Improving the Performance of Reliable Transport Protocols in Mobile Computing Environments

Ramon Caceres and Liviu Iftode

IEEE Journal on Selected Areas in Communications, Vol. 13, No. 5, June 1995

Outline

- Introductory Comments on Mobile Networks
- Wireless Networking Test bed
- The Effects of Motion on TCP
- Approaches for Alleviating the Effects of Motion on TCP
- Wireless Transmission Errors
- Conclusion

Wireless and Mobile Networks

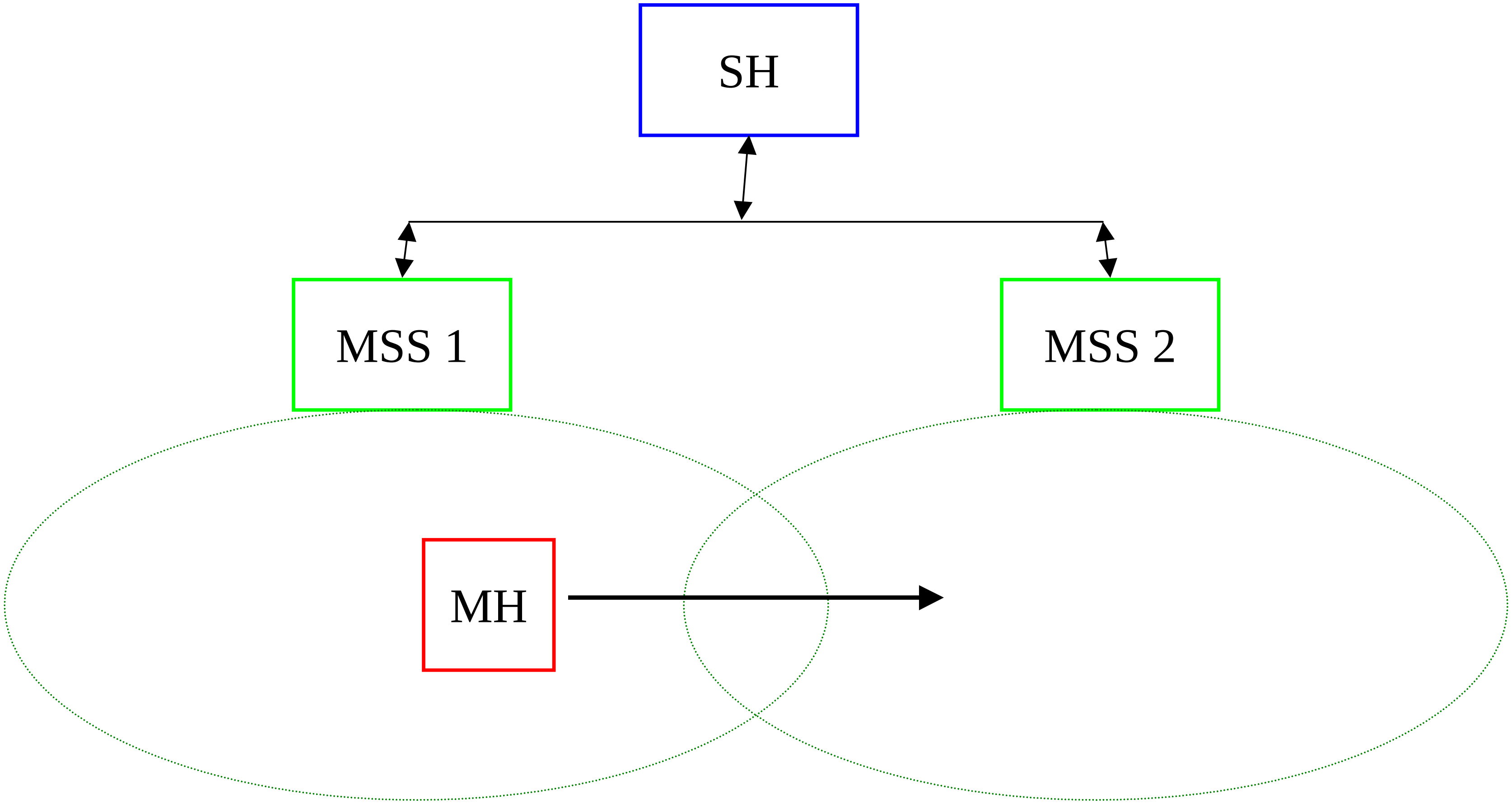
- This paper focuses on performance issues when a wireless, mobile station moves between between two wireless ***micro cellular*** networks.
- The authors consider carefully TCP communication **pauses** when the mobile host [MH] is handed off between two wireless cells.
- Since TCP uses packet loss as an *implicit* indicator of congestion, motion can easily be mistaken for congestion. This results in a significant reduction in throughput and unacceptable delays.

Objectives of the Paper

- To ***quantify*** the effects of motion on throughput and delay for mobile TCP hosts.
- To identify the <u>factors</u> that contribute to the performance loss.
- To suggest an "end-to-end"approach to alleviate the performance degradation.

Wireless Networking Test bed

- The test bed consists of mobile hosts [MH], mobile support stations [MSS], and stationary hosts [SH] deployed in an “ordinary” office environment.
- SH’s connect to a 10 Mbps wired Ethernet.
- MH’s connect to a 2 Mbps WaveLAN wireless LAN.
- MSS’s connect to both networks with one MSS per wireless cell. The MSS is responsible for the MH’s in its cell.
- All stations in the test bed use 4.3 BSD Tahoe.



Wireless Networking Testbed

Cellular Handoff Procedures

- The MSS's route packets from a MH in the wireless cell to SH's in the wired part of the network.
- MSS's make their presence known by broadcasting a beacon signal periodically over the wireless network.
- An MH switches cells either when it receives a stronger beacon signal from a new MSS or when it receives the first beacon signal from a new MSS after failing to receive a beacon signal from the old MSS.

Procedure for an MH to Switch Cells

The MH ::

- sends a ***greeting*** to the new MSS.
- updates its routing table to make new MSS its default gateway.
- sends the *identity* of the old MSS to the new MSS.

The new MSS ::

- ACKs the MH's ***greeting***.
- Adds the MH to its list of MH's.
- Begins to route the MH's packets.

Procedure for an MH to Switch Cells

The new MSS (cont.) ::

- Informs the old MSS that the MH has moved and can be reached via the new MSS.

The old MSS ::

- adjusts its routing table to forward packets for the MH through the new MSS.
- ACKs the handoff to the new MSS.

The new MSS:

- ACKs the completion of the handoff to the MH.

Experimental Methodology

- The primary action of the experiments is to initiate a reliable TCP data transfer between an MH and an SH; cause the MH to cross a cell boundary with the TCP connection active; and measure the performance of the connection.
- The ***motion across cell boundaries*** is **simulated.** This provides precise control over the instant that the handoff begins.
- This test bed setup permits the study of overlapping and non-overlapping cells and the full range of handoff scenarios.
- They only report results when data packets flow from the MH to the SH with ACKs going in the opposite direction.

The Effects of Motion

Four motion scenarios are simulated:

1. The MH does not move.
2. The MH moves between ***overlapping*** cells.
3. The MH moves between non-overlapping cells and receives a beacon from the new MSS at the instant it leaves the old cell (0 sec. rendezvous delay).
4. The MH moves between non-overlapping cells and receives a beacon from the new MSS one second after leaving the old cell (1 sec. rendezvous delay).

Loss of Throughput

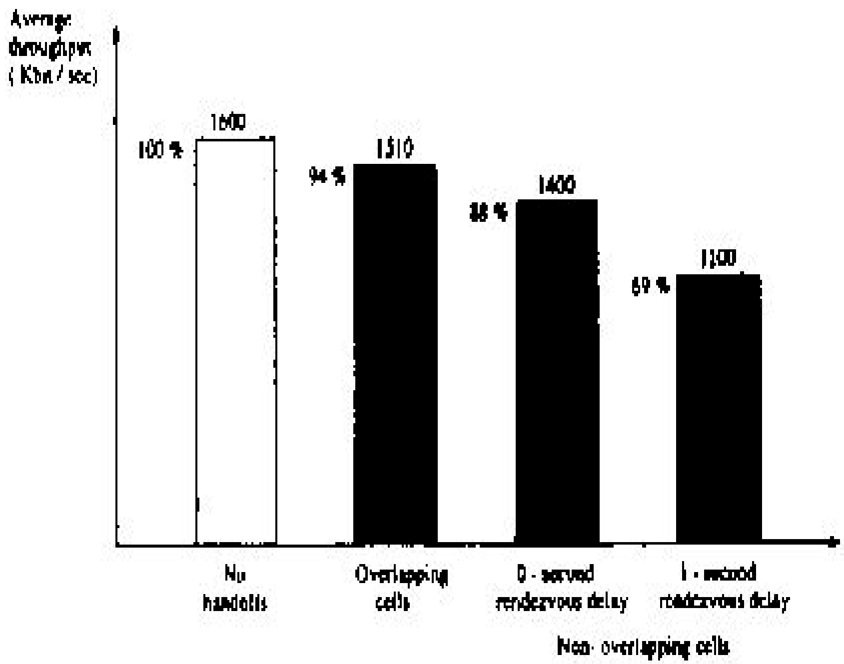


Fig. 2. Loss of throughput due to host motion.

- throughput degrades ***substantially*** in the presence of motion across non-overlapping cells.
  12% drop for zero rendezvous delay
  31% drop for 1-sec. rendezvous delay

Loss of Throughput

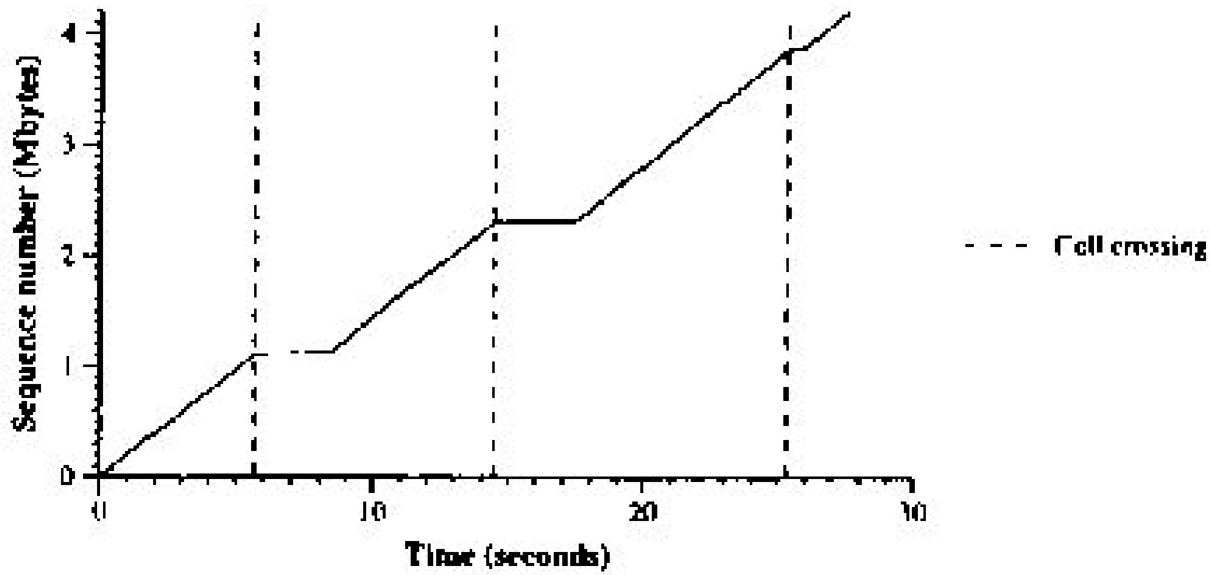


Fig. 3. Behavior of TCP sequence number in response to cell boundary crossings.

1 sec rendezvous delay scenario

Note – 3 second pause after first two handoffs.

TCP comes to a halt and does not transmit data during these pauses.

Packet Loss

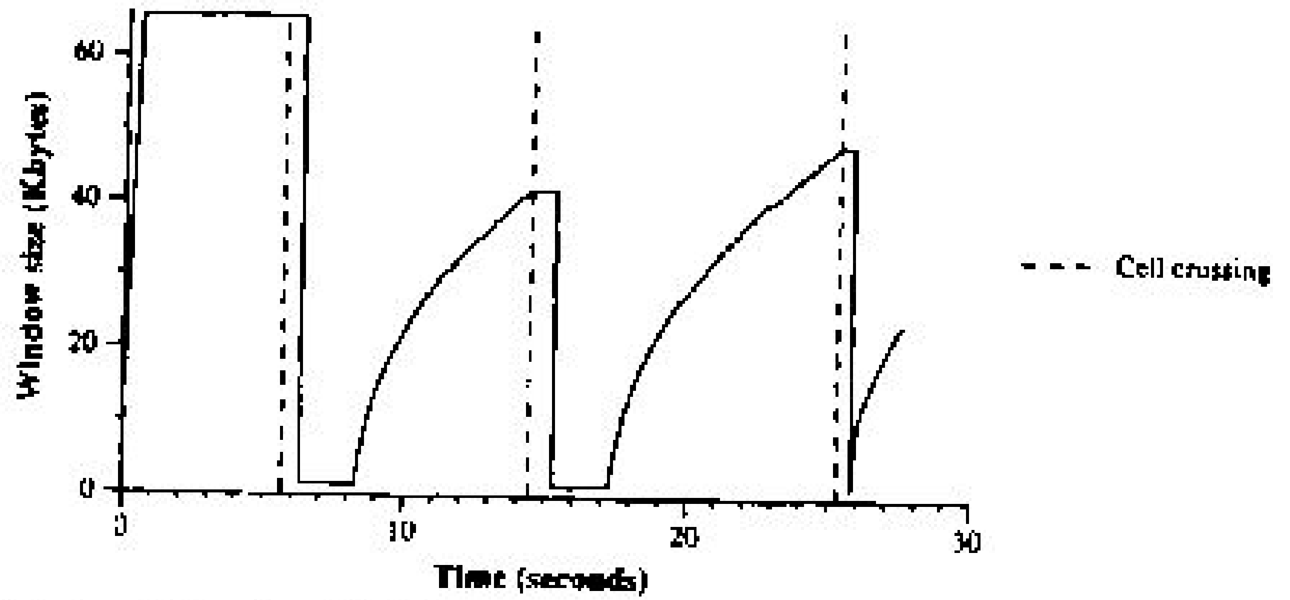


Fig. 4. Behavior of TCP congestion window in response to cell boundary crossings.

1 sec rendezvous delay scenario

The congestion window ***stops growing*** with each cell crossing.

There is a ***pause*** before slow-start begins.

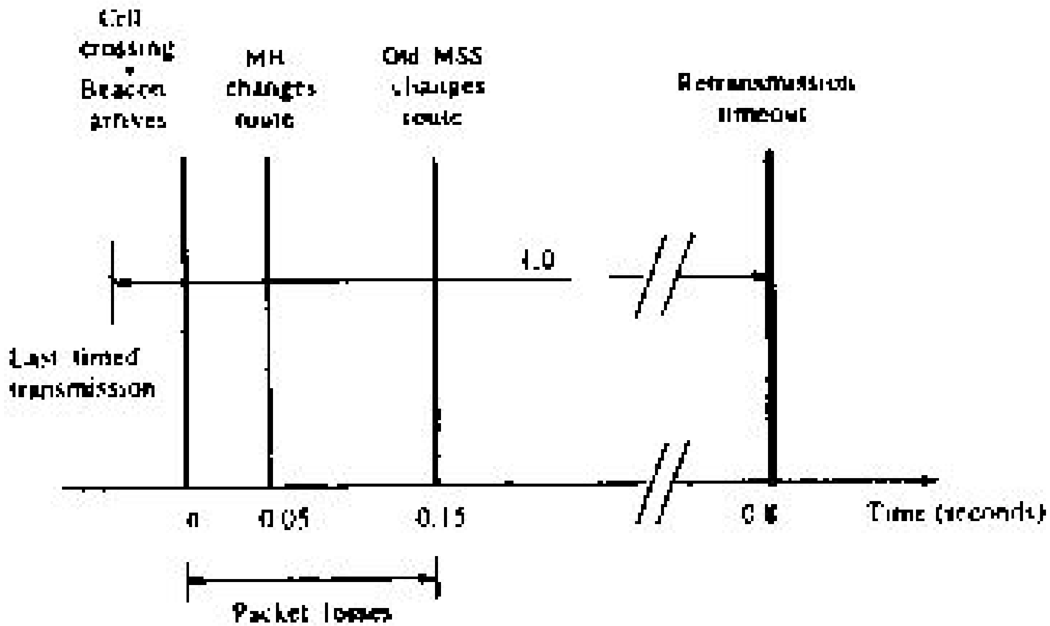


Fig. 5. Handoff latency and related packet losses with a 0-s rendezvous delay.

0 sec rendezvous delay scenario

Handoff latency (which causes packet loss) cannot be completely eliminated because <u>at least</u> two packet exchanges are needed to notify both the new and the old MSS that the MH has changed cells.

An active TCP connection loses up to a full transmission window's worth of packets and related ACKs during each handoff.

Packet Loss

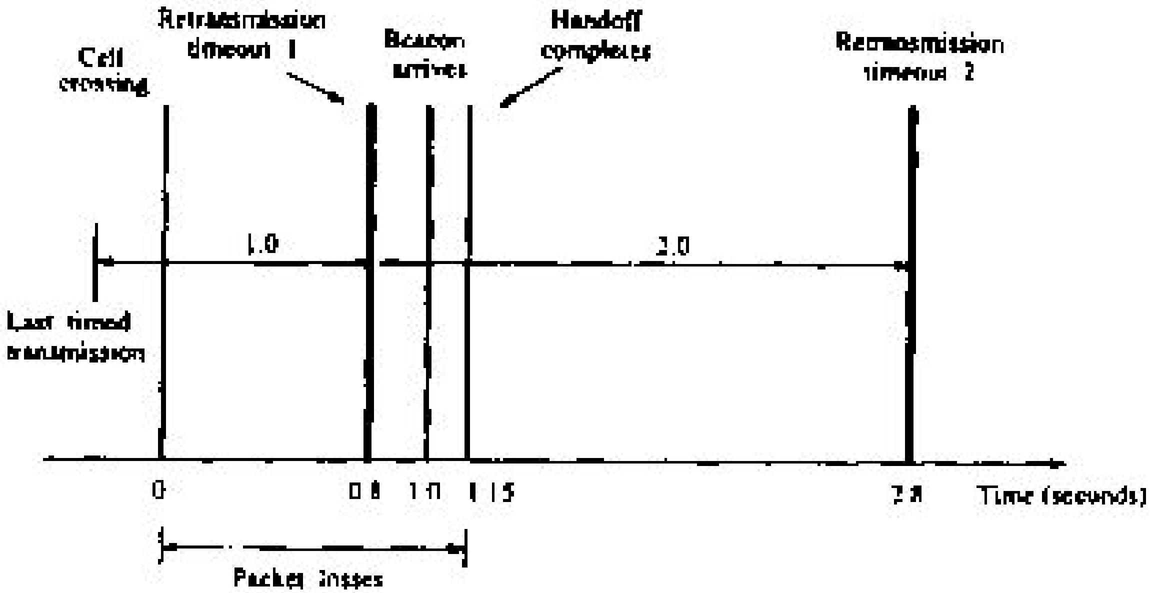


Fig. 6. Handoff latency and related packet losses with a 1-s rendezvous delay.

1 sec rendezvous delay scenario

Two consecutive TCP timeouts are typical for a 1 sec rendezvous delay.

The timeout freezes TCP for 0.8 sec. or more with each cell crossing.

Slow Recovery and Unacceptable Interactive Delay

- TCP slow-start contributes to slow recovery after a cell crossing.
- This, in turn, contributes to loss of throughput {authors claim: only moderate throughput loss}.
- Pauses grow exponentially with growing rendezvous delays.
- Pauses persist from 650 ms to several seconds after the handoff completes.

Alleviating the Effects of Motion

Two ***general classes*** of solutions are proposed:

1. Hiding motion from the transport layer.

- *Provide smooth handoffs during cell crossings.*

2. Adapting the transport layer to react better to motion.

- *Employ more accurate retransmission timers.*
- *Use fast retransmissions.*

Smooth Handoffs

- Implement ***"make then break" handoffs*** ::

  handoffs that are completed before an MH loses contact with the old MSS.

- Have the MSS buffer enough packets recently send by an MH to handle the maximum handoff latency between two cells.

  When the MSS knows that the MH has moved to another cell, the MSS sends the buffered packets for that MH to the new MSS.

Reasons for Little or No Cell Overlap

- To use the same portion of the spectrum in nearby cells → higher aggregate capacity.
- To support low-powered mobile transceivers that meet stringent power consumption requirements.
- To provide accurate location information.

More Accurate Retransmission Timers

- Historically TCP implements coarse timers with a 300-500 ms resolution.
- Minimum RTO is twice the RTT.

  With coarse 500 ms timer, minimum RTO = 1 sec.

- Real RTT's can be less than 1 ms!!
- However higher-resolution timers, can increase the probability of ***multiple timeouts*** while a handoff completes.

More Accurate Retransmission Timers

Effects of multiple timeouts::

1. **ssthresh gets near 1 → early congestion avoidance region → many RTTs to reach high cwnd.**
2. **Higher potential for RTO backoff algorithm.**
3. **During handoff multiple retransmissions may be futile!**

Conclusion:: *abrupt changes in RTT delay due to cell handoffs is hard to handle via timer adjustment strategies.*

Fast Retransmissions

- Rather than wait for a timeout due to a handoff, modify the handoff procedure to trigger ***fast retransmit***.

Allow **IP on MH** to signal **TCP on MH** when a ***greeting ACK*** arrives from the new MSS. **TCP on MH** uses this signal to ***trigger a fast retransmit.***

Fast Retransmission

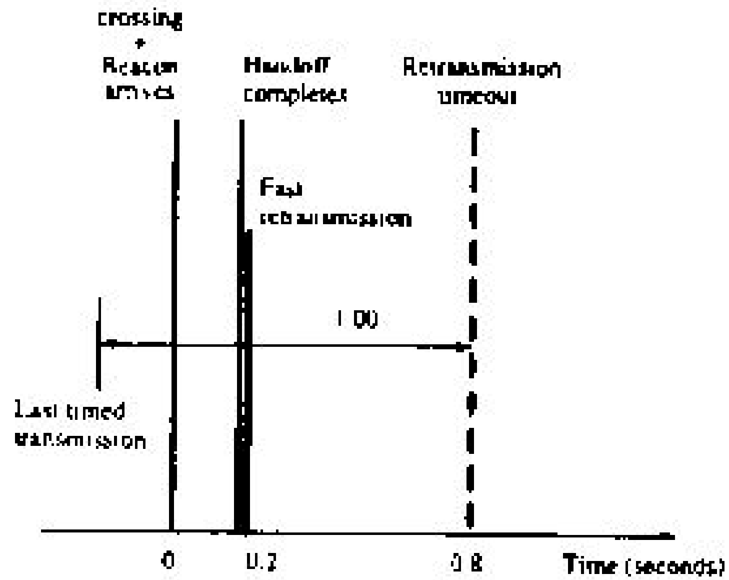


Fig. 7 Fast retransmission after a handoff with a 0-s rendezvous delay.

Fast retransmit with 0 sec rendezvous delay

Figure shows the advantage of triggered fast retransmit compared to retransmission time out.

Fast Retransmissions

- For the fast retransmit signal to be recognized by the **SH**, it needs to be informed that the handoff has been completed.

  Mobile **IP on MH** signals **TCP on MH** completion of the handoff. **TCP on MH** forwards signal to **TCP on SH.** **(This special signal can be specially marked TCP ACK or <u>three</u> ordinary TCP ACK's.)** Once **TCP on SH** receives the signal, it can invoke fast retransmit procedure. [Basically, SH now knows that next packets sent by MH will be retransmits due to handoff at cell crossing.]

Fast Retransmission

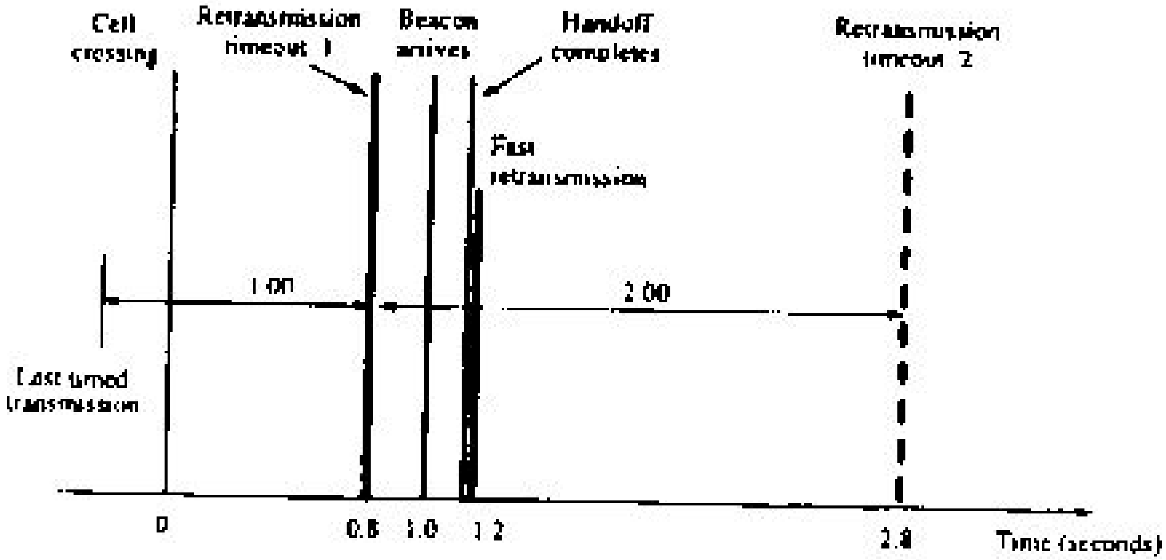


Fig. 8. Fast retransmission after a handoff with a 1-s rendezvous delay.

Fast retransmit with 1 sec rendezvous delay

1 sec delay causes two timeouts with backoff for second timeout.

In this case, fast retransmit saves 1650 ms!!

Fast Retransmissions

***Advantages of triggered fast retransmit* ::**

1. Need end-of-handoff signal in Mobile IP and invoking of triggered fast retransmit when end-of-handoff signal arrives.
2. No special support required within mobile networking environment.
3. Claim:: fast retransmit does not congest new cell with packets.

- [Authors argue that fast retransmit not needed when network can guarantee smooth handoff. Not sure how you can do this??]

Latency Improvements

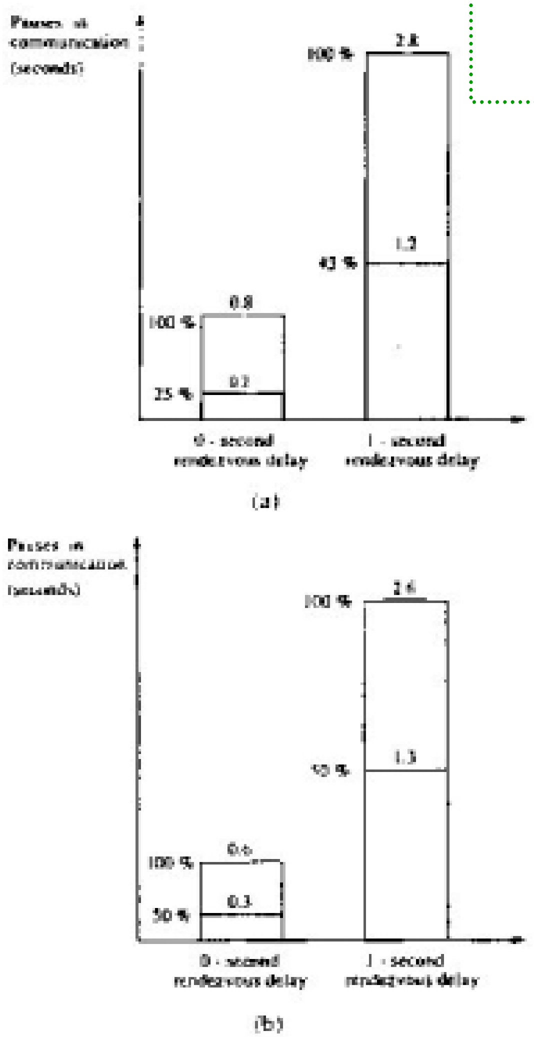


Fig. 9 Improvements in latency due to fast retransmissions. (a) Transmitter on the MH. (b) Transmitter on the SH.

Fast Retransmit scheme reduces interactive delays to 200-300 ms beyond the rendezvous.

Throughput Improvements

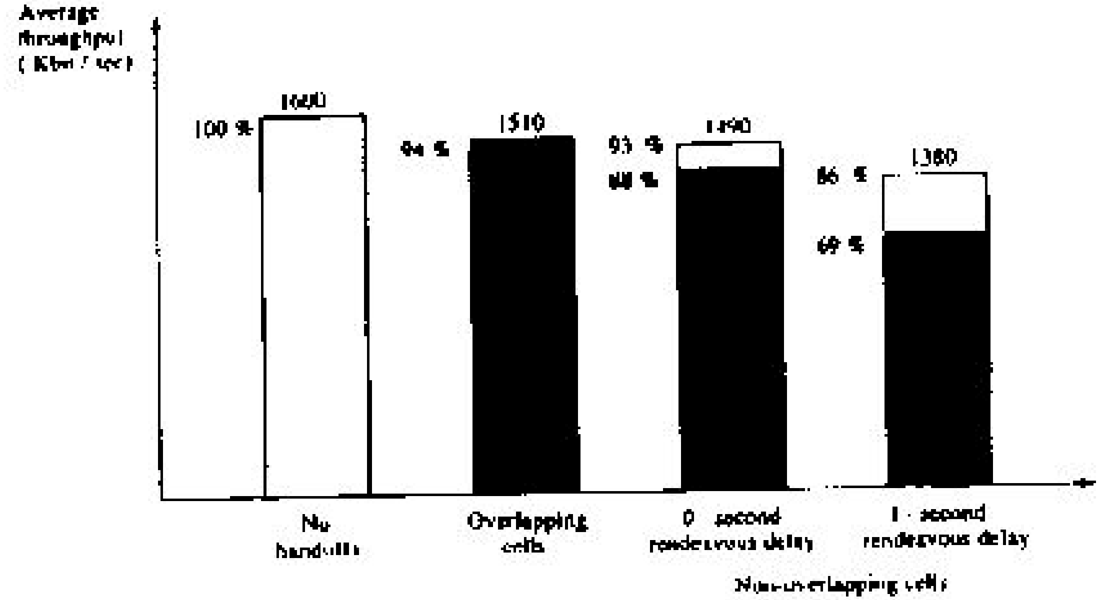


Fig. 10. Improvements in throughput due to fast retransmissions.

For 0 sec rendezvous delay throughput improves
From 1400 to 1490 kbps.

For 1 sec rendezvous delay throughput improves
From 1100 to 1380 kbps.

Wireless Transmission Errors

- WaveLAN network suffered from packet losses due to physical transmission errors.
- Need to be careful with data link layer retransmissions because other research shows that ***competing retransmission strategies*** can interact to reduce throughput while increasing link utilization.
- They suggest ***selective retransmissions.***
- For these experiments, these errors were <u>negligible</u>.

Conclusions

- Mobile, wireless LANs suffer from delays and packet losses unrelated to congestion.
- Waiting for retransmission timeouts due to cell crossings can cause significant pauses in communication.
- Paper presented a new triggered fast retransmit mechanism that reduces delay and increases throughput.