Published in final edited form as:

Int J Auton Adapt Commun Syst. 2011 ; 4(2): 122–144. doi:10.1504/IJAACS.2011.039720.

From biological and social network metaphors to coupled bio-social wireless networks

Christopher L. Barrett,

Department of Computer Science and VBI, Virginia Tech, Blacksburg, VA 24061, USA

Karthik Channakeshava,

Department of Electrical and Computer Engineering and VBI, Virginia Tech, Blacksburg, VA 24061, USA

Stephen Eubank,

Department of Computer Science and VBI, Virginia Tech, Blacksburg, VA 24061, USA

V.S. Anil Kumar, and

Department of Computer Science and VBI, Virginia Tech, Blacksburg, VA 24061, USA

Madhav V. Marathe

Department of Computer Science and VBI, Virginia Tech, Blacksburg, VA 24061, USA

Christopher L. Barrett: cbarrett@vbi.vt.edu; Karthik Channakeshava: kchannak@vbi.vt.edu; Stephen Eubank: seubank@vbi.vt.edu; V.S. Anil Kumar: akumar@vbi.vt.edu; Madhav V. Marathe: mmarathe@vbi.vt.edu

Abstract

Biological and social analogies have been long applied to complex systems. Inspiration has been drawn from biological solutions to solve problems in engineering products and systems, ranging from Velcro to camouflage to robotics to adaptive and learning computing methods. In this paper, we present an overview of recent advances in understanding biological systems as networks and use this understanding to design and analyse wireless communication networks. We expand on two applications, namely *cognitive sensing and control* and *wireless epidemiology*. We discuss how our work in these two applications is motivated by biological metaphors. We believe that recent advances in computing and communications coupled with advances in health and social sciences raise the possibility of studying coupled bio-social communication networks. We argue that we can better utilise the advances in our understanding of one class of networks to better our understanding of the other.

Keywords

bio-inspired; biomimetics; bio-social wireless networks; cognitive radios; cognitive sensing; wireless epidemics

1 Introduction

Biological and social analogies have been often applied to human-engineered systems. Examples of these metaphors include: regulation, adaptation, replication, regeneration and repair, memory, learning and optimisation and population-level interactions and dynamics such as competition and goal-seeking behavior. The field, called *Biomimetics – applying*

designs drawn from observation of natural systems, or inspired by natural systems, to solve problems in engineering, computing, materials science, etc. – is by now a well established branch of science with a number of important contributions. Inspiration from biological solutions (a Biomimetic perspective) to many problems found among living organisms are seen in engineered products and systems ranging from Velcro to camouflage to robotics to adaptive and learning computing methods, to name just a few.

In this paper, we present an overview of recent advances in understanding biological systems using a combination of network science and algorithms, and use this understanding to design and analyse wireless communication network applications. We will call this as *biomimetic wireless networks*. The broad view of a program in biomimetic wireless networks is to represent biosystems using co-evolving graphs. This representation has two important features. Firstly, it views biosystems as dynamic networks as a way to capture the local interaction structure inherent in such systems. Secondly, it uses an algorithmic (aka mechanistic or executable) representation of individual components to represent how the component evolves as a result of its internal state and the state of its neighbours in the dynamic network. This networked and algorithmic representation forms the basis of translating biological analogies and principles for design, control and evolution of wireless networks. This viewpoint represents a departure in the way biosystems are usually studied and ways in which biological principles are used for designing engineered systems. Specifically, we are looking for ways to apply system level principles that govern biosystems to engineered systems, rather than designing or controlling smaller individual components. Further, the inspiration from biological systems can be derived at different levels – modelling approach, control or algorithms can be bio-inspired. The case studies we present here illustrate the application of system level principles and highlight aspects that are bio-inspired.

We make three main points in this paper. Firstly, we describe how a combination of network and algorithms based biological principles/analogies can be applied to better understand, design and study emerging wireless networks. The focus is not so much to naively apply these paradigms, but to translate these ideas through a mathematical theory comprising networks, games, dynamical systems and algorithms (Carlson and Doyle, 2002). General biological principles cut across various levels of hierarchy, for example, memory and control are not a centralised process but it is a part of various levels of the ecosystem; thus, we see individual components within a cell exhibiting these traits, but we also see traits of this kind appear at the level of social networks. We can see corresponding traits within communication networks as well: individual cognitive radios (CRs) can learn and control how frequencies are accessed and power levels are controlled. Network level memory, adaptation and control can also be seen in routing mechanisms and transport protocols. The plan we follow is comprised of the following steps:

1. Understand biological processes as information processing and communication systems using networks and algorithms;
2. Advance mathematical/computational theory for abstract representation of these processes;
3. Develop description of communication systems and processes using the mathematical and computational abstractions developed above; the specific focus is on finding representations in communication systems that are biology-like; and
4. Develop communication networks through iterative modelling, simulation, experimentation and analysis.

The exemplars we choose here obey these general guiding principles.

The second point we would like to make is cautionary in nature. We believe that the use of the terms *cognitive radios* and *cognitive networks* are a bit unfortunate.¹ Cognitive scientists agree that we do not really completely understand cognition at the current time (Fodor, 2000). Moreover, there is a certain level of circular argument that we are faced with. Specifically, a computational theory of mind provides a rich and a reasonable starting point in our understanding of cognition. This theory is based on a large part of Turing's proposal that cognitive mental processes are operations defined on syntactically structured mental representations that are language like (Fodor, 1983). Cognitive wireless networks can be essentially viewed as certain forms of computational devices. Thus, it seems to us that the right way to study cognitive networks is not so much as to use cognitive metaphors, but to jointly investigate a network-based algorithmic representation of both the cognitive and wireless systems.

Finally, we argue that recent advances in wireless networks, computing and communications, social networks, systems biology and medical informatics raise an intriguing situation in which bio-social and wireless networks need to be studied together. This further strengthens our view outlined above using algorithmic network science as transdisciplinary conceptual framework to represent as well as understand these coupled and coevolving systems.

1.1 Why study biomimetic wireless networks now?

Current-generation computing and communications technology have evolved from the elaboration of engineering possibilities. For example, the advent of integrated circuits enabled the development of smaller devices, which in turn motivated development of new battery technologies, etc. However, there is a point beyond which incremental elaboration and refinement of available engineering technology are insufficient or even maladaptive. We believe that the confluence of advances in biology, computing and networks make a convincing case to undertake a study of biomimetic wireless networks (Nature Special Issue, 2006; NSF Workshop, 2009). Three specific advances underscore our thesis:

1. *Emergence of network science.* Over the last ten years, the scientific community has embraced the importance of networks and has begun active research programs to understand them.
2. *Information processing, description and analysis of bio-systems as networked systems.* Until recently, almost all large-scale biological systems (social and neural networks are two exceptions) were primarily described using differential equations and making strong simplifying assumptions about symmetry and influence. These simplified descriptions were necessary for obtaining analytically tractable solutions. Recent advances have shown that formal network-based models of biological processes and systems can be constructed from composing subsystem behaviours and yet yield natural representations of the composed dynamics (Brent and Bruck, 2006). Such network-based algorithmic representations are generative as well as locally algorithmically tractable. The advent of high-performance-computing systems and simulations means that network-based models can now be constructed and studied using computational tools and an algorithmic perspective, thereby fundamentally changing the kind of questions that can be asked.
3. *Meso-scale ubiquitous and nomadic computing.* Some of the most significant technological advances in the past decades have been in the area of wireless networks. Communication and information networks are rapidly approaching the complexity in size, scale and diversity of multi-scale multi-species biological networks. Advances in communication and computing have reached a stage where the distinction between today's PCs and mobile phones is blurred – we are

witnessing the beginning of nomadic and ubiquitous computing and communications, wherein devices are simultaneously transceivers, computers and sensors. They form an integral and increasingly indispensable part of today's social fabric and world economy. These will be heterogeneous, consisting of devices that are small, have limited energy and computing ability, such as sensors, to powerful CRs and wireless devices. They will be owned by various entities and governed by various policies responsive to diverse and often competing social purposes. They will be dynamic in various ways: individual devices fail all the time, the devices move around at various speeds, they shut themselves intermittently to save energy. All this causes changes to network topology and structure that spans various time scales. The network comprises software, hardware and social components – each component interacts with the others and this constant interaction plays a central role in how the network is formed, evolves and sustains itself. The result is a very complex multi-scale coupled and co-evolving social, information, computing and communication system. The system is not so much designed as evolving.

1.2 Organisation

In this paper, we discuss two exemplars of bio-social network metaphors applied to the study of wireless networks – cognitive science and epidemiology. We have chosen to discuss these metaphors for a couple of reasons.

Metaphors from cognitive science are now being used in design of *CRs* to *cognitive radio networks* – wireless networks that are usually comprised of CRs; although the concept of cognitive networks in general need not be restricted to networks of CRs only. The second metaphor we discuss is epidemiology. Epidemiology aims to study the spatiotemporal patterns of the spread of infectious diseases through a population (humans as well as animals). We discuss epidemic algorithms for information diffusion and replicated databases and epidemic models for worm propagation. In both examples, biological metaphors are used in a non-trivial manner for designing the network or dynamics on these networks.

In the last part of this paper, we re-examine some of our applications and discuss a potentially interesting direction for future research. We argue that recent advances in computing and communications coupled with advances in health and social sciences raise the possibility of studying coupled bio-social and communication networks. Study of such networks will require that we understand these coupled networks in a spiral fashion wherein advances in our understanding of class of networks feed into a better understanding of the other network.

2 A formal framework to study networked systems

A formal theory for discrete dynamical systems on networks consists of the following components (Barrett et al., 2005b):

1. a collection of entities with state values and local rules for state transitions
2. a structure to represent interaction between individual entities with its neighbouring entities
3. an update sequence or schedule such that the causality in the system is represented by the composition of local mappings.

We will use coevolving graphical discrete dynamical systems (CGDDS) here; the specific choice of the formal framework depends on their ability to naturally represent the desired phenomenon and the computational complexity of reasoning about objects represented using these formalisms.

A CGDDS S over a given domain $\mathbb{D}$ of state values is a triple $(G(V,E), \mathcal{F}, W)$, whose components are as follows:

1. $G(V,E)$ is the underlying network on a set V of entities – in the case of a wireless network, V (henceforth, referred to as vertices) is the set of devices. The vertices have states which take values from the domain $\mathbb{D}$. Also, we have an edge modification function g_i associated with each vertex v_i , whose application results in an indexed sequence of graphs.
2. $\mathcal{F}$ represents a set of local transition functions f_{v_i, d_t} corresponding to each vertex v_i , where d_t is the degree of node v_i at time t . The function f_{v_i, d_t} acts on the state sub-configuration induced by v_i and its neighbours.
3. W is a schedule of updates and it is represented as a string over the alphabet $\{v_1(s), v_2(s), \dots, v_n(s), v_1(g), \dots, v_n(g)\}$. It specifies the order in which the state of a vertex or the possible edges incident on the vertex will be updated. Both f_i and g_i can be probabilistic.

Computationally, each transition in the CGDDS involves updating either a state associated with a vertex or modifying the set of incident edges on it. f_i is used to decide if the worm is transmitted between two devices. g_i depends on the individual behaviour and specific interventions applied. The effect g_i has on the system is to either remove certain edges or change the probability on the edges in E . CGDDS can be viewed computationally by representing each local function as a local algorithm and interactions between these local algorithms to produce global dynamics. CGDDS are quite general; graphical models of games, inference and dynamical systems are well represented using this formal framework (Barrett et al., 2000).

3 Wireless systems inspired by cognitive science

3.1 Cognitive radio networks

Several studies have shown that the licensed spectrum bands are extremely underutilised. For example, the utilisation in some bands can vary between 15% and 85% (Akyildiz et al., 2008). This has led the FCC to consider proposals to open up some bands to unlicensed users (henceforth, referred to as *secondary users*), as long as they do not disrupt the operation of the licensed users (henceforth, referred to as *primary users*). Recent advances in technology have led to ‘Cognitive Radios (CR)’, which are devices with improved capabilities to sense the spectrum usage and detect ‘holes’ (which refers to unused bands) and adaptively switch channels and power levels with low overheads. Motivated by cognition in living beings, cognitive radio networks have been proposed to allow secondary users to effectively utilise spectrum (Akyildiz et al., 2008). Since a key requirement of these networks is to ensure the operability of primary users, secondary users have to follow a so-called *cognitive cycle* comprising the following steps: *spectrum sensing*, *spectrum analysis* and *spectrum decision* (Mitola, 2000). As the names suggest, the spectrum sensing and analysis steps involve determining the spectrum holes and their characteristics. The third step of the spectrum decision involves scheduling the secondary users among the detected holes in an optimal manner. We refer the reader to MacKenzie et al. (2009) for a more recent survey of CR and networking research.

Other biological metaphors have also been used in different aspects of the above cognitive cycle. For instance, influenced by the movement patterns followed by a flock of birds or a school of fish to minimise the danger of predators as a group, Doerr et al. (2007) propose an approach for CRs to minimise outside interference, while maintaining good communication environment within a swarm. Depending on the environment, CRs select parameter values for frequency, transmission power, modulation, encoding, etc. Motivated by swarm

intelligence, Renk et al. (2008) propose a construction of sub-networks to select communication frequencies among radios within a sub-network, while conveying this information to other neighbouring sub-networks. To account for the dynamic behaviour of CRs, they employ a decay factor for this selection similar to pheromones laid by ants to mark the routes to the food source. A similar approach is proposed in Chen et al. (2008) for allocating control channels to CRs.

3.2 Case study 1: cognitive sensing and control

As an illustrative use of a bio-inspired paradigm for query and monitoring, we describe a cognitively inspired service and model architecture (CoSMo) for sensing and control in a very general setting (Phalak, 2006; Tupe, 2006); the specific application of this architecture, we describe in this paper, is to provide situational assessment and monitoring of urban traffic. A number of transportation departments and online tools provide instantaneous traffic updates by aggregating data from a large collection of sensors and providing an online interface. However, these tools do not allow a user to pose more complex queries, for example, traffic status at some point in future, detect accidents and other events, and estimate emissions. There are a number of challenges to efficiently support such queries: for a large urban area, such a system would involve a large number of sensors and a huge amount of real-time data to be collected from them. Additionally, many kinds of queries mentioned above require some form of reasoning and cannot be supported simply by static analysis of past data. CoSMo is a cognitively inspired architecture, which consists of an internal engine that collates raw sensor data and other inputs and builds a model for the real world, and uses this to reason about the sensor network.

There have been several proposals for such monitoring systems. Examples of such systems include CarTel (Hull et al., 2006), PRESTO (Li et al., 2006) and Directed Diffusion (Intanagonwiwat et al., 2000). However, there are significant differences between CoSMo and such systems. The focus of CarTel, for instance, is more on providing a distributed mobile sensor network, upon which other applications can be built. PRESTO (Li et al., 2006) uses a feedback-based model-driven push approach directed towards capturing data trends, supporting archival queries and incorporating an adaptive design to user and data needs. It has the ability to support queries on current sensor data as well as on historical data using interpolation, prediction and indexing into local repositories. PRESTO is similar in spirit to CoSMo, but our system is endowed with the ability to learn, decide, react and reconfigure itself. It is also equipped with a simulator component to take over predictions and more importantly with dynamic and static representations of the measured and measuring entities. This allows CoSMo to support much more sophisticated queries, with varying levels of (user-specified) accuracy. Directed diffusion (Intanagonwiwat et al., 2000), developed by Estrin et al. is another important data-centric architecture for sensor networks. It uses named data in the form of attribute value pairs and gives importance to what the data are rather than from where it originates. It uses an application-aware paradigm to facilitate efficient aggregation and a robust data delivery technique back to the sink. An important distinction between this system and CoSMo is that the latter is model driven.

3.3 CoSMo architecture

CoSMo is a situational awareness and monitoring of vehicular traffic in urban transportation systems using a network of wireless sensors. The system architecture comprises two main components:

1. a cognitively inspired internal representation for analysing and answering queries concerning the observed system

2. a service-oriented architecture that facilitates interaction among individual modules, of the internal representation, the observed system and the user.

The Cognitively Inspired Model Architecture allows effective deductive as well as inductive reasoning by combining simulation-based dynamic models for planning with traditional relational databases for knowledge and data representation. On the other hand, the service-oriented design of interaction allows one to build flexible, extensible and scalable systems that can be deployed in practical settings.

Figure 1 provides a high-level schematic diagram of CoSMo. The client-side supports users with PDA, laptops or PCs. Client-side is responsible for issuing the desired query through a user interface. The external world also consists of the real world sensor network which is a mesh of sensors which will be ultimately queried for the measurement of the desired phenomenon. Processing components are the core of the system which handles the transformation and interpretation of the query within the system. Some of the constituent components are situation and knowledge aware and add to the cognition of the system. Communication between various modules is done by means of XML because of the machine independence and expressiveness it provides.

The data processing components include a database module and a traffic simulator module that maintains and refines a model for current traffic. The database contains the sensor network state, which is updated periodically, or on demand (at a high cost). It keeps track of the sensor states, their reliability and other information that can help to answer a query. The simulation module uses Transportation Simulation (TRANSIMS), which is a detailed agent-based tool for the simulation of urban traffic. This tool uses past statistical information (such as census, land use, activities, etc.) to make a synthetic model for the movement of individuals in an urban region. It is refined and calibrated by the real-time sensor data from the database periodically. TRANSIMS can be used to get approximate information about the traffic state.

The cognitively inspired element of the system is the *State Manager* that determines how to process a given user query. It currently uses a rule-based scheme to determine whether to use the database or the simulator; these decisions can be based on the accuracy needed for the user, the current state of the database and the simulator and the amount of processing time available. Depending on the question, response within a specific level of accuracy can be obtained by running the simulator, instead of a more expensive query to the sensor network. This organisation is inspired by the current view of mental processes as comprising perceptive processes, intuitive processes (fast, associative) and rational reasoning processes (slow, and deliberate) (Kahneman, 2003).

We use the SOAR (Lewis, 2001) model to evaluate the cognitive capabilities of the CoSMo architecture and list the components that fulfil the requirement. SOAR is a widely known and accepted cognitive architecture in AI and serves as a benchmark for the CoSMo system. The cognitive capabilities are summarised in Table 1.

4 Wireless systems inspired by epidemiology

Computational epidemiology is the development and use of computer models for the spatio-temporal diffusion of disease through populations. The basic goal of epidemiological modelling is to understand the dynamics of disease spread well enough to control it. Potential interventions for controlling infectious disease include pharmaceuticals for treatment or prophylaxis, social interventions designed to change transmission rates among individuals, physical barriers to transmission, and eradication of vectors. Efficient use of these interventions requires targeting sub-populations that are on the critical path of disease

spread. Computational models can be used to identify those critical sub-populations and to assess the feasibility and effectiveness of proposed interventions.

The spread of infectious disease depends both on properties of the pathogen and the host. An important factor that greatly influences an outbreak of an infectious disease is the structure of the interaction network across which it spreads. Descriptive models are useful for estimating properties of the disease, but the structure of the interaction network changes with time and is often affected by the presence of disease and public health interventions. Thus, generative models are most often used to study the effects of public health policies on the control of disease.

Aggregate or collective computational epidemiology models (Bailey, 1975; Diekmann and Heesterbeek, 2000; Dietz and Heesterbeek, 2002; Kermack and McKendrick, 1991; Macdonald, 1952) often assume a population is partitioned into a few sub-populations (e.g. by age) with a regular interaction structure within and between sub-populations. The resulting model can be typically expressed as a set of coupled ordinary differential equations. Such models focus on estimating the number of infected individuals as a function of time and have been useful in understanding population-wide interventions. For example, they can be used to determine the level of immunisation required to create herd immunity.

In contrast to the uniform-mixing models, researchers including Elveback et al. (1976); Eubank et al. (2004); Ferguson et al. (2006) and Longini et al. (2005) have developed network models of infectious disease transmission. In a network model, the entities interact via a network to spread disease. Each potential host is represented by a vertex, and edges between two vertices represent contact between the hosts that creates a non-zero probability of transmission. Associated with each vertex are two things:

1. A label that includes those features of the host relevant to the spread of disease.
2. A model that represents the host's state of health. This may be as simple as a finite state machine distinguishing the susceptible state from the infectious state.

Associated with each edge is a weight giving the probability of transmission across the edge conditioned on the source being infectious. The probability of transmission depends on the susceptibility and infectiousness of the host involved, which may be drawn from a distribution conditioned on any of their attributes. The resulting network, in general, is a directed, weighted graph and is time dependent. Time dependence may be incorporated into edge labels or it may be represented by constructing an explicit unfolding of a sequence of graphs. Further, network models can represent many kinds of heterogeneity (and correlations among them) that can be accommodated in compartmental models only with difficulty, if at all.

Mathematically, epidemics are naturally represented using

1. a CGDDS that captures the diffusion of information across the time dependent, directed, weighted graph; and
2. a partially observable Markov decision process (POMDP) that captures various control and optimisation problems formulated on the phase space of this dynamical system.

4.1 Epidemic algorithms for information diffusion

Information diffusion is a common objective in a number of network applications, for example, peer-to-peer networks and sensor networks; depending on the scenario, the specific goal might be:

- send specific information from a source s to a destination t (unicast)
- send information from a source s to a subset of nodes $T \subseteq N$ (multicast)
- send information from a source s to all nodes in the network (broadcast)
- send information from a node s to all or a subset of nodes within a small distance of the node (spatial gossip or information sharing).

The first class of problems is referred to as epidemic routing (Barrett et al., 2005a; Ganesan et al., 2002), the second class of problems is generally studied under the name of epidemic algorithms for updating replicated databases (Demers et al., 1987; Holliday et al., 2003), the third class is usually viewed as rumor spreading or worm propagation and the final type is really a variant of the second and third with spatial constraints.

Epidemic algorithms are specifically designed to work in environments when network connectivity is highly variable, due to mobility or faults. Thus, these algorithms can be often viewed as extreme forms of distributed algorithms:

1. they assume no global knowledge and very little knowledge even about their neighbours
2. the decision to (re)-transmit information is based usually on the local state at a node and the information (packet) received from its neighbours (e.g. the time packet has been alive, the number of hops, etc.). Specific applications then motivate additional constraints. We briefly discuss two applications to illustrate this.

4.1.1 Updating replicated database—Demers et al. (1987) were the first to consider the use of epidemic algorithms for updating databases. (see Bonnet et al. 2001; Demers et al. 1987; Madden et al. 2002 for additional discussion). In this setting, we are given a distributed replicated database, for example, one formed by a sensor network (Madden et al., 2002) and the underlying network is often dynamic. This makes it much more attractive to use asynchronous methods, for example, epidemic algorithms, to manage such distributed databases. In such algorithms, every time a value of an entry changes at a node, it is locally propagated. Maintaining consistency and serialisability are important challenges, and techniques such as version numbers and exclusive write access are used to address them. Agrawal et al. (1993) outline a more sophisticated approach to ensure consistency and serialisability. The effectiveness of update algorithms can be measured by the total amount of information that is moved in the network to ensure consistency and serialisability, the time it takes for updates to be complete, kind of global information assumed (global clocks) and nature of constraints placed on the databases (only commutative write operations allowed).

4.1.2 Epidemic routing—Probabilistic local routing is another class of algorithms motivated by biological principles (see Epidemic Routing Bibliography for recent papers and discussion). In contrast to the database update problem, the primary goal is to ensure that packets originating at a node s reach a given node t . *Adhoc* networks wherein nodes are moving and sensor networks wherein nodes might be turned off (to conserve power) or fail, provide different kinds of challenges than database maintenance problems. Measures of effectiveness include: time taken to deliver the packet, the total number of packets generated to deliver one packet, the size of individual packets and the information they contain and the amount of global and local information that algorithms assume. Our recent work, Barrett et al. (2005a), demonstrates that such algorithms are very competitive in highly mobile environments such as vehicular *adhoc* networks (VANETs).

Drawing inspiration from social insects such as ants, a class of optimisation algorithms called the Ant Colony Optimisation (ACO) algorithms has been developed (Bonabeau et al., 1999). An ACO algorithm optimises a problem instance by generating multiple solutions to the problem and uses the *goodness* of these to influence future solution construction. This approach has been used for load balancing (Schoonderwoerd et al., 1996) and routing in mobile *ad hoc* networks (MANETs) (Bouazizi, 2002).

4.2 Computer network immunology

The immune system in a biological organism is a collection of processes that protect against diseases by identifying and eliminating pathogens, including bacteria, viruses, parasites and toxins. The natural immune system identifies *self* and *non-self* antigens (Janeway, 2001) generates detectors to bind to non-self antigens and eliminate them. The analogy between a computer's security to the immune system is exploited by artificial immune systems (AIS) to help translate the processes of the natural system to computational systems, engineering and information technology. When applied to computer security, the problem is identifying behaviours as 'self' and 'non-self' actions, both within and outside the host. The connection between immune system and computer security was first introduced in Forrest et al. (1994), although this analogy existed since the term 'computer virus' was coined by Adelman (Cohen, 1987). The multiple facets in the security domain include host- and network-based security systems. Host-based systems can identify self and non-self actions from users' inductive pattern generation (Teng et al., 1990) or process' behaviours (Forrest et al., 1996).

Network-based immune systems protect the computer from external hosts misbehaving over the local area network (LAN) or the internet. Since there exist a number of network applications, defining self and non-self actions for activities over the network is challenging. Further, the immune system can be applied at the different OSI layers and protocols – TCP/IP layer (Hofmeyr and Forrest, 1999), network profile (Heberlein et al., 1990). In wireless networks, there are other complications introduced due to the lack of a centralised authority. This makes applying existing intrusion detection approaches impossible. A survey of artificial immune systems can be found in Drozda and Szczerbicka (2006).

Bit-strings (or detectors) represent the self and non-self actions and can be generated using protocol parameters, such as, sender address, sender port or receiver port (Hofmeyr and Forrest, 1999). When detectors are generated, these systems use a negative-selection algorithm (Kim and Bentley, 2001) to avoid auto-immune reactions (detectors that bind to self actions).

4.3 Case study 2: wireless epidemiology

The past few years have seen the wide-spread and discomfiting frequent occurrences of internet worms such as CodeRed, Code red II, Nimda, SASSER and Slammer (Moore et al., 2006). These worms have caused widespread distributed denial of service attacks resulting in significant network disruptions and staggering economic losses (CERT). Several researchers have developed computational models for representing and understanding internet worms – (see UCSD Internet Epidemiology for recent papers and discussion on internet epidemiology).

The increased ubiquity of smart digital devices and their capability to stay connected, has amplified opportunities for malware attacks in a completely different domain – mobile devices. With over 20 million smart phones shipped in 2007, in North America alone (Kairer, 2008), the malware attacks are expected to increase manifold. If recent worms such as Cabir (Ferrie et al., 2004) and CommWarrior (Lactaotao, 2005) are any indication, future

attacks will target mobile devices frequently. These worms use the bluetooth interface and MMS messages to replicate in susceptible phones.

In contrast to worms such as Code Red, wireless worms, and in particular Bluetooth worms have a spatial nature – an infected Bluetooth-enabled device can infect another Bluetooth-enabled device if the devices are in close proximity (about 10 m) to one another. Due to human mobility, the devices form dynamic networks. Unlike human epidemics where the individual's social network plays an important role in the spread, wireless epidemics can potentially spread through multiple networks:

1. a proximity network created due to the physical distance between devices
2. a social contact network usually consisting of the infected devices' phonebook and/or e-mail contacts.

Thus, as argued in Kleinberg (2007), the short range wireless networks can potentially bring epidemiological models in biological systems and communication networks closer. Recently, researchers have studied the spread of worms in mobile devices (Fleizach et al., 2007; Su et al., 2006; Yan and Eidenbenz, 2006).

In Channakeshava et al. (2009a), we proposed the EpiNet simulator and a framework for studying wireless worms. This framework is based on Simdemics (Barrett et al., 2008; Eubank et al., 2004), a simulation infrastructure for human epidemics on large networks (consisting millions of nodes). Simdemics uses an agent-based model and maintains each individual's attributes, behaviours and activities. It involves the following steps:

Step 1: Create a synthetic urban population by synthesizing several databases from commercial and public sources, while preserving their privacy and maintaining statistical indistinguishability (Beckman et al., 1996).

Step 2: Use activity templates of individuals to create an activity-based mobility model. This generates the social network for individuals using the US census, survey data and time-use surveys (Bowman and Ben-Akiva, 2001).

Step 3: Assign detailed route plans to individuals based on, the locations where activities are performed and the road network that connects the locations.

Step 4: Construct detailed movement patterns using a cellular automata-based micro simulation for individuals over the transportation infrastructure.

Thus, the above four steps generate – via a combination of simulation and data fusion techniques – demographic information for each (synthetic) person and location, and a minute-by-minute schedule of each person's activities and the locations where these activities take place. In this paper, we use the framework we propose in Channakeshava et al. (2009a) to study the dynamics of Bluetooth worms on large, dynamic and realistic networks. We apply only Step 1 and 2 of the Simdemics framework to generate the individuals' activities; these are augmented with the construction of a simple random geometric network within locations to obtain the Bluetooth network. The network so constructed consists of digital devices (as vertices) and wireless links with non-zero probability of transmission of the epidemic (as edges). The probability of transmission (depends on the susceptibility and infectiousness) of the device involved and it is drawn from a distribution conditioned on the attributes of the Bluetooth protocol. For example, inquiry phase or infection phase would have different susceptibility or infectiousness. We embed a mathematical model for the Bluetooth worm in each device and study the spread of wireless epidemics. We briefly describe the network construction, worm model and the simulation study below (see Channakeshava et al. 2009a for more details).

4.3.1 Constructing the bluetooth network—We use *sublocation modelling* to construct the network. Since Bluetooth is constrained by a limited range, not all devices in a particular activity location are connected. The activity information provides the co-location data which is further processed to construct the Bluetooth network. Each activity location in the study is assigned an area based on occupancy and divided into grid cells of 10m side; each such grid cell is viewed as a sub-location. Each person (or device) arriving at an activity location is assigned a random position, which determines its sub-location. The (temporal) bluetooth network is constructed by putting an edge between two devices that are within range. At this time, we do not consider within location mobility of devices.

4.3.2 Mathematical modelling of the bluetooth worm—We propose an abstraction of the Bluetooth protocol and worm characteristics in Channakeshava et al. (2009a) and use it for this study. Figure 2 shows the worm modelled as a probabilistic timed transition system (PTTS), which is directly inspired by diffusion in biological systems. This specific worm model builds on the protocol description in Yan and Eidenbenz (2006), but it is considerably simpler, with fewer Bluetooth specific states. Also, it is easily extensible to other protocols and malware characteristics. Table 2 defines terms used in the worm model description.

The worm model in Figure 2 contains abstract aspects of the Bluetooth protocol for establishing a connection with a neighbouring device. It also includes aspects of the worm protocol including the idle time between infection cycles and time taken to transmit the worm from an infected to a susceptible device after connection establishment. For the purposes of this study, we assume that all devices are susceptible and remain in the *susceptible* state until they become infected through the interaction with infected devices. The worm model has been calibrated and validated through detailed simulation studies (see Channakeshava et al. 2009a for more details).

A Bluetooth connection is established after a successful inquiry phase when an inquiring device – inquiry is performed both by a normal Bluetooth device and an infected device – identifies other neighbourhood devices. The time taken to complete the inquiry phase varies depending on various factors. The *inquiry*, *timeout* and *infected* states handle this phase and the possible factors that affect them. An inquiry may timeout in some cases and is successful in others. The $T_{\text{inq}}^{\text{timeout}}$ represents the time spent by the device when an inquiry request times out and is handled in the timeout state. The T_{inq} histogram for the inquiry state indicates the non-deterministic inquiry time and is randomly chosen from an inquiry time distribution, which is computed by detailed NS2 experiments (Yan and Eidenbenz, 2006).

In addition to the above Bluetooth protocol aspects for establishing connections, the worm exchanges packets with the susceptible devices before replicating the infection. These stages are handled in the incubation and infectious states. The idle state handles the time spent idling before a subsequent infection cycle. T_{idle} represents this idling time. The *incubation* state represents the time taken to infect a new susceptible device and is incurred only by a newly infected device. This is represented in the T_{inf} histogram and is decided randomly following the T^{inf} distribution. The infectious states have several sub-states each denoting the probability of infecting a device after a given number of time steps. This is derived from the T_{inf} histogram. The branches in the infectious states indicate the duration for which the device remains infectious in a particular infection cycle. For example, the first branch is taken with probability p_1 and remains infectious for 1s, with p_2 for 2s, etc.; $p_1, p_2, \dots p_n$ are derived from the T_{inf} histogram.

4.3.3 Simulation study—For the simulation study, we select a region in Chicago downtown with a synthetic population of about 30,000 individuals (Beckman et al., 1996) and about 2,000 locations. We extract their activities for an 8 hr duration, from 8 to 4 pm, at these locations. The people with ages in the range of 20–50 were selected to carry digital devices. We make the following assumptions for this study:

1. We assume that device penetration is 100% in the age group we have selected;
2. Devices are identical and all devices are susceptible;
3. All devices are both *discoverable* and *connectable*, implying that the Bluetooth inquiry and connection requests succeed; and
4. The activities of the people were rounded to 300 sec (or 5 min.) intervals, making arrival and departure occur at those intervals.

Each device has a worm model representing the state of the device while the infection manifests in a device. We consider 5% and 10% of the devices to be initially infected and study the spread. Each experiment was conducted for five seeds and we present the average of the five seeds in the plots. We conduct studies to

1. identify the dynamics of the spread
2. effect of static degree-based interventions
3. effect of early or late interventions
4. effect of increased idle time between infection cycles.

Figure 3 shows the infection spread with 5% of the devices randomly selected to be initially infected. From the spread dynamics, it is seen that most infections take place in the first half an hour after the seed at 8 am. The increase is 100% of the initial seed size at 8:30 am and steadily increases till 12 pm. There is another spike in the infections just after 12 pm, indicating that the devices not exposed in the initial wave are exposed after lunch time. The final infection size is 40% of the susceptible devices. The infection does not spread to all devices during the 8 hr of simulation indicating that the network structure play an important role.

Next, we study the effect of applying interventions to control the spread of the worm – examples of such interventions include putting software patches, isolation or turning off infected or vulnerable devices. Motivated by interventions in the case of human epidemics that use structural measures (e.g. degree and centrality) of nodes quite effectively (Eubank et al., 2004), we study interventions based on node degree. For computing degree of each device, we construct a graph for every 5 min. and then compute the union of all the graphs. We determine the degree of each device simply as the size of the neighbour set of each device in the union graph.

In the study, we evaluate the effect of early (1 hr after seed) and late (4 hr after seed) interventions, on the spread. We also alter the number of devices patched, more (50%, 25%) and less (5%) to compare them. Figure 3 shows the variation in the spread. When only 5% (or 1,500) devices are patched, early or late interventions have little change in the spread. Further, choosing a small set of devices for patching does not impact the spread characteristics. Choosing more devices to patch slows the worm spread considerably – almost by the same factor as the percentage of devices patched. The effect is more in case the intervention is applied earlier. These results clearly indicate that patching devices after the seed is only effective if the patches are applied very early (earlier than 1 hr after seed). This also indicates the impact of worms that are released before the patch is implemented

and incorporated in susceptible devices. Patching more (25 or 50%) devices early slows the spread of the worm, only slightly, with the impact being more for larger percentages.

We now evaluate the effect of idle time on the spread. Idle time is an important parameter that can define the potency of the worm. Larger idle times imply that the worm operates in stealth mode without giving warning signs – faster battery discharge or non-responsive devices – of the infection. So, we are interested in observing the spread when this parameter is varied. By modelling the Bluetooth worm at an abstract level (Figure 2) we can conduct these studies and observe the effect of certain worm parameters on the spread. Figure 4 shows the impact of the idle time on the spread. As expected, the worm spread slows down when the idle time is increased with the significant impact arising due to the slower initial spread. With an increase in the idle time to 1,200 sec, the worm spread reduces to $\approx 50\%$ of the original (idle time = 10 sec). Idle time affects the initial infection speed, causes a slower initial spread. Further, once a certain number of devices get infected, it takes off with the same speed, although, infects less devices within the simulation duration of the study. Furthermore, this indicates that slowing the speed initially can result in a reduced and slower infection spread providing more time to apply patches.

4.3.3.1 Summary of our results: Building on our simulation tool, EpiNet (Channakeshava et al., 2009a), we have done several studies on the dynamics of Bluetooth worms, and static and adaptive strategies for controlling them. The above simulation results along with other results from Channakeshava et al. (2009a,b) suggest that Bluetooth worms spread very fast, reaching the timescales of human mobility, corroborating the hypothesis put forth by Kleinberg (2007). This makes the problems of detection and control very crucial. Control strategies that select nodes for antivirus updates based on static graph measures do not seem to be very effective, as shown above. Adaptive control strategies, which involve choosing nodes based on the level of infection within local neighbourhoods seem much more effective. The underlying mobility model, which is a result of the evolution of the social contact network, seems to have a significant effect on the worm dynamics, motivating the need for further research on coupled social and wireless networks.

5 From biological and social metaphors to coupled bio-socialwireless networks

The above discussion has illustrated the use of biological and social metaphors in understanding a wide variety of network systems. With the advent of mesh, vehicular, *ad hoc* and sensor networks, it is becoming increasingly important to represent the social networks (networks where entities are individuals) and their interaction with the communication networks. Important application areas include: social networking, health and body networks and infrastructure services. For instance, vehicular *ad hoc* networks (VANET) operate on the fabric of urban mobility, which depends on socio-economic aspects, among many others. As shown in Camp et al. (2002), the underlying mobility model and the structure induced by it has a significant impact on the performance of protocols on the resulting *ad hoc* wireless networks. To analyse the operation of VANETs and develop efficient protocols for them, it is crucial to study them in conjunction with realistic detailed mobility models, of the sort generated by Transportation Simulation (TRANSIMS). Health monitoring systems (see Emil et al. 2005 and references there-in) need to take this one step further and require modelling biological phenomena at an intra-organism level, in addition to social mobility. Therefore, the study of such systems necessitates coupling biological, social and wireless systems.

Our thesis is that *design and analysis of efficient next generation wireless networks* will be improved by considering the social and biological context in which these networks are

deployed. Current modelling tools do not represent the social, biological and urban context and thus are of limited use in achieving this objective. Our Simdemics based approach to wireless epidemiology that we outlined earlier provided an example of how these coupled networks can be studied using high-performance computer models. Design of markets to trade radio frequencies for optimal use of this scarce resource represents another example of the need to study coupled social, economic and wireless networks. Our group in conjunction with our collaborators, S. Das and H. Gupta at State University of New York at Stony Brook, and M. Buddhikot at Alcatel-Lucent Bell Labs have recently begun a project to develop practical modeling tools for the market-driven design of future wireless networks. Designing such systems requires knowing which radio frequencies can be reused concurrently over spatial regions without causing signal degradation; understanding the social network that generates the temporal and spatial demand patterns, and having a handle on the market mechanisms that control admission to, and distribution through, the communication network. In other words, one needs to understand the coupled social network that creates the demand and supply, the market that allocates resources and the communication network that ultimately carries the traffic.

Biographies

Christopher L. Barrett is a Professor of Computer Science and Director of the Network Dynamics and Simulation Science Laboratory, Virginia Bio-Informatics Institute. He received his PhD from California Institute of Technology in 1985. His research interests are in theoretical foundations of simulation, simulations of very large systems, computing and dynamical systems, computational problems in epidemiology, cognitive science and computationally aided reasoning.

Karthik Channakeshava is a PhD Student at the Bradley Department of Electrical and Computer Engineering. He received his MS in Computer Engineering from Virginia Tech in 2004. He worked as a Research Intern at R&D Labs of Hitachi America Ltd. in the summer of 2006. Before coming to Virginia Tech, he worked at the Center for Development of Advanced Computing (CDAC), Bangalore, India. His research interests are in performance evaluation of protocols, modelling and simulations of wireless networks and network security.

Stephen Eubank is the Deputy Director of the Network Dynamics and Simulation Sciences Laboratory at the Virginia Bioinformatics Institute at Virginia Tech. He received his BA in Physics from Swarthmore College in 1979 and his PhD in Theoretical Particle Physics from the University of Texas at Austin in 1986. His research interests are in developing advanced technology for the study of large socio-technical systems and understanding the dynamics and structure of social networks.

V.S. Anil Kumar is a Senior Research Associate in the Network Dynamics and Simulation Science Laboratory at the Virginia Bioinformatics Institute, and an Assistant Professor in the Department of Computer Science at Virginia Tech. He received his PhD in Computer Science from the Indian Institute of Science in 1999 and was a Postdoctoral Associate at the Max-Planck Institute and at the Los Alamos National Laboratory, where he became a Technical Staff Member in 2003. His current interests are at the interface of theoretical computer science and modelling and simulation of social and infrastructure systems, epidemiology, and mobile computing.

Madhav V. Marathe is a Professor of Computer Science and Deputy Director of the Network Dynamics and Simulation Science Laboratory, Virginia Bio-Informatics Institute. He obtained his BTech in 1989 in Computer Science and Engineering from the IIT, Madras, and

his PhD in 1994 in Computer Science from the University at Albany. Before coming to Virginia Tech, he was a Team Leader in the Computer and Computational Sciences Division at Los Alamos National Laboratory. His current interests are in theoretical computer science, high-performance computing, modelling large urban societal infrastructures, network and data science and computational social science.

Acknowledgments

We thank members of the Network Dynamics and Simulation Science Laboratory; our work on wireless networks described here was done in conjunction with them. The work on Cognitive Sensing and monitoring is a part of MS thesis written by Kashmira Phalak and Sameer Tupe. We also thank Jeff Reed, Tamal Bose, Ashwin Ammana, Allen Mackenzie, Martin Drozda, Stephanie Forrest for discussions on topics related to this paper. The authors were supported partially by NSF Nets Grant CNS-0626964, NSF HSD Grant SES-0729441, CDC Center of Excellence in Public Health Informatics Grant 2506055-01, NIH-NIGMS MIDAS project 5 U01 GM070694-05, DTRA CNIMS Grant HDTRA1-07-C-0113, NSF NETS CNS-0831633, and CNS-0845700.

References

- Agrawal D, El Abbadi A, Singh AK. Consistency and orderability: semantics-based correctness criteria for databases. *ACM Transactions on Database Systems*. 1993; Vol. 18(No. 3):460–486.
- Akyildiz I, Lee W-Y, Vuran M, Mohanty S. A survey on spectrum management in cognitive radio networks. *Communications Magazine, IEEE*. 2008; Vol. 46(No. 4):40–48.
- Bailey, NTJ. *The Mathematical Theory of Infectious Diseases and its Applications*. London: Griffin; 1975.
- Barrett, CL.; Bush, BW.; Kopp, S.; Mortveit, HS.; Reidys, CM. Sequential dynamical systems and applications to simulations. *SS '00: Proceedings of the 33rd Annual Simulation Symposium*; Washington, DC, USA: IEEE Computer Society; 2000. p. 245
- Barrett CL, Eidenbenz SJ, Kroc L, Marathe MV, Smith JP. Parametric probabilistic routing in sensor networks. *Mobile Networks and Applications*. 2005a; Vol. 10:529–544.
- Barrett, CL.; Eubank, S.; Marathe, M. *Interactive Computation: The New Paradigm*. Springer Berlin Heidelberg: 2005b. Modeling and simulation of large biological, information and socio-technical systems: an interaction based approach; p. 353-392.
- Barrett, CL.; Bisset, KR.; Eubank, SG.; Feng, X.; Marathe, MV. EpiSimdemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks. *SC '08: Proceedings of the 2008 ACM/IEEE conference on Supercomputing*; IEEE Press; 2008. p. 1-12.
- Beckman RJ, Baggerly KA, McKay MD. Creating synthetic base-line populations. *Transportation Research Part A – Policy and Practice*. 1996
- Bonabeau, E.; Dorigo, M.; Theraulaz, G. *Swarm Intelligence From Natural to Artificial Systems*. 1st ed.. New York, NY, USA: USA: Oxford University Press, Inc.; 1999. ISBN: 0195131592
- Bonnet, P.; Gehrke, J.; Seshadri, P. *MDM'01*. London, UK: Springer-Verlag; 2001. Towards sensor database systems; p. 3-14. ISBN 3-540-41454-1
- Bouazizi, I. *ICPPW '02*. Washington, DC, USA: IEEE Computer Society; 2002. ARA - The Ant-Colony Based Routing Algorithm for MANETs. p. 79 ISBN 0-7695-1680-7. doi <http://dx.doi.org/10.1109/ICPPW.2002.1039715>
- Bowman JL, Ben-Akiva ME. Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research. Part A, Policy and Practice*. 2001; Vol. 35(No. 1):1–28.
- Brent R, Bruck J. 2020 Computing: can computers help to explain biology? *Nature*. 2006; Vol. 440:416–417. [PubMed: 16554784]
- Camp T, Boleng J, Davies V. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing*. 2002; Vol. 2(No. 5):483–502.
- Carlson J, Doyle J. Complexity and robustness. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; Vol. 99 Suppl. 1:2538–2545. [PubMed: 11875207]
- CERT. Available at <http://www.cert.org/>

- Channakeshava, K.; Chafekar, D.; Bisset, K.; Kumar, VSA.; Marathe, M. EpiNet: a simulation framework to study the spread of malware in wireless networks. Simutools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques; 2009a. p. 1-10.
- Channakeshava, K.; Kumar, VSA.; Marathe, MV. Technical report. Virginia Tech; 2009b. Strategies for controlling mobile epidemics: insights from using realistic mobile networks.
- Chen, T.; Zhagn, H.; Katz, MD.; Zhou, Z. Swarm intelligence based dynamic control channel assignment in CogMesh. Communications Workshops, 2008; IEEE International Conferenes on ICC Workshops '08; 2008. p. 123-128.
- Cohen F. Computer viruses theory and experiments. Computers and Security. 1987; Vol. 6:22-35.
- Demers, A.; Greene, D.; Hauser, C.; Irish, W.; Larson, J.; Shenker, S.; Sturgis, H.; Swinehart, D.; Terry, D. PODC'87. NewYork, NY, USA: ACM; 1987. Epidemic algorithms for replicated database maintenance; p. 1-12. ISBN 0-89791-239-4. doi:http://doi.acm.org/10.1145/41840.41841
- Diekmann, O.; Heesterbeek, JAP. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. John Wiley; 2000.
- Dietz K, Heesterbeek J. Daniel bernoulli's epidemiological model revisited. Journal of Mathematical Biosciences. 2002; Vol. 180:1-21.
- Doerr, C.; Sicker, DC.; Grunwald, D. WICON '07. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering); 2007. What a cognitive radio network could learn from a school of fish; p. 1-10. ISBN 978-963-9799-12-7
- Drozda, M.; Szczerbicka, H. Artificial immune systems: survey and applications in ad hoc wireless networks. Proceedings 2006 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS 2006); 2006. p. 485-492.
- Elveback LR, Fox JP, Ackerman E, Langworthy A, Boyd M, Gatewood L. An influenza simulation model for immunization studies. American Journal of Epidemiology. 1976; Vol. 103(No. 2):152-165. [PubMed: 814808]
- Emil J, Aleksandar M, Chris O, de Groen PC. A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. Journal of NeuroEngineering and Rehabilitation. 2005; Vol. 2(No. 1)
- Epidemic Routing Bibliography. Available at:<http://roland.grc.nasa.gov/weddy/biblio/epidemic/>
- Eubank S, Guclu H, Kumar VSA, Marathe MV, Srinivasan A, Toroczkai Z, Wang N. Modeling disease outbreaks in realistic urban social networks. Nature. 2004; Vol. 429:180-184. [PubMed: 15141212]
- Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, Burke DS. Strategies for mitigating an influenza pandemic. Nature. 2006; Vol. 442(No. 7101):448-452. [PubMed: 16642006]
- Ferrie P, Szor P, Stanev R, Mouritzen R. Security responses: Symbos.cabir. Symantec Corporation. 2004
- Fleizach C, Liljenstam M, et al. Can you infect me now?: Malware propagation in mobile phone networks. WORM '07: Proceedings of the 2007 ACM Workshop on Recurring Malcode. 2007:61-68.
- Fodor, J. The Modularity of Mind: An Essay on Faculty Psychology. Cambridge, MA, USA: The MIT Press; 1983. ISBN: 0262560259
- Fodor, J. The Mind Doesnt Work that Way: The Scope and Limits of Cognitive Psychology. 1st ed.. Cambridge, MA, USA: The MIT Press; 2000. ISBN: 0262561468
- Forrest, S.; Perelson, AS.; Allen, L.; Cherukuri, R. Self-nonsel self discrimination in a computer. SP '94: Proceedings of the 1994 IEEE Symposium on Security and Privacy; IEEE Computer Society Press; 1994. p. 202-212.
- Forrest, S.; Hofmeyr, SA.; Somayaji, A.; Longstaff, TA. A sense of self for unix processes. SP '96: Proceedings of the 1996 IEEE Symposium of Security and Privacy; IEEE Computer Society Press; 1996. p. 120-128.
- Ganesan, D.; Krishnamachari, B.; Woo, A.; Culler, D.; Estrin, D.; Wicker, S. An Empirical Study of Epidemic Algorithms in Large Scale Multihop Wireless Networks. Technical report. 2002.
- Heberlein L, Dias G, Levitt K, Mukherjee B, Wood J, Wolber D. A network security monitor. RISP'90. 1990:296-304.

- Hofmeyr, S.A.; Forrest, S. GECCO '99. Morgan Kaufmanns; 1999. Immunity by design: an artificial immune system; p. 1289-1296.
- Holliday J, Steinke R, Agrawal D, ElAbbadi A. Epidemic algorithms for replicated databases. IEEE Transactions on Knowledge and Data Engineering. 2003; Vol. 15(No. 5):1218–1238. ISSN 1041–4347. doi <http://dx.doi.org/10.1109/TKDE.2003.1232274>.
- Hull, B.; Bychkovsky, V.; Zhang, Y.; Chen, K.; Goraczko, M.; Miu, A.; Shih, E.; Balakrishnan, H.; Madden, S. SenSys '06. New York, NY, USA: ACM; 2006. CarTel: a distributed mobile sensor computing system; p. 125-138. ISBN 1-59593-343-3. doi:<http://doi.acm.org/10.1145/1182807.1182821>
- Intanagonwiwat, C.; Govindan, R.; Estrin, D. MobiCom '00. ACM; 2000. Directed diffusion: a scalable and robust communication paradigm for sensor networks; p. 56-67.
- Janeway CA. How the immune system protects the host from infection. Microbes and Infection. 2001; Vol. 3(No. 13):1167–1171. [PubMed: 11709297]
- Kahneman, D. Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice. In: Frangsmyr, T., editor. Les Prix Nobel 2002. 2003.
- Kairer, R. Smartphones Showed Strong Growth in 2007. 2008. Available at:<http://www.palminfocenter.com/news/9617/smartphones-showed-strong-growth-in-2007>
- Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics-I,II,III. Bulletin of Mathematical Biology. 1991; Vol. 53(No. 1–2):33–118. [PubMed: 2059741]
- Kim, J.; Bentley, P.J. GECCO '01. Morgan Kaufmann; 2001. An evaluation of negative selection in an artificial immune system for network intrusion detection; p. 1330-1337.
- Kleinberg J. Computing: the wireless epidemic. Nature. 2007; Vol. 449:287–288. [PubMed: 17882205]
- Lactaotao M. Security Information: Virus Encyclopedia: Symbos comwar.a. Trend Micro Incorporated. 2005
- Lewis, R.L. Cognitive theory, SOAR. In: Smelser, N.J.; Baltes, P.B., editors. International Encyclopedia of the Social and Behavioral Sciences. Oxford: Pergamon; 2001. p. 2178-2183.
- Li, M.; Ganesan, D.; Shenoy, P. NSDI'06. Berkeley, CA, USA: USENIX Association; 2006. PRESTO: Feedback-driven Data Management in Sensor Networks; p. 23-23.
- Longini J, I M, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DA, Halloran ME. Containing pandemic influenza at the source. Science. 2005; Vol. 309(No. 5737):1083–1087. [PubMed: 16079251]
- Macdonald G. The analysis of equilibrium in malaria. Trop Dis Bull. 1952; Vol. 49(No. 9):813–829. [PubMed: 12995455]
- MacKenzie A, Reed J, Athanas P, Bostian C, Buehrer R, DaSilva L, Ellingson S, Hou Y, Hsiao M, Park J-M, Patterson C, Raman S, daSilva C. Cognitive radio and networking research at Virginia Tech. Proceedings of the IEEE. 2009; Vol. 97(No. 4):660–688.
- Madden S, Franklin MJ, Hellerstein JM, Hong W. Tag: a tiny aggregation service for ad-hoc sensor networks. SIGOPS Operating Systems Reviews. 2002; Vol. 36(SI):131–146. ISSN 0163-5980. doi:<http://doi.acm.org/10.1145/844128.844142>.
- Moore D, Shannon C, Brown DJ, Voelker GM, Savage S. Inferring internet denial-of-service activity. ACM Transactions on Computer Systems'. 2006; Vol. 24(No. 2):115–139.
- Mitola, J, III. PhD thesis. Stockholm, Sweden: Royal Institute of Technology (KTH); 2000. Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio.
- Mitola, J, III. PhD thesis. Stockholm, Sweden: Royal Institute of Technology (KTH); 2000. Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio.
- Nature Special Issue. Future of Computing. Nature Special Issue. 2006
- NSF Workshop. NSF Workshop on Molecular Communication/Biological Communications technology. 2009 February. <http://netresearch.ics.uci.edu/mc/nsfws08>.
- Phalak, K. Master's Thesis. Virginia Tech; 2006. Cognitively-inspired architecture for wireless sensor networks: a model driven approach for data integration in a traffic monitoring system'.

- Renk T, Kloeck C, Burgkhardt D, Jondral FK, Grandblaise D, Gault S, Dunat J-C. Bio-inspired algorithms for dynamic resource allocation in cognitive wireless networks. *Mobile Networks and Applications*. 2008; Vol. 13(No. 5):431–441.
- Schoonderwoerd R, Bruten JL, Holland OE, Rothkrantz LJM. Ant-based load balancing in telecommunications networks. *Adaptive Behaviour*. 1996; Vol. 5(No. 2):169–207. ISSN 1059-7123. doi:<http://dx.doi.org/10.1177/105971239700500203>.
- SOAR Available at: Online www.soartech.com/projects/16%20SoarOverviewWP.pdf
- Su J, Chan KKW, Miklas AG. A preliminary investigation of worm infections in a bluetooth environment. *WORM '06: Proceedings of the 2006 ACM Workshop on Recurring Malcode*. 2006:9–16.
- Teng, HS.; Chen, K.; Lu, SC-Y. *AIA '90*. IEEE Press; 1990. Security audit trail analysis using inductively generated predictive rules; p. 24-29.
- TRANsportation SIMulation. Available at: TRANSIMS <http://transims.tsasa.lanl.gov>
- Tupe, S. Master's Thesis. Virginia Tech; 2006. A cognitively inspired architecture for wireless sensor networks: a service middleware for a traffic monitoring system.
- UCSD Internet Epidemiology. Available at:<http://ccied.ucsd.edu/>
- Yan, G.; Eidenbenz, S. *ACSAC '06*. Washington, DC USA: IEEE Computer Society; 2006. Bluetooth worms: models, dynamics, and defense implications; p. 245-256.doi <http://dx.doi.org/10.1109/ACSAC.2006.18>

Figure 1.
CoSMo architecture (see online version for colours)

Figure 2.
Bluetooth worm manifestation diagram (see online version for colours)

Figure 3.
Infection CDFs of the study with and without static interventions (see online version for colours)

Figure 4.
Effect of T_{idle} on the infection propagation (see online version for colours)

Table 1

Summary of cognitively inspired features in the CoSMo architecture using SOAR cognitive architecture as a reference

Cognitive features based on SOAR	CoSMo components and their functionality
Recognition and categorisation	State manager stores the rules and decisions for any query. Any new query is matched with the rule based on the attributes and conditions captured by the rule. New plans and rules are learned when no query is recognised
Decision making and choice	State manager decides, based on the query, whether the query can be answered with available information with confidence. If querying the sensor is required, then the data manager determines the right sensor to be queried
Perception	Aggregates data obtained from heterogeneous sensors, maintains information regarding confidence levels of each sensor and for later queries, decides on a subset of the sensors based on this measure
Prediction and monitoring	The traffic simulator provides CoSMo the capability to predict future situations and values, and any new data are archived for later use
Planning	The state manager plans a course of action based on a query taking into account the query's requirements. It implements this plan by utilising the existing modules
Remembering, reflection and learning	The state manager stores newly learnt rules and decisions and case-base learning is implemented in this module. The data model uses the confidence level of sensors to retrieve information based on the confidence and kind of data required by any query
Multiple, simultaneous goals	The CoSMo architecture has support for selecting balance between accuracy, speed and resource usage while responding to the user queries
Learning from experimentation	With every query the data module learns about the sensors, the best suited for certain data and their reliability. This information is used for later queries and also makes query decisions based on this learning.

Table 2

Bluetooth protocol and worm parameters

<i>Terms</i>	<i>Definition</i>
$N_{\text{inq}}^{\text{to}}$	Maximum inquiry responses expected by the inquiring device during an inquiry process
N_{inq}	Number of inquiry responses actually received by the inquiring device
$T_{\text{inq}}^{\text{to}}$	Inquiry timeout value (inquiring node receives 0 or N_{inq} responses at the end of $T_{\text{inq}}^{\text{to}}$)
T_{inq}	time taken for the inquiry phase
T_{inq}^n	inquiry time for the n th inquiry request
T_{page}	time taken for the page response
$T_{\text{page}}^{\text{to}}$	page request time out value
T_{inf}	time taken for the infection
T_{idle}	idle time between infection cycles