Model-Driven Dynamic Control
of Embedded Wireless Sensor Networks

Paul G. Flikkema[1], Pankaj K. Agarwal[2], James S. Clark[2], Carla Ellis[2],
Alan Gelfand[2], Kamesh Munagala[2], and Jun Yang[2]

[1] Northern Arizona University, Flagstaff AZ 86001 USA
[2] Duke University, Durham, NC USA

Abstract. Next-generation wireless sensor networks may revolution-
ize understanding of environmental change by assimilating heteroge-
neous data, assessing the relative value and costs of data collection, and
scheduling activities accordingly. Thus, they are dynamic, data-driven
distributed systems that integrate sensing with modeling and prediction
in an adaptive framework. Integration of a range of technologies will al-
low estimation of the value of future data in terms of its contribution
to understanding and cost. This balance is especially important for en-
vironmental data, where sampling intervals will range from meters and
seconds to landscapes and years. In this paper, we first describe a general
framework for dynamic data-driven wireless network control that com-
bines modeling of the sensor network and its embedding environment,
both in and out of the network. We then describe a range of challenges
that must be addressed, and an integrated suite of solutions for the de-
sign of dynamic sensor networks.

1 Introduction

The technology of wireless sensor networks is now becoming a mature research
field. As a result, the discipline is undergoing a natural evolution into cross-
cutting subtopics and applications, and the original vision of smart dust—
myriads of tiny, minimalist sensors—is now just one niche. Given the physical
reality of many applications, especially the size and expense of physical transduc-
ers and the cost of deployment, we are now seeing a movement toward networks
that are comparatively lower in population and density. However, integral to this
paradigm is the notion that the sensor nodes should be much smarter. We will
argue that a key aspect of this intelligence is that these sensors are, surprisingly,
even *more embedded* into their environment. This tight coupling between the
sensed and the sensors results from the need for the network to have adequate
explanatory power under significant energy consumption constraints.

Our particular application is the monitoring of physical environments and
ecosystems, with the goal of predicting biodiversity and carbon accumulation
based on scenarios of change in the atmosphere and disturbance. Changing tem-
perature, moisture availability, and atmospheric CO_2 concentrations determine
the rates at which leaves gain carbon and, thus, the rates of tree growth. While it
is clear that species respond differently to climate and CO_2, heterogeneous data,

combined with modeling challenges, have frustrated efforts to scale such relationships in ways that would permit reasonable estimates of impacts on biodiversity (see e.g. Clark et al. 2001).

We believe that *dynamic* wireless sensor networks promise to revolutionize understanding of such complex phenomena by using statistically-aware dynamical models that guide their responses to changing sampling and communication needs in the context of evolving data acquistion and communication costs. The underlying simulation models are computationally intensive and ingest data with disparate natural scales. Current wireless networks are capable of meeting many of the data needs, but network control algorithms must be developed to steer sampling and collaboration based on complex multi-level models. By injecting state-of-the art dynamical models and statistical computation into the network, complemented by out-of-network modeling, the network can acquire data at relevant scales and drive models that aid understanding and anticipate change.

2   Dynamic Sensor Networks

If sensor networks are to revolutionize ecological forecasting, they must become autonomous instruments that employ dynamic model-driven control of the sensing process, including sampling, communication, estimation, prediction, and model inference. A framework that encompasses these ideas must also accomodate the limited resources available in the network, complementing them with out-of-network capability. Figure 1 depicts such a framework, showing the two key functions, modeling and control, distributed between the network and out-of-network (i.e., in the lab) computational resources. The elements are in-Network Inference and Prediction (NIP) for real-time assessment of the system, Dynamic In-Network Control (DINC) for real-time adaptation, and Dynamic Out-of-Network Control (DONC) for network analysis, adaptive deployment (reconfiguration), and trouble shooting. To manage complexity, computationally
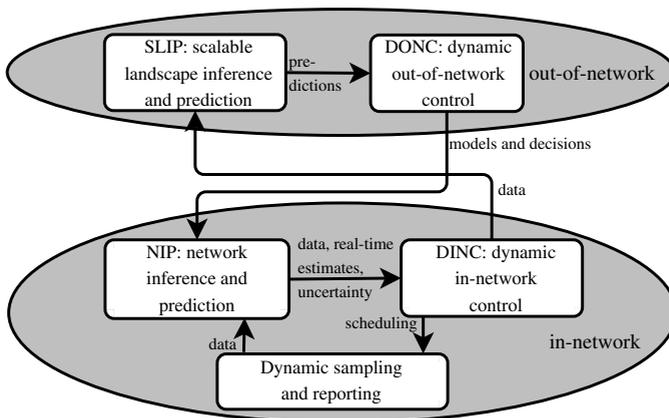

Fig. 1. Components of network control

demanding and long-term tasks are handled by a Scalable Landscape Inference and Prediction (SLIP) model operating outside the network along with DONC processing.

The SLIP model of forest stand dynamics focuses on 1) data assimilation from environmental variables, tree growth, and tree demography and 2) algorithm development to permit efficient computation for change over large landscapes. Data range from long-term (up to 15 yr) experiments involving manipulating canopy characteristics, CO_2 (FACE experiments), herbivores, remote sensing, and wireless networks of environmental sensors. Models involve mass (water, CO_2), and energy exchange, including basic tree physiology, and growth dynamics of trees. To simulate large landscapes, the SLIP model makes use of recent developments in algorithms, data structures, and hardware (Govindarajan et al. 2004, 2005). The SLIP model is also the basis for in-network NIP modeling (Figure 1) and as input to DINC and DONC.

In numerous applications, complex models will actually be networks constructed of component models. Both within and outside the *in situ* network, we expect that these component models will execute in real time with varying degrees of collaboration and competition. Given that communication is dramatically more expensive than computation in wireless networks, in-net and out-of-net model designs will be strongly influenced more by comparative communication costs than computational constraints. Consequently, in-network models will rely heavily on local, inexpensive information, informed by out-of-net models that fully exploit evolving global datasets.

A critical aspect of both the SLIP and NIP models is the integration of estimation and prediction within a Bayesian inference framework. The advantage is predictive intervals that are directly anchored by data. Hierarchical Bayes modeling (see e.g. Smith and Gelfand 1992) has emerged as a powerful new tool for inference, prediction, and decision in high-dimensional systems, and has been further developed for environmental applications (Agarwal et al. 2002) and specifically for modeling tree stand dynamics (Clark 2005). The approach provides a basis for integrating the extensive data from a range of scales. It allows modeling a complex system as a network of relationships among elements, both known and unknown. If we have set of parameters R specifying a particular process model Φ, data assimilation within the hierarchical Bayes framework is based on the decomposition

$$[\text{data}, \Phi, R] = [\text{data}|\Phi, R][\Phi|R][R].$$

Each of the three conditional relationships might be further decomposed, depending on complexity. Data can enter into different stages of the model, with each data type related to one or more elements of the process tailored to the appropriate scale, measurement errors, and so on. Parametric modeling at the various stages provides for context, which may be spatial, temporal, or involve relationships among subsets of observations.

In the context of this research, the SLIP simulator serves as the out of network model (Figure 1). It makes use of sophisticated algorithms to capture large

numbers of interactions across broad landscapes in reasonable time. Because of computational and memory demands, SLIP will run on a server or cluster outside the wireless network itself. DINC responds to NIP predictions in real time, with updated scheduling based on optimization. NIP also passes information to Dynamic Out-of-Network Control (DONC), which provides for more detailed assessment and control.

3   Sensor Network Design and Implementation

Wireless sensor network architectures and implementations should reflect the specific opportunities and constraints of the environmental sensing application. First, research-quality scientific data requires external probes, flexible and quiet analog circuit design, and a weatherproof package. Secondly, for scientific-quality sensing applications, large-scale redundancy (to aid battery lifetime or allow disposability) is economically infeasible. For example, the cost of probes is approximately 40% of the total, so that the networking and sensing infrastructure for a single node can exceed $1,000 at the prototype stage. Moreover, installation costs are significant due to the careful probe deployment procedures required. Hence optimization of spatial sampling and model-aided redeployment, in addition to dynamic temporal sampling and reporting, will be a growing concern in ecosystems sensing applications.

The prototype WiSARDNet (wireless sensing and relay device network) sensor network technology (Yang 2005, www.wisardnet.nau.edu) targets applications in biological and ecological field research. A WiSARDNet consists of two classes of nodes; 1) many nodes that provide sensing and multihop networking capabilities, and 2) relatively few gateways to the internet for SLIP and DONC functions. WiSARD nodes use a dual-processor design, with the labor divided between a brains board that provides communication and networking services and a probe data acquisition board that handles the details of the sensing tasks. The hardware design is a three-board stack; the data acquisition board used in WiSARDs is replaced in gateways by a board that provides communication interfaces and non-volatile memory for data archival.

The WiSARDNet protocol stack design is driven by the sensing application, wherein communication can be scheduled because of the low temporal sampling rates for environmental phenomena. These rates are not necessarily static: adapting sampling to hot or cold spots in time or space allow for improved data quality at a fixed average sampling rate. The protocol stack is based on proactive coordination, or scheduling, of communication transactions. We have implemented a media access control (MAC) algorithm (Flikkema 2003) that employs local proactive coordination in a slotted-time framework so that a node is awake—and consuming energy—only when it has scheduled communication with one of its cliques (sets of neighboring nodes). The MAC layer is coupled with lower and higher layers in the service of energy efficiency, scalability, and robustness.

Dynamic sensor node design.   Most current wireless sensor nodes target minimum energy consumption and 8-bit processors that use little power but have

very limited computational capability. The most obvious requirement driven by dynamic sensing is a hardware/software design that is capable of significantly higher computational power while maintaining low energy consumption. Next-generation 16- and 32-bit embedded processors are being introduced that provide order-of-magnitude increases in computational throughput over the 8-bit microcontrollers while using far less energy than 32-bit architectures now used in easily rechargeable devices such as PDAs. New devices provide for energy management with programmable subsystem clocks and a low-power instruction cache. On-chip memory management units allow separate logical program and data spaces, enabling protection critical in systems with dynamic software reconfiguration.

Support for over-the-network dynamic reconfiguration is also required, including loading of new tasks and programs and replacement of existing components. Progress in this direction includes the module registration approach in the SOS operating system (Han 2005) to provide protection from calls to nonexistent or outdated components. Efficient implementation of code and data protection can use hardware memory management.

Sensor network OS's should be extended to include application-level support of energy-consumption management. The Energy Centric Operating System (ECOSystem) (Zeng et al. 2002, 2003) has demonstrated how to schedule computation and communication to satisfy a specified power budget to guarantee a specified battery lifetime.

4   Algorithm and Software Development

Data service layer. Casting model-driven adaptive sensing and reporting as a data management problem may ultimately broaden the functionality of sensor networks (overview in Kumar 2003). We are currently building a *data service layer* for WiSARDNet that supports a high-level programming interface for specifying data acquisition tasks. These tasks may acquire information about both the environment as well as the system itself (e.g., remaining battery power and storage capacity at each sensor). High-level specifications of tasks simplify programming and enable greater optimization opportunities.

The data service layer is jointly implemented by a central server and the nodes in the WiSARDNet. Given a sensing task, the central server carries out optimization based on models with previously assimilated data, and obtains an initial acquisition plan to be deployed in the network. This plan consists of sub-plans to be executed at individual nodes, which specify when and what to sample and report. The fundamental building blocks of plans are *dynamic continuous queries*: They are continuous in that they continuously execute over streaming input data and produce results; they are dynamic in that they admit dynamic external control—these two features are essential in supporting NIP, DINC, and DONC. Together, the dynamic continuous queries form data/control flows over the network, where one query's output may be another query's data or control input. Each node in the WiSARDNet runs a data service component

that interprets the deployed plan and executes the local sub-plans. The data service component coordinates the execution and adaptation of sub-plans and their interaction with WiSARDNet hardware probes and communication layer, tags data reports with meta-data about sampling and measurement conditions, and provides multi-resolution data storage within the network.

Model-driven sensing and reporting. Environmental and system modeling offers many opportunities for optimization. With the out-of-network models of ecological processes (SLIP), we can quantify the utility or fidelity of data (in both accuracy and timeliness) and use it to control sampling and reporting in the WiSARDNet (DONC). As noted earlier, the utility of an observation depends on how it is to be used. For example, even if the local estimate of soil moisture is precise enough for water balance modeling, a model for plant mortality risk may require a more precise estimate. Thus, DONC must consider the requirements and needs of multiple models.

Inside the network, as new sensor readings are acquired, dynamic continuous queries running on the nodes collectively maintain local models of the environment (NIP). Within the fidelity requirements set by DONC, we use the in-network models to control sampling and reporting adaptively (DINC) in order to achieve energy savings. In particular, we may choose not to acquire a sensor reading if, according to DONC, its utility is low or its value can be confidently predicted by NIP to within the required accuracy. Furthermore, instead of reporting every raw sensor reading to the central server, the network only needs to report updates to local model parameters or deviations of readings from local model predictions.

Finally, our system also maintains reliability models of sensors and communication links, and uses them with statistical inference techniques to reason about errors and failures, which routinely occur in wireless sensor networks. Models of battery lifetime provide us with the basis for energy-based optimization.

Algorithmic challenges. In order to fully realize the potential of the model-driven, dynamic-system approach to understand complex environmental phenomena, one has to address the algorithmic challenges at various levels. One of the basic challenges is to define the right utility and fidelity measures based on the underlying models, which can provide the framework for developing clever algorithms that maximize fidelity with as little resources as possible. These measures will define the (often conflicting) optimization criteria, guide various trade offs, etc.

One of the main algorithm challenges is to develop algorithms that provide various trade-offs in a dynamic sensor network, including: fidelity vs. accuracy, data fidelity vs. sensing and reporting cost; model maintenance cost vs. cost-saving potential; cost vs. benefit of dynamic adaptation; and cost vs. benefit of failure protection. We have recently begun research in this area. Silberstein, Braynard, and Yang (2006) investigate the trade-off between a purely temporal model and a spatio-temporal model for suppressing change reporting in a sensor network; the former only exploits temporal coherence while latter exploits both temporal and spatial coherence. We are studying similar trade-offs for the

forest growth model (Chakraborty, et al., 2006). In the same vein, there has been preliminary work (Silberstein, Munagala, and Yang 2006) on developing power-efficient distributed local control algorithms to exploit spatio-temporal correlations to perform query-specific model-driven data acquisition. Some recent work (Goel, Guha, and Munagala 2006; Flikkema 2006) presents a theoretical framework, algorithms (both adaptive and non-adaptive), and analysis of the trade-offs between acquisition cost and information gained in model-driven resolution of uncertain data.

A related challenge is to develop an optimization framework that spans across multiple software layers and executes both inside and outside the network. Each software layer—application, data service, OS, and networking—continuously carries out optimization and adaptation at runtime, and a coordination among them is needed to maximize the performance. Similarly, DINC and DONC must coordinate with each other to ensure overall efficiency and stability of the system. Furthermore, optimization could be expensive. For example, when optimizing sensor data acquisition tasks, reasoning with complex statistical models can be a challenging computational task, especially when complicated by energy costs and network topology. Our recent work (Silberstein, et al. 2006) involves developing computationally efficient techniques based on sampling and linear programming with both energy and network considerations.

Finally, the availability of sensor data provides many opportunities to significantly improve environmental modeling. For example, one can run environmental models at a finer time scale and thus incorporate variations in the values of parameters (instead of drawing them from a distribution). But these opportunities cannot be realized without addressing a number of algorithmic challenges in the simulation of complex models. For example, one may have to use an event based simulation (e.g., using kinetic data structures (Agarwal et al. 2002)) and update the information locally and only when it is necessary. In some of our previous work, we have exploited spatial coherence but did not pay enough attention to temporal coherence (Govindrajan et al. 2004, 2006). We are currently developing algorithms for the SLIP model that exploits both temporal and spatial coherence.

5 Conclusion

One of our goals is to experimentally evaluate whether a network can deliver better results than the standard fixed sampling and out-of-network modeling in the context of changing real-world environmental conditions. Ultimately, results from these experiments will inform the design of an accurate, dynamic, energy-efficient production networks tailored to the specific demands of the application.

Our effort to understand the biodiversity and carbon consequences of environmental change is broad enough to encompass many of the types of challenges faced by dynamic data-driven application systems. The research results will be applicable to many environmental sensing/decision applications, including micrometeorological sensing, pollution monitoring and environmental remediation,

and public security/safety. In particular, results from this effort will inform the planning and implementation of future distributed in situ monitoring, modeling, and decision systems.

References

1. Agarwal, P. K., Guibas, L. J., et al: Algorithmic issues in modeling motion, ACM Comput. Surv. 24 (2002), 550–572.
2. S. Chakraborty, P. K. Agarwal, and J. Clark, The gap light model, manuscript, 2006.
3. Clark, J.S., et al: Ecological forecasts: an emerging imperative, Science 293:657-660, 2001.
4. Clark, J.S.: Why environmental scientists are becoming Bayesians, Ecol. Lett. 8:2-14, 2005.
5. Flikkema, P., West, B.: Clique-Based Randomised Multiple Access for Energy-Efficient Wireless Ad Hoc Networks, in Proc. 2003 IEEE Wireless Communications and Networking Conference (WCNC '03), New Orleans, March 2003.
6. Flikkema, P: The precision and energetic cost of snapshot estimates in wireless sensor networks. Submitted for publication (2006)
7. Goel, A., Guha, S., and Munagala, K.: Asking the right questions: Model-driven optimization using probes. Submitted for publication (2006)
8. Govindrajan, S., Dietze, M., Agarwal, P. K., and Clark, J., A scalable simulator for forest dynamics, in Proc. 20th Sympos. Comput. Geom. (2004).
9. Govindrajan, S., Dietze, M., Agarwal, P. K., and Clark, J., A scalable algorithm for dispersing population, J. Intelligent Information Systems, in press.
10. Kumar, V. ed.: Special Section on Sensor Network Technology and Sensor Data Management (Part I). SIGMOD Record, **32**(4) (2003)
11. Silberstein, A., Braynard, R., Ellis, C., Munagala, K., and Yang, J.: A sampling-based approach to optimizing top-k queries in sensor networks. Proc. of the 22nd Intl. Conf. on Data Engineering, Atlanta, Georgia (2006)
12. Silberstein, A., Braynard, R., and Yang, J.: Constraint-chaining: on energy-efficient continuous monitoring in sensor networks. Proc. of the 22nd Intl. Conf. on Data Engineering, Atlanta, Georgia (2006)
13. Silberstein, A., Munagala, K., and Yang, J.: Energy efficient monitoring of extreme values in sensor networks. Submitted for publication (2006)
14. Smith, A.F.M., and Gelfand, A.E.: Bayesian statistics without tears: a sampling-resampling perspective. American Statistician, 46, 84-88, 1992.
15. Yang, Z., et al.: WiSARDNet: A system solution for high performance in situ environmental monitoring, 2nd International Workshop on Networked Sensor Systems (INSS 2005), San Diego, 2005.
16. Heng Zeng, Carla S. Ellis, Alvin R. Lebeck, and Amin Vahdat. Ecosystem: Managing energy as a first class operating system resource. In Proc. Tenth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS X), pages 123–132, October 2002.
17. Heng Zeng, Carla S. Ellis, Alvin R. Lebeck, and Amin Vahdat. Currentcy: A unifying abstraction for expressing energy. In Usenix Annual Technical Conference, pages 43–56, June 2003.