

# Mobility: A Double-Edged Sword for HSPA Networks

## A Large-scale Test on Hong Kong Mobile HSPA Networks

Fung Po Tso<sup>†</sup> Jin Teng<sup>‡</sup> Weijia Jia<sup>†</sup> Dong Xuan<sup>‡</sup>

<sup>†</sup>City University of Hong Kong, Tat Chee Avenue, Kowloon, HKSAR, China

<sup>‡</sup>The Ohio State University, Columbus, OH, USA, 43202

### ABSTRACT

This paper presents an empirical study on the performance of mobile High Speed Packet Access (HSPA, a 3.5G cellular standard) networks in Hong Kong via extensive field tests. Our study, from the viewpoint of end users, covers virtually all possible mobile scenarios in urban areas, including subways, trains, off-shore ferries and city buses. We have confirmed that mobility has largely negative impacts on the performance of HSPA networks, as fast-changing wireless environment causes serious service deterioration or even interruption. Meanwhile our field experiment results have shown unexpected new findings and thereby exposed new features of the mobile HSPA networks, which contradict commonly held views. We surprisingly find out that mobility can improve fairness of bandwidth sharing among users and traffic flows. Also the triggering and final results of handoffs in mobile HSPA networks are unpredictable and often inappropriate, thus calling for fast reacting fallover mechanisms. We have conducted in-depth research to furnish detailed analysis and explanations to what we have observed. We conclude that mobility is a double-edged sword for HSPA networks. To the best of our knowledge, this is the first public report on a large scale empirical study on the performance of commercial mobile HSPA networks.

### Categories and Subject Descriptors

C.4 [Performance of Systems]: Measurement techniques;  
C.2.3 [Network Operation]: Network monitoring

### General Terms

Measurement

### Keywords

Mobile HSPA Networks, Performance Evaluation, Bandwidth Sharing

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*MobiHoc '10*, September 20–24, 2010, Chicago, Illinois, USA.  
Copyright 2010 ACM 978-1-4503-0183-1/10/09 ...\$10.00.

## 1. INTRODUCTION

Wireless communication networks have seen a large increase in capacity and rapid progress in real world deployment [10]. The recent High Speed Packet Access (HSPA) [2], which is commonly referred to as 3.5G cellular standard, is conceived as a natural evolution of the existing Wideband CDMA (WCDMA, known as 3G). HSPA is a combination of High Speed Downlink Packet Access (HSDPA) and High Speed Uplink Packet Access (HSUPA). It has emerged as one of the most dominant beyond-3G technologies and has been serving many regions and countries to support broadband applications.

### 1.1 Motivation

HSPA offers data rates up to 14.4 Mbps in downlink and 5.76 Mbps in uplink for stationary users. Therefore, stationary users can easily access broadband applications in most areas with good signal quality through HSPA [9]. Nevertheless, the most important advantage of wireless communication systems is to support users' mobility. Mobility is crucial in daily life as it provides end users with flexibility and opportunities to easily connect with people world wide. Many companies often adopt mobile data access to expand their business coverage, save time and improve productivity. The most common forms of mobility obviously include walking, driving and traveling via public transport. Passengers on public transport are more likely to entertain themselves via forthcoming mobile data access than people who are driving vehicles or walking. In Hong Kong, for instance, millions of people use public transport that traverses the city everyday. Many, virtually each, of these passengers own a mobile phone [1]. Among them, more and more have Smartphones and Netbooks. With these on-body small devices communicating over 3G and beyond networks, people can entertain themselves on the move using broadband applications such as YouTube [26]. They can also share live video captured by mobile phone cameras like Livecast [15], or Veodia [22]. All these broadband applications require the upload and/or download of realtime multimedia contents, for which there is support for stationary users [9]. But that is not necessarily feasible when users are on the move.

Despite its importance, the performance of HSPA networks in the mobile environments has not been deeply investigated before in a large-scale operational network via on-site tests. Extensive efforts have been devoted to theoretically analyze network performance or merely to study performance for VoIP. The empirical study of WCDMA networks has been performed in stationary environments [21],

but this provides no information about the performance for mobile users. The performance of mobile HSPA remains highly unclear in practice, as radio conditions vary with different speeds, locations and time in a mostly unpredictable fashion. Hence, an extensive empirical study for mobile HSPA is needed to resolve these performance uncertainties.

## 1.2 Contributions

We have studied the performance of two largest commercial HSPA networks in Hong Kong with extensive field tests. Our tests cover virtually all possible mobile scenarios in urban areas - from land scenarios (e.g. buses) to sea scenarios (e.g. ferries) and from ground scenarios (e.g. trains) to underground scenarios (e.g. subways) - totaling over 100 km distance during a 3 months' period. We investigate their data throughput, round trip time (RTT), and packet loss rate under different network traffic loads and various traffic patterns. We also examine the behavior of network resource allocation mechanisms and the fairness of the radio-link schedulers in allocating the bandwidth resources to multiple data calls in a HSPA network. Here we present our findings and conclusions on a very high level as follows:

First, we have confirmed with solid data and detailed analysis that mobility has largely negative impact on the performance of HSPA networks. Fast changing physical wireless environment and frequent handoffs both contribute to the degradation of overall HSPA service provisioning. In particular, we make the following two observations:

- Mobile HSPA performance is greatly different from static HSPA performance. We record a large spread of HSPA throughput (e.g., 0 kbps to 3500 kbps) and signal quality (-1 dBm to -32 dBm) for mobile User Equipments (UEs). In spite of the degradation, our data show that subways not only have the highest signal quality and coverage but also give the highest throughput among all mobile scenarios.
- We measure and report HSPA's performance in transitional regions. Handoffs, which normally happens in the transitional regions, are long thought to have been properly handled since the time of 2G networks. However they are still a major problem restricting the traffic volume of HSPA networks. We show that during a handoff, the service is virtually interrupted and the throughput of end users decreases to almost zero.

Second, our field experiment results have shown several unexpected new findings and thereby exposed some misunderstandings commonly held for the HSPA networks:

- *Common View: Mobility is irrelevant, if not detrimental, to the fairness in HSPA bandwidth sharing among users*

We observe that the bandwidth sharing practice in stationary HSPA environments is unfair. In contrast, mobility surprisingly improves fairness of bandwidth sharing among users by diminishing the chance of a consistent unfair scheduling strategy in one single cell. The sharing is improved with mobility such that all UEs get nearly equal service.

- *Common View: Mobility affects all flows equally. And TCP flows suffer more than UDP ones*

Due to TCP's rigid control strategy, TCP flows are in great disadvantage competing for bandwidth with UDP ones. Intuitively TCP flows should see worse service provisioning with more packet losses and larger delays caused by mobility. However, the analysis of our experiments data reveals that TCP flows unexpectedly see much better performance during mobility than UDP flows. In mobile scenarios, the round trip time (RTT) of TCP flows, as well as its deterioration due to mobility, is far less than UDP ones. It compensates for TCP's disadvantage in competing with UDP traffic for network access, especially when the network traffic is very heavy.

- *Common View: Handoffs are triggered in the transitional region between cells and always result in a better wireless connection*

We have found out that the triggering and the final result of handoffs are often unpredictable. Sudden deep fading can force a UE to make a premature handoff. And in nearly 30% of all handoffs, selection of a base station with poorer signal quality can be witnessed. This discovery calls for fast reacting fallover mechanisms, when UEs in HSPA networks have undergone an unsuccessful handoff.

In short, we have concluded that mobility has both favorable and unfavorable impacts on the performance of HSPA, just like a double-edged sword. In order to gain better understanding and control of this blade, we have conducted in-depth research to furnish analysis and explanations to what we have observed later in this paper, in hope of exposing the fundamental tradeoffs in HSPA network implementation.

To the best of our knowledge, this is the first public report on a large scale empirical study on the performance of commercial mobile HSPA networks. Our experimental results are useful for network planners/designers. They can plan better cell coverage to alleviate harmful transition from a already poor to a poorer cell. Moreover, they can improve network capability to address users' demands while mobile and further. Furthermore, they can change their network control policies to make themselves more competitive on the market. Our findings are also helpful for application developers in designing mobile buffering algorithm and uplink flow control to eliminate side effect of handoff. Application developers can use a more aggressive algorithm for mobile users to better utilize network resources based on the fact that mobility improves the fairness of bandwidth sharing.

The remainder of this paper is organized as follows. In Section 2, we give a brief introduction of HSPA technologies. In Section 3, we describe our measurement methodology. In Section 4, 5 and 6, our experimental observations and analysis are presented. We conduct a survey of related work in Section 7. And Section 8 concludes the paper.

## 2. HSPA BACKGROUND

HSPA networks are now in operation in many regions and countries around the world. It is commonly referred to as 3.5G and comprises two components, namely High Speed Downlink Packet Access (HSDPA) and High Speed Uplink Packet Access (HSUPA).

HSDPA was standardized as part of 3GPP Release 5 in 2002. HSUPA then was included as part of 3GPP release 6

in 2004. HSPA is deployed on top of the WCDMA network so that it shares all the network elements in the core network and in the radio network. The upgrade from WCDMA to HSPA requires only new software packages and some new pieces of hardware in the base station to support higher data rate and capacity. On the other side, HSUPA, more correctly called enhanced dedicated channel (E-DCH), is based on the enhancement of dedicated channels rather than of common channels (as is the case for HSDPA). Like HSDPA, HSUPA also achieves higher data rates, primarily through rapid scheduling of resources, fast retransmissions and channel adaptation, but adds the possibility of using lower spreading factors. HSUPA theoretically has uplink speeds up to 5.76 Mbps under ideal radio conditions [3].

### 3. MEASUREMENT METHODOLOGY

We have performed our measurements over two major commercial HSPA networks in Hong Kong. In this section, we introduce the measurement routes, setup as well as metrics in details.

#### 3.1 Measurement Routes

In our measurements, we have covered virtually all possible mobile scenarios in urban areas spreading from land to sea and surface ground to underground, including subways, trains, city buses, ferries and self-driving vehicles. The trace map is shown in Fig. 1. These kinds of mobile transport systems have different velocities, directions, traces and mostly cover all the popular places in Hong Kong. The characteristics of different types of transportation are as follow:

*Trains* - They are on surface ground. Stations are sparsely distributed so that sometimes the trains are moving with speed up to 100kmh. They have average moving speed of 40kmh. We select East Line for carrying out the experiments is simply because it is the busiest line among train lines.

*Subways* - They are in underground. Stations are densely distributed so that the average moving speed is about 30kmh and the highest speed is 80kmh. We most frequently use Tsuen Wan line because it consists of both surface ground and underground stations, also a long harbor crossing tunnel.

*Self-driving Vehicles & Buses* - They are on surface ground

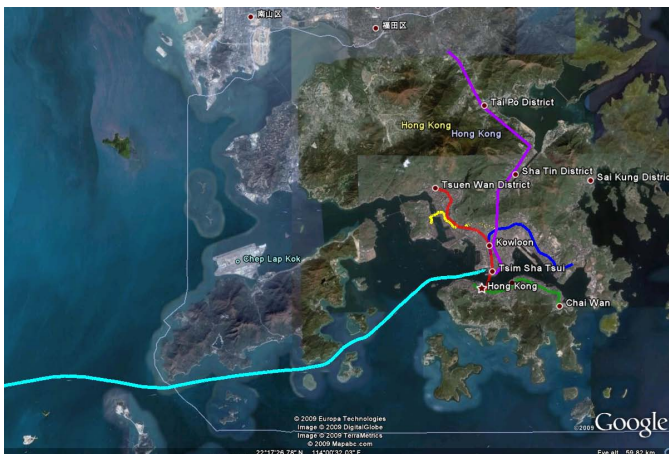


Figure 1: Measurement traces in Hong Kong

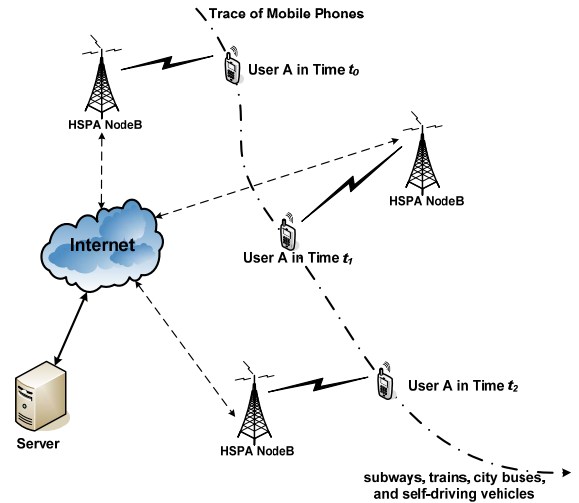


Figure 2: Measurement Setup

Table 1: Adopted bit rate and packet size

Packet Size	Bit Rate	
256Bytes	64 kbps	512 kbps
512Bytes	128 kbps	1024 kbps
1024Bytes	256 kbps	2048 kbps

too and move around open urban areas. The average moving speeds are about 50kmh and 30kmh respectively. We choose a bus route that cover urban roads and a highway, and then we follow the same route using a self-driving vehicle.

*Ferries* - They are of course in the sea and it has the highest average moving speed of 80kmh among four types of listed transportation. The ferry line we select is a regular transportation connecting Hong Kong and Macau.

Our main data come from train and subway lines, since they feature stable speed and constant traces. Measurements acquired on buses, ferries are used as complementary data and hence the reported results are combination of above mentioned transport modes. The speeds of railway (subways & trains), buses and off-shore ferries average respectively 40kmh, 30kmh and 70kmh.

#### 3.2 Measurement Setup

Our measurement setup is shown in Fig. 2. We used a set of client-and-server programs to conduct such mobility experiments. The servers are PCs running Linux with Kernel 2.6.20. Two identical servers are separately hosted in City University of Hong Kong and a commercial data center to ensure the results will not be affected or biased by anything specific to one server and its path. We have also developed a set of data collection programs in C++ and Java, and data analyzing tools in C++ and MATLAB script. Our UEs include four laptops equipped with four MC950D HSPA modems, two HTC G1 smartphones, and two iPhones.

We have conducted three types of evaluations: download only, upload only and simultaneous download & upload. These three types can almost cover all kinds of our frequently used applications. Downloading communications typically occur with FTP transmissions and website brows-

ing, while uploading communications are closely associated with file sharing or P2P services. For simultaneous download & upload communications, bidirectional sessions, such as video calling/conferencing and online gaming, are the most typical. These tests were done by sending bursts of data packets back-to-back from UEs to the servers via HSPA links and measuring throughput based on the inter-arrival times between packets in a burst. Table 1 illustrates the bit rates and packet sizes adopted in our experiments using both TCP and UDP. We purposely choose different bit rates for emulating different types of forthcoming services including realtime light and bulky UDP data calls and non realtime light and bulky TCP data calls. Different packets are also used in order to find out the network behaviors due to various packet sizes. For example, 2048 kbps throughput has been used to intentionally maximize the cell's throughput so that we can determine the behavior of bandwidth sharing under saturated network.

We used throughput, round trip time (RTT) and loss rate of UDP traffic as primary metrics for comparison throughout this study because they are directly related with QoS of traffic sessions. Throughput is always the first users' impression for perceiving the network performance. But is it a sufficient indicator of the actual network performance? UDP's connectionless mechanism allows constant bit rate regardless of actual network condition so that its throughput is highly biased in comparison with that of TCP. When network is experiencing congestion, TCP's reliable transmission mechanism will refrain the terminal from sending too much to the congested network. In comparison, UDP keeps sending at original speed leading to potentially many packet losses. Therefore, when users want to learn more, RTT, which is referred to as application layer level round trip time, is often used to estimate the degree of network congestion. Less congested (smaller RTT) network path is capable of providing better QoS for certain types of services, e.g. VoIP. Along with the metrics, we have also logged the Cell ID, as well as the  $E_c/I_o$  of the UEs, queried via AT commands, while the measurements were being conducted. Thus we can establish relationship between the HSPA performance data with the physical location and the cell information. Through cell information, many impacts of mobility on performance, such as handoffs between cell towers and signal fluctuation, can be observed more clearly. So it is possible for us to evaluate the impact level of different mobility-related features. The above method is not only applied to network saturation measurements, but also applied to throughput measurements under different traffic modes.

At the beginning of measurement, we calibrated our GPS receiver for logging the moving speed and tracks of the transport. We then checked the associated Cell ID of all UEs to make sure they are connected to the same cell. During the measurement, all UEs sat next to one another to ensure they have identical, at least very similar, test environment. For each test journey, half of UEs ran TCP tests and the other half ran UDP tests at the same packet size and bit rate setup. All tests have been carried out at off-peak hours so as to alleviate the impact of background traffic to the tests and vice versa.

#### 4. GENERAL IMPACT OF MOBILITY

Mobile HSPA performance is very different from static HSPA. As was stated in Section 1.2, we record a large spread

of HSDPA bit rates (Fig. 3) and signal quality (Fig. 4) because the HSPA UEs pass through a variety of dynamic and unpredictable radio conditions along the different mobility routes.

Fig. 3 depicts CDF of mobile HSPA throughput under various mobile scenarios for both Operator A and Operator B. It manifests that the maximum achievable bit rate for downlink is 3500 kbps and 1400 kbps for uplink. It also shows that stationary UEs have throughput falling into a certain short range while mobile UEs have wide spread of throughput. Nonetheless, the recorded throughput for mobile UEs mainly falls into less than 2000 kbps in downlink and less than 600 kbps in uplink, which are far lower than those of static environments. Particularly, there are incredibly high, about 10% to 20% of the time, UEs would experience no throughput at all in all mobile environments compared with less than 4% in stationary environments.

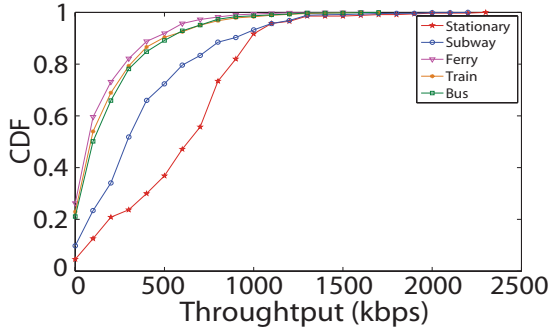
Since HSPA bit rate can adapt quickly to the different radio conditions, stationary UEs perform best, because they see more stable and better radio conditions. In mobility, the rapid change of environment results in poorer radio condition on average and thus causes poorer throughput for UEs. This phenomenon aligns with our data. However sometimes we saw quite good throughput even UEs were in high speed mobility. To this extent, bit rate during acceleration or deceleration should also be dominated by radio conditions that UEs experienced at the moment.

The signal quality for different transport means are shown in Fig. 4. We find that UEs in all mobile scenarios have signal quality spreading from -1 dBm (very good) to -32 dBm (no signal) but rarely have signal quality spreading range more than 5 dBm in stationary environments.

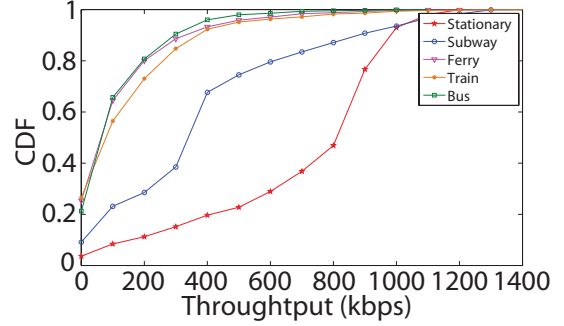
One point to mention is that we are aware that normally the signal quality is often measured by average HS-DSCH signal-to-interference-plus-noise ratio ( $SINR$ ) [11]. Nevertheless, we employ another commonly used metric, the ratio of received pilot energy,  $E_c$ . Since this metric is traditionally used in all CDMA systems and is easier to get than  $SINR$ .

Interestingly, CDF of mobile HSPA throughput (Fig. 3) and PDF of signal quality (Fig. 4) turn out to converge on one common phenomenon. Both figures show that subways conspicuously has the best performance among all selected mobile scenarios. Meanwhile, all other open area mobilities surprisingly have very close (i.e. similar performance) CDF of mobile HSPA throughput and PDF of signal quality curves. Fig. 3 illustrates that overall throughput for UEs in subways apparently outperform those in other mobility situations as subways has the closest CDF with stationary one. On the other hand, despite the fact that open space mobilities such as trains, buses and ferries have larger differences with stationary one, they surprisingly have nearly overlapped CDF of throughput in both operators' network. More interestingly, PDF of mobility signal quality matches strongly with CDF of mobile HSPA throughput. Among them, as shown in Fig. 4, subways has the highest good signal quality coverage of 92% of the journey compared with only 76%, 68% and 58% (fairly close) on trains, buses and ferries respectively for Operator A. Likewise in Operator B, subways also have 85% good signal coverage compared with 67%, 72% as well as 63% (very close) on trains, buses and ferries respectively.

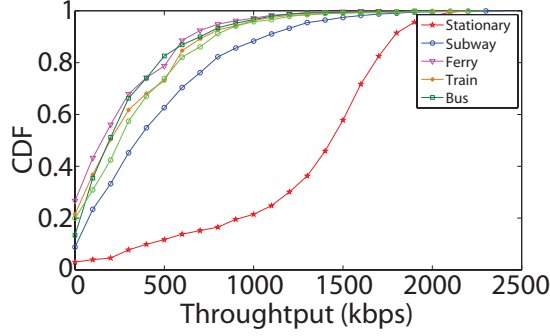
The similarity among CDF of mobile HSPA throughput and PDF signal quality evidently provides us a hint that



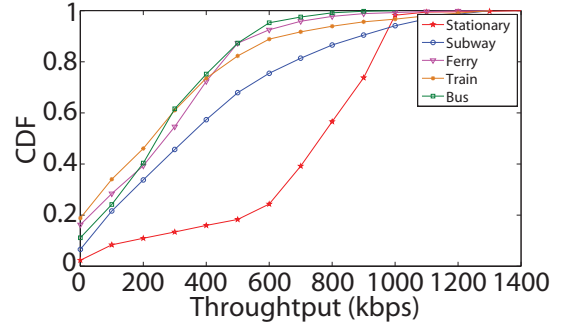
(a) Downlink of Operator A



(b) Uplink of Operator A

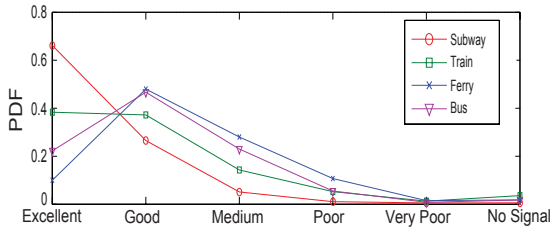


(c) Downlink of Operator B

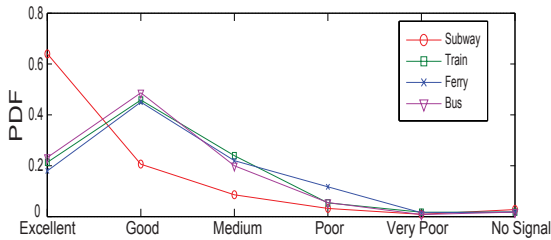


(d) Uplink of Operator B

Figure 3: CDF of HSPA Network Throughput



(a)  $E_c/I_o$  of Operator A



(b)  $E_c/I_o$  of Operator B

Figure 4: PDF of  $E_c/I_o$  for Mobile HSPA Network

there exists a fundamental difference between of underground and surface ground network. The distinct difference is due to different transmission media in use. Open areas are mainly covered by normal cell radio as electromagnetic (EM) wave can propagate efficiently from cell tower to end user UEs at sufficiently low cost. Nevertheless, all subway stations are covered by leaky cables in tubes because EM wave can not be efficiently transmitted through tunnels. Thus, the leaky cables, though the cost is higher, are purposely used to guide and radiate EM wave from station into the tubes to avoid underground communication blind spot and also to provide a highly optimized underground radio coverage.

## 5. MOBILITY IMPACT ON BANDWIDTH SHARING

Mobility hurts the general performance of HSPA networks. However, our experiment data reveal that in some cases, mobility does help some aspects of the networking functionalities in HSPA networks. In this section, we explore the useful side of this two-edged sword.

### 5.1 Bandwidth Sharing among Users

Bandwidth sharing is more discussed together with scheduling or call admission. It is not straightforward to link bandwidth sharing directly with mobility. People may think that the dynamic and unpredictable mobile environment will further aggravate the unfairness of bandwidth allocation among users. However, this is not true. In fact, we have observed that the bandwidth sharing is not as fair as expected in stationary HSPA environment, while mobility

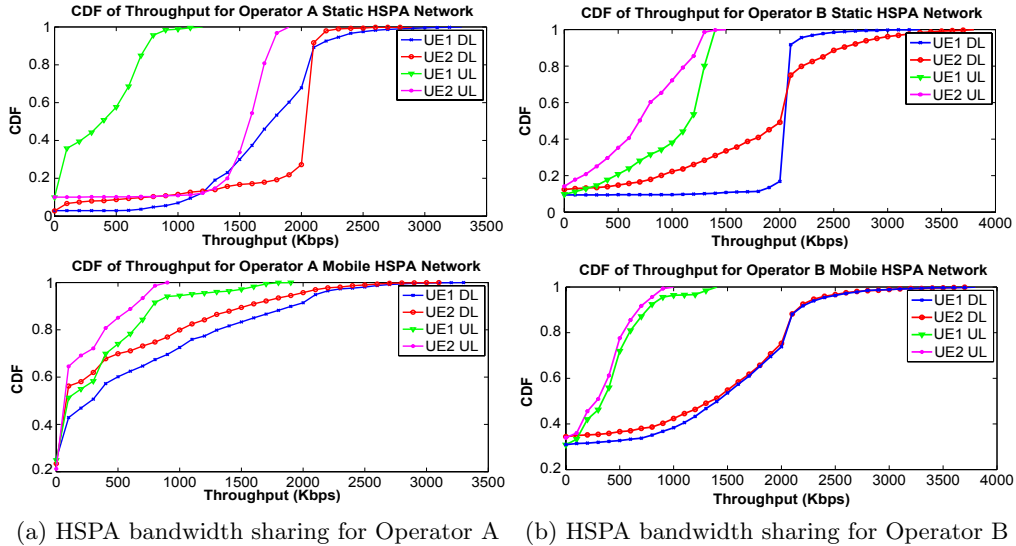


Figure 5: Mobility leads to fairer bandwidth sharing.

actually improves the fairness of bandwidth sharing among users, though it also introduces a plenty of side effects on link performance.

In stationary environment, we observed that a single UE is more likely to occupy most of available bandwidth and leave only a little portion for other UEs. For example, in Fig. 5a, UE2 has throughput at about 2Mbps during 60% of all the time, but UE1, within the same time period, has throughput averagely distributed from 1.3Mbps to 2Mbps. In the static scenario of Fig. 5, the bandwidth allocation is not fair both in the uplink and downlink directions. The behavior somehow violates HSPA bandwidth allocation mechanism, because if UEs were placed in a similar environment, each of them should have had the equal chance to get served by Node B. On the contrary, The improvement can be seen from the Fig. 5b that all UEs have no apparent advantages over one another when moving, though sometime one of them would have no granted bandwidth, but other UEs had equal chance to have the highest bandwidth granted.

There are two reasons why fairness of sharing can be improved with mobility. First, when UEs are on the move, a single UE cannot effectively overwhelm other UEs to take most of available bandwidth due to rapid change of radio condition of the UE, this leaves more available Node-B's bandwidth to be shared by other UEs. Second, the scheduling algorithm is of cell to cell basis, so that the decisions made by one cell would not propagate to next cell for the cell change due to handoff, that means decision maker changes from time to time and thus eventually results in fairer decisions.

We have calculated the auto-correlations and cross-correlations of the signal quality for all UEs in a period of over 20 minutes. The results show that when static, both UEs see almost the same signal quality over the time span and the cross-correlations between UE1 and UE2 is a nearly perfect triangle, which means that their relative value does not change. However, in mobility, there are much 'noises' for all auto-correlations and cross-correlations figures. Auto-correlations normally see a 5% – 20% fluctuation compared

to their static counterpart, and cross-correlations a 10% – 30% fluctuation. The results imply that the radio environment is changing very fast, so an UE can hardly keep its dominance in a cell due to unstable radio signals and ensuing failed transmissions.

## 5.2 Bandwidth Sharing among Traffic Flows

There are different types of traffic flows in HSPA networks, e.g., TCP and UDP flows. As mobility introduces dynamic changes to the fundamental physical environment of end users, all these flows are expected to suffer deterioration caused by mobility. But do all the flows experience the same influence?

At first sight, mobility's impact will not differentiate between high level traffic, e.g., transport layer flows. However, our data point in another direction. Through studying the RTT distribution of TCP/UDP flows, we have found unexpectedly that TCP flows see better performance during mobility. It compensates for TCP's disadvantage in competing with UDP traffic, especially in heavy load scenarios. From Fig. 6, we can see all mobile CDF curves lie to the left of static CDF curves. Though the RTT improvement is marginal, yet it still contradicts the common sense that unstable mobile wireless environment will 'slow down' all flow traffic. In order to account for this, we need to look at the other major category of flows in the network, the UDP flows.

First point to observe regarding UDP flows is that the UDP packets return much slower than TCP packets. From Fig. 6, we can see that almost all TCP packets can return within 2.5 seconds. But for UDP packets, it takes at least 5 seconds for 95% of them to return. In extreme mobile cases, the RTT of some packets can even reach 50 seconds. This is not the case in traditional static networks, as TCP imposes stricter acknowledgement and flow control strategy. We figure out that, besides scheduling difference, the flow control functionality of TCP protocol helps to prevent too much traffic pouring in the network, thus avoiding relatively



long delays, as experienced in UDP traffic. In Fig. 7, we plot the average TCP traffic and UDP traffic pumped into the wireless channel. The figure shows TCP traffic is much constrained and adaptive to the channel condition, while UDP traffic keeps pumping almost the same amount of data regardless of the channel condition, resulting in much poorer performance of UDP traffic than TCP traffic. Our second observation regarding UDP flows is that the UDP RTT degrades significantly in the mobile scenarios. In Operator A's network, RTT of over 15% UDP packets exceeds 40 seconds. In Operator B's network, the average mobile RTT is nearly 1 seconds longer than their static counterparts.

In a wireless network, TCP and UDP traffic coexist. But due to the unconstrained nature of UDP, UDP flows normally eat up the majority of the bandwidth, if the total volume is high. So TCP flows are in a great disadvantage when competing for bandwidth. However, in mobile scenarios, the efficiency of UDP flows dramatically deteriorates, leaving more space for TCP flows to get in. This enhances the fairness in bandwidth sharing between TCP and UDP flows in HSPA networks. We also notice that TCP traffic is normally related to webpage browsing and corresponds to the Interactive QoS class, where the traffic amount is not large and accurate transmission is vital. UDP traffic is more tailored for real time services, such as streaming and conversational, where certain percentage of packet loss can be well tolerated. So given limited bandwidth in HSPA networks, it is reasonable to guarantee the prioritized handling of TCP flows.

Finally, it is also worth mentioning that the RTT graphs disclose the scheduling strategies employed by the operators regarding TCP and UDP traffic. With mobility, the RTTs vary more dramatically, and exhausting all possible scheduling strategies applied to different round trip lag time. We made our conjectures on the Node-B and provider scheduling strategy based on the following two points: 1, we deliberately select off-peak time to carry out our measurement, so it is reasonable to assume that there is no other traffic for our connected base station in most of the time; 2, we believe that the RTT statistics, which are based on a large amount of samples, can fairly represent the buffer time at the base station.

In general, there are four types of QoS classes in HSPA networks, i.e., conversational, streaming, interactive and background. Different scheduling strategy is applied to these classes. In each scheduling strategy, a utility function  $U_n(r_n)$  is defined, where  $n$  denotes a particular HSDPA user and  $r_n$  is the average throughput for the  $n$ th user. The objective of the scheduling algorithm is to find a selection  $n^*$ , such that

$$n^* = \operatorname{argmax}_n \{M_n\}, \text{ where } M_n = d_n \frac{\partial U_n(r_n)}{\partial r_n} \quad (1)$$

Here,  $M_n$  denotes metric,  $d_n$  is the instantaneous data rate that HSDPA user number  $n$  can support in the next TTI. The RTT CDF of TCP traffic in Fig. 6 complies with the features of proportional fair queue scheduling [18]. We can write the metric and the utility functions as  $U_n = \log(r_n)$ ,  $M_n = d_n/r_n$ . The curve grows gradually without abrupt changes. However, for UDP traffic, the service providers apply different scheduling algorithms to meet their own needs. Operator B is more tolerant towards UDP traffic, for it schedules these traffic with relatively high efficiency. But

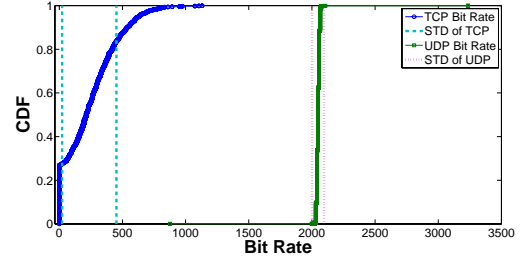


Figure 7: CDF of TCP & UDP sending bit rate in one test. TCP is more constrained and adaptive, while UDP bit rate is almost constant. STD stands for Standard Deviation.

Operator A differentiates between different types of UDP traffic. 80% of the traffic can pass within 20 seconds. However, some of the low-priority packets do not get scheduled until after a long time. But around the 50th second, a modified largest weighted delay first (M-LWDF) scheduling is likely to be used. The M-LWDF aims to achieve

$$Pr(D_i > T_i) \leq \delta_i \quad (2)$$

where  $D_i$  indicates the Head of Line (HOL) packet delay of user  $i$ ,  $T_i$  represents the delay bound, which is set to around 50 seconds, and  $\delta_i$  is the allowed percentage of discarded packets, which is set to a very low value near zero.

## 6. MOBILITY IMPACT IN TRANSITIONAL REGIONS

Frequent handoffs and rapid change of radio environments, which often happen in transitional regions and always result in poor data throughput to UEs, usually exist in the event of mobility. Handoff is a natural procedure when the UE is moving away from the area covered by one cell and entering the area covered by another cell, the call is transferred to the second cell in order to avoid call termination when the phone gets outside the range of the first cell.

Although much effort has been paid for guaranteeing service continuity for handoff since the time of introduction of 2G cellular network, the same feature seems not truly available for mobile HSPA data traffic. Fig. 8a shows a instantaneous sample of HSDPA TCP throughput for a mobile UE. Fig. 8b is about the  $E_c/I_o$  value before and after the changes of cell ID. Despite the fact that handoff can be easily identified when the UE reports the change of cell ID, it can not be readily seen in Fig. 8b. We understand that  $E_c/I_o$  values have to be always smaller or equal to 0, in order to visualize it more clearly, we show the handoff as oscillations between positive and negative values. That means, changing from +ve region to -ve region (or vice versa) is an indication that UE left the original cell with a certain  $E_c/I_o$  value and now entered another cell with new  $E_c/I_o$  value (i.e. handoff). Hence, from Fig. 8, we see that TCP throughput drops sharply when handoff happens. By analyzing 1195 logged handoff events from combination of subway, train, ferry and bus, we also show in Fig. 9 the CDF of average TCP throughput with and without handoffs. Fig. 9 intuitively depicts that handoffs generally induces 50% reduction in throughput since while UEs are experiencing handoff the

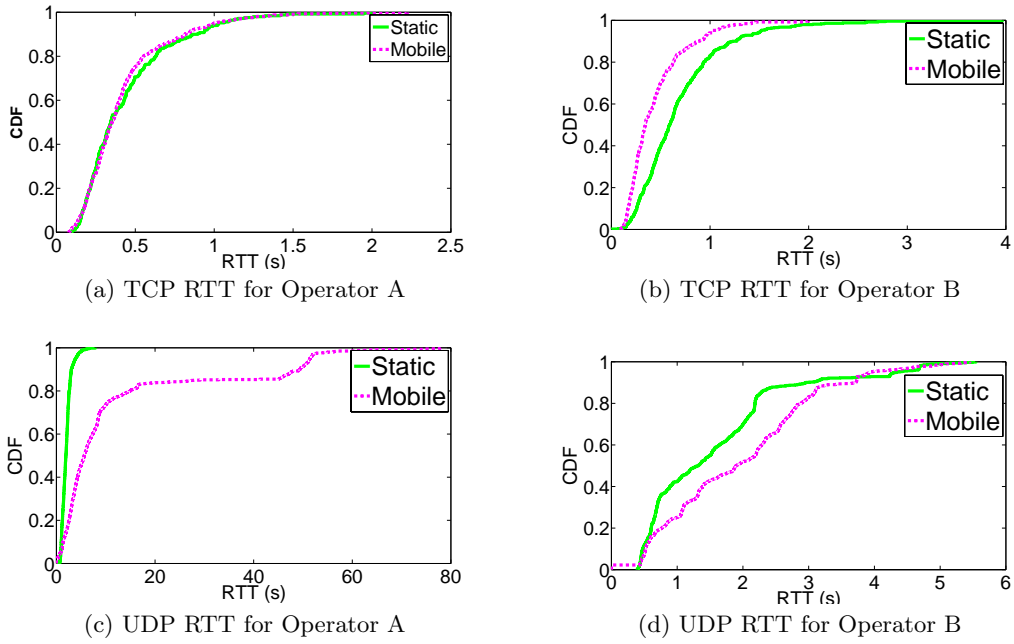


Figure 6: Static Roundtrip time for TCP and UDP connections in two operators' networks

throughput often drops and rebounds quickly. However, the peak throughput is only about 30% of that without handoff. At the moment handoff is being conducted, the link is temporarily lost. Consequently some packets are lost or the returning acknowledgment packets are delayed. Either case will cause TCP mechanism to prolong its retransmission timer and decrease its transmission windows size resulting in reduced throughput for longer interval. This situation persists until the logical link is reestablished and signal quality rebound to acceptable level. As for UDP, we also observed longer RTT and larger loss, both are at least doubled, during handoff period because UDP do not employ reliable flow control mechanism like TCP.

We also observe that UEs undergoing ping-pong effect have even worse performance than in handoff. The end users may not be aware of the temporary service unavailability for delay-tolerant services like web browsing but it is apparently noticeable and destructive to real-time services such as VoIP, video conferencing.

From our experiments, we find that the handoffs happening on fast-moving transport vehicles can also result from sudden strong fading or interference. As we can see from Fig. 10 that almost every handoff comes with a large slump of  $E_c/I_o$ . It is worth noting that a large proportion, namely around 30% in Fig. 10, of the new base stations after handoffs also have low  $E_c/I_o$ . It implies that the general wireless condition at that time is very bad. In fact, none of the base station can satisfy the QoS requirement of the UE. However, the handoff still takes place, which does not help a lot under these situations. So it is recommended that the handoff functionality be turned off when the UE is virtually screened or shadowed, until there is a solid chance that it may be connected to a base station with signal quality good enough.

After handoffs, UE normally connects to a base station with substantially better signal quality, i.e., with higher

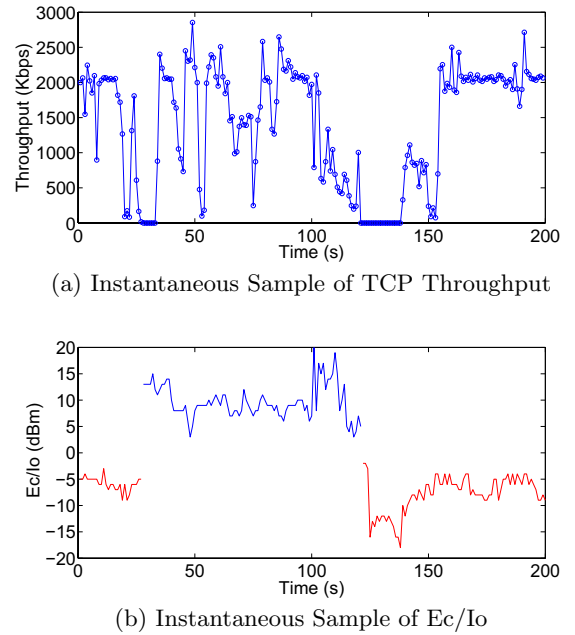


Figure 8: Impact of handoff on mobile HSDPA throughput.

$E_c/I_o$  value. However, it is not always the case. Fig. 11 shows the  $E_c/I_o$  immediately before and after each handoff. From Fig. 11, we can see the  $E_c/I_o$  of the new base stations are statistically better than the original base stations by 10dBm, which is a remarkable improvement for better QoS. But as is shown in Fig. 12, almost 30% of all the handoffs do not end up with a better base stations. This contradicts the basic principle of handoff. We find two reasons to justify this phenomena. First, the wireless signals varies quickly all the time, so that the UE made a wrong handoff decision



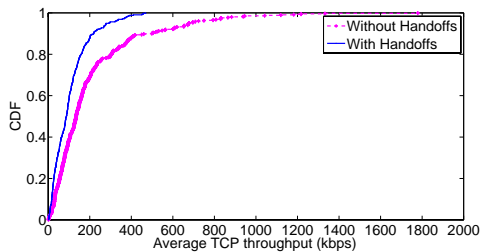


Figure 9: CDF of TCP throughput reduction for handoffs

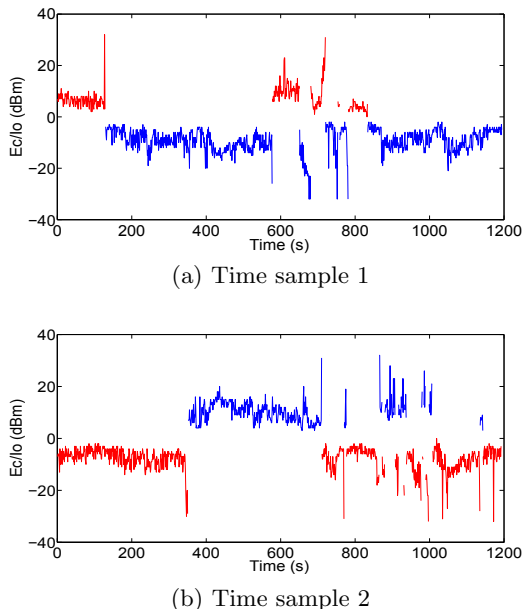


Figure 10: Typical  $E_c/I_o$  during handoffs. All  $E_c/I_o$  values are originally negative. For clarity, we flip the value to opposite, whenever a handoff happens.

based on outdated  $E_c/I_o$  information. Second, the handoff decision during high mobility does not depend only on the reading of  $E_c/I_o$ . Because a fast moving UE may frequently enter or exit different cells with varying traffic volume. So the handoff decision must also take into account the consideration of packet scheduling and admission control. A good base station candidate may reject a UE due to a bursty high traffic, so the UE must connect with another poorer base station, as it has left the original cell.

Another interesting finding worth mentioning is that in underground subways we have far less handoffs than on surface ground trains due to the benefit of leaky cable coverage. Our data shows that, for the same distance journey, number of handoffs recorded in subways is about half of that of on trains. As a result, end users can have more satisfactory realtime service experience on subways.

## 7. RELATED WORK

Many theoretical studies which are more useful for preliminary capacity approximation and network planning purposes are proposed in prior works. These works focus on HSDPA scheduling or performance evaluation by simula-

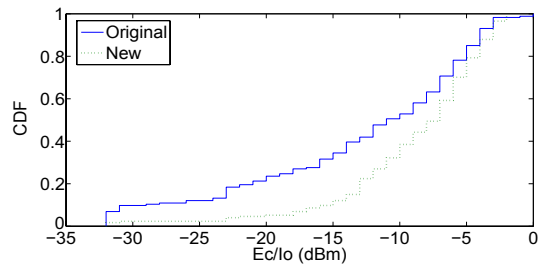


Figure 11: CDF of  $E_c/I_o$  of the original and new basestations during handoffs

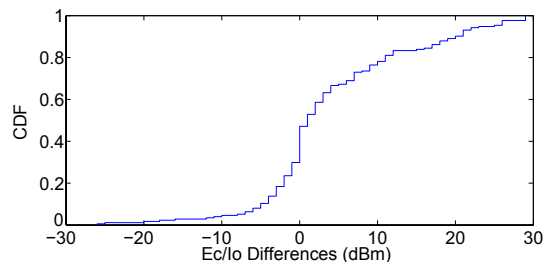


Figure 12: CDF of  $E_c/I_o$  difference between the original and new basestations during handoffs

tions [4] [16] [17] [23] [24]. However in real 3G and HSPA networks, the theoretical model is hard to formulate, especially in mobile environments.

There have also been quite a number of field measurement studies on operational 3G networks, Cano-Garcia [8] and Pentikousis [19] mainly focus on the packet delay behaviour and goodput of pure data traffic under lightly-loaded scenarios respectively. And Tan [21] measures the data throughput and latency of live 3G (WCDMA) networks under saturated conditions, using a mixture of data, video and voice traffic, but which is just in stationary environments under WCDMA. Liu [14] presents an experimental characterization of the physical and MAC layers in CDMA 1xEV-DO and their impact on transport layer performance.

Derksen [9] presents the results of HSDPA measurements made in live, commercial networks supplied by Ericsson. But the paper just gives an average downlink throughput in mobile vehicle and not analyzes the factors impact on the performance. Litjens [13] presents a flow level performance evaluation of data transfer in a UMTS/HSDPA network with a principal focus on the performance impact of terminal mobility. But the experiments are carried out in an experimental setup and small scale in a cell. Arjona [5] gives an empirical study towards high quality VoIP in 3G and HSPA Networks. Arjona [6] evaluates the VoIP performance of the HSDPA network and the proprietary FLASH-OFDM network in Finland. Kara [12], Ruser [20] and Banitsas [7] evaluate the performance of video streaming in HSDPA networks in simple scenarios. Yao [25] gives an empirical study of bandwidth predictability in the mobile environments. The authors repeat trips along a 23km route in Sydney under typical driving conditions and measure bandwidth from two independent cellular providers implementing the popular HSDPA technology in two different peak access

rates (1.8 and 3.6 Mbps). But they only investigate the bandwidth and examine download traffics.

## 8. CONCLUSION

We have designed and performed a large-scale empirical study on the mobility performance of two largest commercial HSPA networks in Hong Kong. We report our data and discuss the unexpected findings in HSPA network and their corresponding causes with details. Interestingly, mobility turns out to be a double-edged sword. Although it generally degrades HSPA services, mobility surprisingly improves some aspects of networking performance, such as fairness in bandwidth allocation among users and traffic flows. We also note that the communication characteristics in HSPA transitional regions are very complicated, so more intelligent handoff algorithms are needed for seamless service provisioning. As our future work, we plan to research into the performance of real networking applications on 3G Smartphones. As Smartphones are gaining more and more popularity and fit more deeply into 3G wireless networks, a lot of networking problems emerge and call for further studies.

## 9. ACKNOWLEDGMENTS

We thank our shepherd Dr. Erran Li and anonymous referees whose comments helped us improve the presentation of this paper. This work was supported in part by Hong Kong SAR RGC General Research Fund (GRF) No. (CityU 114908), (CityU 114609), CityU Applied R & D Centre (ARD(Ctr)) No. 9681001 and ShenZhen Basic Research Fund No. JC200903170456A; the US National Science Foundation (NSF) under Grant No. CNS0916584, CAREER Award CCF-0546668 and the Army Research Office (ARO) under grant No. AMSRD-ACC-R 50521-CI. Any opinions, findings, conclusions, and recommendations in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

## 10. REFERENCES

- [1] <http://www.gov.hk/en/about/abouthk/factsheets/docs/telecommunications.pdf>.
- [2] <http://www.gsmworld.com/technology/hspa.htm>.
- [3] HSPA technical update. <http://www.wraycastleusa.com/news.php?id=10>, 9, Sep 2008.
- [4] K. Aho, J. Aijanen, and T. Ristaniemi. Impact of mobility to the VoIP over HSUPA system level performance. pages 2091–2095, May 2008.
- [5] A. Arjona, J. Kerttula, and A. Yla-Jaaski. Live network performance challenge: FLASH-OFDM vs. HSDPA. pages 918–925, March 2008.
- [6] A. Arjona, C. Westphal, A. Yla-Jaaski, and M. Kristensson. Towards high quality VoIP in 3G networks - an empirical study. pages 143–150, June 2008.
- [7] K. Banitsas, S. Tachakra, E. Stefanidis, and K. Boletis. Using HSPA to improve the telemedical links on a moving ambulance. pages 739–742, Aug. 2008.
- [8] J. M. Cano-Garcia, E. Gonzalez-Parada, and E. Casilari. Experimental analysis and characterization of packet delay in UMTS networks. volume 4003 of *Lecture Notes in Computer Science*, pages 396–407. Springer, 2006.
- [9] J. Derksen, R. Jansen, M. Maijala, and E. Westerberg. HSDPA performance and evolution. *Ericsson Review*, (3):117–120, 2006.
- [10] F. Fitzek, A. Kopsel, A. Wolisz, M. Krishnam, and M. Reisslein. Providing application-level QoS in 3G/4G wireless systems: A comprehensive framework based on multirate CDMA. *IEEE Wireless Communications*, 9(2):42–47, April 2002.
- [11] H. Holma and A. Toskala. *HSDPA/HSUPA for UMTS*. John Wiley & Sons Ltd, 2006.
- [12] N. Kara and V. Planat. Performance analysis of IP multimedia services over HSDPA mobile networks. pages 1–5, Dec. 2007.
- [13] R. Litjens. HSDPA flow level performance and the impact of terminal mobility. volume 3, pages 1657–1663 Vol. 3, March 2005.
- [14] X. Liu, A. Sridharan, S. Machiraju, M. Seshadri, and H. Zang. Experiences in a 3G network: interplay between the wireless channel and applications. In *MOBICOM*, pages 211–222, 2008.
- [15] Livecast. <http://www.livecast.com/corporate/index.html>.
- [16] M. López-Benítez, F. Bernardo, N. Vucevic, and A. Umberto. Real-time HSPA emulator for end-to-edge QoS evaluation in all-ip beyond 3G heterogeneous wireless networks. In *SimuTools*, page 8, 2008.
- [17] P. Lunden, J. Aijanen, K. Aho, and T. Ristaniemi. Performance of VoIP over HSDPA in mobility scenarios. pages 2046–2050, May 2008.
- [18] M. Malkowski, A. Kemper, and X. Wang. Performance of scheduling algorithms for HSDPA. pages 1052 – 1056, Aug 2007.
- [19] K. Pentikousis, M. Palola, M. Jurvansuu, and P. Perala. Active goodput measurements from a public 3G/UMTS network. *Communications Letters, IEEE*, 9(9):802–804, Sep 2005.
- [20] H. Ruser, P. Halvorsen, C. Griwodz, and B. Hestnes. Performance measurements and evaluation of video streaming in HSDPA networks with 16QAM modulation. pages 489–492, 23 2008-April 26 2008.
- [21] W. L. Tan, F. Lam, and W. C. Lau. An empirical study on the capacity and performance of 3G networks. *IEEE Transactions on Mobile Computing*, 7(6):737–750, June 2008.
- [22] Veodia. <http://veodia.com/index.php>.
- [23] V. Vukadinovic and G. Karlsson. Video streaming in 3.5G: On throughput-delay performance of proportional fair scheduling. pages 393–400, Sept. 2006.
- [24] X. Yan, J. Khan, and B. Jones. Impact of a radio access network capacity on the HSDPA link performance. pages 220–224, 30 2007-Oct. 3 2007.
- [25] J. Yao, S. S. Kanhere, and M. Hassan. An empirical study of bandwidth predictability in mobile computing. In *WINTECH*, pages 11–18, 2008.
- [26] YouTube. <http://www.youtube.com>.