Holistic Operations in Large-scale Sensor Network Systems: a probabilistic peer-to-peer approach

Indranil Gupta and Ken Birman*
Dept. of Computer Science
Cornell University
Ithaca NY 14853
{gupta,ken}@cs.cornell.edu

Large-scale Sensor Networks

Smart sensor nodes [14, 15] integrate multiple sensors (e.g., temperature, humidity, accelerometers), processing capability (microprocessor connected to the sensors using I2C technology), wireless communications, and a battery power source. Sensor nodes range around a few millimeters or smaller in size [11, 12].

Sensor nodes have low processing, memory and wireless communication capabilities. Processor and bus clock speeds of a few Mhzs, RAM sizes of a few KB, flash memory sizes of a few MB, and wireless bandwidths of a few hundred Kbps are typical [14, 15]. These ranges are often determined by application domain constraints such as cost, power consumption, and deployment space. Thus, they are likely to stay in spite of improvements in fabrication techniques. For example, retinal sensor sizes depend on ganglial nerve separation [11]. Space limits on the wireless antenna restricts maximum frequency and range of transmission. Availability of power is limited due to the difficulty or impossibility of recharging nodes in inhospitable terrains, e.g., military applications [12].

On the other hand, sensor nodes are becoming easier to produce in large numbers [14, 15]. An autonomous system consisting of a large number of sensor nodes deployed over an area and integrated to collaborate through a network (wireless or wired) network, would encourage several novel as well as existing applications. Examples include high fidelity image processing in retinal prosthesis chips [11], battlefield applications such as vehicle tracking, environmental observation and forecasting systems (EOFS) [13], etc.

Holistic Operations

In such sensor network systems, operations spanning a large number of sensor nodes (perhaps only a subset of all the nodes) assume greater importance than the exact reading at a specific node [2, 4, 10, 11, 13]. A user monitoring the area is usually more interested in collecting data from sensor nodes, disseminating commands to them, and in general being able to control the system *holistically.* An autonomous system-wide application (e.g., one that takes corrective action when the average humidity in the area crosses a threshold) would require streaming updates of global aggregates of sensor node readings. Moreover, individual sensor node readings are inaccurate, making a system-wide estimate more valuable.

* The authors were supported in part by DARPA/AFRL-IFGA grant F30602-99-1-0532 and in part by NSF-CISE grant 9703470, with additional support from the AFRL-IFGA Information Assurance Institute, from Microsoft Research and from the Intel Corporation.

We term these *holistic operations*, because they involve large parts of the sensor network rather than individual or small groups of sensor nodes. We are interested in finding scalable and reliable solutions to holistic operations such as reliable multicast, data aggregation, leader election, group membership, distributed indexing of files and data, etc. We believe that implementing holistic operations *scalably* and *reliably*, along with the ability to trade off between the two properties, is an important step towards the development of truly autonomous large-area sensor network systems. We limit our discussion in this paper to reliable multicast and data aggregation. Data aggregation can be used to calculate global aggregates across sensor nodes (e.g., average, variance, or maximum, of a measured quantity such as temperature), for collaborative signal processing [1], etc. A multicast protocol can be used to disseminate commands or data into the sensor network.

Solving these problems over an asynchronous network is known to be difficult, and often a provably impossible task [5]. Nevertheless, protocols for sensor network systems are required to be tolerant to failures of individual sensor nodes (e.g., from discharge of power; external causes of destruction) and packet losses within the network (e.g., due to interference in the wireless medium, inhospitable terrain). They have to scale in terms of overhead (computation and communication) and power consumption at nodes, network load (which affects per-node overhead under multi-hop routing), and variation of these loads with system size.

Properties of these protocols should be applicable even when the assumed sensor network model differs from the real-life one. Protocol behavior must degrade gracefully due to interference from other concurrently present strategies or protocols in the sensor network. Examples of such influences include anonymity of sensor nodes (i.e., lack of unique identifiers), presence of power saving strategies such as periodic sleeps, etc. This makes protocol design a challenge, since such constraints are determined by models and standards for the sensor node, network, basic protocols and applications, all of which are constantly evolving.

Why A Peer-to-Peer Approach?

Environmental monitoring and biomedical systems [11, 13] appear to use either a centralized architecture, where sensor node data (e.g., measurements) are centrally processed. Sensor nodes communicate data either (a) directly to a single or small number of base stations, or (b) through a cluster-based overlay network formed among the nodes (often structured hierarchically). Data compression and application-dependent optimization (e.g., segmentation in image processing [11]) are used to reduce the power used by communication. Cyclic redundancy checks are used to counter packet loss or corruption. Cluster-based systems reduce the effect of node failures through fault-tolerant backup within and across clusters.

Using a centralized architecture to aggregate or multicast data to the sensor network has the disadvantages of (a) high communication overhead to a distant base station (due to the absence of the ability to set up a denser base station infrastructure, a characteristic property of isolated application domains), and (b) a high turnaround time in the aggregate calculation (latency that grows linearly with the number of nodes). Using a cluster-based technique alleviates some of these problems but introduces a fault-tolerance problem [7], e.g., failure of cluster leaders can introduce incompleteness into the calculation of global aggregate functions (until a new leader is elected). The high overhead on cluster leaders could be eliminated by migrating leadership across nodes in a cluster, but this would entail frequent runs of a leader election protocol. Often, it is also difficult to mathematically guarantee or prove any properties about the turnaround time between initiation and completion of a one-shot holistic operation in the system. A quick turnaround time is necessary to achieve soft real-time properties that are essential to an self-managing system.

The Probabilistic Approach

One promising approach to holistic operations in large sensor networks is based on the use of probabilistic protocols [3, 6–10]. These protocols are peer-to-peer in nature, and do not require infrastructure such as base stations or central points of control. They impose low computation and communication overhead on sensor nodes, and this load is equally distributed across all sensor nodes, avoiding hot-spots. The protocols are simple, and thus have a small code footprint. They provide probabilistic guarantees on correctness or reliability of the operation. A tradeoff between scalability and reliability is achieved through tuning of the overhead at individual sensor nodes. High probability reliabilities can be obtained even in the presence of node failures and high packet loss rates, and within turnaround times that increase slowly with system size. Finally, deterministic correctness is provided by a concurrent protocol that inexpensively backs up the probabilistic protocol.

We briefly discuss probabilistic protocols for multicast and aggregation.

Multicast: Epidemic- (or gossip-) based protocols for data dissemination in an ad-hoc network employ a probabilistic style of dissemination, where each node independently makes probabilistic decisions about how it forwards (gossips about) received data to its neighboring and remote nodes [6, 8, 9]. Target number (per message) and choice strategy can be tuned to trade off between the reliability of dissemination and communication overhead. Epidemic-style of data dissemination is an efficient alternative to data flooding.

Epidemic protocols can be imparted such properties as topology awareness and adaptivity by using weak overlays and topological mappings [6–8]. Topological awareness causes packets to traverse fewer hops, reducing routing overhead. Adaptivity lowers overhead at small failure rates in the system.

The dissemination protocols in [6] require sensor nodes to construct a virtual weak overlay called the Leaf Box Hierarchy (or the Grid Box Hierarchy, which is an instance of the former). [1] Hierarchical epidemic algorithms within the Leaf Box Hierarchy impose an overhead per process that grows as the square of the logarithm of system size, and disseminate multicast data to interested nodes with very high probability. The dissemination latency grows very slowly with system size (a sublinear variation).

A scheme that adapts per-node overheads to the failure rate in the network is also presented in [6]. When there are few failures in the network, the aver-

[1] Other examples of such overlays include the Amorphous Computing hierarchy [2].

age per-node cost is a constant independent of system size. The overhead rises slowly with failure rate. The adaptive scheme is based on a hybrid of the gossip-style dissemination and a tree dissemination algorithm; the construction of this tree is also probabilistic, and so retains the scaling properties of gossip-based algorithms, e.g., diffusing protocol overhead among nodes.

Data Aggregation: In [7], we have reasoned about traditional approaches to the problem of calculating composable global functions (such as average, variance, maximum etc.) over a large set of readings from sensor nodes - considered system sizes ran into thousands of nodes. Schemes based on traditional leader election of nodes as cluster heads have inherent drawbacks with respect to tolerance of the final estimate to node failures or packet losses, e.g., during a protocol run, failure of a node holding an aggregate of a large subset of sensor nodes would cause exclusion of all such sensor node readings in the global aggregate. Our work in [7] proposes a new protocol that uses gossiping within the Leaf/Grid Box Hierarchy to disseminate an estimate of the global aggregate to *all* sensor nodes. Evaluation of the protocol shows that with a per-node message overhead varying with the square of the logarithm of system size, the protocol produces global aggregate estimates with very high completeness (i.e., fraction of sensor node readings included in the estimate).

Continuing Work and Future Directions

Probabilistic protocols appear to not only match the hardware specifications of sensor nodes but also solve a variety of holistic operation specifications essential to the proliferation of truly autonomous sensor network systems. Future directions for research in this class of solutions include the following.

- *Overlay Self-Assembly:* Protocols for constructing this overlay in a distributed fashion, even in the absence of location services such as GPS.
- *Overlay Self-Management and -Reconfiguration:* Overlay reconfiguration might be (a) application initiated, e.g., a change in the set of sensor nodes involved in the operation, or (b) required automatically, e.g., due to failures and mobility of nodes. This assumes importance for long-running holistic operations.
- *Energy Efficiency:* Studying the increase in power consumption introduced by redundant transmissions of the same information in probabilistic protocols, and comparison with centralized approaches.
- *Applicability under Altered Models:*
  - *Anonymous Nodes:* An overlay such as the Leaf Box Hierarchy could be used to provide a coarse addressing scheme in a network with anonymous sensor nodes.
  - *Interference from other concurrently present strategies:* Our approach accommodates a range of concurrently present strategies in the network. For example, a concurrent power-saving strategy, where each node sleeps periodically for a while, causes graceful degradation because of the fault-tolerance properties of probabilistic protocols.
- *Performance Tuning:* Variation across applications of tradeoff between sensor node overhead (e.g., power consumption) and required reliability of the holistic operation.

– *Interaction with Ad-hoc Routing Protocols:* Adapting to variation in route length and quality due to the multi-hop ad-hoc routing protocol used, arbitrary placements of nodes causing large disparity between routing and geographical distances.
– *Improving The Proposed Algorithms:* For example, the aggregation protocol can be used to calculate streaming aggregates, or for an arbitrary-sized region, or to accommodate non-composable aggregate functions.

Broader research goals include exploration of:

– Promising areas of applicability of probabilistic solutions to holistic operations, e.g., environmental observation systems.
– Combination of collaborative in-network signal processing with protocols for holistic operations.
– Holistic operations that are atypical in Internet-based process group computing, yet assume significance in sensor networks.
– New sensor network applications that might be enabled by efficient and scalable solutions to holistic operations.

References

1. S.R. Blatt, "Collaborative signal processing", BAE Systems' presentation at *SensIT PI Meet*, DARPA IXO, Apr 2001.
2. D. Coore, R. Nagpal, R. Weiss, "Paradigms for structure in an Amorphous Computer", A.I. Memo No. 1614, A.I. Laboratory, MIT, Oct 1997.
3. A. Demers, et al, "Epidemic algorithms for replicated database maintenance", *Proc. 6th ACM PODC*, Aug 1987, pages 1-12.
4. D. Estrin, R. Govindan, J. Heidemann, S. Kumar, "Next century challenges: scalable coordination in sensor networks", *Proc. 5th ACM/IEEE MobiCom*, Aug 1999, pages 263-270.
5. M.J. Fischer, N.A. Lynch, M.S. Paterson, "Impossibility of distributed consensus with one faulty process", *Journ. ACM*, 32:2, Apr 1985, pages 374-382.
6. I. Gupta, A.-M. Kermarrec, A.J. Ganesh, "Efficient epidemic-style protocols for reliable and scalable multicast", To appear in *Proc. 21st SRDS*, Oct 2002.
7. I. Gupta, R. van Renesse, K. P. Birman, "Scalable fault-tolerant aggregation in large process groups", *Proc. 2001 DSN*, Jul 2001, pages 303-312.
8. D. Kempe, J. Kleinberg, A. Demers. "Spatial gossip and resource location protocols", *Proc. 33rd ACM STOC*, Jul 2001, pages 163-172.
9. L. Li, Z. Haas, J.Y. Halpern,"Gossip-based ad hoc routing", *Proc. 21st IEEE INFOCOM*, Jun 2002.
10. R. van Renesse, K. Birman, "Scalable management and data mining using Astrolabe", *Proc. 1st IPTPS*, Mar 2002.
11. L. Schwiebert, S.K.S. Gupta, J. Weinmann, "Research challenges in wireless networks of biomedical sensors", *Proc. 7th ACM/IEEE MobiCom*, Jul 2001, pages 151-165.
12. Smart Dust Project, `http://robotics.eecs.berkeley.edu/~pister/SmartDust`
13. D.C. Steere, A. Baptista, D. McNamee, C. Pu, J. Walpole, "Research challenges in environmental observation and forecasting systems", *Proc. 6th ACM/IEEE MobiCom*, Aug 2000, pages 292-299.
14. `http://www.cs.berkeley.edy/~culler/cs294-8`, `http://www.cs.rutgers.edu/~mini`
15. Crossbow Technology Inc., `http://www.xbow.com`