

RAP: A Real-Time Communication Architecture for Large-Scale Wireless Sensor Networks

Chenyang Lu Brian M. Blum Tarek F. Abdelzaher John A. Stankovic Tian He

*Department of Computer Science
University of Virginia
Charlottesville, VA 22903
{chenyang, bmb5v, zaher, stankovic, th7c}@cs.virginia.edu*

Abstract

Large-scale wireless sensor networks represent a new generation of real-time embedded systems with significantly different communication constraints from traditional networked systems. This paper presents RAP, a new real-time communication architecture for large-scale sensor networks. RAP provides convenient, high-level query and event services for distributed micro-sensing applications. Novel location-addressed communication models are supported by a scalable and light-weight network stack. We present and evaluate a new packet scheduling policy called velocity monotonic scheduling that inherently accounts for both time and distance constraints. We show that this policy is particularly suitable for communication scheduling in sensor networks in which a large number of wireless devices are seamlessly integrated into a physical space to perform real-time monitoring and control. Detailed simulations of representative sensor network environments demonstrate that RAP significantly reduces the end-to-end deadline miss ratio in the sensor network.

1. Introduction

With the advances in MEMS devices and embedded processors and radios, it will soon be feasible to deploy large-scale sensor networks to perform distributed micro-sensing and control of physical environments [1]. For example, a surveillance system may use a large network of acoustic sensors to detect and track vehicles in a security area. Similarly, biometric sensors can be deployed in airports to detect harmful bio-agents and issue alarms to command and control centers during potential bio-attacks. These smart sensors and actuators are equipped with low-power processors and short-range radio transceivers [9]. They will automatically form multi-hop ad hoc networks to communicate both among themselves and to remote base stations (e.g., PDA's).

Because distributed micro-sensing involves direct interaction with a physical environment, data communication in sensor networks often has timing constraints in the form of *end-to-end deadlines*. Surveillance may require the position of an intruder be reported to a command center within 15 sec so that pursuing actions can be initiated in time. Data in a system may have different deadlines due to different validity intervals. The validity intervals (and hence, update deadlines) of the locations of different intruders such as pedestrians and motor vehicles may depend on their movement speeds. For example, locations of tanks have shorter update deadlines than those of pedestrians. Similarly, the location of an intruder should have a shorter update deadline than the temperature measurement of a region because the former can change faster than the latter. Sensor network protocols should support *real-time communication* by minimizing the packet *deadline miss ratio*, i.e., the percentage of packets that meet their end-to-end deadlines.

While sensor networks share the notion of timing constraints with more traditional embedded systems, they differ in two respects. First, individual sensors are typically very small in size and resource capacity. Hence, the philosophy of sensor networks relies on resource dedication rather than sharing. In other words, individual sensor devices and nodes are likely to be dedicated for individual tasks, thereby eliminating much of the need for sophisticated CPU scheduling in a multitask environment.

Second, it is envisioned that sensor nodes will operate in groups, since individual nodes are too limited and unreliable to perform useful activities from the application's perspective. Group activities require coordination and communication among member nodes. Sensing results of groups need to be sent back to base stations through multi-hop communication. Thus, the main schedulable resource becomes the wireless com-

munication channel. Progress of user-level activities and their ability to meet end-end deadlines are therefore determined by scheduling of the communication medium rather than scheduling of the processor. Towards that end, new real-time communication architectures are required for ad hoc wireless environments.

Multi-hop wireless communication scheduling differs from CPU scheduling in that it has an inherent notion of distance. In sensor networks, the distance is determined by the physical locations of source and destination. These locations impose distance constraints on messages, in addition to time constraints, calling for communication scheduling policies that are cognizant of both time and space.

The first contribution of this paper is RAP, a real-time communication architecture for large-scale wireless sensor networks. RAP provides a set of convenient, high-level query and event services to real-time distributed micro-sensing applications. Query and event services are based on novel location-addressed communication models supported by a scalable and lightweight network stack.

The second contribution of RAP is a novel *Velocity Monotonic Scheduling* (VMS) policy suitable for packet scheduling in sensor networks. VMS is based on a notion of packet *requested velocity*. Each packet is expected to make its end-to-end deadline if it can move toward the destination at its requested velocity, which reflects its local urgency. Compared with non-prioritized packet scheduling, VMS improves the deadline miss ratios of sensor networks by giving higher priority to packets with higher requested velocities. VMS can outperform deadline-based packet scheduling because velocity more accurately reflects the local urgency at each hop when packets with the same deadline have different distances to their destinations. Assuming that each sensor knows its own location (using GPS or other location services [8]), the requested velocity can be determined locally. This property enables VMS to scale well in large-scale sensor networks.

The final contribution of this paper is a detailed simulation study of the real-time performance of several routing protocols and packet scheduling algorithms in a typical sensor network scenario. Our simulation experiments demonstrate that, for sensors far away from their base station, RAP reduces the deadline miss ratio from 90.0% to 17.9%, compared to existing wireless communication protocols (DSR over 802.11b). To our best knowledge, ours is the first detailed performance study on deadline issues in multi-hop wireless sensor network settings under overload conditions.

In the following sections, we discuss the key characteristics of sensor networks, present the design of RAP, report a set of simulation experiments with sensor network configurations, and conclude the paper by summarizing our key results and future work.

2. Real-time Communication in Sensor Networks

In this section, we describe the characteristics of sensor networks and communication models on sensor networks. This analysis serves as a basis for our design of real-time communication protocols.

Sensor networks are an instance of mobile ad hoc networks (MANET) [10] that have recently attracted a lot of interest and visibility due to their flexibility and the feasibility of their deployment at low costs. In general, mobile ad hoc networks depend on peer-to-peer communication protocols that do not require a fixed infrastructure such as centralized servers and access points. Sensor networks are different from their traditional ad hoc wireless counterparts (composed of laptops and PDA's) in that they have a larger scale, higher density, smaller devices, and a tighter interaction with a physical environment. Energy conservation is critical in sensor networks because of their long lifetime and form-factor constraints, which preclude the use of large batteries or power supplies.

In most envisioned sensor network applications, a large number of sensors are deployed in an area and a small number of more powerful nodes (such as PDA's with Internet connections) form possibly mobile interface stations which serve as the entry points to the sensor network. In the following, we shall call such interface stations, *base stations*. A user may query the physical environment through such base stations. Alternatively she may also register for an *event*. The occurrence of the event will automatically trigger a specified query. A query can specify timing requirements including rate, start time, duration, and end-to-end deadlines. For example, a user can register for a *virus_found* event in a rectangular area with coordinates (10,10,20,20), and specify a query on the event to report the density of the detected virus. If a virus is found, the density of the viruses should be reported to the base station from where they are found every 1.5 sec for a duration of 30 min. Every reading should reach the base station within an end-to-end deadline of 5 sec.

Communication in a sensor network can be divided into two categories: *local coordination* and *sensor-base communication*. Before sending information to the base station, sensors within the local area coordinate among

themselves to aggregate data and generate a reliable result. For example, acoustic sensors may need triangulation among multiple nodes to decide the location of a tank. Local coordination often occurs within a distance of one or a few radio radii. Sensor-base communication is responsible for reporting the aggregated data to the base station, which often spans many (e.g., tens of) hops. Consider a communication radius of 30 m of short-range radios transceivers, it is conceivable to have more than 10,000 nodes and tens of hops of communication in a coverage area of several square kilometers. Since sensor-base communication typically travels a much larger number of hops than local coordination messages, in this paper we focus on the former type of communication.

Unlike IP networks, sensor-base communication directly uses *location* as the target address. Instead of querying a sensor with ID 1002, a user or application queries a geographic region. The identities of sensors that happen to be located in that region are not important. Any sensors in that region that receive the query may initiate local coordination to aggregate the requested data. A leader may be elected to send the query result back to the base station. If continuous monitoring is required, the query may report the desired measurement periodically through the multi-hop ad hoc network. The base station can attach its location to the query message so that the query results can also be addressed by location (assuming no two base stations are at a same location).

Communication in sensor networks can suffer from “hot regions”, i.e., areas where the network is seriously congested. Hot regions are often caused by numerous related events that synchronously trigger a large number of data flows toward the base station. Examples of related events include correlated measurement of the same environmental activity, or correlated environmental activities such as a group of new targets simultaneously entering a security area, or a bio-attack on a part of an airport. Maximizing the number of packets that make their deadlines in overload conditions is critical in sensor networks.

3. Design of RAP

We now present the design of RAP to support real-time communication in large-scale sensor networks. Given the unique characteristics of sensor networks, the goal of RAP includes the following:

- Provide general service APIs that are suitable for distributed micro-sensing and control in sensor networks

- Maximize the number of packets meeting their end-to-end deadlines
- Scale well with large number of nodes and hops
- Introduce minimum communication and processing overhead.

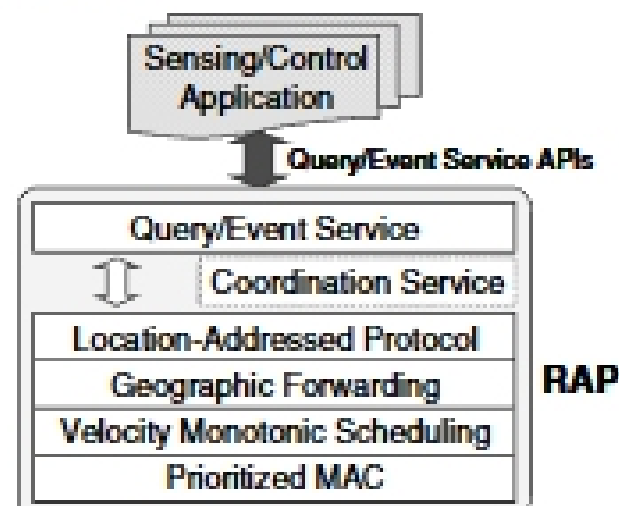


Figure 1 The RAP communication architecture

The architecture of RAP is shown in 1. Sensing and control applications interact with RAP through a set of Query/Event Service APIs. A Query/Event Service layer submits the query or event registration to an area. The Query/Event Service at the sensors in that area then (periodically or aperiodically) sends query results back to the base station. If an event is registered, the query is started only if the registered event happens. The sensor-base communication is supported by a network stack including a transport-layer Location-Addressed Protocol (LAP), a Geographic Forwarding (GF) routing protocol, a Velocity Monotonic (packet) Scheduling (VMS) layer, and a prioritized MAC. This network stack embodies a set of efficient and localized algorithms to reduce the end-to-end deadline miss ratio of sensor-base communication. This network stack is the focus of this paper.

The coordination service is responsible of dynamic group management and data aggregation among sensors (e.g., multiple sensors coordinate to determine the location of a target through triangulation). The coordination services are part of our on-going research and not addressed in this paper.

We now describe the Query/Event service APIs and the network protocol stack in detail in the following subsections.

3.1. Query/Event Service APIs

Applications may submit queries or register for events through a set of query/event service APIs. The API provides a high-level abstraction to applications by hiding the specific location and status of each individ-