Achieving Real-time Guarantees in Mobile Ad Hoc Wireless Networks

Barbara Hughes and Vinny Cahill
Distributed Systems Group,
Department of Computer Science, Trinity College, Dublin, Ireland
{barbara.hughes, vinny.cahill}@cs.tcd.ie

ABSTRACT

Timely wireless communication is essential to allowing real-time mobile applications, such as communication between mobile robots or inter-vehicle communication to be realized. The real-time event-based communication paradigm has been recognized as an appropriate high-level communication scheme to connect autonomous components in large distributed control systems [1]. We investigate whether real-time event constraints can be guaranteed in a mobile ad hoc wireless network.

In this work in progress paper we present our analysis of the impact of mobile ad hoc wireless networks on achieving real-time guarantees. We introduce our ongoing work on the use of a proactive routing and resource reservation protocol using mobility awareness and prediction to reduce the unpredictability of a dynamic mobile ad hoc wireless network.

1. INTRODUCTION

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 or infrastructure. Ad hoc wireless networks are inherently self-creating, self-organizing and self-administering [2].

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 real-time event-based communication paradigm has been recognized as an appropriate high-level communication scheme to connect autonomous components in large distributed control systems [1]. In this paper, the impact of mobile ad hoc wireless characteristics particularly dynamic mobility, dynamic connectivity and limited resource availability on real-time guarantees is analysed.

The structure of the paper is as follows: in the next section we discuss the main characteristics of mobile ad hoc wireless networks that impact guaranteed real-time event-based communication. We follow this by an analysis of how real-time guarantees may be achieved in a mobile ad hoc wireless network and finish with an introduction to our ongoing work, a novel proactive routing and resource reservation protocol that uses prediction and mobility awareness to reduce the dynamics of the mobile wireless environment.

2. IMPACT OF MOBILE AD HOC WIRELESS NETWORKS

Real-time event-based communication protocols must guarantee the timeliness and reliability constraints of real-time events by minimizing the packet deadline miss ratio, i.e. the percentage of packets that miss their end-to-end deadlines [3].

2.1 DYNAMIC MOBILITY

The absence of a fixed infrastructure means that nodes in an ad hoc network communicate directly with one another in a peer-to-peer fashion. The mobile nodes themselves constitute the communication infrastructure – a node acts as both a real-time event router and an end host. As nodes move in and out of range of other nodes, the connectivity and network topology changes dynamically [4].

Unlike fixed infrastructure networks where link failures are comparatively rare events, the rate of link failure due to node mobility and varying signal strength is the primary obstacle to routing in ad hoc networks [5]. Since the rate of link failure is directly related to node mobility, greater mobility increases the fluctuations in link quality, the volume of topological updates (e.g. for route discovery protocols), the time spent processing the updates, and congestion due to increased update transmissions and retransmissions. Unpredictable latency, for the discovery and maintenance of real-time routes may be catastrophic depending upon the constraints of the real-time event class. Link failures may result in network partitions, and the potentially critical situation that a hard real-time event cannot be propagated to all relevant[1] nodes.

The topology changes introduced by node mobility and wireless link failures must somehow be communicated to other nodes. Topology updates throughout an ad hoc network cannot happen instantaneously. Nodes may have inconsistent views of the network that may never be accurate [2]. Current QoS routing algorithms [6-8], require accurate link state (e.g., available bandwidth, packet loss rate, estimated latency etc.) and topological information. The time-varying capacity of wireless links, limited resources and node mobility make maintaining accurate real-time routing information very difficult, if not impossible, in ad hoc wireless networks [9]. Routing for real-time event-based communication must ensure resource availability (e.g. bandwidth) whilst maintaining minimum latency [10].

Minimizing end-to-end latency is critical to achieve the timeliness requirements of real-time event-based

[1] Relevant nodes are interested in the event and reside in the area where the event is applicable.

communication. Collisions cause unpredictable latency for medium access that is unacceptable in real-time event-based communication, where each mobile node must have time-bounded access to the wireless medium to transmit a real-time event. 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 and therefore transmission range of wireless mobile nodes, means that wireless nodes are designed to serve as routers if needed. The result is a distributed multi-hop network with a time-varying topology where routes are typically short-lived [11]. The latency involved in route determination might be quite significant and may be increased by the use of incomplete network information. The unpredictable latency for route determination and medium access (encountered at each hop) makes the estimation of end-to-end delivery latency, which is critical in real-time event-based communication, very difficult with a high probability of inaccuracy.

Any wireless node participating in real-time communication requires guaranteed medium access and routing latency. In section 3, we analyse the requirements for real-time event-based communication from the perspective of a real-time event producer, e.g. the propagation of real-time events to geographically dispersed consumers within a known time bound.

2.2 LIMITED RESOURCE AVAILABILITY

In mobile ad hoc wireless networks the available bandwidth is very limited and some wireless devices have severe energy constraints, relying for example on battery power [12]. Hence, communication is an expensive operation in mobile ad hoc wireless networks in terms of bandwidth and energy consumption and therefore any additional control packet overhead (e.g. resource reservation, routing and scheduling) must be kept to a minimum. Additional control packets increase the competition for network resources (e.g. bandwidth, medium access etc.) for all (control and data) transmissions. In addition, the routing and resource reservation protocol for guaranteed real-time class constraints might be limited by the capacity and power limitations of the wireless device. A trade-off may exist where the ability to guarantee real-time constraints is limited by the overhead involved. For example, the benefit of proactive routes and resource reservations for guaranteed event transmission might be reduced by the additional overhead to discover and maintain the routes, particularly in a limited resource environment. We discuss the impact of proactive decision-making on the probability of guaranteeing real-time class constraints in the next section.

To achieve real-time event-based communication in wireless ad hoc networks the impact of the network characteristics described must be reduced. In the next section we analyse how to reduce the impact of these characteristics to guarantee real-time class constraints.

3. ANALYSIS OF REAL-TIME EVENT CLASSES

We distinguish three classes of real-time event: hard real-time (HRT), soft real-time (SRT) and non real-time (NRT). Our objective is to guarantee the timely delivery of hard real-time events with a known probability, guarantee soft real-time events only if non-detrimental to hard real-time guarantees and provide best-effort delivery guarantees for non real-time events. The application scenario, i.e. the mobility, geographic dispersion and density of wireless nodes in the proximity bound for real-time event transmission, impact the real-time guarantees achievable. In our future work we will investigate constraining the application scenario, e.g. limiting mobility, bounding density, to analyse the impact on the real-time guarantees achievable.

In this section we analyse how to achieve these real-time class guarantees in a highly dynamic, resource constrained mobile wireless network. To enhance our analysis we extended traditional real-time classifications (i.e. hard, soft and non real-time) to include node mobility and event periodicity. Figure 1 is an example for the hard real-time (HRT) class, with event periodicity at the first level and mobility of the real-time event producer at the second. The extended classification accommodates a fine-grained analysis of the key characteristics that influence achieving real-time guarantees in a highly dynamic, mobile wireless ad hoc network.

Our analysis assumes an event-based communication model for ad hoc networks such as STEAM [13] is available for event propagation. STEAM exploits proximity bounded event transmission, involving both geographical and functional aspects, to allow interested wireless nodes to interact.

Our analysis is based on the availability of a TDMA style predictable MAC layer, such as TBMAC [14], to provide time-bounded and predictable medium access latency via time slots. In TBMAC a mobile node contends for a time slot and if successful has a time slot allocated for predictable and contention-free event transmission. The number of slots available is finite, and represents a resource limitation, in terms of bandwidth availability. Similar to [14], we assume a virtual cell structure overlays the wireless network and all nodes within a cell are within one-hop transmission range of each other and receive an event transmission within a known time bound. Intra-cell slots are used for transmission within a cell and inter-cell slots for transmissions between cells.

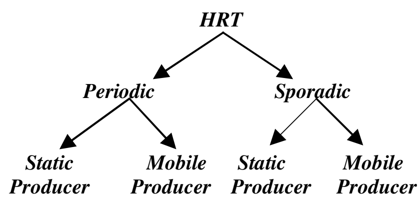

Figure 1: Real-time event classifications

3.1 HARD REAL-TIME (HRT) EVENTS

A hard real-time event must guarantee timeliness constraints for event propagation, implying time-bounded medium-access and routing latency. Guaranteed timeliness is a critical requirement of hard real-time applications, e.g. propagating changing traffic conditions to automated vehicles.

We first consider a static HRT event producer, for example a traffic light propagating a traffic signal change to all vehicles within the proximity of the traffic light. To guarantee predictable event transmission within a virtual cell the HRT event producer contends for a time slot allocation. A HRT event producer is allocated a slot only if there are sufficient unallocated slots available to satisfy the slot request, i.e. a simple admission control procedure must be passed. A prioritised slot allocation scheme guarantees HRT event producers have the highest probability of successful slot allocation. Time-bounded HRT event transmission in a single cell is guaranteed following a successful slot allocation i.e. a proactive resource (slot) reservation has been made.

The cost associated with this proactive slot reservation scheme is the potentially high volume of failed HRT slot requests (and therefore failed real-time event transmissions) because of excessive slots allocations for infrequently transmitted events, i.e. frequency of usage of an allocated slot is not considered. Thus, the trade-off of using proactive resource reservation to guarantee HRT events in a limited resource network is the potential reduction in the total volume of real-time events that satisfy deadlines.

We address this trade-off by adopting a passive reservation approach to proactive resource reservation, similar to MRSVP [15], i.e. reserved resources are available for use by other nodes until required by the original node. In real-time event-based communication, additional information about resource usage is available via the periodicity of the real-time event. A reserved resource is available for use in the interval between deadlines of a real-time event, when it is reserved but not being used. Using temporary transfer of ownership, other HRT or lower priority event producers may transmit an event in this interval, only if the resource owner's transmission is still guaranteed.

This proactive resource reservation scheme guarantees time-bounded event transmission to all wireless nodes in the same virtual cell, i.e. one-hop neighbours of a static event producer. We now focus on extending these guarantees to a mobile HRT event producer, e.g. an ambulance transmitting "yield right of way" events to vehicles within its' vicinity.

Following the same event transmission protocol, a mobile event producer must contend for a slot prior to event transmission. Mobility is a key factor to determine when this slot contention should be initiated. One possibility is to initiate slot contention when the event deadline is reached, i.e. the mobile HRT event producer contends with all other wireless nodes for a slot in the current virtual cell. Due to the minimum laxity available for successful slot contention and allocation in this case, the probability of sufficient slot availability (even using a passive reservation scheme) prior to the deadline of the HRT event is very low.

An alternative is to use mobility information, e.g. velocity, trajectory, and location information[2] to initiate proactive slot contention in a cell on the predicted route of the mobile event producer. This is a similar approach to MRSVP [15]. Mobility and particularly velocity, may affect the urgency of slot reservation. The faster the event producer is moving, the greater the physical distance traversed prior to event transmission. To proactively reserve a slot in a predicted cell requires transmission of a slot reservation request over a possibly great distance. Incorporating Velocity Monotonic Scheduling [3], priority can be attributed to this slot request which reflects the velocity of the mobile event producer.

Accurate mobility information is essential to reduce wrong predictions and excessive proactive resource reservation, and is achievable by combining the periodicity of event transmission with mobility information of the event producer. The time interval between event transmissions and the velocity of the event producer bounds the movement possible, thus limiting the candidate cells for predicted movement with increasing accuracy.

[2] Using GPS for example.

Our analysis so far has considered single cell (one-hop) HRT event propagation. Real-time multi-cell (multi-hop) routing requires time-bounded routing latency for all hops on the route. In terms of the event transmission protocol inter-cell slots are required for predictable and time-bounded event transmission to all cells on the route.

We propose a proactive routing protocol using prediction for proactive route maintenance and adaptation. Unlike the proactive routing protocols of [11] or [5], we establish proactive routes between multiple cells. The complete route for an HRT event within a defined proximity bound [13], is established in advance. We assume that if a route for guaranteed HRT event transmission can be discovered, and the route is proactively maintained, HRT event transmission will be guaranteed when the deadline occurs. This assumption relies on the continued population of all virtual cells on the proactive route, i.e. the inter-cell slots must remain allocated, and the real-time constraints of the route are maintained, i.e. an alternative route can be discovered that guarantees the HRT real-time constraints.

Proactive multi-hop route and resource reservation in a network with limited resource availability reduces the resources available at each hop. In this case, the slots available in each virtual cell on the route are reduced. Using event periodicity or event information accumulated over time, a passive reservation scheme similar to the single cell scenario is available with an additional constraint that slots must remain allocated in all cells on the route.

The volume of real-time events competing for limited resources also affects guaranteed HRT event transmission. HRT events should never fail because of lower priority events. Our objective is to guarantee HRT event constraints. In the remainder of this section we discuss the restrictions and limitations placed on soft and non real-time event transmissions to meet this objective.

3.2 SOFT REAL-TIME (SRT) EVENTS

Soft real-time events must satisfy timeliness constraints that may be violated under load and fault conditions without critical consequences, e.g. video streams for video-on-demand.

SRT events do not have the same criticality, and therefore priority, as HRT events. Prioritised slot allocation guarantees precedence to HRT events. A slot request by a SRT event producer will only be considered when all HRT event producers in the same virtual cell have a slot allocation. If a SRT event producer is successfully allocated a slot, event transmission using this slot is still not guaranteed. A temporary transfer of slot ownership is initiated if there are any HRT event producers in the virtual cell, without a slot allocation and with a HRT event deadline approaching. Dynamically transferring slot ownership to higher priority real-time events may mean that SRT events are continually pre-empted by HRT events and subsequently never transmitted.

Proactive routing and resource reservations are not feasible for SRT events, as the decision to dynamically transfer slot usage can only be determined when the SRT event deadline is reached and the real-time class of other event producers without slots are known. Proactive slot contention to minimise the slot contention latency of a mobile SRT event producer is also not beneficial. A mobile SRT event producer will only have a slot allocated if there are no HRT event producers contending for slots when the SRT deadline is reached. The event transmission protocol dynamically transfers temporary resource ownership to any higher priority event producers contending for a slot allocation. The real-time class of the event producers contending for slots when the SRT event deadline would occur could not be known in advance. Real-time event periodicity has reduced importance for SRT events and is now superseded by real-time event priority.

Maximising prioritised real-time event-based communication is our objective. HRT events always take precedence and pre-empt lower priority events, if there are insufficient resources remaining in the network to satisfy the HRT request.

3.3 NON REAL-TIME (NRT) EVENTS

Non real-time events do not have timeliness guarantees, e.g. the propagation of weather reports to moving vehicles. There is no guarantee that NRT event transmission will occur at all due to the prioritised slot allocation mechanism and the temporary transfer of slot ownership to higher priority events. NRT events will have a "best-effort" transmission policy, with the assumption that NRT event transmissions will never compromise higher priority events. We do not consider NRT events any further here.

4. FUTURE WORK

We have completed our analysis of achieving real-time class guarantees in mobile wireless ad hoc networks. We have proposed a prioritised event transmission protocol with temporary transfer of resource ownership. We have highlighted the role of proactive routing and resource reservation to increase the probability of satisfying hard real-time class guarantees.

We are currently finalising the design of a proactive routing and resource reservation protocol, based on mobility awareness and prediction, to achieve the guarantees of real-time classes we have analysed. The protocol combines proactive discovery of real-time constrained routes with proactive route maintenance and route adaptation to maximise robustness of the routes discovered. The protocol assumes predictable medium-access using TBMAC [14] and an event-based middleware for ad hoc networks, STEAM.

Following the completion of the protocol design, we will proceed to implement the protocol, initially in a simulated environment, with a real-world implementation using 802.11b to follow. We are interested in evaluating the protocol for metrics such as: route discovery latency, percentage of missed real-time deadlines and percentage of missed hard real-time deadlines.

5. CONCLUSION

We have outlined our work-in-progress on achieving real-time guarantees in a highly dynamic mobile wireless ad hoc network. We presented our analysis of achieving real-time constraints, which lead to the introduction of our proactive routing and resource reservation protocol. Our future work will include a detailed discussion of this new protocol with a simulated and real-world implementation to evaluate the ability of this protocol to guarantee timeliness constraints for real-time applications.

6. 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 these colleagues at Trinity College Dublin including René Meier, Raymond Cunningham and Stefan Weber for their valuable input.

REFERENCES

1. R. Rajkumar and M. Gagliardi, "The Real-Time Publisher/Subscriber Inter-process communication Model for Distributed Real-Time Systems: Design and Implementation". in *IEEE Real-time Technology and Applications Symposium*, May 1995.
2. S. Chakrabarti and A. Mishra "QoS Issues in Ad Hoc Wireless Networks". *IEEE Communications Magazine*, *39* (2). pp. 142-148. February 2001
3. Chenyang Lu, Brian M. Blum, Tarek F. Abdelzaher, John A. Stankovic and Tian He, "RAP: A Real-Time Communication Architecture for Large-Scale Wireless Sensor Networks". in *'8th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS02)*, San Jose, California, September 25-27, 2002.
4. 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.
5. 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
6. D. B. Johnson, D. A Maltz, Y-C. Hu and J. G. Jetcheva "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)," IETF Internet Draft. 21st February, 2002
7. C. Perkins, E. Royer and S. R. Das. "Quality of service for Ad Hoc On-Demand Distance Vector (AODV) routing," Internet Draft, July 2000.
8. C. E. Perkins and P. Bhagwat "Highly Dynamic Destination Distance-Vector Routing (DSDV)". *Computer Communications Review*. pp. 234-244. October, 1994
9. D. D. Perkins and H. D. Hughes "A survey on QOS support for mobile ad hoc networks". *Wireless Communications and Mobile Computing*, *2*. pp. 503 - 513. 2002
10. 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.
11. 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
12. M. S. Corson and A. T. Campbell, "Towards Supporting Quality of Service in Mobile Ad hoc Networks". in *1st IEEE Conference on Open Architecture and Network Programming (OPENARCH'98)*, San Francisco, April 3-4, 1998.
13. R. Meier and V. Cahill, "Exploiting Proximity in Event-Based Middleware for Collaborative Mobile Applications". in *Proceedings of the 4th International Conference on Distributed Applications and Interoperable Systems (DAIS'03)*, Paris, France, 2003, Springer-Verlag Heidelberg.
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. A. Talkudar, B. Badrinath and A. Acharya "MRSVP: A resource reservation protocol for an integrated services network with mobile hosts". *The Journal of Wireless Networks*, *7* (1). 2001