adapt their response over time according to the spatial distribution of resources and varying complexity of the environment. The robust, adaptive and scalable computation realized by biological systems makes them suitable models for addressing problems that require distributed computation that is adaptive and scalable (Moses et al. 2013).

Five RADAR principles proposed by Banerjee and Moses (2011) are common in complex systems and relevant to all systems that seek to adapt to information from dynamic environments. These principles are (1) Robustness, achieved by redundancy, flexible diversity and probabilistic response to partial information, (2) Adaptation to environment signals, (3) Decentralized control for search, (4) an Automated Response that is as distributed as the search, and (5) are scalable to millions of agents, conferring the ability to act in parallel. In previous work (Banerjee and Moses 2011, Moses et al. 2013), we discussed how each of these principles is evident in ant colonies and immune systems. Extending this work here, we consider how the principle of decentralized control and search adapts to a dynamic distribution of resources, and scales to territory or body size increases. We start by describing biological distributed detection and search systems. Drawing from the immune system and ant colonies, we discuss the strategies that support the scaling of search and detection to large areas, and to dynamic distributions of resources by three adaptations: trafficking, or movement of agents through space; functional specificity or spatial memory; and hubs or temporary resident structures. Finally, we discuss the application of these concepts to a bio-inspired approach for a distributed, adaptive and scalable biosurveillance system.
Biological systems

Immune systems and ant colonies use a distributed approach that adapts to dynamic distributions of pathogens or resources. Millions of ants in a colony, and trillions of cells in the immune system detect food or pathogens locally and scale efficiently with increase in size. Scale invariance is relevant to any problem where distributed detection can positively affect the efficiency of the system. Although ants and immune systems are spatially constrained, with behavior clearly adapted to their environment and the use of movement, memory, and local and global information balance, these systems accomplish their tasks efficiently regardless of organism or territory size (Banerjee and Moses 2011).

The immune system

It is unusual for biological systems to be scale invariant. Metabolic scaling theory proposes that most biological rates systematically slow as a function of body size due to the overhead of centralized transport and energy (West et al. 1997, Banavar et al. 2010). Whether immune response times systematically slow with body size is an important theoretical question (Weigel and Perelson 2004, 2009; Althaus 2015). There is some evidence of differences in immune cost, replication rates, and resulting duration of infectivity due to body size (Blaze et al 2017, Althaus 2015, Banerjee, Perelson and Moses 2017). However scaling theory would predict that humans (10,000 times larger than mice) would have immune response times 10 times slower than mice; this has not been observed. We hypothesize that the apparent scale invariance of immune response is due to RADAR principle 3, decentralized search, an example of decentralized processing of information. We propose that scalability is achieved in the immune search for pathogens through three factors, trafficking, memory or functional specificity, and communication.

Mammals rely on their immune system to detect and react to invading pathogens whose distribution in the body varies over time and space. Despite this dynamic environment, and widely varying body sizes, immune response times are nearly scale invariant (Banerjee and Moses 2011). When communication and actions are executed locally, each cell can respond quickly regardless of the size of the system. Distributed processing and the absence of central control lead to immune system computation that is highly scalable, through a combination of strategies such as T cell trafficking, functional specificity, and a balance of local and global communication. Trafficking and specificity allow for better homing to a particular tissue (Wong et al. 2016), thus enabling search strategies to adapt to local context and larger areas. All these strategies are all supported by dynamic structures in the immune system that support the adaptations of search to local context, and larger areas.

Immune cells traffic through the body via a partially decentralized infrastructure: the lymphatic network transports immune cells to local regions where they either identify pathogens in lymph nodes or kill pathogens in tissue. In contrast to the systemic, centralized movement of cells though the cardiovascular network, immune cells traffic to particular tissues and recirculate between tissues and the local lymph nodes, through the lymphatic network in a way that balances
local and global movement of immune cells (Moses 2011, Banerjee & Moses 2011). The trafficking of T cells is a dynamic process. Following their development in the thymus, naïve T cells continually circulate throughout the body until they encounter foreign pathogens. When naïve T cells recognize a pathogen, they divide and express molecules that help fight infection. Of this new population, 90–95% undergoes apoptosis, the rest remain close to the site of infection.

As immune cells circulate, some stay close to the tissues where a pathogen is likely to reside. Human T cells have unique phenotypes with different degrees of tissue specificity (Wong et al. 2016). Surviving T cells give rise to long-lived memory populations (Nolz et al. 2011). Resident memory T cells mediate immune memory, which generates long-lived non-recirculating cells that reside within the originally infected tissue. These cells are superior to circulating T cells at providing rapid long-term protection against re-infection in specific tissues (Jiang et al. 2012). In contrast to re-circulating T cells, resident T cells are positioned for rapid detection and response. Once a virus is detected, resident T cells respond to an infection by using the local tissue environment to recruit immune cells (Rosato et al. 2016). To ensure communication of learned pathogens throughout the body, a proportion of memory T cells, like naïve cells, circulate throughout the body until they are needed (Omilusik and Goldrath 2017). Memory T cells are maintained by continual recruitment of new cells from the circulation, suggesting a dynamic memory in the immune system that depends on a systemic source (Ely et al. 2006).

**Local and dynamic structures** throughout the body support efficient search and response strategies. Search is focused in small lymph nodes where antigen-bearing cells are concentrated. Immune cells conduct efficient parallel search in lymph nodes where immune cells are introduced to potential pathogens in a small search space. Immune cell movement is evolved to sample multiple pathogens quickly (Fricke et al. 2016). Immune cells are guided by chemokines (Banerjee et al. 2011, Levin 2016) and structural cues (Mrass et al. 2017) in tissues. Guidance to sites of infection particularly speeds up search in large animals more than in small animals, decreasing time to clear infection in humans by orders of magnitude more than in mice (Banerjee et al. 2010).

To support dynamic specificity of local search and response strategies, the immune system evolved temporary resident structures that allow the immune system to dynamically adapt to its environment. One such structure is the inducible bronchus-associated lymphoid tissue (iBALT), an immune system structure that develops in lung tissue in response to tissue inflammation. Present in larger numbers when local inflammation is chronic, iBALTs provide a local site for T cell priming and B-cell education to clear future infections in nearby tissue and enhance protective immunity against future respiratory pathogens (Foo et al. 2010).

An efficient and proportionate response derives from having distributed memory realized through circulating and resident memory T cells. Antibodies are produced faster and more efficiently where the body experiences the same disease in the same general location. Thus a memory and movement dependent efficiency of response and proportionate response (not overreacting to the threat) are essential features of well-trained immune systems.
Distributed processing and the absence of central control in the immune system lead to scalable processing and response. Scalability is realized through a combination of T cell trafficking, functional specificity, and a balance of local and global communication. All of these factors, which make the immune system scalable, are possible because the immune system uses decentralized recognition of self from others that may be potential pathogens, and remembers previously encountered pathogens (Von Boehmer 1990) so that, for example, T cells are able to kill cells in tissue without any centralized input: negative selection ensures that T cells attack only non-self cells. Negative selection is the hallmark of the adaptive immune system enabling encoding of self and other. In scaling, these strategies, supported by dynamic structures, adapt search to local context, and larger areas.

Ant colonies

Ant colonies rely on individual foragers to search for food sources and bring them back to the colony. The distribution of these resources varies in space and time, so ant colonies use diverse foraging strategies (Lanan 2014) that emerge in response to direct and indirect social cues (Gordon 2010). Distributed foraging and lack of central control in ant colonies lead to scalable foraging and response. Scalability in ant foraging is achieved through a combination of foraging behaviors that involve distributed movement of individual ants, learning, remembering, and communicating the location of food resources. These foraging strategies are all supported by dynamic colony structures, nests and trails, and adapt to local context and territory size. We hypothesize that scalability of foraging in ant colonies, like the immune system, is due to RADAR principle 3, decentralized search, a means to process distributed information. This is possible through three strategies, autonomous movement of individual ants, memory, and communication that allows the colony to learn from individual sampling of information.

The movement of individual ants in a colony reflects different strategies to retrieve food for the colony. The repertoire of foraging behaviors reflects the distribution of resources (Levin 2015). In prior modeling work we demonstrated that ant colonies effectively use different collective foraging strategies that respond to these distributions by combining a small set of simple behaviors tuned for a particular environment (Letendre et al. 2013, Hecker et al. 2015). These combined strategies make for effective search among large numbers of individuals connected by a distributed communication network. The resulting behaviors are not directed by any individual ant but, rather, emerge from interactions among individuals and from the interaction of the individuals with its local environment, where ants perceive information about the distribution of resources in their territory using only local sensing. An individual ant can learn and memorize information about the location of resources only from a small portion of its environment and respond to local conditions. However, the sampling of environmental information by ants through their movement, individual sensing, and communication among them through local interactions tends to overcome individual errors, improving collective function on average. The combination of movement, memory and local perception paired with communication increases the repertoire of responses to varying quality of foraging sites.

As colonies and their territories grow, search and communication strategies must vary accordingly. Ant colonies use additional behaviors and dynamic structures that allow them to
retain efficiency when searching. In large ant colonies, energy constraints prevent moving resources to one central nest. In response, ant colonies choose to deploy temporary structures (nests or trails) closer to locations where resources are more likely to occur. In a strategy analogous to the immune system, ant colonies distribute their nests, making their foraging spatially specific to a smaller area and allowing them to use behaviors adapted to the local environment. This is the case in polydomous ant colonies, which have evolved strategies to deal with the diminishing returns of central place foraging by establishing multiple interconnected nests with decentralized foraging. Argentine ants support distributed, adaptive foraging through dynamic foraging structures (nests and trails) that exist only when needed. Ephemeral trails connect to persistent trails, providing efficient routing, just as with virtual networks like cell phone towers (Flanagan et al. 2013).

The ability of ant colonies to scale foraging to large territories using distributed, dynamic and adaptive structures leads to scalable processing and response. Like the immune system, ant colonies can efficiently scale to large territories through a combination of movement, local foraging, and a balance of local and global communication; they support these strategies by establishing dynamic support structures, temporary nests and ephemeral trails.

Scale invariant biosurveillance

From these principles and scaling strategies used by biological systems we posit that memory and movement are fundamental properties of adaptive biosurveillance. To maximize representation accuracy and minimize detection time, a combination of trafficking and spatial specificity is necessary. This can be achieved by sampling adaptively and locally, matching the dynamic distribution of information in space and time, through deployment of distributed motile sensors, which can become specific and reside in the local environment, and through structures that create, and allow sensors to learn and respond in close proximity. The number, spatial distribution, functionality, and behavior of sensors will depend on the distribution of information and local context. To maximize representation accuracy, trafficking sensors are necessary to detect information that is randomly distributed in space or time. Once information clusters are found, more communication between sensors can lead to an efficient response. To minimize detection time, the use of individual memory, resident detectors, and physical structures analogous to temporary ant colony structures, memory cells, and iBALT can quickly respond to localized events or information, regardless of scale.

Ants and immune systems have evolved strategies that solve distributed search and communication problems (Prabhakar 2012, Dorigo 2006). Some of these strategies mirror or inspire engineered approaches. The multi-place foraging algorithm for robot swarm by Lu et al. (2017) is an example of the efficiency provided by dynamic structures in an engineered system inspired by ant colony behaviors. The study demonstrates how using robotic depots that dynamically adapt to local information in their environment generates more flexible and scalable swarms.

Due to the expansive size, anticipated growth rate and extent of modern biosurveillance data feeds, any potential approach must lend itself well to distributed computation. A balance
between local and global information processing can achieve detection that is appropriate
geographically and produces optimal response times. Scale invariance in decentralized
information processing systems is a must for information systems that operate over a large
geographical extent, such as national biosurveillance. As with a partially decentralized immune
system, we propose that a scalable design should follows RADAR principles with the addition of
memory and movement for scalability, a system that thinks locally, but can act globally.
Following these principles will result in a system that is:

(1) robust, redundant, flexible, and stochastic responses to partial information by using
sensors that utilize stochastic and inferential change detection able to predict, rather than
only statistically describe;
(2) adaptive to dynamic environmental information through sensors that are capable of
processing multi modal information;
(3) decentralized through local sampling and detecting, aggregating and analyzing
information locally, using temporary or permanent local nodes as support structures, and
increasing the spatial extent for aggregation of data according to the severity of the
signal;
(4) automated for efficient responses, distributed according to local detection; and
(5) scalable to millions of agents through the use of a balance of autonomous trafficking
through space to search for distributed information and to distribute information
remembered locally, utilizing specificity, and resident structures to minimize detection
and response time in a dynamic environment.

A biosurveillance system consists of detection and analysis of harvested information. Detectors
that are motile, able to learn from local context, use diverse information streams, along with
temporary resident nodes that provide local aggregation can support scalability in a
biosurveillance system. In previous work (Levin 2017), we described the implementation of an
anomaly detector for health data based on the human immune system. Our negative
selection algorithm detects anomalies in the large, complex data from modern health monitoring data
feeds. The parallelized version of the algorithm demonstrates the potential for implementation on
a scalable distributed architecture. Using strategies analogous to distributing search into lymph
nodes of the immune system, these anomaly detectors have the potential to be motile, to be able
to distinguish self and remember encounters with non-self, and to act as trafficking or resident T
cells, making them the perfect detector for a robust, adaptive, and scalable national
biosurveillance system. Figure 1 illustrates our concept of a National Immune Network.
Information nodes can form dynamically according to local information gathered by detectors.
By varying cluster size and number of connections, we achieve an optimal global detection time
and immediate local detection $O(\log(n/c))$, where $n$ is the number of nodes, and $c$ is the number
of nodes in a cluster. This ‘densification’ is an emergent property of technological networks
(Kleiberg 2004).

There are a number of biosurveillance scenarios that would improve using dynamic, adaptive
detection, local training and residence, and global sharing of information. Biosurveillance efforts
need to balance memory of events that occur in a specific spatial context while recognizing the
motility of both people and pathogens that require the motility of detectors. For example, as a
result of sporadic cases, New Mexico hospitals are trained to detect bubonic plague in patients, while most other hospitals in the nation are not. This resulted in the death of a person from New Mexico when a hospital in South Carolina was not able to make a timely detection (Valentine 1983). In a more recent incident (DePalma 2013), a case of bubonic plague was initially suspected of being a bio-terrorist attack because it was discovered in New York, a location with a different epidemiological context than New Mexico, the location where the infection occurred. Although the spatial incidence of disease may be more likely in some locations than others, human motility can cause disease spread to span large areas. While local biosurveillance nodes can develop specialized detectors, ensure the presence of locally trained detectors that work in combination with specialized nodes, detector motility provides a powerful tool to share detected information on a global scale. In contrast, Zika is an example of a local outbreak of a virus endemic to specific regions and recently detected in Florida. The Zika outbreak spurred a national effort to expand our detection ability. In this case, detectors informed by the local context can improve the efficiency of biosurveillance by acting locally and communicating globally only when needed, following the spatial patterns of the spread from the center of an outbreak to unexpected locations.

Using a distributed, adaptive biosurveillance system we can also address questions about health behaviors such as opioid abuse. How does opioid abuse manifest itself in different regions of the country? Regional detectors can adapt to the regional context and behavior indicators specific to patterns of opioid abuse and spread. We can detect behaviors within those regions, and compare rural versus urban behaviors. A detector for one region may not work as well as a detector for another, but motile detectors would be a way to combine local context with global communication and eventually adapt to different regional contexts.

The application of RADAR principles is not limited to the implementation of detection and analysis, it can be used to complement organizational practices in national agencies invested in biosurveillance. Managing biosurveillance data requires discriminating access to information, for example due to privacy, national security and other data sharing limitations, but requires transparency at the same time. RADAR suggests that information sharing can be effective when information is communicated locally. Regionalization, benchmarking, and sharing best practices can be seen as organizational analogies for keeping institutional memory/modeling robust in biosurveillance initiatives.

Further studies would benefit from an extended mathematical framework for distributed perception. To characterize perception accuracy, an approach that perceives environmental information conditional on the position of perceptual nodes and detectors, and introduces inter-agent communication with the quantification of associated costs and contributions to accuracy. This elaboration could be used to understand the relationship between information variability and distribution/specialization as an optimal detector design, as a function of relevant costs and communication designs. Immune systems and public health networks reward early detection of non-stationary processes. To characterize perception delays, penalties associated with delayed response may be captured as misclassification of stationary processes. Longer-term non-stationarity is also of interest. We are interested not only in the optimal organization of perceptual networks, but also in the ability of that organization to efficiently track dynamics of the environmental signal over time. Studying the properties of optimal solutions will be helpful,
however attention to the processes driving structural dynamics will be needed for a design approach that utilizes different perspectives.

Figure 1. A bio-inspired disease surveillance system with distributed information processing in information nodes (lymph nodes), shown as black ovals. The flow of information towards and between nodes is shown as black lines. Blue circles represent population density according to shade; darker shades of blue are more densely populated than lighter shades. Each node concentrates and processes information.

Acknowledgements

We gratefully acknowledge funding from a Sandia National Laboratories Academic Alliance LDRD Award. TPF and MEM also acknowledge funding from the McDonnell Foundation Complex Systems Scholar Award. We also thank Judy Cannon for her insightful guidance on immune system memory cells, Stephanie Forrest for insightful discussions about negative selection and its relevance to biosurveillance, and Louise Maffitt for helpful editing suggestions.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy’s National Nuclear Security Administration
under contract DE-NA0003525. The views expressed in the article do not necessarily represent
the views of the U.S. Department of Energy or the United States Government.

**Literature cited**

Althaus, Christian L. "Of mice, macaques and men: scaling of virus dynamics and immune

Banavar, Jayanth R., Melanie E. Moses, James H. Brown, John Damuth, Andrea Rinaldo,
Richard M. Sibly, and Amos Maritan. "A general basis for quarter-power scaling in animals."

Banerjee, Soumya, and Melanie E. Moses. "Scale invariance of immune system response rates
and times: perspectives on immune system architecture and implications for artificial immune

Banerjee, Soumya, Alan S. Perelson, and Melanie Moses. "Modelling the effects of phylogeny
and body size on within-host pathogen replication and immune response." *Journal of The Royal

Banerjee, Soumya, Drew Levin, Melanie Moses, Frederick Koster, and Stephanie Forrest. "The
value of inflammatory signals in adaptive immune responses." In *International Conference on

Brace, Amber J., Marc J. Lajeunesse, Daniel R. Ardia, Dana M. Hawley, James S. Adelman,
Katherine L. Buchanan, Jeanne M. Fair, Jennifer L. Grindstaff, Kevin D. Matson, and Lynn B.
Martin. "Costs of immune responses are related to host body size and lifespan." *Journal of

DePalma Anthony. “Reliving the nightmare of plague, 10 years on”. *The New York Times,
Health*, January 7, 2013. URL: https://www.nytimes.com/2013/01/08/health/when-the-plague-
came-to-new-york.html. Accessed on 3/26/2018. A version of this article appears in print on
January 8, 2013, on Page D5 of the New York edition with the headline: Reliving Nightmare Of
Plague, 10 Years On.

Donaldson Matasci, Matina C., Carl T. Bergstrom, and Michael Lachmann. "The fitness value

Optimization and Swarm Intelligence: 5th International Workshop, ANTS 2006, Brussels,


