Towards Real-time Event-based Communication in Mobile Ad Hoc Wireless Networks

Barbara Hughes
Distributed Systems Group
Department of Computer Science,
Trinity College, Dublin Ireland
Barbara.Hughes@cs.tcd.ie

Vinny Cahill
Distributed Systems Group
Department of Computer Science,
Trinity College, Dublin, Ireland
Vinny.Cahill@cs.tcd.ie

Abstract

Most previous work on real-time event-based communication has assumed infrastructure-based networks with stationary components. In contrast, ad hoc wireless networks comprise sets of mobile nodes connected by wireless links that form arbitrary wireless network topologies without the use of any centralized access point. Such highly mobile, dynamic networks do not satisfy the design assumptions for previous real-time event-based communication.

In this paper we propose a conceptual model for real-time event-based communication in mobile ad hoc wireless networks. Our model is the first to address the issue of achieving timeliness and reliability for real-time event-based communication in dynamic mobile ad hoc wireless networks.

1. Introduction

The inherent loose coupling that characterizes applications in a wireless ad hoc network promote event-based communication as a natural design abstraction for a growing class of software systems [1]. The event-based communication model is well suited to addressing the requirements of wireless mobile computing applications [2] and it is supported by several middleware services that use the event paradigm as a high-level communication abstraction. The underlying assumption of most of these services is that application components are stationary and that a fixed network infrastructure exists to facilitate communication[1-4]. The complexities introduced by the mobile ad hoc wireless network model, e.g. instantaneous topology changes, are not considered. For event-based communication to scale to mobile ad hoc wireless networks it is important that the designs of middleware services are not based on many of the assumptions made for fixed infrastructure networks, such as low latency, abundant bandwidth, static topology, and most importantly centralized control [1].

With the increased research in ad hoc networks in recent years new application domains such as communication between mobile robots and inter-vehicle communication have evolved. Timely communication is essential to allow applications in these domains to be realized. The challenge remains to extend real-time event-based communication to mobile nodes in a dynamic wireless ad hoc network where the assumption of a fixed infrastructure is not applicable.

In this paper we discuss the assumptions upon which existing (real-time) event-based communication models rely. We identify the characteristics of ad hoc wireless networks and how these characteristics inhibit real-time event-based communication. We propose a conceptual model for real-time event-based communication that relies on predictive techniques to alleviate the impediments to real-time event-based communication that are characteristic of a mobile ad hoc wireless environment.

In the next section we review the assumptions and subsequent design decisions upon which existing real-time event-based communication models rely.

2. Fixed Infrastructure Assumptions

For real-time event-based communication models to be used in mobile ad hoc wireless networks it is important for their designs not to be based on many of the assumptions made for infrastructure-based networks, both wired and wireless. The assumptions of particular interest are:

2a. Accessibility

In event-based communication an implicit assumption of known connectivity, in the absence of the failure of network components underlies the development of intermediate components such as event channels, [3, 4] or event dispatchers, [1, 2] which are often run independently and remotely to event producers and consumers. In these models the accessibility of the intermediary components is both assumed and critical, for all entities participating in the event-based communication.

A serious impediment of ad hoc wireless networks is the limited area that can be covered by mobile application components using a wireless transmitter. In an ad hoc wireless network entities may be distributed over a potentially large geographical area and thus are unlikely to be able to maintain a permanent communication link to the intermediary [5]. Also the mutual accessibility of event producers and consumers cannot be assumed. The design of an event-based communication model for mobile ad hoc wireless networks necessitates the omission of intermediate components providing system-wide services.

2b. Known upper bound on the number of participating nodes

Real-time event-based communication models for fixed infrastructure networks often assume a known maximum number of nodes connected to the physical medium [4]. In contrast, ad hoc wireless networks are created ‘on-the-fly’. This feature of wireless communication, coupled with the dynamic mobility of wireless nodes to move within range of another wireless node, means there is potentially no upper bound on the number of nodes participating in an ad hoc wireless network. The unbounded network size leads to scalability issues due to the increased computational load and difficulties of propagating network topology updates within given time bounds, increasing the unpredictability of wireless connections and timely and accurate route and resource reservation decisions all of which effect time-bounded event transmission and propagation.

2c. Known resource requirements

In addition to the previous assumption of a known upper bound on the number of participating nodes, it is assumed that there are known resource requirements for event-based communication among the participants. For example, the TAO Real-time Event Service [6] and the Real-time Event Channel Model for the CAN-Bus [4], use this assumption to perform real-time medium-access scheduling off-line using a reservation-based scheme to avoid collisions by statically planning the transmission schedule. In contrast, mobile ad hoc wireless networks require a dynamic resource reservation scheme to handle the effects of dynamic mobility where resource requirements are not known in advance.

Summary The dynamic mobility of nodes in ad hoc wireless networks renders the assumptions of event-based communication for infrastructure networks inappropriate. STEAM [5] addresses some of the fundamental issues arising for event-based communication among mobile ad hoc wireless nodes. Our work will extend STEAM to provide real-time capabilities.

3. Impact of Ad Hoc Networks

In this section we discuss the impact of wireless characteristics on real-time event-based communication.

3a. Dynamic connectivity

The absence of a fixed infrastructure means that mobile nodes themselves constitute the communication infrastructure. As nodes move in and out of range of other nodes, the connectivity and network topology changes dynamically [7].

Communication between mobile nodes requires the received signal strength (RSS) to be adequate to connect to another mobile node. The RSS is significantly affected by the movement of wireless nodes, terrain configuration [8] which may include hilly or mountainous areas, wooded or forested rural areas, urban areas with multistory buildings or low-density suburban areas and transmission power which may depend on battery life [9]. The changes in RSS lead to highly unpredictable connections between mobile nodes.

Unlike fixed infrastructure networks where link failures are comparatively rare events, the rate of link failure due to node mobility is the primary obstacle to routing in ad hoc networks [10]. Greater mobility increases the fluctuations in link connectivity, the volume of topological updates, the time spent processing the updates (e.g. for route discovery protocols), and congestion due to increased update transmissions and retransmissions. Link failures may also result in network partitions.

The topology changes introduced by node mobility and wireless link failures must somehow be communicated to other nodes. Since wireless and computation resources, for example bandwidth and battery power, are limited in wireless ad hoc networks, any overhead must be kept to a minimum and additional communication delays due to an increase in the volume of topological updates must be avoided.

Topology updates throughout an ad hoc network cannot happen instantaneously. The time-varying capacity of wireless links, limited resources and node mobility make maintaining accurate routing information very difficult, if not impossible, in ad hoc wireless networks. Routing for real-time event-based communication must ensure resource availability (e.g. bandwidth) whilst maintaining minimum latency [11]. Routing decisions may be compromised by inaccurate network information and time-bounded route determination, where optimal routes may not be found within the time available [12]. Decisions based on inaccurate information have unpredictable consequences that may be critical for real-time event-based communication.

3b. Unpredictable latency

Minimizing end-to-end latency is critical to achieve the timeliness requirements of real-time event-based communication. We pay particular attention to medium access and routing latency, as both are specifically affected by the characteristics of ad hoc wireless networks. Wireless transmissions in ad hoc networks are broadcast through a shared physical communication channel. Collisions in wireless communications can be caused by simultaneous transmissions by two or more wireless nodes sharing the same frequency band, or the hidden terminal problem [13]. Collisions cause unpredictable latency for medium access that is unacceptable in real-time event-based communication. Time-bounded access is not achievable with a high probability in the presence of unpredictable collisions and retransmissions.

The lack of a fixed infrastructure and the limited power of wireless mobile nodes that limits the transmission range, means that wireless nodes are designed to serve as message forwarding nodes or routers if needed. The result is a distributed multi-hop network with a time-varying topology with typically short-lived routes. The unpredictable latency for route determination and medium access (encountered at each hop) makes an estimate of end-to-end delivery latency that is critical in real-time event-based communication, very difficult and with a high probability of being inaccurate.

The TBMAC protocol [14], reduces the probability of collisions by providing each wireless node with time-bounded access to the medium with a high probability. We will use TBMAC protocol to provide predictable medium access latency for real-time event-based communication and investigate the open issue of predictable multi-hop routing latency using TBMAC.

3c. Limited resource availability

In mobile ad hoc wireless networks the available bandwidth is very limited and some wireless devices have severe energy constraints. Hence, communication is an expensive operation in terms of bandwidth and energy consumption and any additional control packet overhead (e.g. resource reservation) must be kept to a minimum. Additional control packets increase the competition for network resources (e.g. bandwidth) for all transmissions. The relationship between transmission power and battery life has been investigated in [9], where Campbell et al. identify that a critical design issue for wireless ad hoc networks is the development of suitable communication architectures, protocols and services that efficiently reduce power consumption thereby increasing the operational lifetime of the wireless device. Increasing battery life will help reduce the number of link breaks caused by node failure, and thus the volume of topological updates, the competition for scarce resources, and communication unpredictability that impede real-time event-based communication.

Wireless characteristic	Caused by	Impact on real-time event-based communication	Benefit of prediction
Dynamic connectivity (frequent link failure and possible partition)	*Mobility* *Variable RSS* *Terrain* *Power*	*1. Increased volume of topological updates* *2. Increased use of wireless and computational resources* *3. Inaccurate network information*	*Mobility prediction, partition anticipation and predictive routing and resource reservation reduce reaction time to topological change by finding new routes in advance of the failure of existing routes. Coverage estimation is used to calculate the accuracy of network information.*
Unpredictable latency	*Collisions* *Multi-hop routing*	*1. Unpredictable medium access latency* *2. Unpredictable route discovery and end-to-end event delivery latency*	*Preemptive routing and proactive resource reservation, coupled with time-bounded medium-access control, e.g. TBMAC.*
Limited resources	*Limited Bandwidth* *Limited power*	*1. Unpredictable medium access* *2. Limited event packet size* *3. Increase in link failures due to node failure*	*Power-aware and preemptive routing coupled with in-band signaling and piggybacking to reduce control overhead impact on the network*

Table 1: Using prediction to reduce the impediments of ad hoc wireless network

Summary To achieve real-time event-based communication in a mobile ad hoc wireless environment the impact of the ad hoc wireless network characteristics previously identified must be limited. We propose that prediction is essential to limit these wireless ad hoc characteristics. Table 1 introduces how predictive techniques limit the impediments of the wireless ad hoc environment for real-time event-based communication.

4. Proposed Conceptual Model

To achieve real-time event-based communication in a dynamic mobile ad hoc wireless network, the unpredictability inherent in the environment must be reduced. In this section we outline a conceptual model based on prediction for timely event-based communication between mobile nodes. Figure 1 identifies the components and high-level interactions among them.

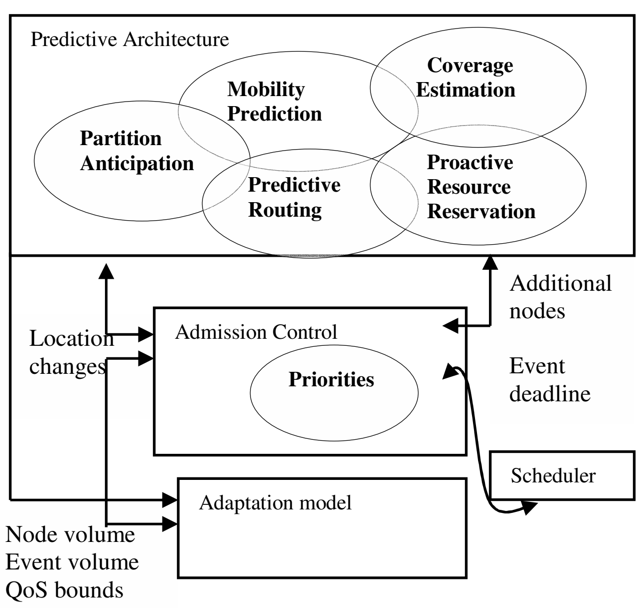

Figure 1. Conceptual model

Bounding the area of interest in an ad hoc network makes a large network appear smaller, but more importantly for real-time communication, it makes a highly dynamic topology appear less dynamic. Our approach is to partition the network into dynamically organized zones, similar to proximity groups [15], which bound the number of participants, the area for maintaining topology information and the area within which event information is propagated. The components of the conceptual model cooperate to maintain the timeliness and reliability requirements within a proximity-bounded zone.

The focus of our design is to reduce reaction to dynamic mobility and topological change by prediction. The admission control and adaptation components interact with the predictive architecture to make proactive decisions in advance of network change.

Predictive Architecture: predictive techniques are used to reduce the impact of dynamic topological changes within a zone. Location-awareness is key to determining the mobile nodes within a zone at a point in time. In our opinion location-aware routing [16] is central to achieving proximity-bounded communication in a mobile network. We plan to extend this work to predict the future location of mobile nodes. Using this information future node movement into a zone, the impact on routing, resource reservation, and guarantees for timeliness and reliability for the zone can be predicted in advance.

The ability to predict node movement contributes to achieving probabilistic guarantees of path availability due to link failure caused by node mobility. Other reasons why a link may fail, such as environment conditions or battery usage, must also be considered to avoid or anticipate network partitions. Using partition anticipation based on [15] coupled with proactive and preemptive routing [10, 11, 17] and resource reservation, we aim to improve the re-routing process by attempting to find new paths prior to the failure of existing ones. To obtain mobility independent real-time guarantees, a mobile host would need to make advance resource reservations at predicted locations they may visit during the lifetime of the communication. Accurate mobility and location prediction is critical for limiting the overhead of excessive resource reservation.

Admission Control: bounding the area of interest for real-time event-based communication implicitly limits the number of participating nodes to those within the bounded area. We apply explicit admission control policies within the zone to further reduce the number of participating nodes and the hop distance for routes within the zone [17]. The admission control policies reflect the impact of the number of participating nodes in the zone on the timeliness and reliability guarantees for a real-time event given the resources available when the real-time event is raised.

Using predictive techniques to detect future node movement is essential for deciding the admission policy to use. For example, if resource usage is nearing maximum capacity what temporal and reliability guarantees can be made for future nodes moving into the zone and what impact does the class of real-time event have on admission control decisions?

Adaptation Model: an important aspect of achieving timeliness constraints is dependable QoS adaptation [18]. However in contrast to [18] mobility is a critical consideration. The predictive architecture detects topological changes and initiates proactive routing and resource reservation. QoS adaptation may be necessary to reflect the new routes and resources available. The speed of node movement and the class of event for delivery impact the urgency of time-bounded delivery of a real-time event and impacts any QoS adaptation. Information from the predictive architecture is essential for limiting the reactive QoS adaptation required.

We propose a conceptual model to make the dynamic topology of mobile ad hoc networks less dynamic and therefore more suitable to real-time event communication. We propose that prediction is essential to reduce the reaction to dynamic node mobility and therefore essential for real-time event-based communication in wireless ad hoc networks.

5. Conclusion and Future work

A critical requirement for our future work is to determine the cost of prediction, in terms of increased overhead and the consequences of wrong predictions on real-time event-based communication. Our future work must also determine the impact of the wireless application on the ability to predict. For example, a wireless car has a known speed and trajectory limited by the vehicle and the road, which are known parameters that can increase the probability of correct prediction.

We have outlined our approach to the complex problem of achieving real-time event communication in infrastructure-free wireless networks. We outlined the limitations of previous event-based and real-time event communication models when applied to the ad hoc wireless domain. We described a predictive architecture for limiting the unpredictability of wireless communication by predicting node mobility, link failure and for anticipating partitions. Using this predictive architecture and QoS adaptation strategies we have proposed a novel approach to achieving real-time event-based communication in ad hoc wireless networks.

Acknowledgements

The work described in this paper was partly supported by the Future and Emerging Technologies programme of the Commission of the European Union under research contract IST-2000-26031 (CORTEX – CO-operating Real-Time sentient objects: architectures and EXperimental evaluation).

The authors are grateful to past and current colleagues at Trinity College Dublin including René Meier, Raymond Cunningham and Stefan Weber for their valuable input.

REFERENCES

1. A. Carzaniga, "Design and Evaluation of a Wide-Area Event Notification Service". *ACM Transactions on Computer Systems*, *19* (3). pp. 332-383, August 2001.
2. G.Cugola, E. D. Nitto and A. Fuggetta, "The JEDI event-based infrastructure and its application to the development of the OPSS WFMS". *IEEE Transactions on Software Engineering*, *27* (9). pp. 827-858, 2001.
3. "CORBAservices:Common Object Services - Event Service Specification v1.1", Object Management Group, March 2001.
4. J Kaiser and M. Mock, "Implementing the Real-Time Publisher/Subscriber Model on the Controller Area Network (CAN)". in *2nd IEEE International Symposium on Object-oriented Real-Time Distributed Computing*, Saint-Malo, France, May 1999.
5. R. Meier and V. Cahill, "STEAM: Event-Based Middleware for Wireless Ad Hoc Networks". in *Proceedings of the International Workshop on Distributed Event-Based Systems (ICDCS/DEBS'02)*, Vienna, Austria, 2002.
6. D. Schmidt, D. Levine and S. Mungee, "The Design of the TAO Real-Time Object Request Broker". *Computer Communications Special Issue on Building Quality of Service into Distributed Systems* (Elsevier Science), 1998.
7. K. Wang and B. Li, "Efficient and Guaranteed Service Coverage in Partitionable Mobile Ad-hoc Networks". in *IEEE Joint Conference of Computer and Communication Societies (INFOCOM'02)*, New York City, New York, June 23-27 2002, pp. 1089-1098.
8. H. A. Karimi and P. Krishnamurthy, "Real-Time Routing in Mobile Networks Using GPS and GIS Techniques". in *34th Hawaii International Conference on System Sciences*, Maui, Hawaii, January 3-6, 2001.
9. J. Gomez, A. T. Campbell, M. Naghshineh and C. Bisdikian, "Conserving Transmission Power in Wireless Ad Hoc Networks". in *9th International Conference on Network Protocols (ICNP 2001)*, Riverside, California, November 11 - 14 2001.
10. A. B. McDonald and T. Znati, "A Mobility Based Framework for Adaptive Clustering in Wireless Ad-Hoc Networks". *IEEE Journal in Selected Areas in Communications (JSAC)*, *17* (8). pp. 1466-1487, August 1999.
11. T. Goff, N. B. Abu-Ghazaleh, D. S. Phatak and R. Kahvecioglu, "Preemptive Routing in Ad Hoc Networks". in *7th Annual International Conference on Mobile Computing and Networking*, July 2001, pp. 43-52.
12. S-B. Lee and A. T. Campbell, "INSIGNIA: In-band Signaling Support for QOS Mobile Ad Hoc Networks". in *5th International Workshop on Mobile Multimedia Communications (MoMuC, 98)*, Berlin, Germany, October , 1998.
13. F. Tobagi and L. Kleinrock, "Packet Scheduling in Radio Channels: Part ii-The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy-Tone Solution". *IEEE Transactions on Communications*, *23* (12). pp. 1477-1433, December, 1975.
14. R. Cunningham and V. Cahill, "Time bounded Medium Access Control for Ad Hoc Networks". in *Principles of Mobile Computing (POMC'2002)*, Toulouse, France, October 30-31, 2002.
15. M. O. Killijian, R. Cunningham, R. Meier and V. Cahill, "Towards Group Communication for Mobile Participants". in *Principles of Mobile Computing (POMC'2001)*, Newport, Rhode Island, USA, 2001, pp. 75-82.
16. Y. B. Ko and N. H. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks". in *International Conference on Mobile Computing and Networking (MobiCom'98)*, Dallas, Texas, USA, 1998, ACM/IEEE.
17. M.R. Pearlman and Z.J. Haas, "Determining the Optimal Configuration for the Zone Routing Protocol". *IEEE Journal on Selected Areas in Communication*, *17* (8). pp. 1395-1414, 1999.
18. A. Casimiro and P. Verissimo, "Using the Timely Computing Base for Dependable QoS Adaptation". in *20th IEEE Symposium on Reliable Distributed Systems*, New Orleans, USA, October 2001, pp. 208-217.