

Principles of Genomic Robustness Inspire Fault-Tolerant WSN Topologies: a Network Science Based Case Study

Preetam Ghosh*, Michael Mayo[†], Vijender Chaitankar*, Tanwir Habib[†], Ed Perkins[†] and Sajal K Das[‡]

*School of Computing, The University of Southern Mississippi, Hattiesburg, MS 39406.

Email: preetam.ghosh@usm.edu, vijender.chaitankar@eagles.usm.edu

[†]Environmental Laboratory, US Army Engineer Research and Development Center, Vicksburg, MS 39180.

Email: Michael.L.Mayo@usace.army.mil, Tanwir.Habib@usace.army.mil, Edward.J.Perkins@usace.army.mil

[‡]Dep. Comp. Sci. & Eng., The University of Texas at Arlington, Arlington, TX 76019. Email: das@uta.edu

Abstract—Wireless sensor networks (WSNs) are frameworks for modern pervasive computing infrastructures, and are often subject to operational difficulties, such as the inability to effectively mitigate signal noise or sensor failure. Natural systems, such as gene regulatory networks (GRNs), participate in similar information transport and are often subject to similar operational disruptions (noise, damage, etc.). Moreover, they self-adapt to maintain system function under adverse conditions. Using a PBN-type model valid in the operational and functional overlap between GRNs and WSNs, we study how attractors in the GRN—the target state of an evolving network—behave under selective gene or sensor failure. For “larger” networks, attractors are “robust”, in the sense that gene failures (or selective sensor failures in the WSN) conditionally increase their total number; the “distance” between initial states and their attractors (interpreted as the end-to-end packet delay) simultaneously decreases. Moreover, the number of attractors is conserved if the receiving sensor returns packets to the transmitting node; however, the distance to the attractors increases under similar conditions and sensor failures. Interpreting network state-transitions as packet transmission scenarios may allow for trade-offs between network topology and attractor robustness to be exploited to design novel fault-tolerant routing protocols, or other damage-mitigation strategies.

Keywords—gene regulatory network; attractor; probabilistic boolean network; robustness; network science;

I. INTRODUCTION

Wireless sensor networks (WSNs) are platforms for information transport, possessing sensor protocols that adapt to dynamic changes in channel traffic and noise, sensor failures, and other disruptions, to maximize the probability of signal receipt while minimizing the end-to-end packet delay. A conceptually similar and self-adapting system can be found in gene regulatory networks (GRNs), wherein the interactions between genes couple to cell-signaling networks by synthesizing enzymes and other biomolecules. Genes and sensors possess intriguing operational parallels: signaling molecules synthesized by genes, such as microRNAs, can influence whether other genes are triggered to produce their associated RNA transcript, while sensors decide to interact according to some on-board routing protocol, forwarding information packets to the appropriate neighbor sensor. Both

genes and sensors possess on-board “rules” that determine how incoming signals are processed, and is either encoded in the nucleotide sequence of the gene or in the routing protocol of the sensor.

Here we propose that the functional response of network “attractors”—the terminal states of a network’s state-transition diagram [1]—capture the “robustness” of gene regulatory networks under certain conditions of stress, namely under strong inhibition of gene function. Based on our analysis of the attractor dynamics under this stress, and the generality of the model system we consider, it can be argued that routing protocols of similarly described WSNs can be engineered to provide analogous “robustness” of signal transmission under selective sensor failures.

While others have considered the problem of adaptive routing in WSNs, where, for example, messages may be dropped due to channel errors, node failures or congestion [2], reliable packet transmission remains a current area of active research. While current protocols, such as the MAC protocol, ensure reliable multi-hop data forwarding, NS2-based simulation studies have demonstrated that others, such as IEEE 802.15.4, can have a very low reliability in terms of packet transmission—especially when power management is enabled for energy conservation [3]. Other packet-forwarding approaches have been proposed as transport protocols specifically designed for WSNs [4], and their performance is typically evaluated by simulation.

The operation of sensor networks under real-world conditions, however, often introduces complexities not entirely captured through simulation; a “hands-on” approach is required for evaluating data forwarding solutions. One example is given by the improperly tuned interactions between different layers of the networking stack, leading to very low packet transmission rates [5]. Other solutions focus on an experimental validation of routing strategies for improving general communication reliability [6], [7]. Generally, however, a routing protocol is supported by a link-quality estimation mechanism that evaluates the stability and reliability of routes [8], [9], and locates alternate routes when presented with sensor failures [10]. This is also the

case for the default protocols available in the common sensor platforms: the Motes [11] and the Sunspot [12].

This paper is organized as follows: in Section II we present the concept of attractor-based genomic robustness and discuss existing frameworks for their study; in Section III we define a one-to-one mapping between genes and sensors, and present a model system valid under operating conditions shared by GRNs and WSNs; in Section IV, we study the dynamical aspects of the model’s attractors; in Section V, we analyze the attractor state-space; we discuss the extension of our results to networks of arbitrary size in Section VI; in Section VII we outline promising future research directions.

II. MODELING ATTRACTORS

If the genes of an interacting network occupy one of only two possible states (*on* or *off*), and if transitions may occur between the corresponding network states by selective gene activation/deactivation (*on* to *off* and vice versa), then the network state in the long-time limit is called an attractor [1]. In general, there can be many attractors for a given network, and it may cycle through some of them indefinitely. The network’s state-transition diagram is partitioned into level sets l_j , that include all of the states that terminate with an attractor state in exactly j transitions. Such attractor cycles are mutually disjoint, and the partition class corresponding to an attractor cycle is called the basin of the cycle, with any transient state belonging to a unique basin and level.

The dynamical behavior of a GRN can be represented by its state transitions-along with some degree of fault-tolerance-in which its genes are continually expressed and repressed, and ultimately terminate with attractors and cycles. To quantify these states, two formalisms have been proposed: Boolean and Probabilistic Boolean networks.

A. Boolean Networks

Boolean Network (BN) models were first introduced to study the expression patterns of randomly constructed “genetic nets” [1]. In a BN, a gene can be in either of two possible states, ON/OFF, where the state (or expression level) of each gene is functionally related to that of other genes using logical rules, or “functions”. A BN is defined by a set of nodes, $V = \{x_1, x_2, \dots, x_N\}$, corresponding to a set of genes, together with a list of Boolean functions $F = \{f_1, f_2, \dots, f_N\}$, one for each node, that determine the node-node interactions. The state of the entire network is labeled in the form $x_1x_2\dots x_N$ (e.g. 010...1), wherein the individual gene-states are either 1 (ON) or 0 (OFF). Transitions between them are triggered by adjusting a gene’s state, and are represented by a directed graph (Fig 1A).

As an example, consider a Boolean network with only three genes: $V = \{x_1, x_2, x_3\}$, wherein a Boolean function is associated to each of them according to the truth table shown in Fig 1A. Out of potentially $2^3 = 8$ attractors, this

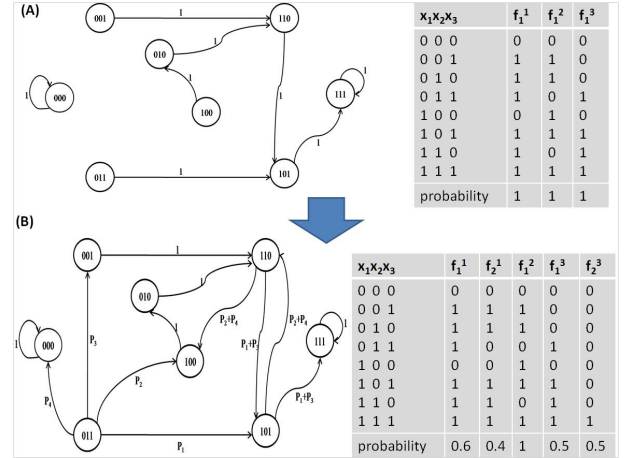


Figure 1. BN vs PBN dynamics for GRNs. The attractor nodes are 000 and 111 in both cases. The truth tables for the BN and PBN are shown alongside

BN has exactly two: 000 and 111. Here, a single Boolean function describes the next state of any gene based upon the wiring rules of the corresponding GRN; the probability to choose these particular functions is always 1.

B. Probabilistic Boolean Networks

The Probabilistic Boolean Network (PBN) formalism was developed out of the need for a network to cope with uncertainty; one challenge is to identify the correct regulatory relationships for a target gene [13], while also preserving the rule-based properties of BNs.

More specifically, PBNs accommodate more than one possible function for each node in the network. A set of functions $F_i = \{f_j^i : j = 1, 2, \dots, l_i\}$ is associated with each gene x_i , where each function associated to a given gene, f_j^i , determines the expression value x_i of that gene, and l_i is the number of possible functions for that particular gene. In Fig 1B, the values for the first and third gene, x_1 and x_3 , are, respectively, determined by two possible functions; i.e., f_1^1 and f_2^1 for x_1 (with others similarly described).

A realization of the PBN at any given instant is determined by a vector of Boolean functions, where the i^{th} element of the vector contains the predictor selected for the gene x_i . This vector function maps one network state onto another, and is referred to as a multiple-output Boolean function. Since each of the N possible realizations (for N genes) is a standard BN, valid for only one time step, a PBN describes the evolution of an ensemble of alternative BN states for each gene. The probability for the i^{th} BN to be selected is written in terms of the individual selection probabilities, shown in the bottom row of the truth table within Fig 1B (see [13] for further details).

III. WIRELESS SENSOR NETWORKS

Transmission inconsistencies often plague WSNs where they suffer from signal disruptions due to sensor failure, or

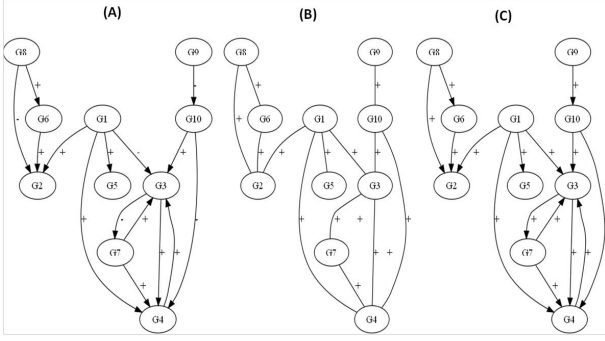


Figure 2. Attractor state space evolution in GRNs, and their equivalent WSNs: (A) A 10-node GRN from the yeast cell cycle; (B) Equivalent 10-node WSN from (A) where each edge is undirected as two sensor nodes within transmission radius of each other (assuming homogeneous nodes) can send packets to each other and there is no down-regulation involved; (C) Restricted WSN from (A) with no down-regulation and an edge from G_i to G_j does not imply the existence of an edge between G_j to G_i . This network abstracts the case for heterogeneous sensors and also the scenario of dynamic routing where a sensor only transmits packets along its outgoing edges although it is capable of sending packets back to the transmitting nodes (incoming edges).

from the absence of routing protocols that are sufficiently insensitive to local as well as global network conditions. As GRNs are subject to operating conditions conceptually similar to WSNs, the evolution of their attractor state-space can be equivalently described within a WSN context according to the following rules: (1) for every gene within the GRN, we replace it with a sensor; (2) sensor interactions are restricted to existing edges between corresponding genes in the GRN. In the following sections, we discuss the implications of these gene-to-sensor mapping rules to WSN packet-transmission characteristics.

A. Packet Transmission along Wireless Channels

The first rule preserves node structure from GRN to WSN, while the second rule defines the interactions between sensors in terms of packet transmission, with respect to the GRN architecture. According to the second rule, a sensor (or equivalently, a gene) at time point t can either send or not send a packet to another sensor—one that either promotes or suppresses another gene—at time point $t+1$ with a certain probability (as a function of the network state, packet error rate, or the distance between corresponding sensors) and destination node break-down probability. Thus, a gene is considered to be in state “1/0”, based on whether it is up- or down-regulated; a sensor will have states “1/0” depending on whether it has received a packet or not.

B. Wireless Sensor Topology

The physical signaling structure of sensors within the WSN must be adapted to reflect the interplay between genes in the GRN. If gene G_1 up-regulates G_2 , then the equivalent interaction in the WSN is that sensor G_1 sends a packet to G_2 according to some probability distribution defined by the originating gene-gene interactions.

For homogeneous sensor nodes, each up-regulation edge (denoted by +) in a GRN is replaced by a bi-directional edge; if we allow sensor G_1 to send a packet to G_2 , then G_2 should also be able to send a packet to G_1 (see Fig 2B). For heterogeneous sensor nodes, however, it is not necessary that both G_1 and G_2 possess the same transmission radii, giving a directed edge from G_1 to G_2 and not vice versa (such a network is shown in Fig 2C).

C. Packet Collisions

According to the second rule, biological down-regulation must also be mapped from a GRN to a WSN. If G_1 down-regulates G_2 , then sensor G_1 prohibits the receipt of packets at sensor G_2 . This phenomenon is interpreted as packet “collisions” in the WSN; data packets simultaneously received by G_2 from both G_1 and G_3 result in effectively zero packets received by G_2 .

D. Attractors and Packet Transmission Scenarios

Using the BN formalism, each attractor will record the transmission history of the packet. Fig 1A depicts two attractor states of an exemplary BN 000 and 111. Here, the evolution of the initial state 001 to an attractor is analogous to one in which gene x_3 transmits a signal to its neighbors; the network then evolves according to a sequence of state-transitions until it terminates in attractor 111. Ultimately, all three genes x_1, x_2 and x_3 of Fig 1A will receive a signal originating from gene x_3 . Similarly, if gene x_2 receives a signal from all the other ones (equivalent to setting node x_2 as the Cluster Head in the WSN), then each of the three initial states of 001, 010 and 100 will ultimately reach the attractor 111 in 3, 3 and 4 state transitions, respectively.

The PBN formulation reproduces the attractors 000 and 111 of this 3-node system, shown in Fig 1B. In both the BN and PBN models of this network, the receipt of a packet by any node is independent of the initial conditions, ensuring that every node participates in packet transport. The PBN formalism allows us to incorporate the packet-drop rates due to “noisy” channel characteristics into the packet transmission state-space by redefining the perturbation probability from [13] for each edge in the WSN. In the next section, we will present a case study to illustrate the evolution of attractors under different network conditions.

IV. THE MODEL

To study attractors under different network conditions, we used GeneNetWeaver [14] to derive a 10-node sub-GRN from the yeast *Saccharomyces cerevisiae* (Fig 2A), from a total of 4441 genes and 12873 interactions. Using the mapping rules defined in Section III, we established a corresponding WSN for homogeneous sensor nodes (Fig 2B); each sensor is “equivalent”, in that they possess similar transmission radii. Each directional edge from the GRN (Fig 2A) is replaced by an undirected one (Fig 2B), assuming

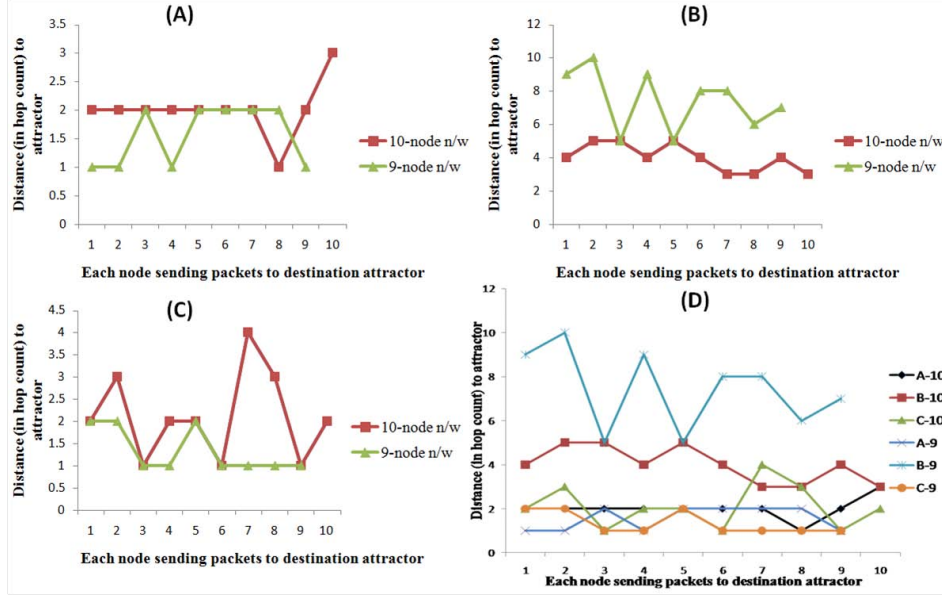


Figure 3. Distances (in units of “hop counts”) to corresponding attractor states for the networks shown in Fig 2 from the sample starting states that mimic each node sending a packet: (A) distances reported for the original 10-node network in Fig 2A (A-10) and that with the corresponding 9-node network with G_3 removed (A-9); (B) distances reported for the original 10-node network in Fig 2B (B-10) and that with the corresponding 9-node network with G_3 removed (B-9); (C) distances reported for the original 10-node network in Fig 2C (C-10) and that with the corresponding 9-node network with G_3 removed (C-9); (D) Combined plot showing the relationship of distances for each of the different networks considered.

that two directly connected nodes are capable of mutual communication. Only up-regulation edges exist in this WSN mapping, because we restrict sensors from destroying any received packets (i.e., a collision-free time slot allocation algorithm is assumed). Ideally, if both sensors G_1 and G_6 transmit packets to G_2 , and are received in the same time slot, then a “down-regulation edge” should be assigned from G_6 to G_2 , and an up-regulation edge is assigned connecting G_1 to G_2 . Packet collision scenarios will be considered elsewhere.

The WSN in Fig 2C allows for the same directed edges (without the $+/-$ signs) assigned in the GRN from Fig 2A, denoting two scenarios: (i) the transmission radii of the nodes are unequal (a heterogeneous WSN); (ii) the transmission of packets from a particular sensor node are restricted to only a subset of its neighbors (following some routing protocol). The resulting network (Fig 2C) therefore contains only directed and up-regulated edges.

A requirement of the PBN-based framework is that it gives the probabilities for an associated gene-activating Boolean function to be realized. To map this GRN onto a WSN-based scenario, a power set must be defined, allowing for one logical Boolean function to be associated to each possible subset of the neighboring sensors. For a sensor with m neighbors there are therefore 2^m possible Boolean functions, and the PBN-based dynamics allow for the selection of the “best” neighbor set (i.e., the one with highest probability) for any sensor in the network. Here, each node’s Boolean function is an “OR” function, relating all the sensors from

a node’s particular subset of neighbors, as if any of the incident sensors send a packet, the node may receive it and transit to a ‘1’ state.

In the toy model we consider, *one* Boolean function is associated with each node, accounting for all incoming edges that set it to $1/0$; the probability for a gene’s state to be altered is set to zero, effectively reducing the PBN model into a BN one. This has the advantage of restricting the network so that each sensor potentially forwards data packets to all of its neighboring nodes.

V. RESULTS

A. The Attractor Space

The undirected network (corresponding to $B-9$ and $B-10$, see figure caption for details) results in the minimum number of attractors (two). Ideally, if all sensor interactions are undirected, then only two attractors are possible (either all 0’s or 1’s), because all sensors eventually participate in packet transmission.

A 9-node network was created by removing gene G_3 from the original 10-node network; subsequently, all of its incoming and outgoing edges were removed. With directed edges, the number of attractors is higher, with 10 and 8 attractors for the 10-node networks shown in Fig 2A and C, respectively, out of potentially $2^{10} = 1024$ attractors. The 9-node networks corresponding to these cases (with G_3 removed) resulted in an increase in the number of attractors (16 for both cases), suggesting a decrease in the robustness of the network facilitated by node deletion. Fig 4 reports

these observations. Comprehensive simulation based studies are required to better understand the trade-offs between attractor cardinality and network structure.

Clearly, the number of attractors must increase only locally with node deletion. On one hand, if a network contains N nodes, then in an ordered sequence of single node “deletions,” the removal of the $N - 1^{st}$ node gives exactly one nontrivial attractor. On the other hand, removing the last node gives the empty set. Identifying the critical node, or number of nodes, that maximize the number of attractors (with respect to the initial network) provides a metric for the system’s robustness to node removal.

Figure 3 shows that the *path length* required to reach an attractor decreases with the number of attractors in the network. For example, the 10-node GRN, as opposed to its 9-node variant (Fig. 2A), maps to a WSN (Fig. 2B) providing the fewest attractors of both WSNs studied, yet gives the maximum path length to its attractor. Removing G_3 generates the 9-node network, resulting in a decrease in the distance to attractors for the following variants (see figure caption for details): from $A - 10$ to $A - 9$ and from $C - 10$ to $C - 9$; whereas the distance increases from $B - 10$ to $B - 9$. This is because the deletion of G_3 increases the number of attractors, and hence the *sink* nodes in $A - 9$ and $C - 9$, resulting in shorter distances, whereas, the number of attractors stay the same in $B - 9$ resulting in larger distances.

A node’s Boolean functions (in terms of the neighbor set it uses to forward packets) are central elements involved in the regulation of both the number of a network’s attractors and the end-to-end delay in packet transmission (which is directly proportional to the path length). If a sensor is allowed to forward packets to all the other nodes within its transmission radius, then such *packet flooding* results in maximum reliability; every other node will ultimately receive the packet, albeit after a larger end-to-end delay (assuming a realistic packet drop rate due to channel errors).

B. Scalability issues and Incremental solutions

Using BNs and PBNs to study packet transmission scenarios invites a combinatorial “explosion” in the state-space, leaving only “smaller” networks tractable to practical study—for a 100-node WSN, the corresponding BN/PBN-type model will have 2^{100} possible states. It is therefore important to discover ways to scale-up results from smaller WSNs to larger ones. Here, we discuss two such potential solutions.

If WSN state-transitions, recording the transport of packets across the network, are viewed as Markov chains, then it may be useful to interpret the influence of the Boolean functions on isotropic packet “diffusion” as diffusional bias. Here, a packet initially hops to neighbor nodes with equal probability, but such probabilities are weighted and renormalized by the Boolean functions controlling the routing of packets at the nodes. An advantage of this description is that its inversion would allow one to reconstruct an appropriate

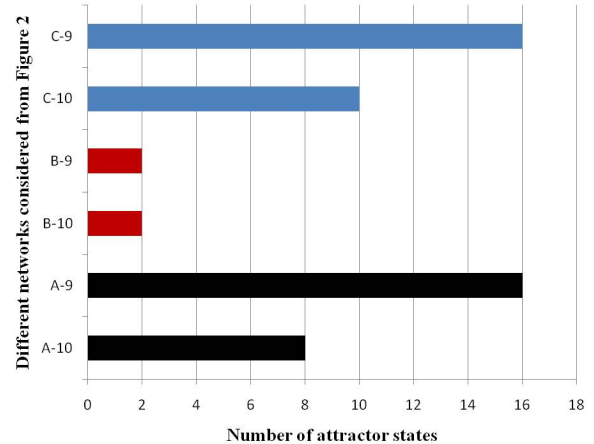


Figure 4. Number of attractor states for the different network scenarios considered in the example

Boolean function for a node (though the inversion may not be unique). A subset of state-transitions terminating with certain attractors, taken as “boundary conditions” in the state-evolution, might be identified by placing bounds on desirable network properties, such as the end-to-end packet delay, providing a mechanism to “prune” wasteful portions of the state-transition diagrams.

Another potential solution involves the “growth” of a WSN according to some protocol, beginning from a much smaller, but well understood, network, termed a *generator*. Beginning from the generator, growth rules determine which nodes receive new connections. For example, using the 10-node network previously discussed, if 0111010101 is an attractor in $A - 10$, the problem is to identify the rules to create $A - 11$ such that one of its attractors is given by 10111010101. Such strategies may also be used to incrementally alter the topology of static WSNs with less overhead, guaranteeing that “larger” networks adopt the robust properties of the “smaller” ones.

C. Topology aware sensor deployment strategies

The transformation rules presented in this paper may potentially be used to identify context-specific sensor deployment algorithms through an examination of the topological properties of an appropriate GRN, taking cues from its interaction structure to maximize the sensor coverage area. GRN-inspired WSN topologies only provide wiring rules between sensors—new sensor deployment algorithms must be designed that maximize the coverage area while simultaneously preserving these wiring rules.

VI. CONCLUSION

We introduced two rules that, when applied to a GRN, map its attractor state properties onto an equivalent WSN. We applied these rules to study the attractor properties of

a toy 10 node network, where we showed that, in a limited context, the packet-transmission characteristics across a WSN, suffering from sensor and communication failure, is deeply connected with the availability of its attractors.

Interestingly, the number of a network's attractors does not monotonically decrease with node removal. In fact, it is possible for the size of the attractor space to locally *increase* with selective removal of the network's nodes. Further work is needed to identify and understand the peculiar relationship between attractor evolution and system stress, for any practical development of WSN topology and routing protocols.

Down-regulation between genes determines the packet collision properties of the equivalent WSN. Understanding how the attractor space evolves when several sensors transmit packets in the same time slot, under a finite channel error rate and sensor breakdown probability, may lead to designs for novel time-slot allocation and synchronization schemes.

Although preliminary results look promising, much work is needed to fully integrate the self-regulation properties of GRNs with the fault-tolerance required of any sensor network expected to successfully operate in practical, real-world applications.

ACKNOWLEDGMENT

This work was supported by the US Army Environmental Quality Research program, MS INBRE funded by the National Center for Research Resources, NIH (award# P20RR016476), NSF-0903787 and NSF-1049661. Permission was granted by the Chief of Engineers to publish this information.

REFERENCES

- [1] S. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," *J. Theor. Biol.*, vol. 22, pp. 437–467, 1969.
- [2] J. Li and P. Mohapatra, "Analytical modeling and mitigation techniques for the energy hole problem in sensor networks," *Pervasive Mobile Computing*, vol. 3, pp. 233–254, 2007.
- [3] G. Anastasi, M. D. Francesco, M. Conti, and V. Neri, "Reliability and energy efficiency in multi-hop IEEE 802.15.4/zigbee wireless sensor networks," in *Proc. of the 15th IEEE Symposium on Computers and Communications (ISCC'10)*, Riccione, Italy, June 2010, pp. 336–341.
- [4] C. Wang, K. Sohraby, B. Li, M. Daneshmand, and Y. Hu, "A survey of transport protocols for wireless sensor networks," *IEEE Network*, vol. 20, pp. 34–40, 2006.
- [5] K. Rome, H. Karl, and F. Mattern, Eds., *Wireless Sensor Networks: Third European Workshop, EWSN 2006*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer, 2006, vol. 3868.
- [6] S. Kim, R. Fonseca, and D. Culler, "Reliable transfer on wireless sensor networks," in *Proc. of IEEE SECON 2004*, Santa Clara, CA, October 2004, pp. 449–459.
- [7] A. Woo, T. Tong, and D. Culler, "Taming the underlying challenges of reliable multihop routing in sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, Los Angeles, CA, November 2003, pp. 14–27.
- [8] D. Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," *Wireless Networks*, vol. 11, pp. 419–434, 2005.
- [9] R. Fonseca, O. Gnawali, K. Jamieson, and P. Levis, "Four bit wireless link estimation," in *Proceedings of the 6th Workshop on Hot Topics in Networks (HotNets)*, Atlanta, GA, November 2007.
- [10] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, Berkeley, CA, November 2009, pp. 1–14.
- [11] (2009) The multihop lqi protocol. [Online]. Available: <http://www.tinyos.net/tinyos-2.x/tos/lib/net/lqi>
- [12] (2009) Link quality routing protocol (lqrp). [Online]. Available: <http://www.sunspotworld.com/docs/Red/javadoc/com/sun/spot/peripheral/radio/mhrp/lqrp/package-summary.html>
- [13] I. Schmulevich, E. R. Dougherty, S. Kim, and W. Zhang, "Probabilistic boolean networks: A rule-based uncertainty model for gene regulatory networks," *BMC Bioinformatics*, vol. 18, pp. 261–274, 2002.
- [14] D. Marbach, T. Schaffter, C. Mattiussi, and D. Floreano, "Generating realistic in silico gene networks for performance assessment of reverse engineering methods," *J. Comp. Biol.*, vol. 16, pp. 229–239, 2009.