Module 6

Mobile Transport Layer

Contents

- TCP over Wireless Networks
- Indirect TCP
- Snooping TCP
- Mobile TCP – Fast Retransmit / Fast Recovery Transmission/Timeout Freezing-Selective Retransmission
- Transaction Oriented TCP
- TCP over 2.5 / 3G wireless Networks

TCP v/s UDP: Traditional Network

- TCP- Transmission Control Protocol
- UDP- User Datagram Protocol
- **Functions of transport layer:**
  - Checksumming over user data
  - multiplexing/demultiplexing of data from/to applications
  - Allow dedicated applications to be addressed rather than only host part like network layer
- **TCP:**
  - Connection oriented
  - Guarantees reliable data transmission using retransmission techniques
  - Manages data transfer in congestion
  - Needs special mechanisms to be useful in mobile

1. Traditional TCP

- TCP stands for Transmission control protocol.
- It is a transport layer protocol responsible for providing mobility support to lower layers up to network layer.
- Data transmission takes place using network adapters, fiber optics, copper wires, special hardware for routers etc.
- Characteristics of TCP for wired network:
  - Congestion Control
  - Slow Start
  - Fast retransmit/ fast recovery

Congestion Control

- Congestion is a situation in which the number of input packets for an output link is more than the capacity of output link.
- Router **drop the packet** in this case- Router's packet buffer is full.
- A dropped packet is lost for the transmission, and the receiver notices a gap in the packet stream.
- Now the receiver does not directly tell the sender which packet is missing, but continues to acknowledge all in-sequence packets up to the missing one.
- The sender notices the **missing acknowledgement** for the lost packet and assumes a packet loss due to congestion.
- **Retransmitting the missing packet** and continuing at **full sending rate** would now be **risky,** as this might only increase the congestion.
- To mitigate congestion, **TCP slows down the transmission rate** dramatically.
- **Cooperation of multiple TCP connections over the internet mitigates congestion.**

Slow start

- **TCP slow-start algorithm- behaviour TCP shows after the detection of congestion**
  - sender calculates a congestion window for a receiver.
  - starts with a congestion window size equal to one segment.
  - The sender sends one packet and waits for acknowledgement.
  - If this acknowledgement arrives, the sender increases the congestion window by one, now sending two packets (congestion window = 2).
  - This scheme doubles the congestion window every time the acknowledgements come back, which takes one round trip time (RTT).
  - This is called the **exponential growth** of the congestion window in the slow start mechanism.

Slow start- Cont..

- The exponential growth stops at the **congestion threshold**.
- **After congestion threshold- linear growth**
- **Linear increase continues until a time-out** at the sender occurs due to a missing acknowledgement, or until the sender detects a gap in transmitted data because of congestion
- In either of the two cases- congestion threshold to half of the current congestion window.

Fast retransmit/fast recovery

- Reduction of the congestion threshold:
  - a sender receiving continuous acknowledgements for the same packet.
- This informs the sender of two things.
  - One is that the receiver got all packets up to the acknowledged packet in sequence.
  - Receiving acknowledgements from a receiver also shows that the receiver continuously receives something from the sender.
- The gap in the packet stream is not due to severe congestion, but a simple packet loss due to a transmission error.
- The sender can now retransmit the missing packet(s) before the timer expires. This behavior is called **fast retransmit.**

- The receipt of acknowledgements shows that there is no congestion to justify a slow start.
- The sender can continue with the current congestion window.
- The sender performs a **fast recovery** from the packet loss.

Implications on mobility

- Many problems arises for packet transmission in wireless network:
  - Higher latency: wireless network has higher latency(is the amt of time required for message to travel between the two points in network.
  - High error rates and unreliability: error rate in wireless is much high→ packet loss is very common
  - Handovers: receiver and sender move away from each other or receiver move away from access point.→packet loss

Implications on mobility

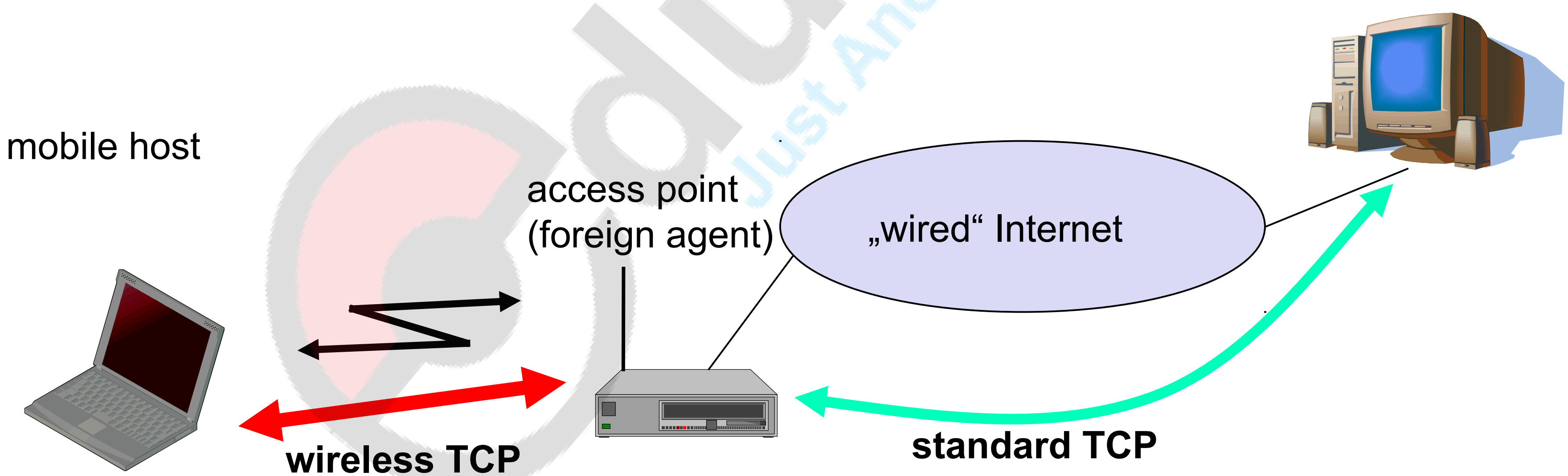
- Mobility itself can cause packet loss. There are many situations where a soft handover from one access point to another is not possible for a mobile endsystem.
- The TCP mechanism detecting missing acknowledgements via time-outs and concluding packet loss due to congestion cannot distinguish between the different causes. This is a fundamental design problem in TCP: An error control mechanism (missing acknowledgement due to a transmission error) is misused for congestion control (missing acknowledgement due to network overload).
- In both cases packets are lost (either due to invalid checksums or to dropping in Routers)

Classical TCP improvements

- approaches:
  - Indirect TCP
  - Snooping TCP
  - Mobile TCP
  - Fast Retransmit/Fast Recovery
  - Transmission/time-out freezing
  - Selective retransmission
  - Transaction Oriented TCP

Indirect TCP (I-TCP)

- Segments a TCP connection into a fixed part and a wireless part.
- hosts in the fixed part of the net do not notice the characteristics of the wireless part.
- The connection between the access point and mobile host can use a special TCP adapted to wireless network.
- Access point is usually implemented on the FA of mobile host because The FA controls the mobility of the mobile host anyway and can also hand over the connection to the next FAwhen the mobile host moves on.

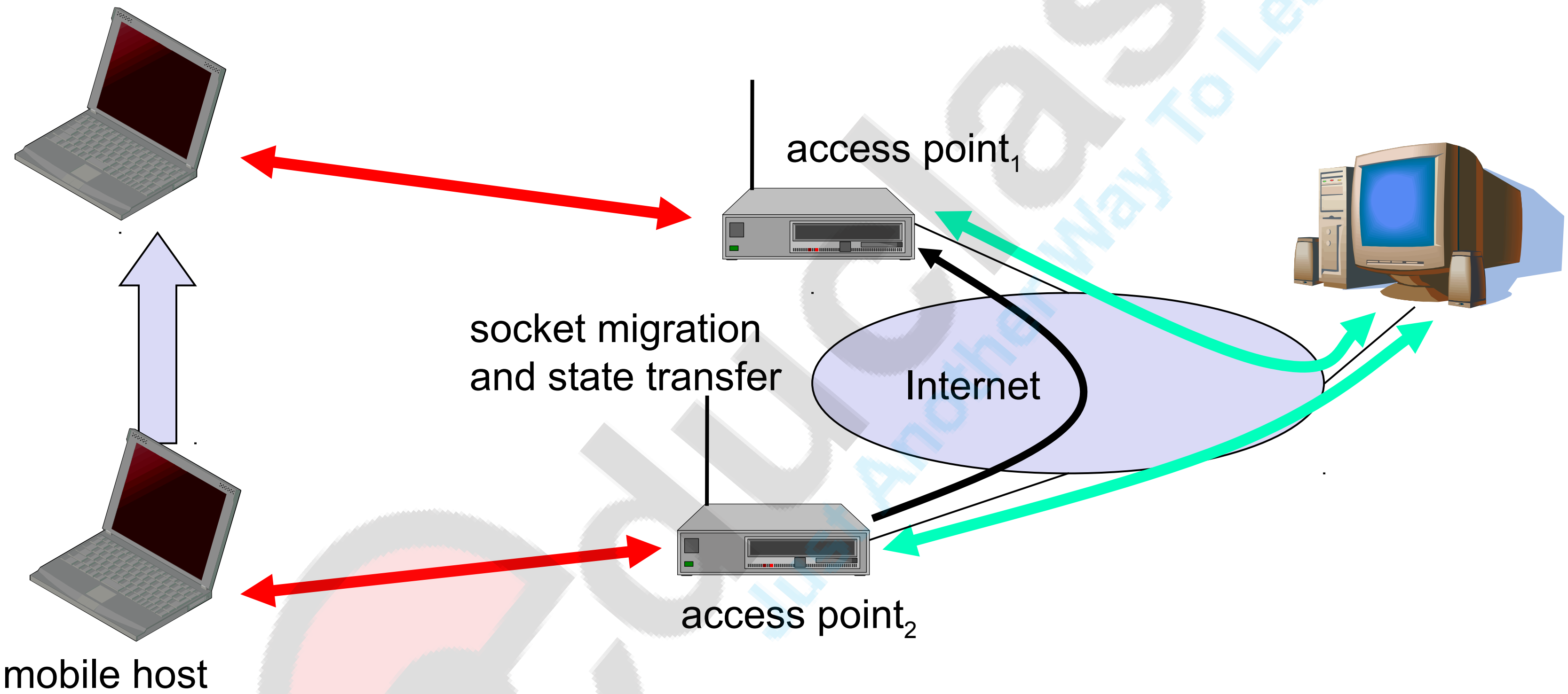


- ***Case I: packet is transmitted from fixed computer to mobile host***
  - The standard TCP connection terminates at access point.
  - The AP receives the TCP packets instead of mobile host and sends ack for received packets to the fixed computer.
  - The AP stores the packets using buffer and tries to forward the packet to mobile host using special TCP.
  - AP act as proxy.
  - If there is transmission error on wireless link local retransmission of packet take place.
  - Once ack for packet is received from mobile host, packet is removed from buffer.

- ***Case II: packet is transmitted from mobile host to fixed computer***
  - As in above case AP receives the packet sent by mobile host and acknowledges it. It also stores the packet.
  - In case packet is loss in wireless link, mobile host notices this and retransmits the packets.
  - The AP then forwards the packet to fixed computer via standard TCP connection.
  - Once ack for packet is received from fixed computer , packet is removed from buffer.

- ***Case III: mobile host moves to a new location and a handover take place***
  - Once mobile host register itself with new FA, the new FA can inform the old FA about its current location for packet forwarding .
  - All packets are buffered by old AP need to be forwarded to new AP as packets in the buffer have already acknowledged by old FA.
  - Besides buffer content, the sockets of the AP, too, must migrate to the new foreign agent.
  - The socket reflects the current state of the TCP connection, i.e., sequence number, addresses, ports etc.
  - This handover is transparent to correspondent host. No new connection is required.

I-TCP socket and state migration



Advantages of I-TCP

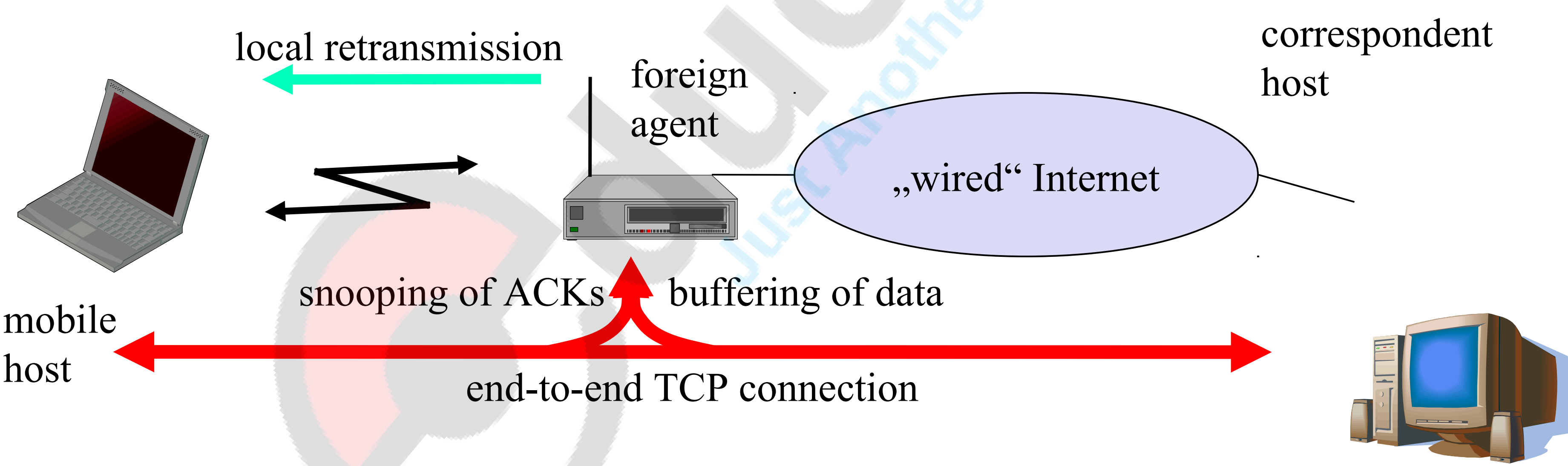
- Due to the strict partitioning into two connections, transmission errors on the wireless link, i.e., lost packets, cannot propagate into the fixed network.
- In case of packet loss in wireless link there is just local retransmission of the packet from access point.
- The wireless TCP can use precise time-out to guarantee retransmission as soon as possible.
- Partitioning into two connections also allows the use of a different transport layer protocol between the foreign agent and the mobile host.

Disadvantages of I-TCP

- The loss of the end-to-end semantics of TCP might cause problems if the foreign agent partitioning the TCP connection crashes.
- Increased handover latency may be much more problematic.
- All packets sent by the correspondent host are buffered by the foreign agent besides forwarding them to the mobile host.
- The foreign agent removes a packet from the buffer as soon as the appropriate acknowledgement arrives.
- If the mobile host now performs a handover to another foreign agent, it takes a while before the old foreign agent can forward the buffered data to the new foreign agent.
- During this time more packets may arrive. All these packets have to be forwarded to the new foreign agent first, before it can start forwarding new packets redirected to it.

Snooping TCP

- The following TCP enhancement works completely transparently and leaves the TCP end-to-end connection intact.
- The main function of the enhancement is to buffer data close to the mobile host to perform fast local retransmission in case of packet loss.



- In this approach, the foreign agent buffers all packets with **destination mobile host** and additionally ‘snoops’ the packet flow in both directions to recognize acknowledgements.
- The reason for buffering packets toward the mobile node is to enable the foreign agent to perform a local retransmission in case of packet loss on the wireless link.
- The foreign agent buffers every packet until it receives an acknowledgement from the mobile host.

- ***Case I : packet is transmitted from fixed computer to mobile host***
  - In this approach, the FA buffers all packets with **destination mobile host** and additionally 'snoops' the packet flow in both directions to recognize acknowledgements.
  - The reason for buffering packets toward the mobile node is to enable the foreign agent to perform a local retransmission in case of packet loss on the wireless link.
  - The foreign agent buffers every packet until it receives an acknowledgement from the mobile host.

- ***Case II: packet is transmitted from mobile host to fixed computer***
  - The FA keep track of sequence numbers of the packets sent by mobile host.
  - The foreign agent snoops into the packet stream to detect gaps in the sequence numbers of TCP.
  - As soon as the foreign agent detects a missing packet, it returns a negative acknowledgement (NACK) to the mobile host.
  - The mobile host can now retransmit the missing packet immediately.

Mobile TCP

- An unmodified TCP is used on the standard host-**supervisory host (SH)** connection, while an optimized TCP is used on the SH-MH connection
- The supervisory host is responsible for exchanging data between both parts similar to the proxy in ITCP
- The M-TCP approach assumes a relatively low bit error rate on the wireless link.
- Therefore, it does not perform **caching/retransmission** of data via the SH.
- If a packet is lost on the wireless link, it has to be retransmitted by the original sender.
- This maintains the TCP end-to-end semantics.

Disconnection Problem

- Once the MH is disconnected, it can no longer receive the packets nor can it send the acknowledgement to the CN.
- The TCP waits for a period of time equal to its retransmission timer and then tries to retransmit the unacknowledged packet
- Initially the retransmission timer is equal to one round trip time, i.e. the time interval between sending the packet and the arrival of the acknowledgement
- Each time the retransmission of the packet fails the retransmission timer is doubled.
- In wireless network, this behavior can degrade the performance of the TCP.

Reaction of M-TCP to the Disconnection Problem

- The SH monitors all packets sent to the MH and ACKs returned from the MH.
- If the SH does not receive an ACK for some time, it assumes that the MH is disconnected.
- It then **chokes** the sender by setting the sender's window size to 0.
- Setting the window size to 0 forces the sender to go into **persistent mode**, i.e., the state of the sender will not change no matter how long the receiver is disconnected.
- In this state the sender does not try to retransmit the data.

Reaction of M-TCP to the Disconnection Problem

- When SH detects the connectivity again it sets the sender’s window back to it’s original value.
- The sender can now start the transmission once again at the same speed as before
- Useless retransmission are avoided.

Reaction of I-TCP and S-TCP to the disconnection problem

- In case of I-TCP after disconnection of the mobile host, the access point would have to buffer more and more data. Thus, no longer the disconnection, larger would be the buffer size is required.
- Also handover follows the disconnection the large amount of buffered packets have to be transferred to the new access point.
- In case of S-TCP snooping is no longer be useful as the mobile host is disconnected and hence the AP can not receive any ack.

The advantages of M-TCP

- It maintains the TCP end-to-end semantics. The SH does not send any ACK itself but forwards the ACKs from the MH.
- If the MH is disconnected, it avoids useless retransmissions, slow starts or breaking connections by simply shrinking the sender's window to 0.
- Since it does not buffer data in the SH as I-TCP does, it is not necessary to forward buffers to a new SH. Lost packets will be automatically retransmitted to the new SH.

Disadvantages of M-TCP

- As the SH does not act as proxy as in I-TCP, packet loss on the wireless link due to bit errors is propagated to the sender. M-TCP assumes low bit error rates, which is not always a valid assumption.
- A modified TCP on the wireless link not only requires modifications to the MH protocol software but also new network elements like the bandwidth manager.

Fast retransmit/fast recovery

- As soon as the mobile host registers at a new foreign agent using mobile IP, it starts sending duplicated acknowledgements to correspondent hosts.
- The proposal is to send three duplicates.
- This forces the corresponding host to go into fast retransmit mode and not to start slow start, i.e., the correspondent host continues to send with the same rate it did before the mobile host moved to another foreign agent.

- As the mobile host may also go into slow start after moving to a new foreign agent, this approach additionally puts the mobile host into fast retransmit.
- The mobile host retransmits all unacknowledged packets using the current congestion window size without going into slow start.
- The **advantage** of this approach is its simplicity. Only minor changes in the mobile host's software already result in a performance increase.
- The main **disadvantage** of this scheme is the insufficient isolation of packet losses approach focuses on loss due to handover.

- Forcing fast retransmission increases the efficiency, but retransmitted packets still have to cross the whole network between correspondent host and mobile host. If the handover from one foreign agent to another takes a longer time, the correspondent host will have already started retransmission.

- The approach focuses on loss due to handover.

Transmission/time-out freezing

- Mobile hosts can be disconnected for a longer time
  - no packet exchange possible, e.g., in a tunnel, disconnection due to overloaded cells
  - TCP disconnects after time-out completely
- TCP freezing
  - MAC layer is often able to detect interruption in advance
  - MAC can inform TCP layer of upcoming loss of connection
  - TCP stops sending, but does not assume a congested link
  - TCP can now stop sending and ‘freezes’ the current state of its congestion window and further timers.
  - MAC layer signals again if reconnected, it UNFREEZES the TCP state

- As soon as the MAC layer detects connectivity again, it signals TCP that it can resume operation at exactly the same point where it had been forced to stop.

- The **advantage** of this approach is that it offers a way to resume TCP connections even after longer interruptions of the connection.

- **Disadvantage**
  - TCP on mobile host has to be changed, mechanism depends on MAC layer

Selective retransmission

- In standard TCP acknowledgements are often increasing
  - ACK are always provided sequentially
  - In case of packet loss, the receiver sends duplicate ACK of the last packets received.
  - The TCP then retransmits ALL the packets starting from the missing one
  - This is waste of bandwidth as the receiver might have already received the packets following the missing one
- Selective retransmission as one solution
  - allows for acknowledgements of single packets, not only acknowledgements of in-sequence packet streams without gaps
  - sender can now retransmit only the missing packets.

- Advantage
  - much higher efficiency
  - As the sender retransmits only the missing packets bandwidth requirements is much lower.
- Disadvantage
  - more complex software in a receiver, more buffer needed at the receiver

Transaction-oriented TCP (T-TCP)

- TCP phases
  - connection setup, data transmission, connection release
  - using 3-way-handshake needs 3 packets for setup and release, respectively
  - thus, even short messages need a minimum of 7 packets
  - Three for connection setup, one for data and then again three for release.
- Transaction oriented TCP
  - T-TCP, describes a TCP version to avoid this overhead
  - connection setup, data transfer and connection release can be combined
  - thus, only 2 or 3 packets are needed

- Advantage
  - Efficiency
  - Reduced overhead
- Disadvantage
  - requires changed TCP
  - mobility not longer transparent

TCP over 2.5 / 3G wireless Networks

- The following characteristics have to be considered when deploying application over 2.5G/3G wireless links:
  - **Data rates:** While typical data rates of today's 2.5G systems are 10–20 kbit/s uplink and 20–50 kbit/s downlink, 3G and future 2.5G systems will initially offer data rates around 64 kbit/s uplink and 115–384 kbit/s downlink. Typically, data rates are asymmetric as it is expected that users will download more data compared to uploading.
  - **Jitter:** Wireless systems suffer from large delay variations or 'delay spikes'.
  - Reasons for sudden increase in the latency are: link outages due to temporal loss of radio coverage, blocking due to high-priority traffic, or handovers.

Parameters to adapt TCP to wireless environments

- **Large windows:** TCP should support large enough window sizes based on the bandwidth delay product experienced in wireless systems. With the help of the windows scale option and larger buffer sizes this can be accomplished

- **Large MTU:** The larger the MTU (Maximum Transfer Unit) the faster TCP increases the congestion window

- **Selective Acknowledgement (SACK):** SACK allows the selective retransmission of packets and is almost always beneficial compared to the standard cumulative scheme.

- **Explicit Congestion Notification (ECN):** ECN as defined in allows a receiver to inform a sender of congestion in the network by setting the ECN-Echo flag on receiving an IP packet that has experienced congestion.

Comparison of different approaches for a "mobile" TCP

Approach	Mechanism	Advantages	Disadvantages
Indirect TCP	splits TCP connection into two connections	isolation of wireless link, simple	loss of TCP semantics, higher latency at handover
Snooping TCP	"snoops" data and acknowledgements, local retransmission	transparent for end-to-end connection, MAC integration possible	problematic with encryption, bad isolation of wireless link
M-TCP	splits TCP connection, chokes sender via window size	Maintains end-to-end semantics, handles long term and frequent disconnections	Bad isolation of wireless link, processing overhead due to bandwidth management
Fast retransmit/ fast recovery	avoids slow-start after roaming	simple and efficient	mixed layers, not transparent
Transmission/ time-out freezing	freezes TCP state at disconnect, resumes after reconnection	independent of content or encryption, works for longer interrupts	changes in TCP required, MAC dependant
Selective retransmission	retransmit only lost data	very efficient	slightly more complex receiver software, more buffer needed
Transaction oriented TCP	combine connection setup/release and data transmission	Efficient for certain applications	changes in TCP required, not transparent

References:

Mobile Communications, Second Edition, Jochen Schiller, Pearson Education- Chapter 9.

University Questions:

- Explain Indirect TCP and Snooping TCP.-May 16, Nov 16 -8M