WIRELESS INFORMATION NETWORK LABORATORY

Towards Better Protocols for 3G Wireless Internet Access

or

how to make TCP and RLP cooperate

Andy T Ogielski **Yong Bai**

IAB September 29, 1999

Outline

- the big picture
- TCP-Radio Link Protocol non-cooperation problem
- approach
    - *RLP reliability classes, interprotocol signaling*
    - *ssf modeling & simulation*
- detailed analysis - *TCP over IS-707 RLP family*
- future work

3G wireless and the Internet

- **3G wireless - after a decade, a tangle of committees, alliances, ETSI-ANSI frictions**
  - *Will it go the way of ATM?*
- **3G wireless Internet initiatives**
  - *new consortia (TIA WIPP, 3G.IP), no agenda yet*
  - *corporate efforts (Ericsson, Nortel, Cisco,...)*
  - *no serious presence in IETF (pilc, mobileip)*
- **Best bet: wireless Internet wins one way or another.**
  - *good time to get the foundations right, redesign the protocols*

Wireless vs Internet

- **wired Internet:**

  very low BER links - congestion losses dominate, little mobility, over-provisioned access, large packets. Application demands grow with Moore's law. IETF mentality - prototype before standard.

- **wireless:**

  high BER links - link AND congestion losses, acess contention, mobility, small packets. Applications limited by battery power. ITU mentality - standard before prototype.

- **wireless access challenge:**

  end2end IP connectivity, pure IP infrastructure claimed to bring 10x cost reduction/megabyte (Nortel)

Wireless Internet - technical challenges

- **mobility:**
  - *fast admission (res. discovery, authentication, self-configuration)*
  - *integration of fast handoff and IP routing, esp. multicast*
    - *full area coverage (cellular); intermittent coverage (infostations)*

- **co-design of RLPs and Internet transport:**
  - *current TCP/IP and RLP - destructive interactions*
  - *QoS mapping & interprotocol signaling - better adaptation*
- **end2end transport:**
  - *encryption - no proxies/agents*

Origins of the TCP - RLP problem

Suppose it's 1980 again.

How would TCP evolve if many access networks were wireless?

– we'd decouple reliable delivery from data rate adaptation,

distinguish link loss and congestion

– TCP, IP and link protocols would communicate cleanly

and cooperate, no pretense of layering (broken anyway),

– TCP would be less rigidly married to IP (host name is NOT address),

Moral: *don't just work around TCP (Berkeley snoop or proxies).*

That's an interim solution. Analyze how to modify TCP, design better RLP

Approach to the TCP - RLP problem

TCP to IP to RLP***: service class selection***

- ***RLP cannot see only a raw bytestream - more intelligence needed***
- ***RLP has to offer multiple service reliability classes, settable on byte boundaries on request from IP (QoS mapping).***
- ***RLP service reliability classes must be mapped onto RLP parameters, such as number of retransmission cycles, etc.***
- ***IP packet classifier selects RLP service class based on IP header***

This alone can substantially improve TCP performance!

Approach to the TCP - RLP problem

RLP to IP to sender TCP: *link loss notification*

- *separate link loss & congestion loss recognition mechanisms: LLN (extension of ELN), ECN, implicit methods*
- *TCP must retransmit, may adapt window or payload size (SMSS) for recurrent link losses - not congestion control*
- *Open problems:*
  - *Do we need a link loss process estimator?*
  - *Is fast retransmission enough? Efficient ACK loss handling?*

Evolutionary changes preferred

Approach - continued

Analyze concrete RLPs: details matter

IS-707, cdma2000 enhancements, winmac,...

Analyze distinct TCP variants: details matter

Propose modifications

Techniques:

- *simulated wireless + wired Internet environment,*
  detailed protocol models, simple to complex scenarios
- *measure real data for validation (TCP over winmac)*
- *find when/if an abstracted analysis is possible:*
  which features matter, which don't

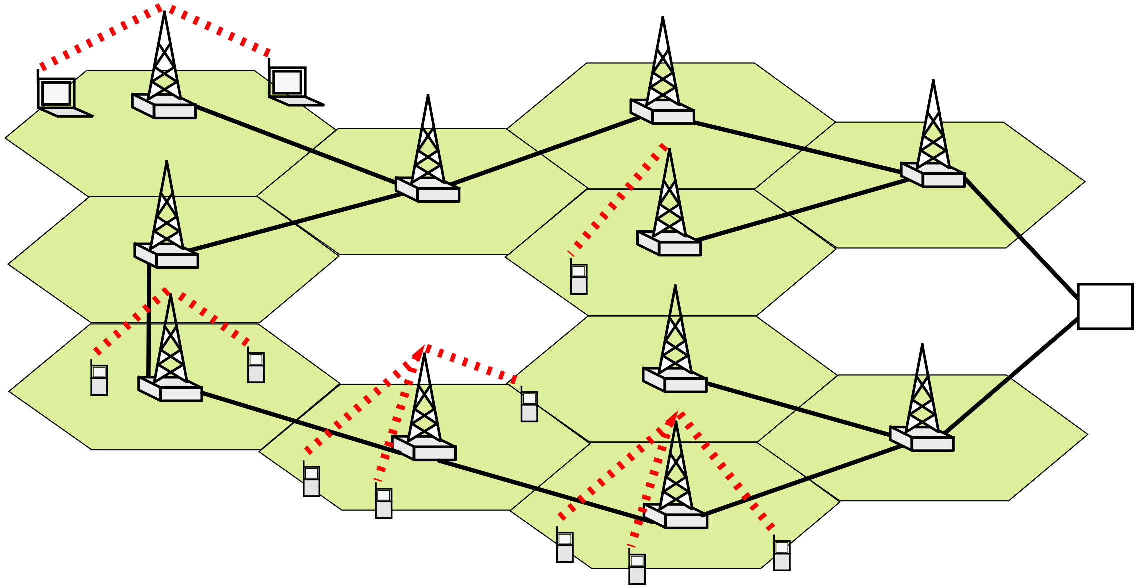
SSF Wireless - cellular 3G system models

Example:

NTT DoCoMo

IMT-2000 WCDMA

system analysis



- fine detail level: DS-CDMA chip waveforms
- multipath propagation, interference, mobility
  - *many basestations, many mobiles*
- transmitter, receiver signal processing - *original DoCoMo code*
- PHY layer - coding, spreading, framing
- signaling protocols - power control, handoff

SSF Internet - scalable modeling & simulation

`www.ssfnet.org`

SSF.OS protocol design framework

detailed protocol models

IP, TCP, UDP, OSPF, BGP-4,...

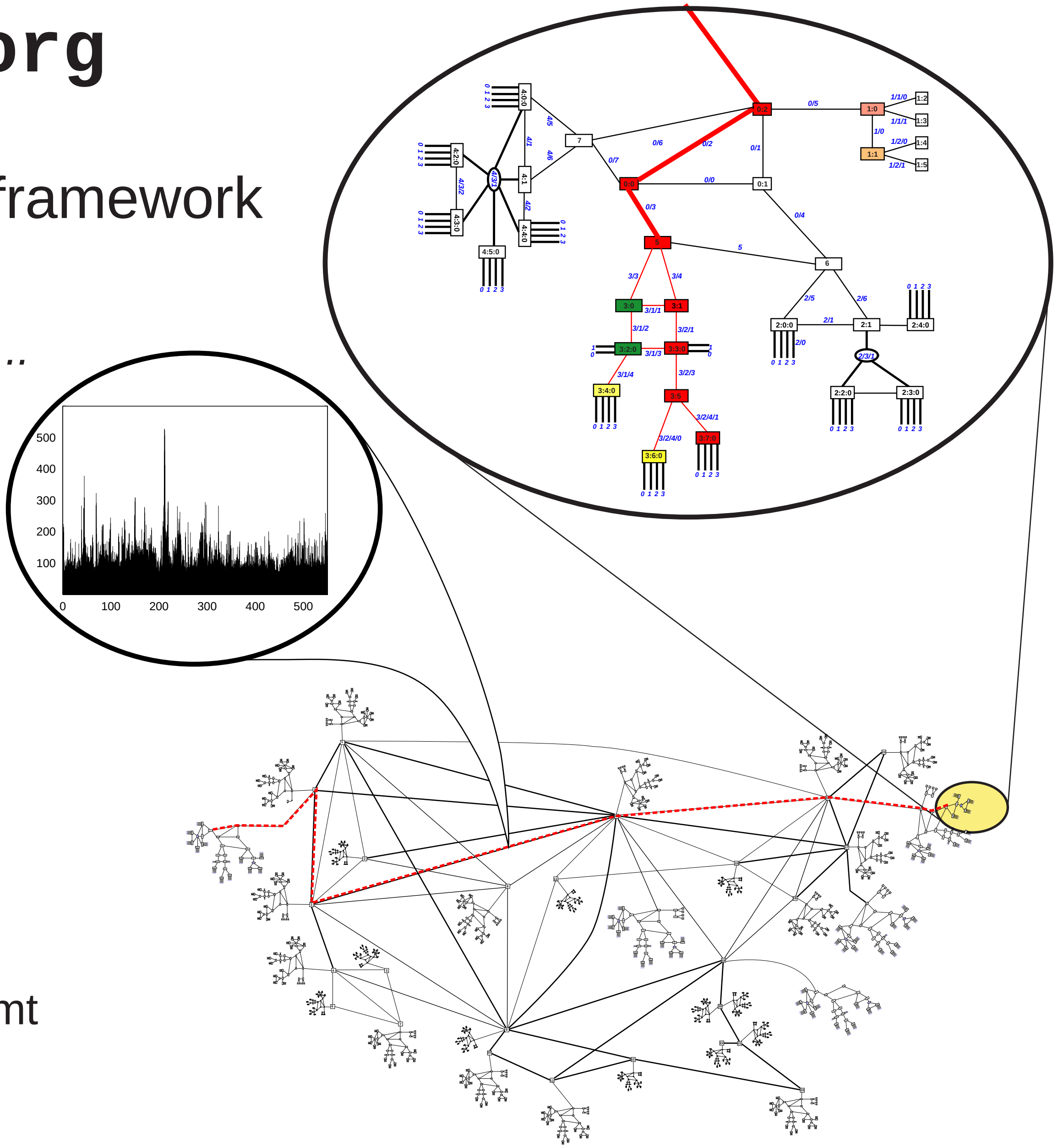
empirical traffic

Web clients/servers

research examples:

- large scale traffic dynamics

- diffserv, MPLS



Package features:

pure Java, parallel execution

db-oriented configuration mngmt

highly scalable and extensible

SSF 2000 - integration of Internet & Wireless models

Scale and Details matter!

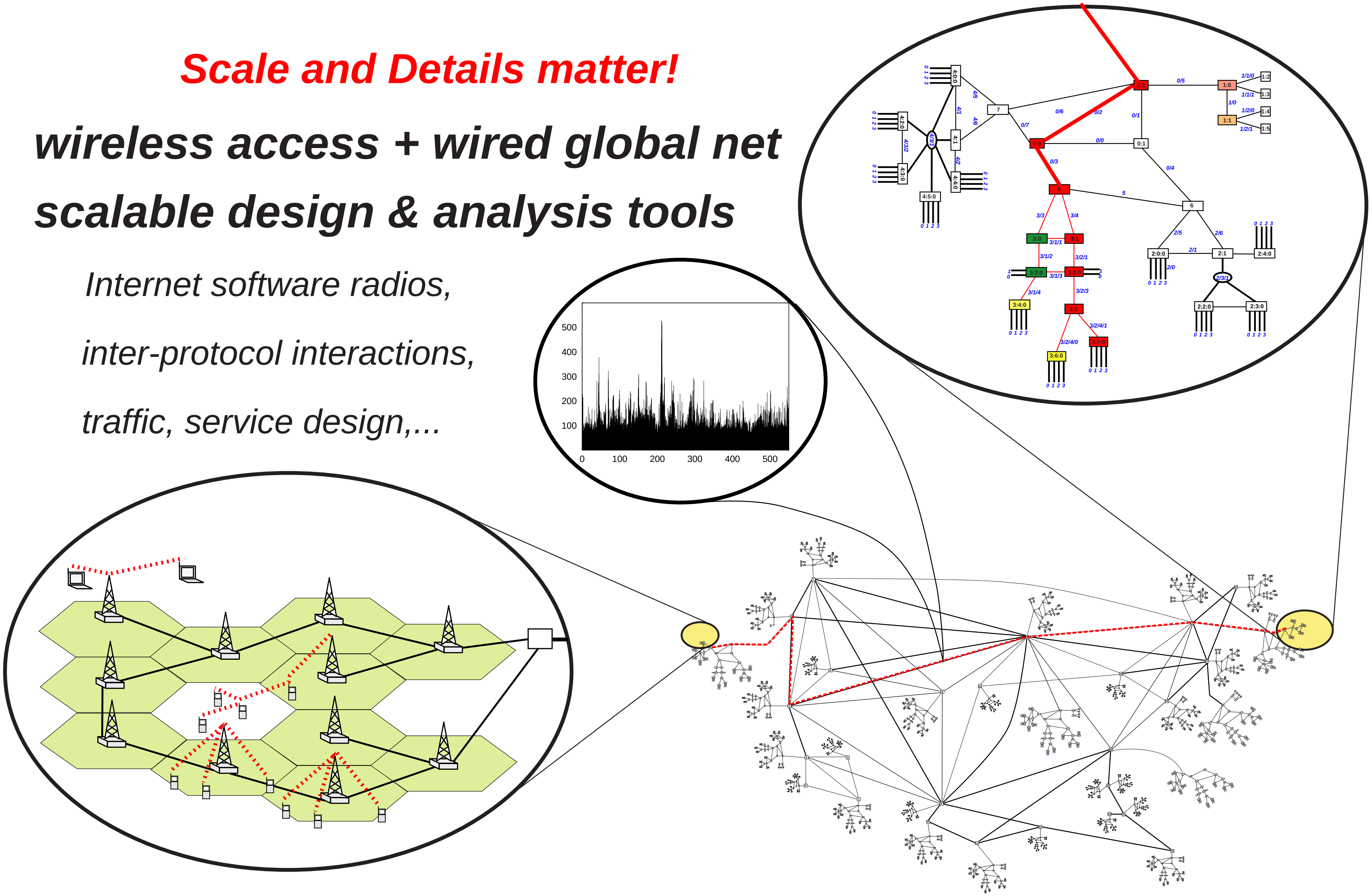
wireless access + wired global net

scalable design & analysis tools

Internet software radios,

inter-protocol interactions,

traffic, service design,...

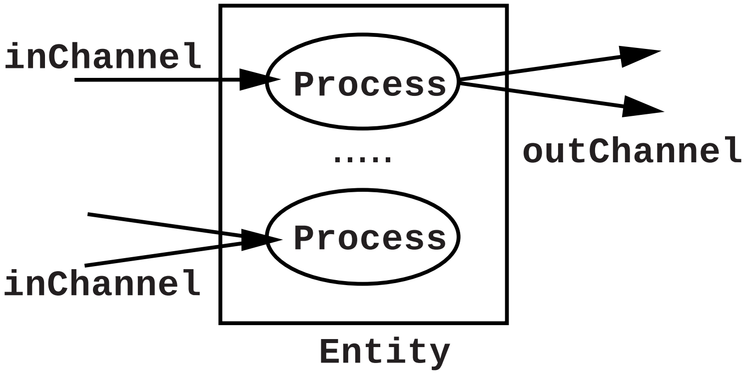


SSF: base modeling API

SSF is built on modeling experience from VHDL to TeD

5 base modeling classes
20+ methods

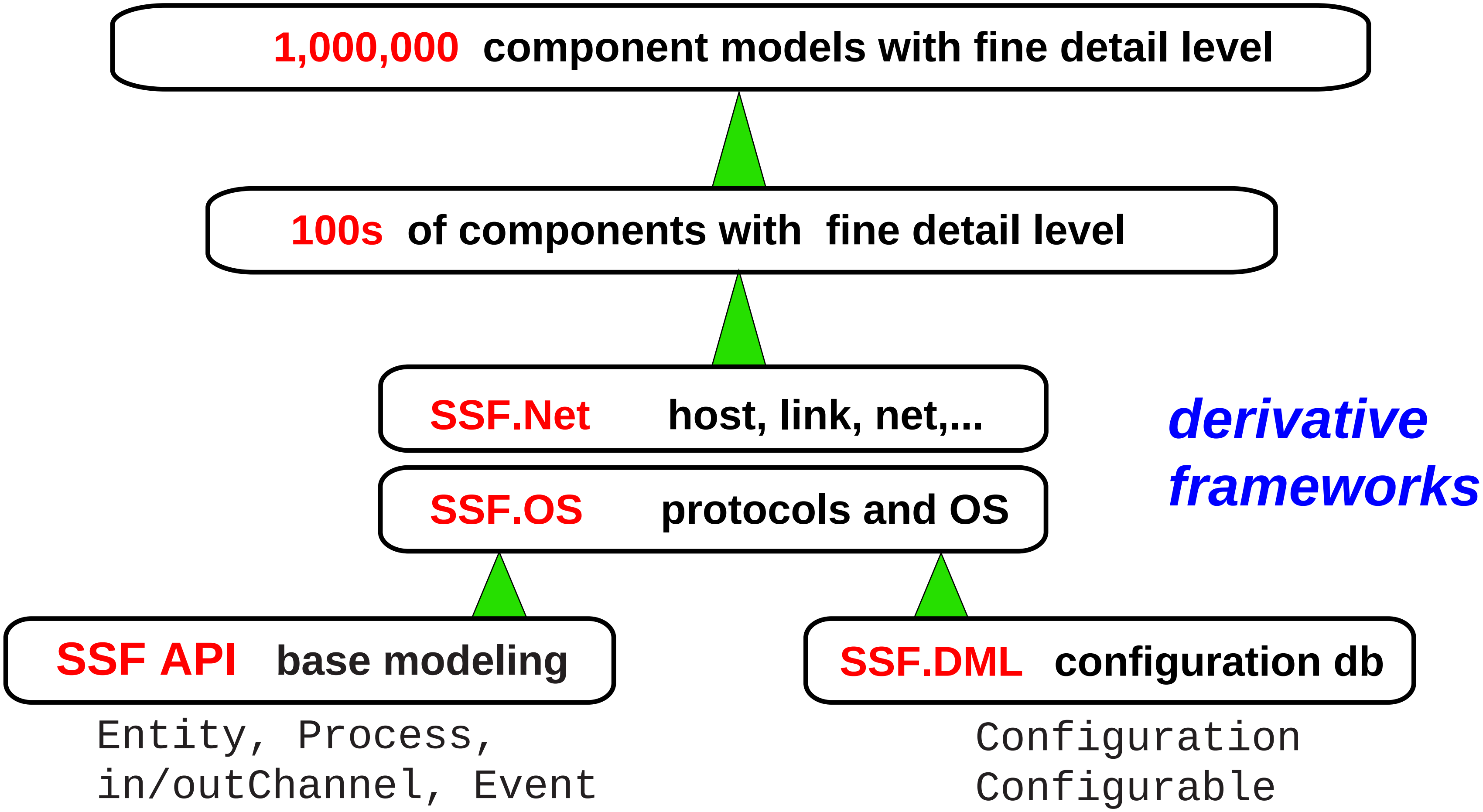
`Entity,Process,in/outChannel,Event`



- process-addressable events, efficient in/out multicast.
- class Entity is a "container" for alignment of Processes
  and in/outChannels with Timelines (Schedulers).
- Usual OO programming - no limitations.
- Java and C++ parallel implementations: JSSF, CSSF, DaSSF

SSFNET modeling layers

Solution for the challenges of very large model construction



1,000,000 component models with fine detail level

100s of components with fine detail level

SSF.Net host, link, net,...

SSF.OS protocols and OS

derivative frameworks

SSF API base modeling

`Entity, Process,`
`in/outChannel, Event`

SSF.DML configuration db

`Configuration`
`Configurable`

Things to know

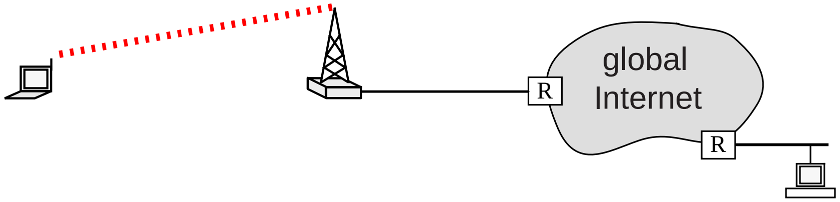
- Cannot predict the future, but at least use correct user and traffic models, empirically valid today:
  - Web TCP packets 60-80% of all IP packets or bytes (1998-9)
  - Very many small, some larger, a few very large TCP data transmissions (either SYN-to-FIN connections, of flows separated by idle state) - ubiquitous Pareto (= long-tailed power law) distributions of transmitted object sizes
  - Inhomogeneous Poisson arrivals of sessions OK
- Large variablity of TCP path lengths:
  - long-tailed distribution of delays (lognormal or Pareto) affects TCP
- Adaptive traffic shaping by interacting TCP connections (*shared links, shared losses*)
- **Keep it simple:** Begin with studies of a single TCP transmission over a lossy radio link.
  Increase complexity of TCP-RLP analysis step by step.

Research on TCP - RLP interactions: towards better protocols for wireless Internet

phase 1: correlated fading losses

1 wireless host, 1 TCP connection,

RLP, fading, simplified IP cloud

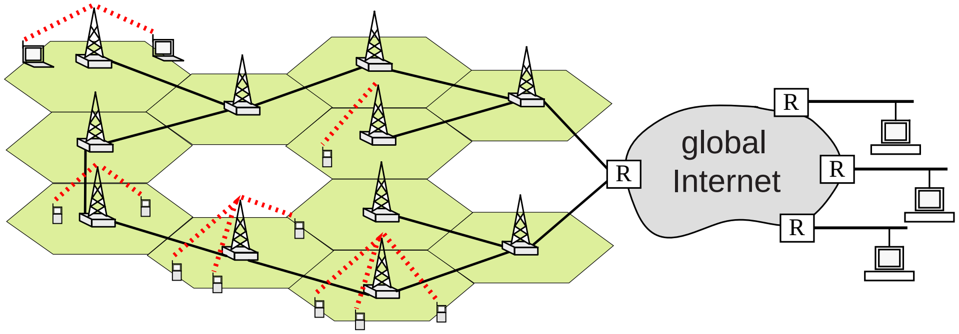


phase 2: add multiaccess

N wireless hosts, many connections,

MAC + RLP, fading, interference,

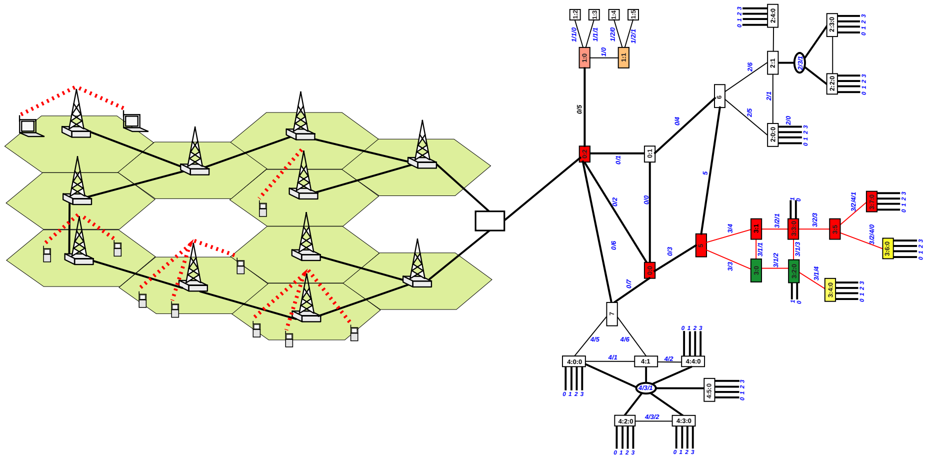
simplified IP cloud



phase 3: add wired congestion

phase 2 plus realistic IP cloud:

correlated wireline delays/losses



TCP over IS-707

How IS-707 works - ***what can be tuned***

How TCP works - ***what can be tuned***

Experiments and results:

- ***TCP window size adaptation for link loss reduction,***
- ***variable number of RLP retransmissions,***
- ***exploitation of idle frames (MAC states)***

Motivation

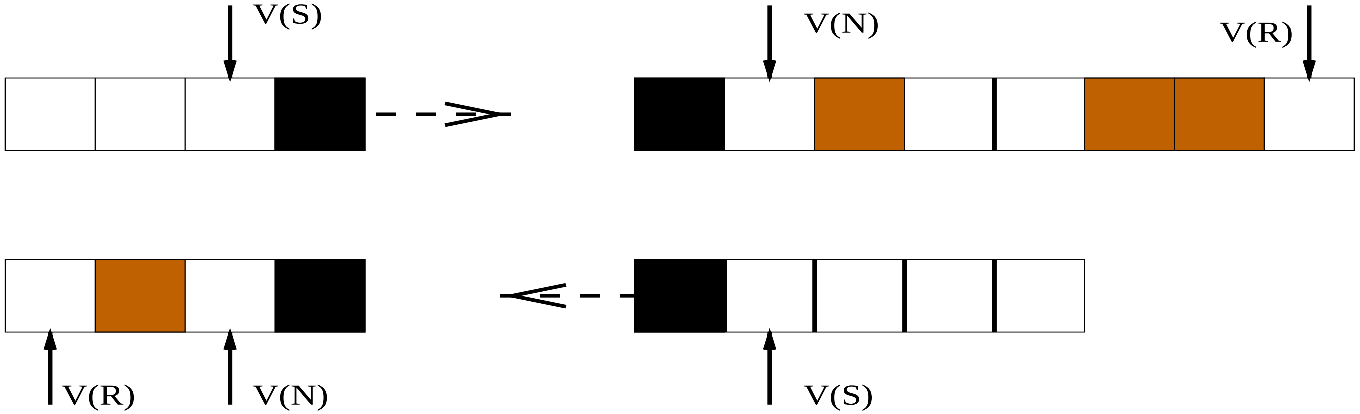
- TCP segment losses on high error rate wireless links significantly degrade TCP throughput because these losses are interpreted as congestion losses and TCP subsequently invokes congestion control mechanisms.
- Radio link layer ARQ protocol can perform frame error recovery, however, the interactions between the separate TCP- and link-level error control mechanisms need to be investigated, especially when the link layer ARQ is a partial error recovery algorithm such as specified in IS-707.

IS-707 RLP Frame Error Recovery

- In the non-transparent mode, IS-707 uses a NAK (negative acknowledegment) selective repeat ARQ protocol to retransmit lost data frames.
- In case of a frame loss, RLP performs a partial error recovery through a finite number of frame retransmissions (2 cycles; three rounds in one cycle; 1-2-3 NAKs).
- The RLP layer provides a storage buffer for resequencing of out-of-sequence RLP data frames.

IS-707 RLP Frame Error Recovery - Con'd

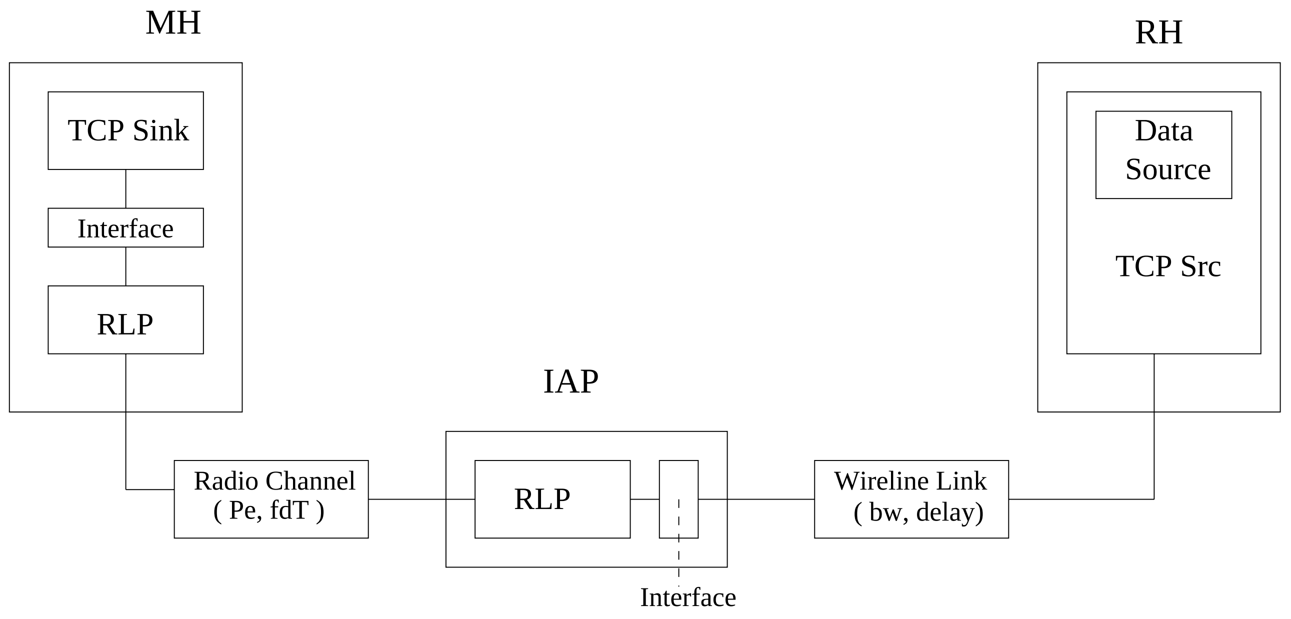
- RLP transmitting buffer and receiving resequencing buffer



$V(S)$ = sequence number of next frame to be sent
$V(N)$ = next frame needed for sequential delivery
$V(R)$ = next frame expected

- The receiving side can notice that a frame is missing by three means, i.e., the arrival of a valid data frame, an idle frame (with $SN = V(S)$), and a NAK control frames (with $SN = V(S)$).

Simulation Model



This simulation model is implemented in a new object-oriented framework called SSF (Scalable Simulation Framework). Each block is constructed as an entity.

Impacts of TCP Sliding Window Size

- When a slow lossy radio link is involved, if the advertised window is too large, TCP packets build up at the Internet Access Point (IAP), occupying large amount of buffer space and perhaps resulting in congestions. Moreover, the retransmitted TCP packet takes a long time to be delivered to the receiver delayed by the IAP FIFO queue.
- If the advertised window is too small, there are no sufficient TCP packets available for delevery. This deficit degrades the overall performance as well.

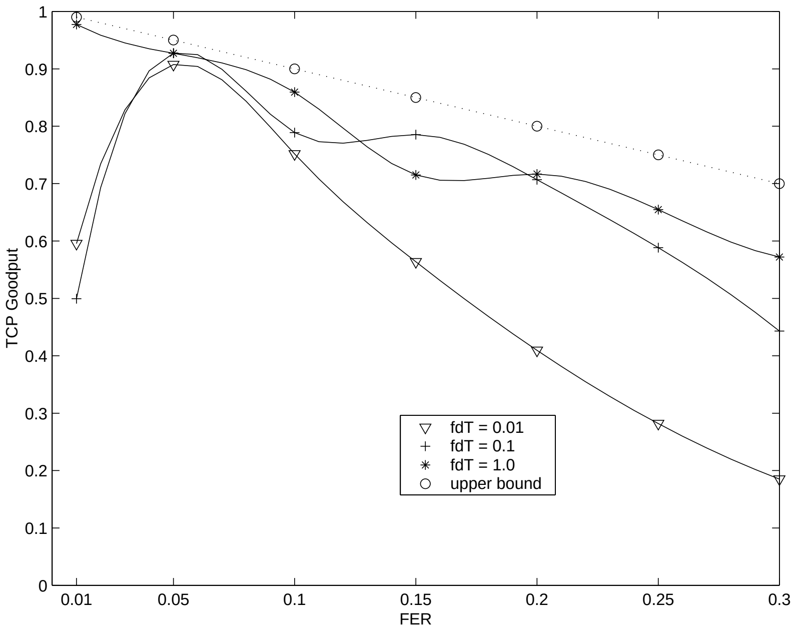


Figure 1: TCP goodput ($awnd = 256, data\ rate = 9600bps$)

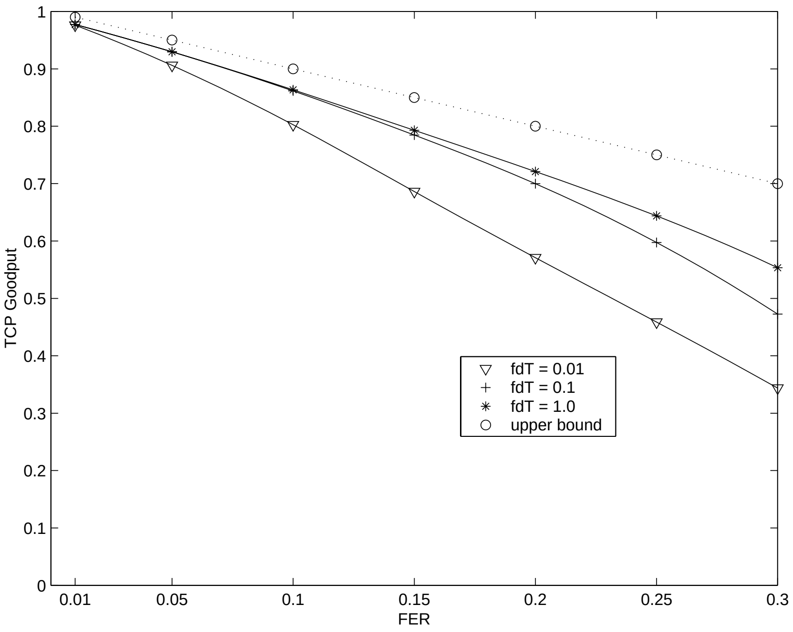


Figure 2: TCP goodput ($awnd = 2, data\ rate = 9600bps$)

Impacts of Persistence on RLP Frame Error Recovery

For best-effort bulk data applications like FTP, persisting longer RLP frame error recovery and subsequently hiding more frame losses to the TCP layer result in better performance.

TCP can accommodate the increased RLP frame transmission latency by updating TCP retransmission time-out (RTO) value adaptively.

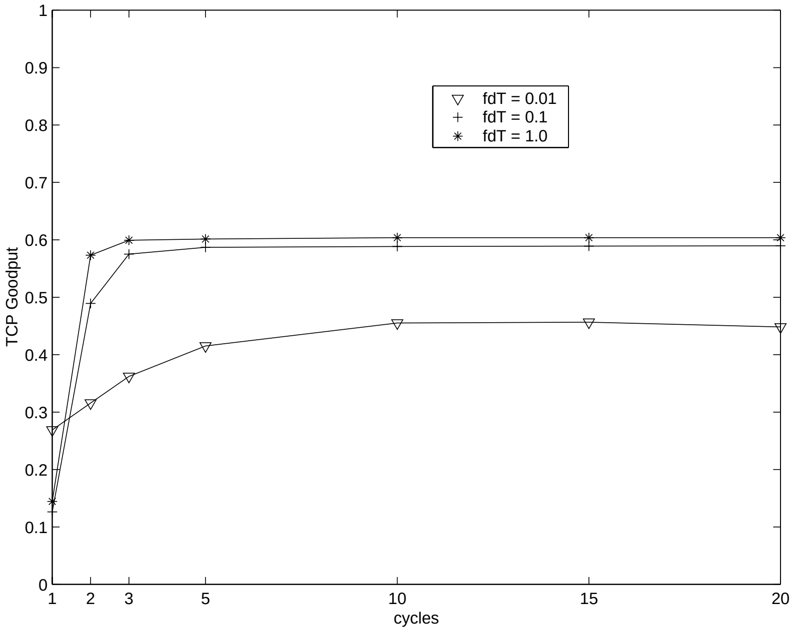


Figure 3: TCP goodput versus RLP retransmission cycles ($P_e = 0.3, awnd = 40$)

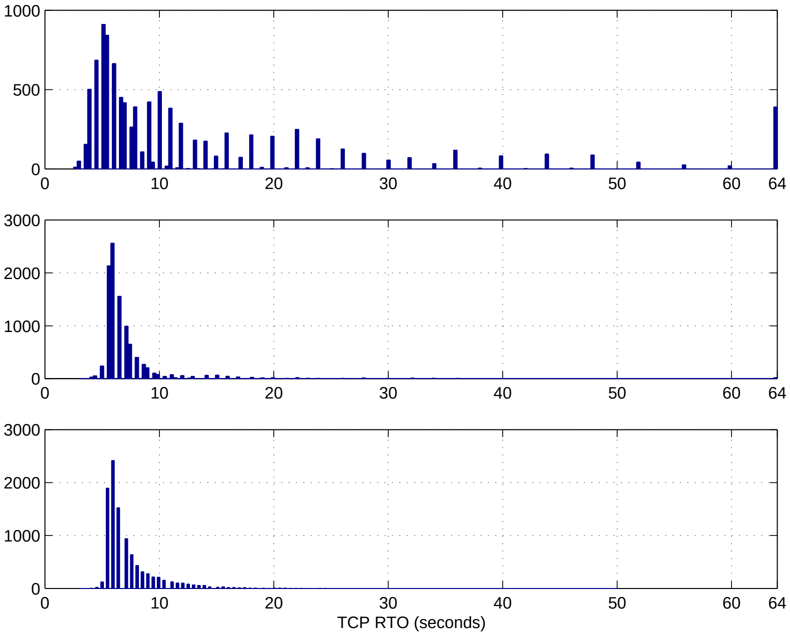


Figure 4: Histogram of TCP RTOs at 1, 2, and 10 RLP retransmission cycles ($P_e = 0.3,\ f_dT = 1.0, awnd = 40$)

Interaction with MAC Multiple States

- Idle frames are sent to the receiver when the channel is idle. They have sender's last-sent frame sequence number information and can advance the receiver's frame-clocked retransmission timer. Thus they can drive fast frame error recovery.

- Without idle frames, in the high FER regime, slow frame loss recovery and unrecovered errors cause successive TCP timeouts resulting in very poor system performance.

- However, in the IS-95-B and cdma2000 proposals, the radio link is released to other users in the Suspended State, so we need to determine appropriate number of idle frames in order to achieve the required performance.

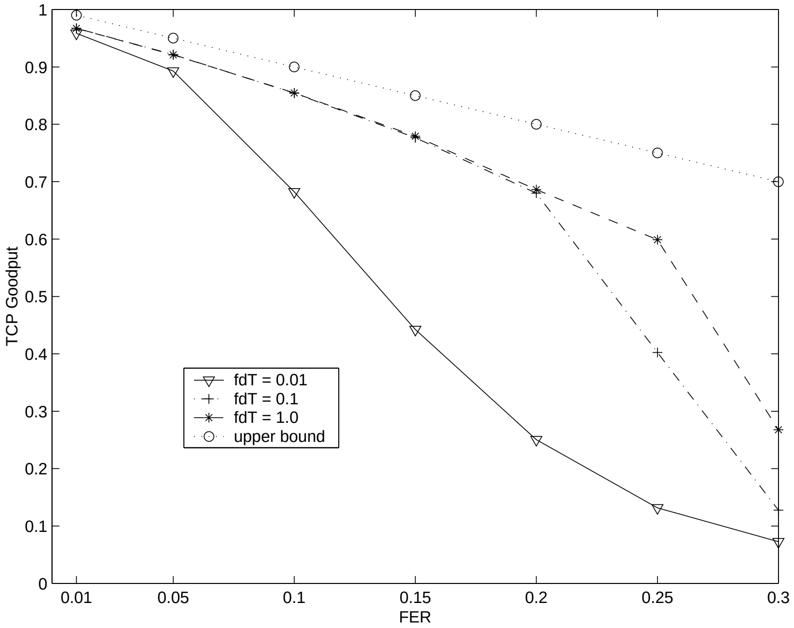


Figure 5: TCP goodput without idle frames ($f_dT = 0.01, P_e = 0.3, awnd = 40$)

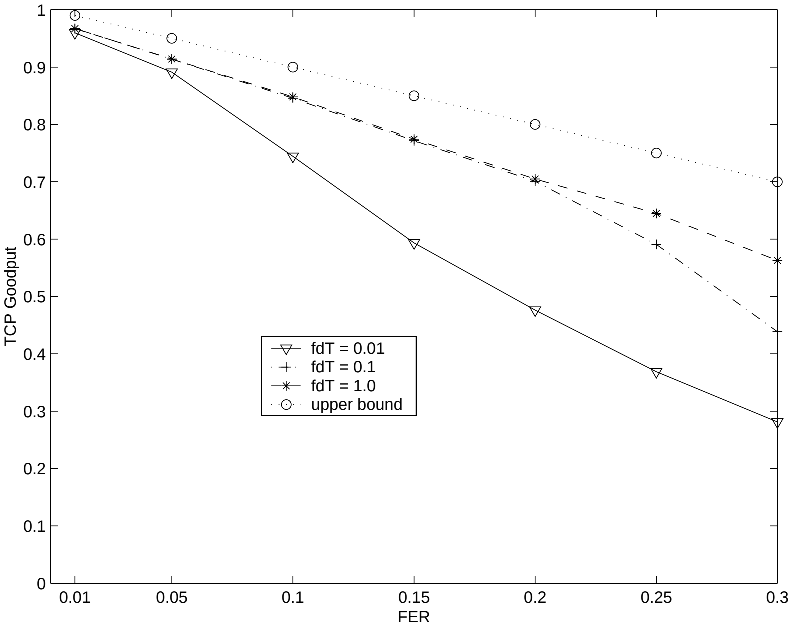


Figure 6: TCP goodput with idle frames ($f_dT = 0.01, P_e = 0.3, awnd = 40$)

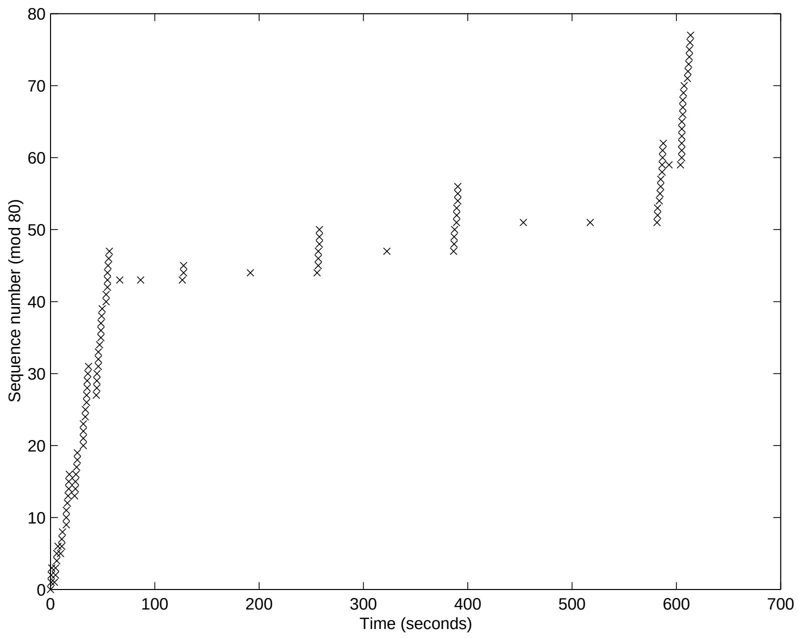


Figure 7: TCP sequence number evolution without idle frames

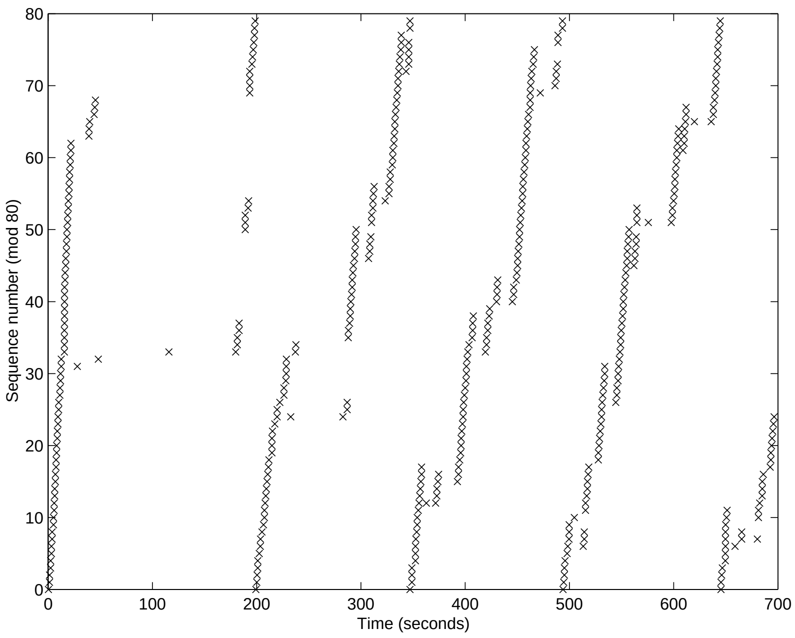


Figure 8: TCP sequence number evolution with idle frames

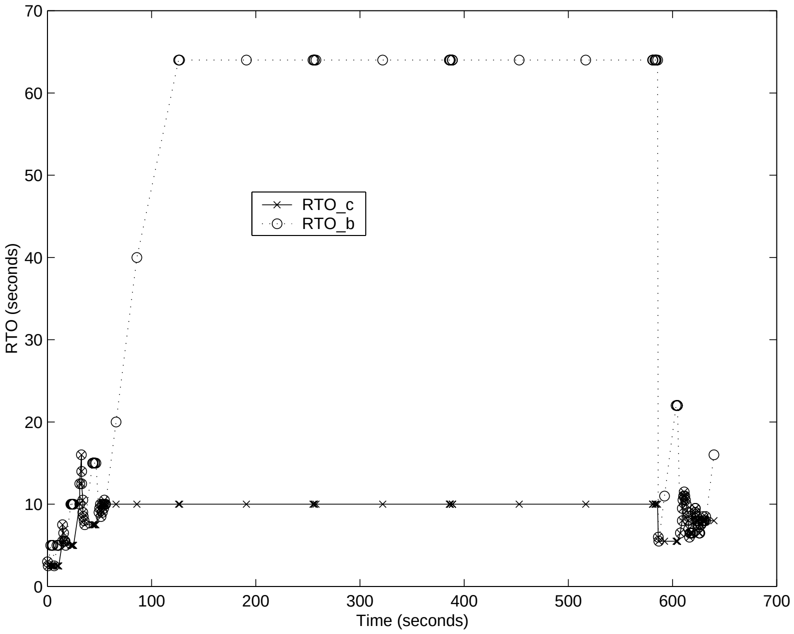


Figure 9: TCP RTO values without idle frames

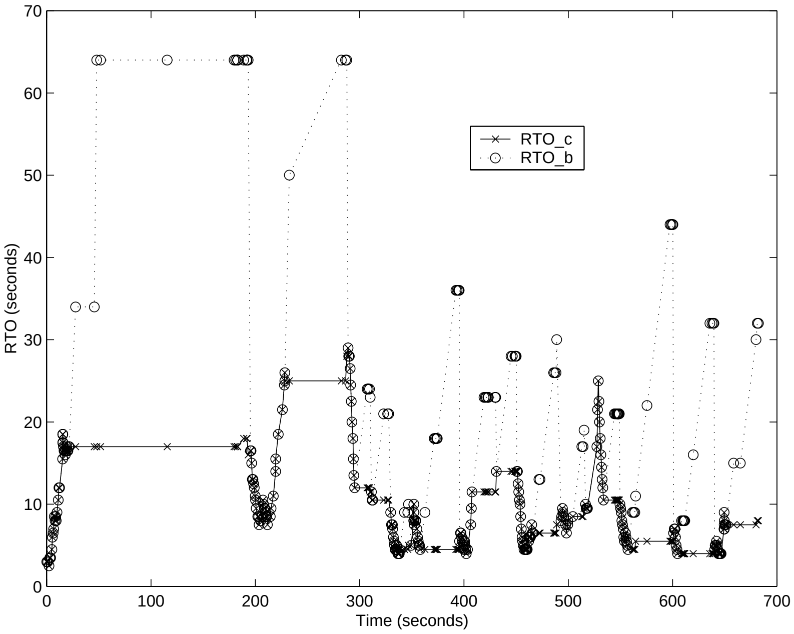


Figure 10: TCP RTO values with idle frames

Simulation Results - Con'd

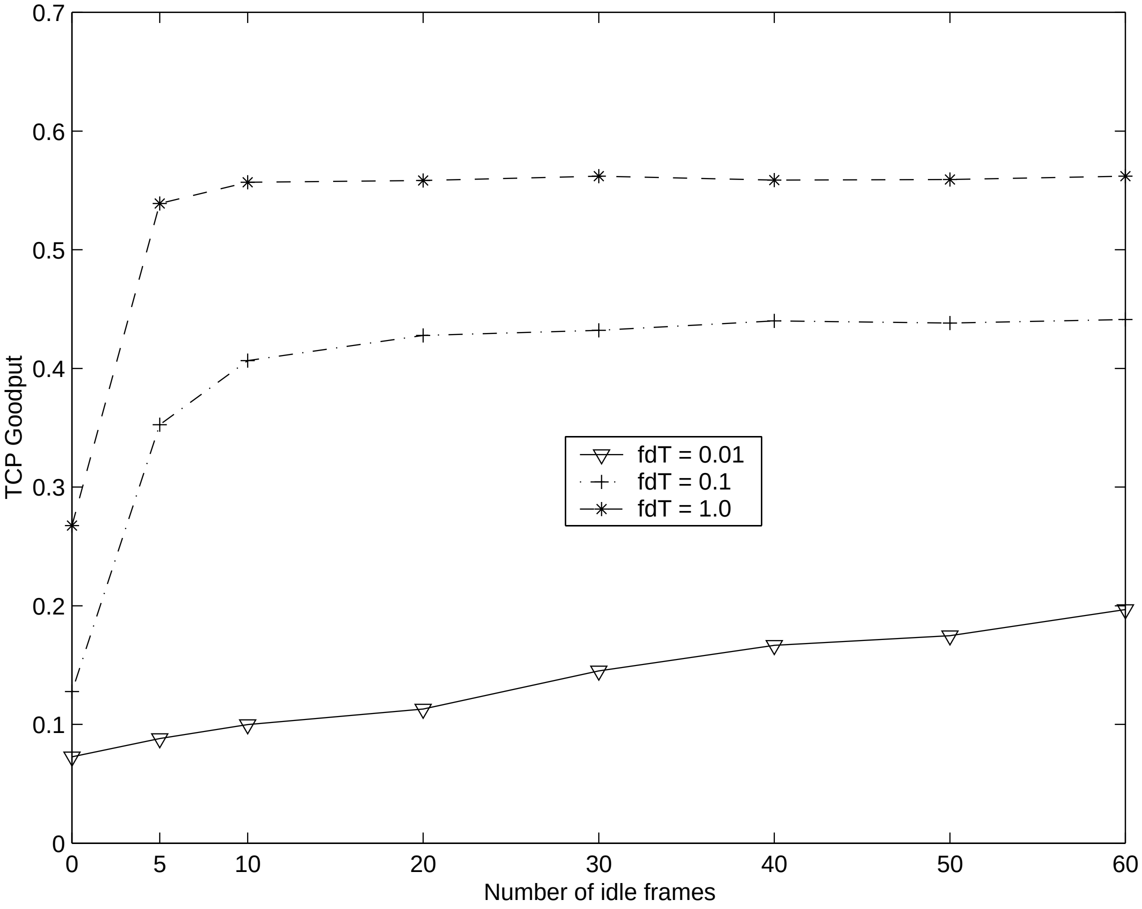


Figure 11: TCP goodput versus idle frames in high FER regime, $P_e = 0.3$.

Conclusions

- The simulation shows that low FER (frame error rate) and fast RLP frame error recovery are crucial to the overall performance.
- A suitable TCP advertised window size should be chosen to achieve desirable performance.
- Longer persistence in radio link frame error recovery results in better performance for best-effort bulk data applications.
- Idle frames specified in IS-707 carries sending frame sequence number and can drive fast RLP frame error recovery. Appropriate number of idle frames needs to be sent before MAC transits to Suspended State.

Current & future work

- Implement and evaluate IS-707 -- TCP signaling
- Multiple wireless terminals, multiple TCP connections, empirical Web workload, multiaccess problems

- Joint occurence of wireless link losses and wired network congestion losses - TCP modifications

- measurements - Windows TCP over winmac
- more future work...

TIA/EIA IS-707, *Data services option standard for wideband spread spectrum digital cellular system* , Feb. 1998.

Y. Bai, G. Wu, and A.T. Ogielski, *TCP/RLP Coordination and Interprotocol Signaling for Wireless Internet* , in Proc. of VTC'99 Spring, pp. 1945 -1951, May 1999.

Y. Bai, A.T. Ogielski, and G. Wu , *TCP over IS-707*, in Proc. of PIMRC'99, (to appear), Sept. 1999.

Y. Bai, A.T. Ogielski, and G. Wu , *Interactions of TCP and Radio Link ARQ Protocol*, in Proc. of VTC'99 Fall, (to appear), Sept. 1999.