Multiuser Diversity in Wireless Networks

David Tse
Department of EECS, U.C. Berkeley

Wireless Communication Seminar
Stanford University

April 16, 2001

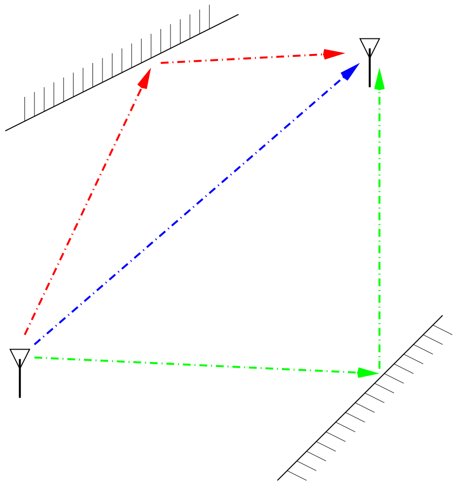
Wireless Communication Theory

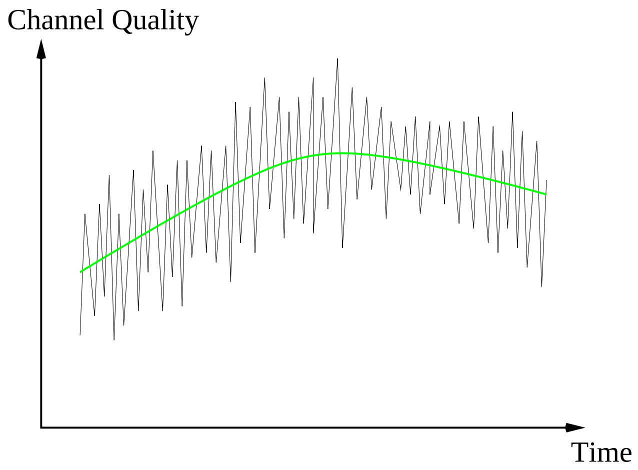
- Wireless communication theory is having a renaissance in the past decade.
- Much of it is centered around how to communicate effectively over fading channels.
- A modern view of communication over fading channels is emerging.
- This view has ramifications to design of not only the physical layer but also higher layers of future wireless data networks.

Multiuser Diversity

- multiuser diversity motivated: information theoretic roots (Knopp and Humblet 95, Tse 97)
- downlink scheduling in Qualcomm's HDR system (Tse 99)
- multiuser diversity in ad-hoc networks (Grossglauser and Tse 2000)

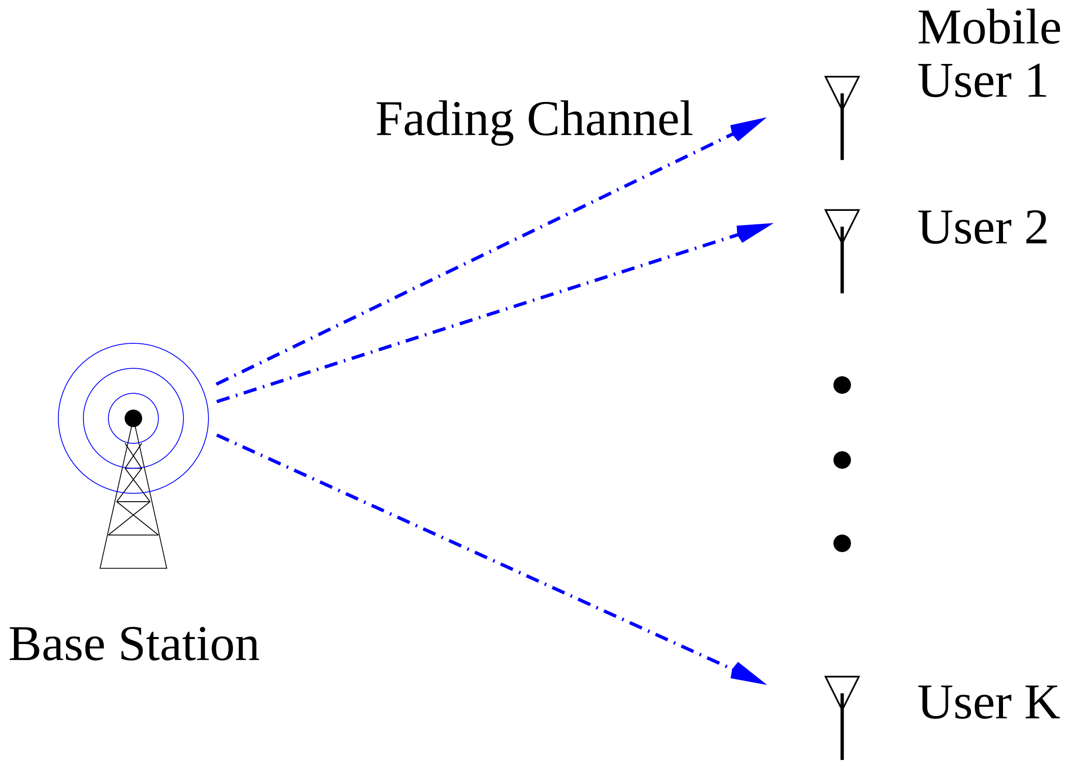
Wireless Fading Channels





- fast time-scale fading due to constructive and destructive interference between multipaths;
- slow time-scale fading due to shadowing effects and varying distances ($r^{-\alpha}$ path loss).

Downlink Multiuser Fading Channel



How should we design the downlink?

Traditional Approach

Two steps:

- orthogonalize the users with respect to time, frequency or codes;
- compensate for fading of individual links via power control, interleaving, frequency hopping, etc.

Goal is to create a set of reliable point-to-point links.

All second-generation wireless systems (GSM, TDMA , CDMA) are designed based on this philosophy.

Let's Step Back

Are there better ways to do things?

What does information theory tell us?

Capacity-Achieving Strategy

Assumption:

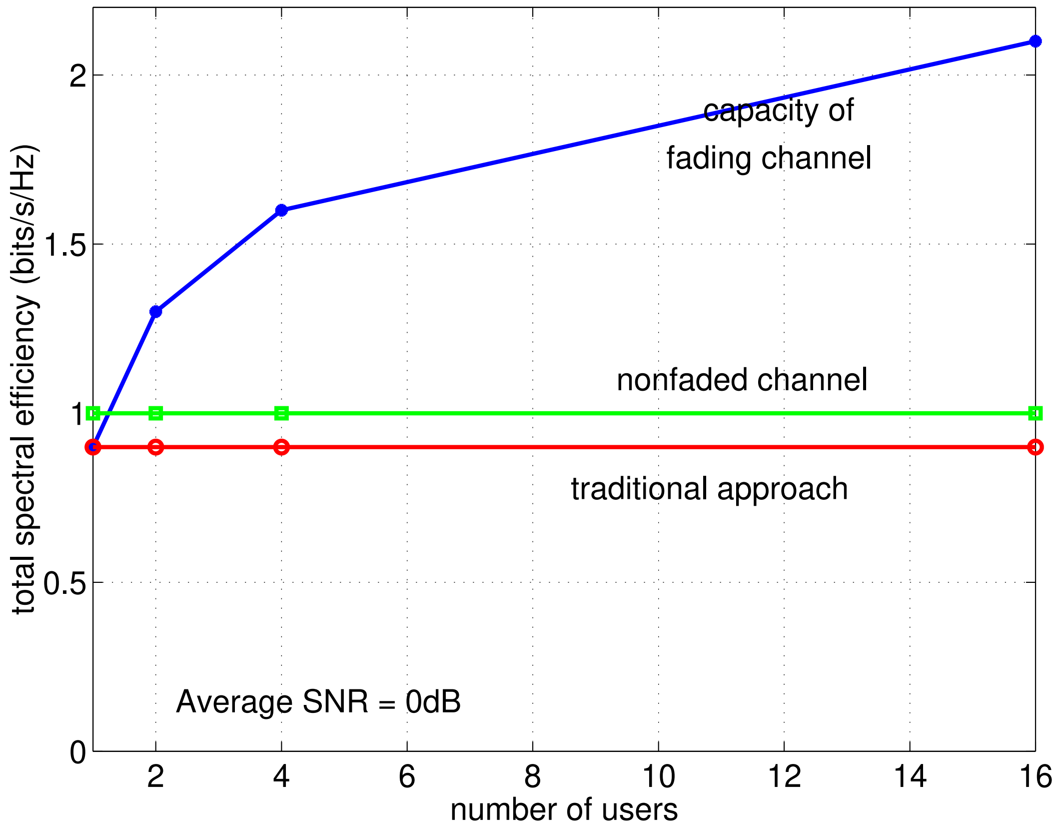
Flat fading channel gains can be perfectly tracked at the receivers and the transmitter.

Result (Tse 97):

Total downlink information capacity is achieved by transmitting only to the user with the best channel at any time.

Users are orthogonalized, but in a channel-dependent way.

To Fade or Not to Fade?



- Capacity of fading channel much larger than non-faded channel!
- In a large system, there is likely to be a user with a very good channel at any time.

Multiuser Diversity

- Diversity in wireless systems arises from independent signal paths.
- Traditional forms of diversity includes time, frequency and antennas.
- Multiuser diversity arises from independent fading channels across different users.
- Fundamental difference: multiuser diversity takes advantage of rather than compensates for channel fading

Talk Outline

- multiuser diversity motivated: information theoretic roots (Knopp and Humblet 95, Tse 97)
- downlink scheduling in Qualcomm's HDR system (Tse 99)
- multiuser diversity in ad-hoc networks (Grossglauser and Tse 2000)

Design of MAC Layer

Information theory result cannot be directly applied!

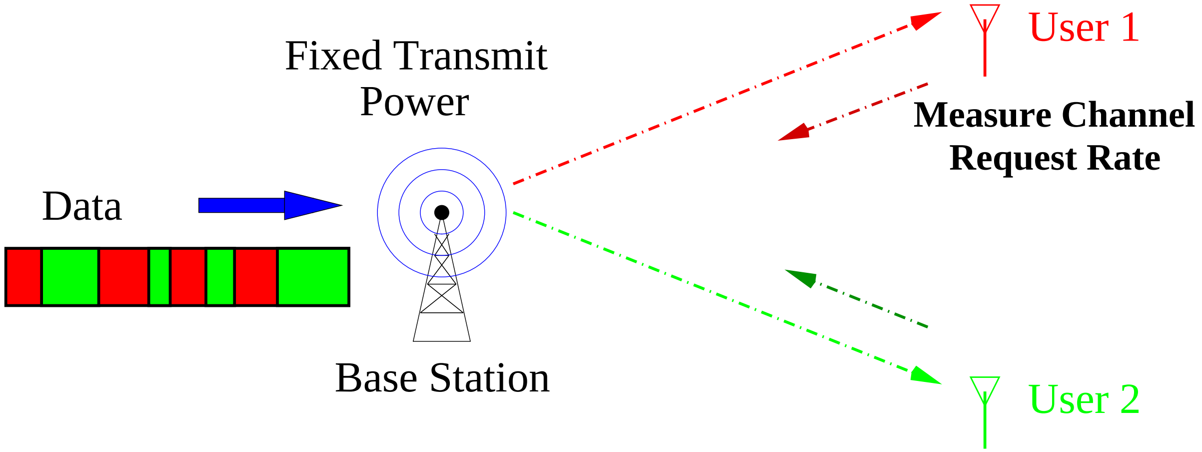
- fairness: users have different channel statistics
- delay: some control of latency is needed, even for data applications.

Challenge:

How to exploit multiuser diversity while sharing the benefits fairly to individual users, within reasonable latency.

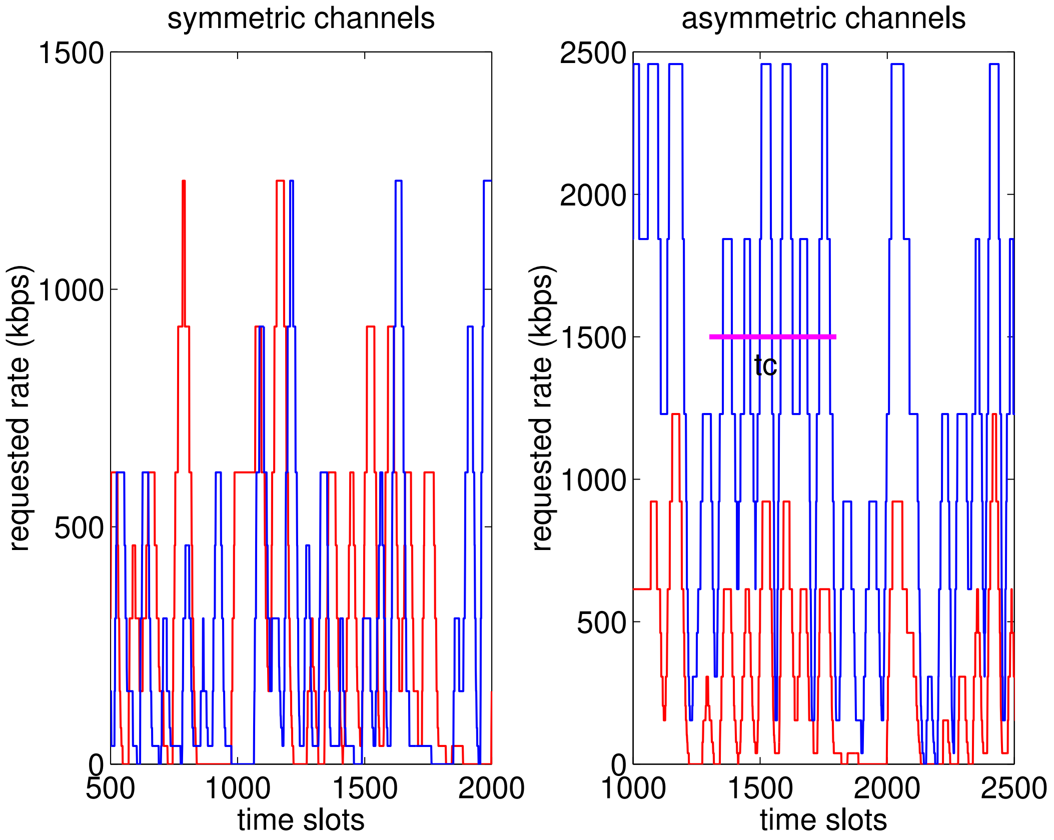
Qualcomm HDR's DownLink

High Data Rate (HDR) : a wireless data system operating on IS-95 band (1.25 MHz)



- HDR downlink operates on a time-division basis.
- Scheduler decides which user to serve in each time-slot.

Multiuser Diversity Refined



At any time there is a user whose channel is near its peak within the latency time-scale t_c.

Proportional Fair Scheduling Algorithm

At time slot t, given

1) users' average throughputs $T_1(t), T_2(t), \ldots, T_K(t)$ in a past window.

2) current requested rates $R_1(t), R_2(t), \ldots, R_K(t)$

transmit to the user k^* with the largest

$$\frac{R_k(t)}{T_k(t)}.$$

Update T_k's for next time slot:

$$T_k(t+1) = \begin{cases} (1 - \frac{1}{t_c})T_k(t) + \frac{1}{t_c}R_k(t) & k = k^* \\ \\ (1 - \frac{1}{t_c})T_k(t) & k \neq k^* \end{cases}$$

Comments

- If users have symmetric channel statistics, this reduces to the greedy policy of transmitting to the mobile with the highest requested rate.
- If channels have different statistics, competition for resource is made fair by normalization
- feedback is built into the metric $R_k(t)/T_k(t)$ to provide a fair bandwidth allocation over the time-scale t_c.

Comparison with Round-Robin Policy

Round-Robin Policy

- Give same number of time slots to all the users in a round-robin fashion, regardless of their channel conditions.

Proportional fair policy:

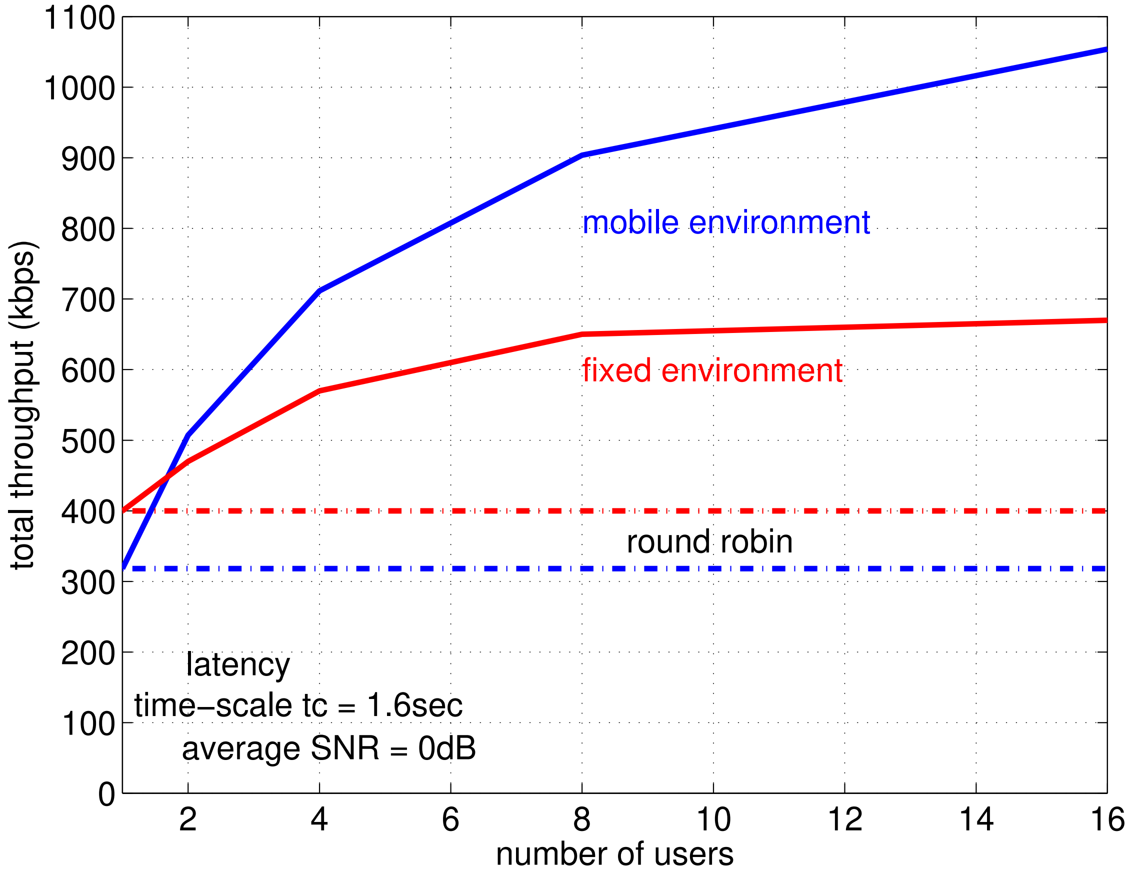
- Give roughly the same number of time slots to all users, but try to transmit to a user when its channel condition is near its peak.

Proportional Fairness

Under stationary assumptions, average throughputs $T_1^*, \ldots, T_K^*$ of the scheduler has the property that the throughput $T_1, \ldots, T_K$ of any other policy must satisfy:

$$\sum_{k=1}^{K} \frac{T_k - T_k^*}{T_k^*} \leq 0.$$

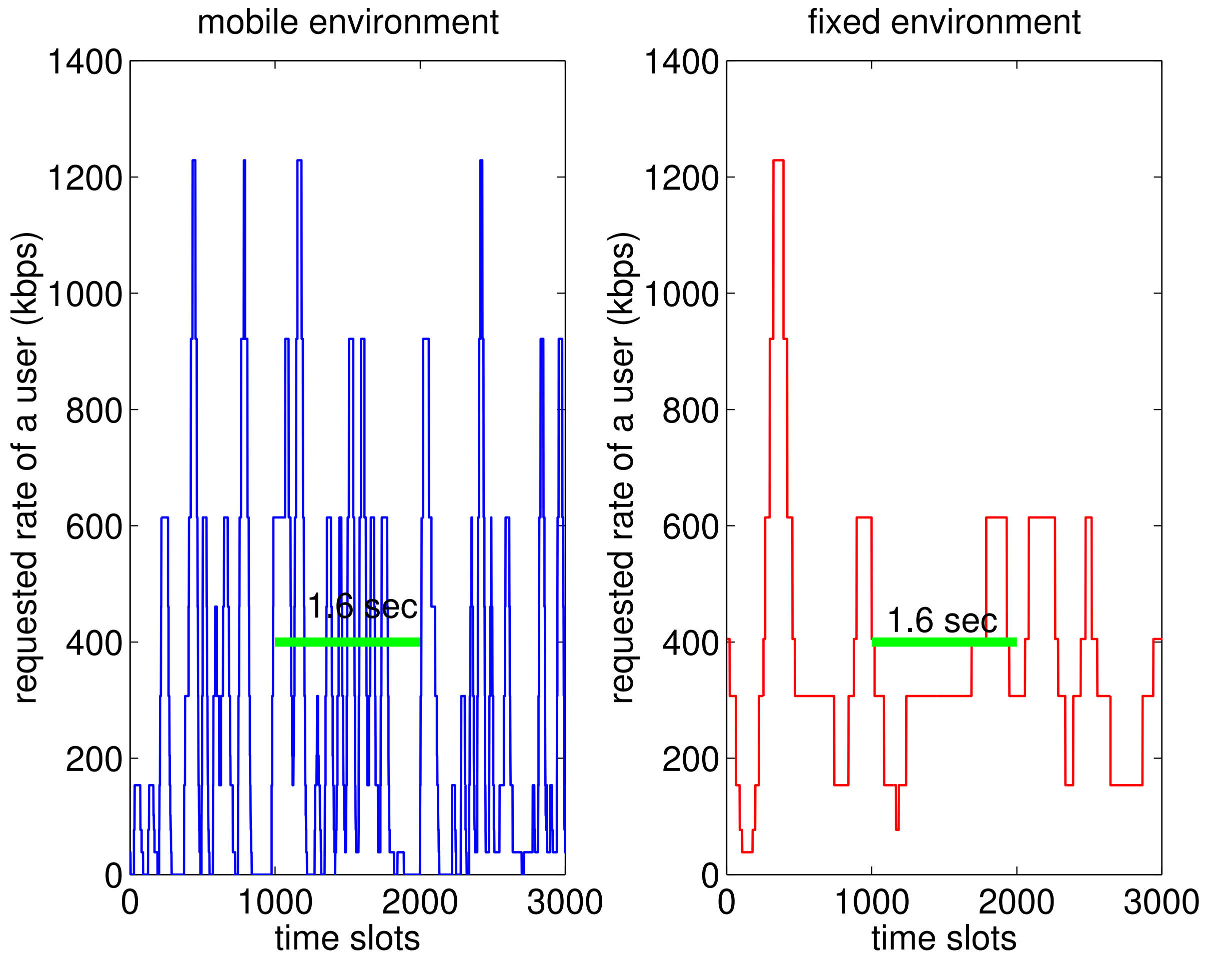
Throughput of HDR Scheduler: Symmetric Users



Mobile environment: 3 km/hr, Rayleigh fading

Fixed environment: 2Hz Rician fading with $K = 5$.

Channel Dynamics



Channel varies faster and has more dynamic range in mobile environments.

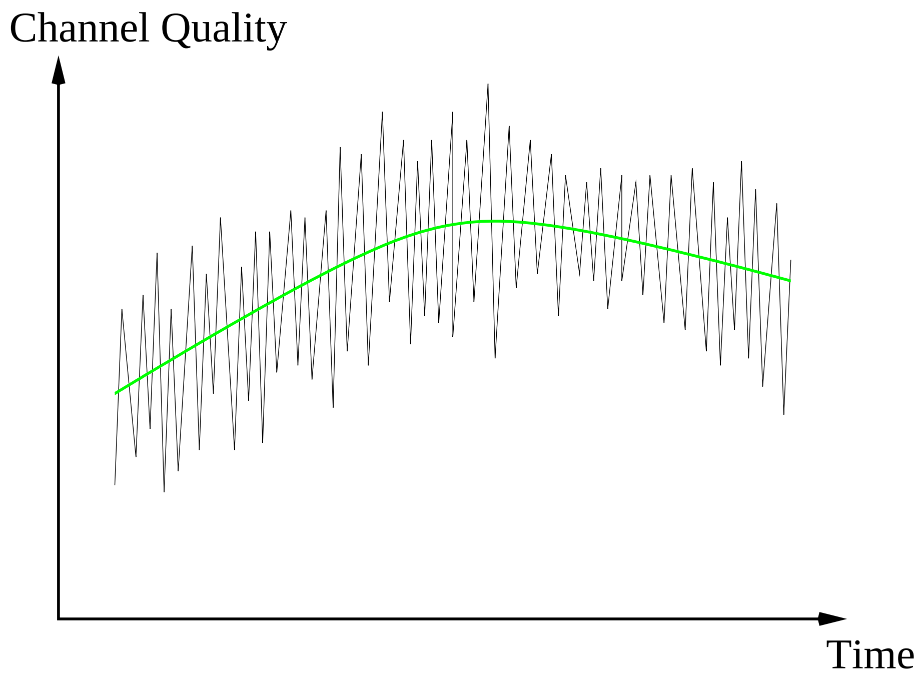
Recap: A Modern View of Fading

Traditional view:

- Fading is bad: a source of unreliability.
- has to be compensated for.

Modern view:

- Fading can be good: a source of randomness.
- the more random the better.
- In a system with large number of users, it is the peak channel condition that governs performance, not the mean.



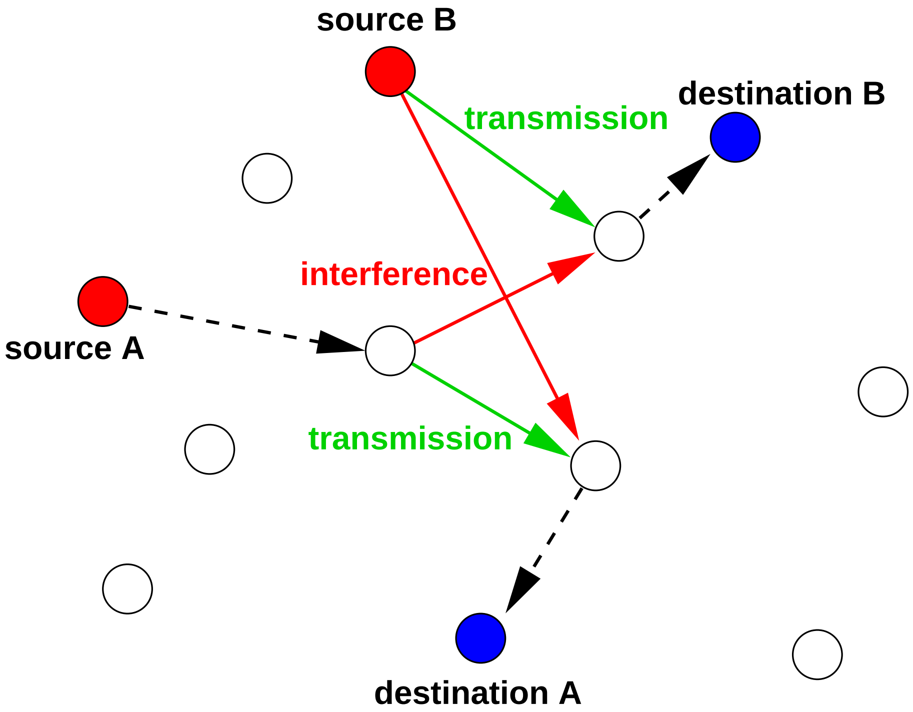
So far we have exploited randomness from fast fading.

How about the slow fading?

Talk Outline

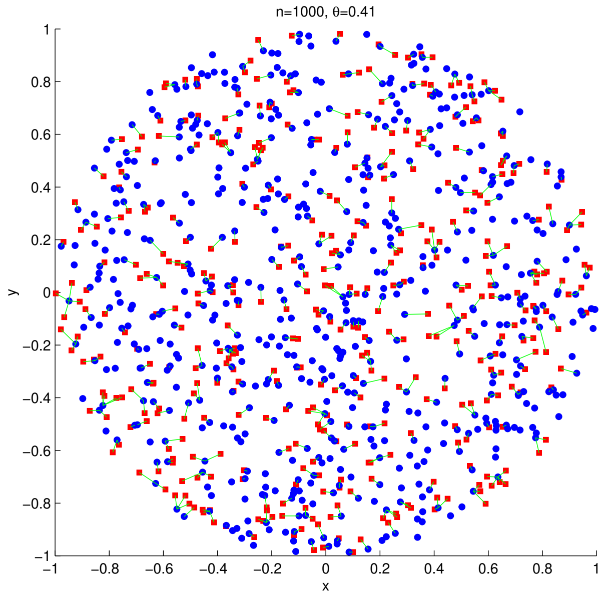
- multiuser diversity motivated: information theoretic roots (Knopp and Humblet 95, Tse 97)
- downlink scheduling in Qualcomm's HDR system (Tse 99)
- multiuser diversity in ad-hoc networks (Grossglauser and Tse 2000)

All-Wireless Ad-Hoc Networks



- no wireline infrastructure
- nodes can be relays for other nodes
- packets can go through multiple hops

Scalability of Ad-Hoc Networks



Point-to-point: Suppose each node has a stream of traffic for a particular destination node.

Question: How does the throughput per source-destination pair scale with the number of nodes n per unit area?

Mobile Ad-hoc Network Model

- Mobility model: Nodes move randomly and independently on a disk of unit area.
- Channel model: path loss factor of $r^{-\alpha}$ at distance r, with $\alpha > 2$ (slow fading);
- Communication model: a packet is successfully received if signal-to-interference ratio is greater than a prescribed threshold.

Baseline: Fixed Nodes Case

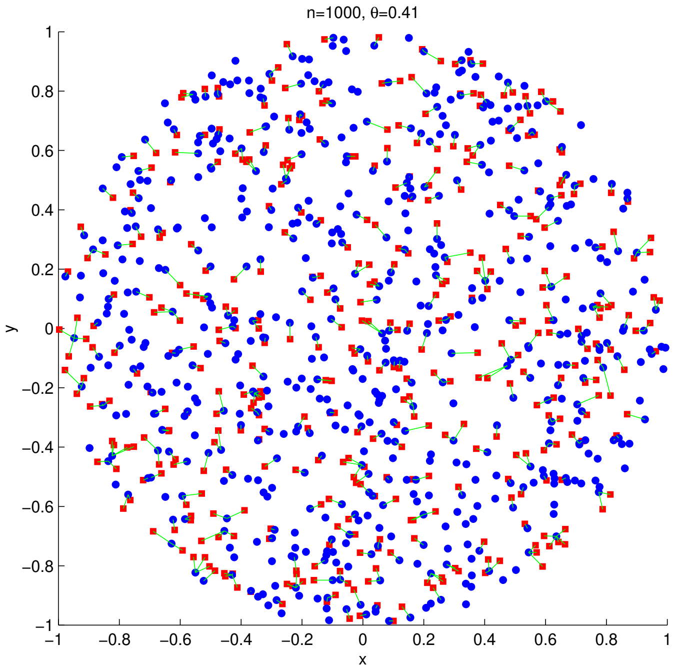
Suppose nodes are randomly placed but immobile.

Gupta and Kumar:

The best achievable throughput per source-destination pair decreases to zero like $O(\frac{1}{\sqrt{n}})$.

This result holds even allowing optimal and centralized scheduling, power control, relaying and routing.

What's the Problem?



- communication confined to nearest neighbors to allow dense spatial reuse.
- a typical route has order $\sqrt{n}$ relay nodes
- each packet is transmitted of the order $\sqrt{n}$ times.

Mobility Can Help!

Main result: (Grossglauser and Tse 00)

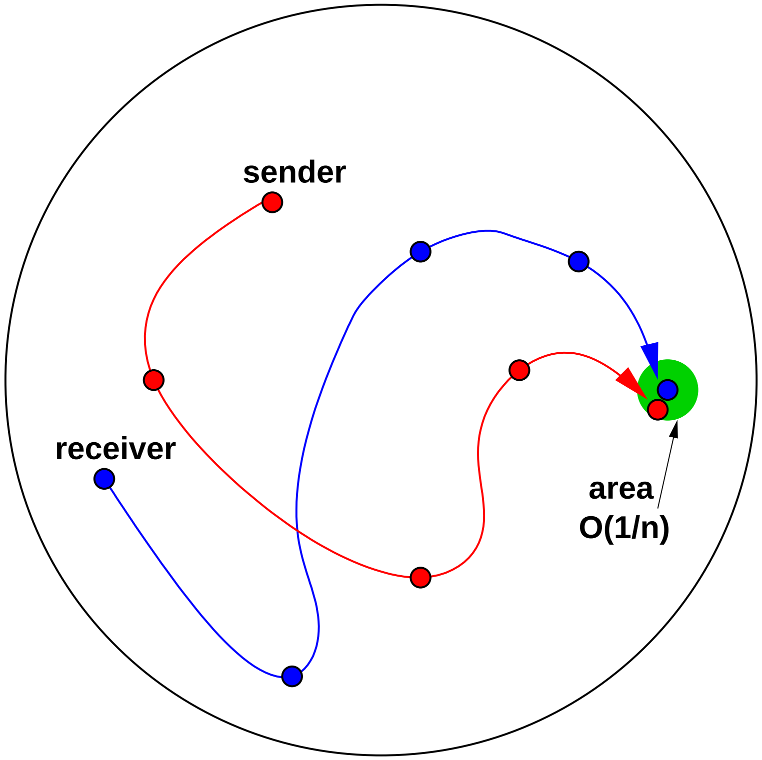
A long-term throughput of $O(1)$ per S-D pair can be achieved when nodes are mobile.

Throughput averaged over the time-scale of mobility.

Idea $\#1$

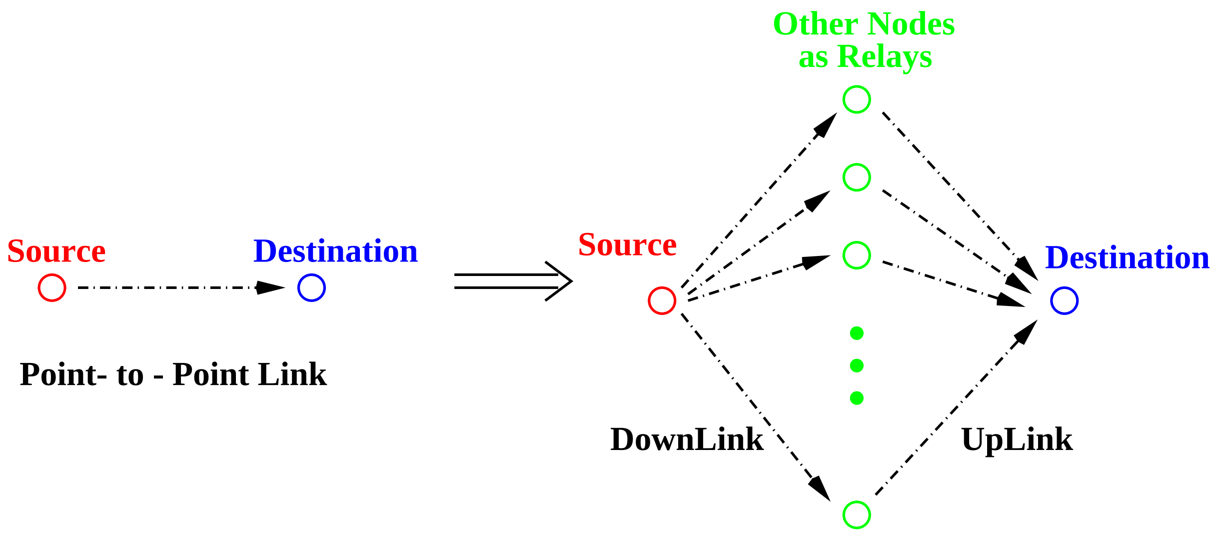
Communicate only when the source and destination are nearest neighbors to each other.

Direct Communication Does Not Work



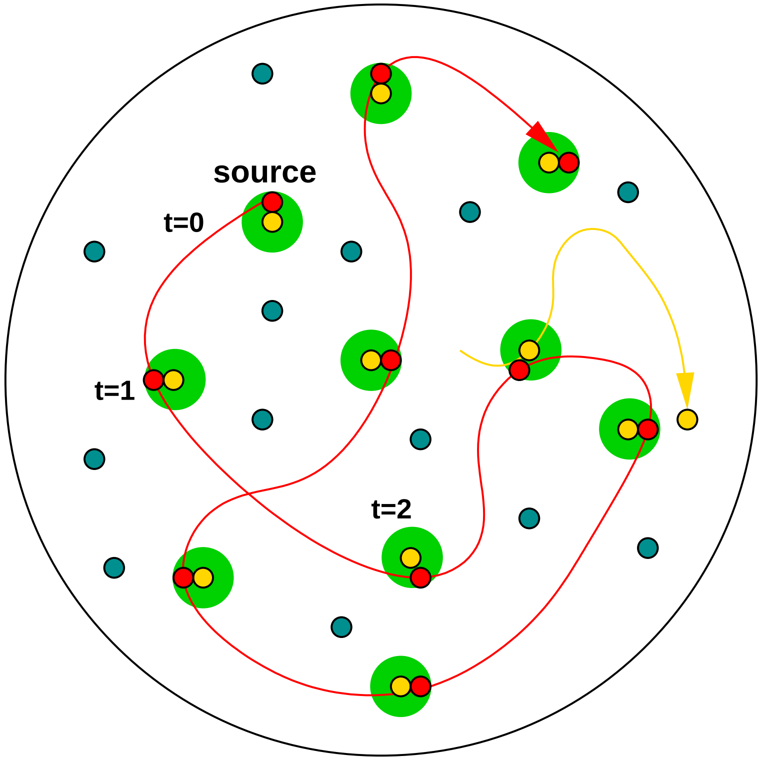
- The source and destination are nearest neighbors only $O(1/n)$ of the time
- In fact, can show S-D throughput is at most $O\left(n^{-\frac{1}{1+\alpha/2}}\right)$ for any policy that does not use relays.

Multiuser Diversity via Relaying



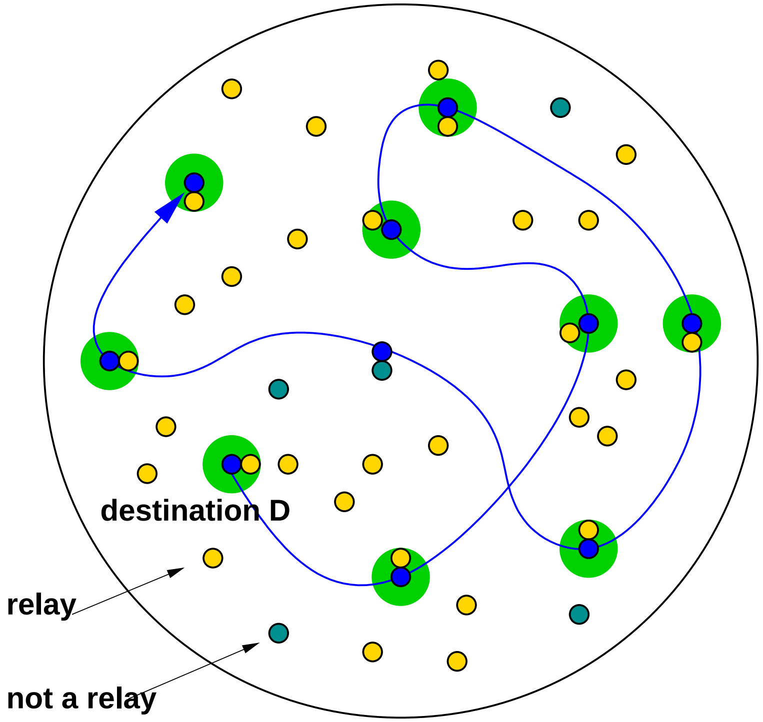
- Multiuser diversity created artificially using all other nodes as relays.
- Channel variation comes from slow rather than fast time-scale fading.

Phase I: Source to Relays



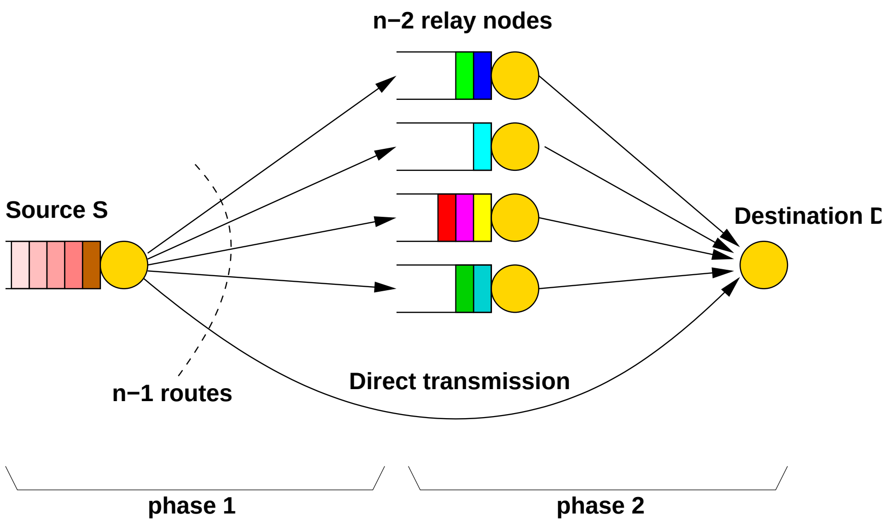
- At each time slot, source relays a packet to nearest neighbor
- Different packets are distributed to different relay nodes.

Phase II: Relays to Destination



- Steady state: all nodes have packets destined for D
- Each relay node forwards packets to D only when it gets close.

Phase I and II Staggered



- $O(1)$ throughput from S to D
- Communication is confined to nearest neighbors, but each packet goes through at most two hops.

Network Capacity

- The above discussion pertains to a single source-destination pair.
- We have to show that every S-D pair can follow the same strategy simultaneously.
- $O(n)$ simultaneous nearest neighbor communication is possible, due to power law decay of the received power from a randomly located node.

Discussion

- Throughput is gained at the expense of delays, of the time-scale of significant changes in network topology.
- Appropriate for delay-insensitive applications, such as email delivery, database synchronization, content distribution, etc.
- The result is extreme in the delay-throughput tradeoff, but suggests that mobility should be exploited in designing scalable architectures for these applications.

Mobile Infostations

- Infostations (WINLAB) : a network of “gas stations” for delay-insensitive data traffic.
- A variation of classical cellular architectures but without universal coverage.
- Our architecture takes it one step further: every node becomes a mobile Infostation; data is distributed over the entire network.

Conclusions

- In a multiuser system, channel fading is a source of randomization that should be taken advantage of, not fought against!
- This point of view can be applied to the design of wireless networks at the physical, networking and architectural levels.