Controlled Mobility for Sustainable Wireless Sensor Networks

Aman Kansal, Mohammad Rahimi, Deborah Estrin, William J Kaiser, Gregory J Pottie, and Mani B Srivastava
University of California, Los Angeles

Abstract—A key challenge in sensor networks is ensuring the sustainability of the system at the required performance level, in an autonomous manner. Sustainability is a major concern because of severe resource constraints in terms of energy, bandwidth and sensing capabilities in the system. In this paper, we envision the use of a new design dimension to enhance sustainability in sensor networks – the use of controlled mobility. We argue that this capability can alleviate resource limitations and improve system performance by adapting to deployment demands. While opportunistic use of external mobility has been considered before, the use of controlled mobility is largely unexplored. We also outline the research issues associated with effectively utilizing this new design dimension. Two system prototypes are described to present first steps towards realizing the proposed vision.

I. INTRODUCTION

A salient feature of sensor networks is that the network elements are very tightly coupled with the physical environment. The resource requirements change dynamically as the environment evolves in space and time. Sustained unattended operation makes it crucial for the system to efficiently schedule its resource usage in both space and time to meet the performance specification. This makes it attractive to physically reconfigure the network at runtime rather than merely relying on protocol level changes to adapt to the external dynamics.

We propose to use a direct method for physical reconfiguration: controlled and coordinated motion of network elements or their subparts. Traditional systems have considered mobility as an extra overhead to which the network must adapt, possibly at a loss of performance. Even when mobility has been considered useful for the network, control over mobility is assumed to rest outside the system. The network protocols in those cases either treat mobility as random [1], [2] or, when possible, predict the mobility patterns [3], [4], [5] but do not modify them. We propose a new method where controlled mobility is autonomously used by the network itself for enhancing functionality. The identifier Morph will be used for our proposed approach in the following discussion.

Some amount of actuation is essential to network operation and is already incorporated into the network architecture, although not yet automated – system deployment requires motion and is carried out manually. Morph extends this manual motion into an autonomous system capability in order to enhance sustainability.

The cost of adding actuated components is envisioned to be less than the alternative cost of manually repairing and replacing a less sustainable system. Controlled mobility is typically considered costly in terms of energy, hardware resources, and navigational needs. We propose alternatives to make controlled mobility tractable. The first alternative relies on infrastructure. Addition of simple infrastructural support such as a cableway or track along which network elements may move drastically reduces the mobility overheads. Navigation becomes trivial, uncertainty associated with localization almost disappears and the energy requirements are reduced. The second alternative relies on limiting the range of motion and thus reducing associated overheads.

II. MOTIVATION

In studying any physical process, an inescapable question is the density of sampling. This can be approached using static nodes by uniformly over-sampling much of the sample volume. However, over-sampling is not an option when the frequency characteristics of the phenomenon to be studied are unknown or when the phenomenon has sharp discontinuities which require an infinite sampling density. Further, it would be desirable to concentrate resources so that the minimum number of sensor nodes can achieve the maximum coverage in any deployment. This calls for an intelligent placement of nodes in space. Working against this placement are two factors: first, the unknown initial state of the process, and thus lack of prior knowledge of where to concentrate sample collection and second, the variability of the environment over time. For either factor mobility provides a solution. Consider the situation in Figure 1, which depicts lines of constant value such as isotherms. Clearly the most critical region for sampling is in the top left quadrant where there is a steep gradient. Using controlled mobility, resources can efficiently be moved to regions where they are required.

Mobile nodes can also transfer data. Apart from moving nodes, movable directional antennae can be used to
III. ADVANTAGES OF CONTROLLED MOBILITY

The key benefits of Morph are performance and sustainability, for deployments at all densities. Sustainability is defined as an attribute of wireless networks that describes their ability to maintain a continuous quality of service (QoS) for their client applications. The ability to move directly translates into enhanced QoS through the mechanisms discussed below.

A. Topology Adaptivity

Mobility provides a level of control on the network topology which is hard to emulate using other mechanisms. This gives us the following advantages:

Run Time Adaptation: The inevitable evolution of the environment changes the sensor data generation and the resultant traffic patterns in the network over time. Adjusting protocol parameters such as coding rates or initiating new routes along the existing topology may not allow the network to meet the new traffic requirements. Physical mobility is required in these situations to adapt to the run time dynamics of system evolution.

Deployment Configuration: An unsolved problem which becomes tractable with Morph is that of initial deployment. A-priori knowledge of the external world where the network is being deployed is typically unavailable and can only be learnt over time. An optimal physical topology is thus hard to create at deployment time. Most systems are deployed based on assumed models for sensing requirements and traffic patterns. While assumed generic models can give probabilistic estimates on the system performance they cannot guarantee QoS for the specific characteristics of a particular deployment.

Controlled actuation can tune the deployment to the exact requirements of the specific application.

B. Increased Network Capacity

Controlled mobility leads to increased traffic capacity. Recent research [6], [8] has shown how random mobility can increase the data carrying capacity of a wireless network. While those results are derived for asymptotic cases controlled mobility can achieve the capacity increase for arbitrary and finite network topologies:

Channel Capacity: The data carrying capacity of a wireless network increases when mobility is used. This was shown in [6] for the case of random mobility. However, transport delay is not bounded in [6]. With controlled mobility, data capacity increases even with bounded delay. An example where the mobility of network elements increases channel capacity occurs in our prototype described in section V-A. The system consists of wireless nodes [9] and a mobile router. In this system, even for the small amounts of data stored at a node (512KB, due to its limited memory capacity) the time taken to transfer it over two wireless hops (at its full 38.4kbps error-free bandwidth with ideal MAC and no protocol overhead) is 36.7% more than using the mobile router to physically carry it (at a modest speed of 200cm/s, about half the maximum capability of hardware [10], and including the time taken to extract data from the static nodes over wireless connections) for the same distance as two wireless hops (20m).

Energy Capacity: The energy capacity of the network can be improved in certain situations where mobile elements can be used to aggregate data from several nodes and hence carry a sufficiently large amount of data to offset the energy cost of physically moving the node. The cost of transmitting a bit over a wireless medium using some common radios is given in the upper rows of Table I. The cost is measured in energy/bit/meter and is derived using the nominal range for the mentioned radios. The lower two rows give the energy cost per-bit-meter when data is carried physically using mobile elements (assuming 250MB data is aggregated). It may be noted that the costs evaluated for the wireless technologies are assuming no bandwidth is wasted in MAC overheads, and so the actual costs will be higher than these estimates. The costs for mobile devices are based on experimental hardware and will be lower for more advanced versions.

Also, mobile nodes can travel to recharge stations, and hence the energy used by a mobile node to save energy at static nodes sets up a virtual flow of energy into the system, increasing its lifetime. Mobility has been found beneficial to both replenish energy sources [11] and reallocate resources [12], again helping overcome battery limitations.

Data Fidelity: Wireless channel errors depend on the number of hops a packet has to travel. When data travels...
TABLE I
Energy cost of some data transmission technologies.

<table>
<thead>
<tr>
<th>Communication Technology</th>
<th>Energy (nJ/bm)</th>
<th>System</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLAN</td>
<td>2</td>
<td>Orinoco 802.11b</td>
</tr>
<tr>
<td>Bluetooth</td>
<td>35</td>
<td>IOgear CF [13]</td>
</tr>
<tr>
<td>Narrowband sensor radio</td>
<td>130</td>
<td>Mica2[9]</td>
</tr>
<tr>
<td>UGV</td>
<td>30</td>
<td>Packbot [10]</td>
</tr>
<tr>
<td>Infrastructure Assisted</td>
<td>5</td>
<td>NIMS (Section V-B)</td>
</tr>
</tbody>
</table>

physically, the number of wireless hops is reduced thus reducing packet errors, simply because storage as a channel is much less error prone than radio communication. This increases the effective goodput of the network by saving retransmission bandwidth and energy.

C. Fault Isolation

The appearance of faults and malicious attacks have an adversarial effect on quality of service. Here, physical mobility can be used for rescuing system integrity. The performance of a suspected faulty node can be compared to that of a known operational node by physically moving the operational node to the suspected node’s location. Similar methods can be used to isolate malicious or compromised nodes. This is extremely critical when the complete network cannot be secured either due to the vast extent of deployment, due to the requirement to operate in a hostile setting, or the nature of environmental coupling which makes the nodes more vulnerable to damage. Securing a small subset of nodes may have to be done in order to enhance global performance. A notion of priority among the tasks to be served may have to be developed. Algorithms which facilitate the coordination of the mobile nodes to optimize network wide performance metrics are desirable but scalability demands may force local greedy approaches to be utilized.

IV. RESEARCH CHALLENGES FOR MORPH

The challenges in using controlled mobility lie in two domains – first, in designing sensor network protocols which exploit mobile components effectively and second, in the navigational problems for the mobile elements. In this paper we focus mainly on the first domain but briefly address the second in section V where we describe our prototype networks based on controlled mobility.

A wide variety of actuation abilities can be added to the network (Table II). Each class has its related trade-offs which make it suitable for solving certain types of problems, and a Morph network aims to effectively utilize the given actuation capabilities. We now discuss the challenges involved in doing this.

A. Spatio-Temporal Scheduling

The network must actively determine and acquire the topology required to best serve the current QoS demands.

1) Mobility Strategies: The network elements at any time will be involved in several functions – measuring the data to be transmitted, relaying data from other sources, aggregating data from multiple sources or distributing user commands to the network. The first scheduling problem is how the nodes should move and position themselves in order to enhance global performance. A notion of priority among the tasks to be served may have to be developed. Algorithms which facilitate the coordination of the mobile nodes to optimize network wide performance metrics are desirable but scalability demands may force local greedy approaches to be utilized.

2) Resource Redistribution: One major performance bottleneck for global system performance comes from non-uniform resource depletion. Extended activity in a small part of the system may deplete batteries or cause other damage leading to coverage or communication loss in local regions. Here, even though the system as a whole may have sufficient resources, the delivered QoS is no longer acceptable due to the existence of localized regions where performance constraints are violated. Efficient algorithms are needed for detecting localized damage and making the performance degradation more graceful. Some initial methods to repair localized damage to sensor coverage using mobility were presented in [17], [12].

3) Evolution in Time: The decisions taken to optimize performance at the current instant should be guarded to avoid performance penalties in future operation. Energy and time spent on mobility to optimize current operation may cause gross suboptimality in the future. Node mobility, though a powerful method to salvage performance, must be used with due consideration of temporal dynamics. For example, in the case of a communication link degradation, if the time taken to actuate a directional antenna into the right direction is larger than the time it takes to finish the transmission at the degraded channel using a higher code rate, clearly actuation would be a waste of energy.

B. Network Layer Adaptivity

Network and transport layer protocols are presented with new design requirements in Morph, since the organization of the network is constantly changing. The changes are affected by the network itself and hence are known to the network but many of them are in response to external variables.
1) Network Organization: Various hierarchical schemes organize nodes into clusters. Such schemes set up aggregation trees where data flows from leaf nodes to the root with application specific concentration at each level of hierarchy. However, in such scenarios cluster-heads are typically stable and changed only for load balancing. In our proposed scenario, the availability and disappearance of cluster-heads is controlled by the network itself, but depends on movement actions taken in response to external parameters. Hierarchical schemes must be modified to operate under such dynamics if the potential advantages due to mobility are to be realized. Suppose an environmental change demands that the nodes modify their locations in order to follow the change in phenomenon. However, a cluster-head node must also maintain its connectivity to relay nodes to which it is transferring the collected data. If the clustering protocol is not aware of network layer control over mobility it will naively wait for a timeout period to detect cluster-head loss, then elect a new cluster-head and then find a new route. Thus, instead of improving the data fidelity, which was one of the motivations for modifying node locations, the clustering protocol will introduce data discontinuities which may actually be worse than lower fidelity data collected by a static network which did not move with the phenomenon.

Several topology management algorithms utilize low power modes in embedded nodes to save energy. This problem may become more involved in the presence of intermittently available links. The active and low power mode usage at static nodes and mobile nodes will have to be coordinated to effectively utilize placement optimizations.

2) Transport Layer Protocols: Several algorithms in the communication stack depend on continuous availability of links and may activate time-out procedures when links are unavailable. It is important to note that in Morph some of the link disruptions are not due to external factors but due to network initiated motion. Thus, old algorithms must be modified to coordinate with the mobility control primitives so that they operate efficiently in fluid topologies. Non-availability of constant connections but availability of good quality connections at programmable times opens up new trade-offs in data latency and storage capacity at nodes. Data transfers and end-to-end ACKs must be scheduled to meet application layer latency constraints along intermittent links.

3) Cross-layer Integration: The system must control its mobility to bridge two orthogonal and potentially conflicting external variables - the quality of the communication channel and the data requirements of the application layer. Knowledge about the channel dynamics lies with the PHY and MAC sub-layers which cannot account for the application requirements. Also, certain applications are meaningful only when viewed in the context of the collaborative behavior of several nodes while the environmental dynamics are available locally at each node. This calls for a joint optimization of not only the various layers at a node but also across nodes. A robust and modular solution must make judicious choices regarding which layer should implement the respective solutions.

V. Case Studies

We now describe two nascent prototypes based on Morph to:

- Show that the above challenges are tractable
- Describe frameworks for development of Morph systems

The first prototype, Autonomous Intelligent Mobile Micro-server (AIMMS), demonstrates routing layer design issues when router mobility is controlled. The second one, Networked Infomechanical Systems (NIMS), demonstrates a feasible approach for making locomotion tractable.

A. Autonomous Intelligent Mobile Micro-Server

This system is developed for an application where numerous static sensors collect data about their environment. Sustainability is an issue because the wireless nodes are too numerous and too deeply embedded for their batteries to be replaced manually. Morph methods help increase the energy capacity of the network. In AIMMS, a mobile micro-server moves across the network to route data from the deeply embedded nodes. Thus, the static embedded nodes need to carry much less traffic than in the situation when there was no mobile micro-server and all the data was transferred over multihop wireless routes. Here, nodes have to relay data for only those nodes which are not in the micro-server’s direct range. This relaying overhead is much less than without AIMMS, as now data is relayed only to the closest node which is accessible from the router’s path rather than to the final destination. The energy used at the micro-server itself is replenishable since the router is not permanently embedded in the environment. The micro-server uses an existing trail in the natural environment, to keep navigation tractable.

The network is assumed to be randomly deployed. At certain locations, more than one node may be in range of the router, and communication bandwidth will be divided among them, suffering further due to MAC collisions. The wireless channel itself is non uniform and may introduce more errors at particular nodes at certain times. The motion of the micro-server is controlled to maximize communication with the embedded nodes, by stopping or slowing down within latency contraints, in regions where the communication channel is constrained [18].
The routing protocol is also changed for the Morph approach. The protocol used is a variant of directed diffusion [19]. Directed diffusion is well suited for sensor networks where data from several sensor nodes is sent to a few destinations, denoted sinks. Briefly, the sink broadcasts a query which is re-broadcast by all nodes hearing it. As the broadcast progresses, each node which hears the query gets to know the reverse path to the sink node. Also, nodes which have data matching the query, reply on the reverse path.

Directed diffusion if used directly in AIMMS will not yield the advantages of Morph. Directed diffusion uses reverse paths set up during query broadcast. Thus, data is routed over multi-hop wireless routes, not exploiting the mobile micro-server. Also, if micro-server generates the query and moves, the reverse paths are no longer valid, causing loss or re-transmission of data, hence reducing QoS and sustainability, as opposed to improving it.

We modify directed diffusion as follows. Queries are transmitted by the micro-server on behalf of the sink. The micro-server repeatedly broadcasts the query as it moves, and hence all nodes which are within communication range from its trail, will realize that they can access the micro-server directly. Other nodes will receive the query along one or more paths. Each node finds out that among the various query messages it received, which one arrived on the shortest path from the micro-server, using number-of-hops-traversed field of the interest message. The key modification is that nodes will respond to and re-broadcast only the query arriving via the shortest path. This routing protocol effectively exploits the controlled mobility and improves sustainability by reducing relay overheads [18].

Third, the controlled mobility method must access complex terrain and surface regions that may be incompatible with conventional surface vehicles or may represent regions where surface vehicle passage may disturb the environment in an undesired fashion.

Fourth, mobility must be sustainable and constantly available since the environment may continue to evolve, requiring repeated and rapid reconfiguration.

Networked Infomechanical System (NIMS) [21] introduces a suspended infrastructure for controlled mobility that addresses most of the above requirements. As shown in Figure 3, NIMS infrastructure allows nodes to move in a sustainable and precise fashion within complex three-dimensional environments.

B. Networked Infomechanical Systems

The requirements of controlled mobility for sustainable wireless networks differ dramatically from past robotics applications.

First, the controlled mobility methods required for sustainable sensor networks must provide precise and reproducible location control over large spanning distances (for example with span ranges of over 100m in our prototype). It is essential to avoid navigation errors or motion limitations that would increase uncertainty of position and degrade performance.

Second, mobility must permit a wide range of node locations and articulated antenna viewing perspectives within the space of large three-dimensional environments. Mobile communication nodes may have to be elevated vertically in a three-dimensional space to avoid the frequently prohibitive path loss that results from communication in surface-to-surface geometries [20]. The capability to choose useful sensor viewing perspectives in cluttered environments is also desirable.

The NIMS architecture exploits multiple hierarchical levels of fixed, mobile, and infrastructure elements to provide a high effective spatial sampling density and sustainable energy logistics. NIMS infrastructure has demonstrated energy harvesting capability through solar photovoltaic systems in outdoor deployments. This energy may be distributed to fixed nodes that are optimally placed to benefit communication, but are not in a position to harvest energy. NIMS also provides logistics for the relocation of fixed nodes.

VI. RELATED WORK

The use of mobility has been explored recently in wireless networks. Mobility can be classified into three categories: random, predictable and controlled. Random mobility has been studied before for improving data capacity [6], [8] and networking performance [1], [22], [23], [2]. However, in such cases the latency of data transfer cannot be bounded deterministically, and delivery itself is in jeopardy if the data is cleared from the buffer at the mobile agent. The use of a predictable mobile agent was considered in [3], [4].
Morph concentrates on the third category of mobility. While some of the advantages of mobility can be realized in the above approaches, controlled mobile elements are required for a more complete exploitation of this design dimension. Examples of systems in this class are now emerging. In [24], a network of UAV’s was proposed to create a reconfigurable network topology. However, this sort of mobility incurs high energy overheads, requires expensive hardware and has complex navigation requirements. We presented prototypes where mobility is significantly more economical [18], [21]. Other examples include [25], [11], [12], [26], [17], [27].

Protocols for networks without constant end-to-end connectivity are also provided in Delay Tolerant Networking [28]. However, Morph is distinct from DTN because in Morph control over mobility allows the network to schedule the disruptions and link availabilities.

VII. CONCLUSIONS

We see that the introduction of controlled mobility has significant benefit for sensor network sustainability. The Morph vision provides a new method for directly addressing several complex problems associated with such networks. However, there are open issues to be addressed before the potential advantages can be realized. These challenges inspire research at all layers of the network stack and application specific exploitation of controlled mobility. We also presented initial prototypes which demonstrate the advantages and feasibility of Morph. We believe it is timely and attractive to develop network functionality for exploiting controlled actuation spanning varying degrees of freedom on different time-scales.

VIII. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. ANI- 00331481.

REFERENCES


