Abstract — Mobile networks represent a promising paradigm to provide a scalable infrastructure for Internet access in metropolitan areas. Smart devices bring us the ubiquitous mobile accessing to Internet, making mobile Internet grow rapidly. Using the mobile traffic data collected at core metropolitan 2G and 3G networks of China over a week, this paper studies the mobile user behavior from three aspects: 1) data usage; 2) mobility pattern; and 3) application usage. We classify mobile users into different groups to study the resource consumption in mobile Internet. We observe that traffic heavy users and high mobility users tend to consume massive data and radio resources simultaneously. Both the data usage and the mobility pattern are closely related to the application access behavior of the users. Users can be clustered through their application usage behavior, and application categories can be identified by the ways to attract the users.

In this paper, a large wireless smart phone deployed in the number of cities is described and experimentation results obtained from the public use that are reported and analyzed. Our analysis provides an comprehensive understanding of user behavior in mobile Internet.

Keywords — Data usage, Mobility pattern, User behavior, Network traffic, Mobile internet.

I. INTRODUCTION

Mobile Web access today still suffers from interoperability and usability problems. Interoperability issues stem from the platform fragmentation of mobile devices, mobile operating systems, and browsers. Usability problems are centered on the small physical size of the mobile phone form factors (limits on display resolution and user input/operating). Despite these shortcomings, many mobile developers choose to create apps using mobile Web. A June 2011 research on mobile development found mobile Web the third most used platform, trailing Android and iOS.

'Mobile Internet' refers to access to the Internet via a cellular telephone service provider. It is wireless access that can handoff to another radio tower while it is moving across the service area. It can refer an immobile device that stays connected to one tower, but this is not the meaning of "mobile" here. Wi-Fi and other better methods are commonly available for users not on the move. Cellular base stations are more expensive to provide than a wireless base station that connects directly to an internet service provider, rather than through the telephone system.

A mobile phone, such as a smartphone, that connects to data or voice services without going through the cellular base station is not on mobile Internet. A laptop with a broadband modem and a cellular service provider subscription, that is traveling on a bus through the city is on mobile Internet.

II. STANDARDS

Standards improve the interoperability, usability, and accessibility of mobile web usage. The Mobile Web Initiative (MWI) was set up by the W3C to develop the best practices and technologies relevant to the mobile Web. The goal of the initiative is to make browsing the Web from mobile devices more reliable and accessible. The main aim is to evolve standards of data formats from Internet providers that are tailored to the specifications of particular mobile devices. The W3C has published guidelines for mobile content, and is addressing the problem of device diversity by establishing a technology to support a repository of device descriptions.
W3C is also developing a validating scheme to assess the readiness of content for the mobile web, through its mobile OK Scheme, which will help content developers to quickly determine if their content is web-ready. The W3C guidelines and mobile OK approach have not been immune from criticism. This puts the emphasis on Adaptation, which is now seen as the key process in achieving the ubiquitous web, when combined with a device description repository.

mTLD, the registry for .mobi, has released a free testing tool called the MobiReady Report (see mobiForge) to analyze the mobile readiness of website. It does a free page analysis and gives a Mobi Ready score. This report tests the mobile-readiness of the site using industry best practices and standards.

III. RELATED WORK

In this section, we can provide an overview of the prior research relevant to traffic analysis in cellular data networks. Especially, we focus on user behavior in Mobile Internet.

Traffic analysis is a critical step to model network traffic. Traffic characteristics change with the increasing network usage demands from individual users as well as business communities. With the development of Internet, the flow based traffic analysis has always been an “academic hotspot”. However, a large amount of work on traffic characteristics accomplished a decade ago might not be suitable for current networks. Early in the 19th century, the necessity of cellular network traffic analysis already grows dramatically.

The authors in studied the traffic composition, the transfer sizes, the performance of TCP transfers and the interaction with radio power management of smart phone traffic, providing valuable information for Internet service provider. Recently, more and more researchers pay attention to understanding the user behavior of smart phone users, and the following topics have been focused on, including mobile audience measurements, mobile-phone-based content, automatically uncover and quantify characteristic behavior patterns in users’ daily lives.

A. TRAFFIC DATA USAGE PATTERN

Investigated the usage patterns of mobile data users by analyzing the characteristics of traffic heavy users and normal users. Research results suggested that a small number of traffic heavy users contributed the majority of traffic in cellular network.

B. APPLICATIONS USAGE

In this, the authors tried to find out the reason users choose and adopted an application in their daily lives. One author investigated the usage patterns of Smart phone apps in terms of physical location, time, user and device. Other author provided an empirical characterization of web use on smart phones, and studied how native applications and a browser were used on smart phones. The performance of smart phone applications was investigated in. It quantified how application performance, in particular web browsing, was impacted by various factors.
C. SEARCH BEHAVIOUR
Church Karen one of the scientist studied the Mobile Internet habits (mobile search especially) of more than 600,000 European Mobile Internet users. In this, the authors investigated mobile Web access patterns. It focused on how, why, where and in what situations people used the Mobile Internet and mobile search.

D. MOBILITY PATTERN
The authors investigated the mobility patterns in mobile cellular networks. It found that both inter arrival time and dwell time distributions could be well approximated by power-law distribution, no matter in daytime, night, rural and urban areas. The authors in examined data service usage and mobility patterns from various perspectives including application breakdown, user roles, device types and diurnal characteristics. Paul Utpal one of the scientist analyzed the network resource usage and subscriber behavior in a large scale 3G data network. Traffic load, mobility and resource efficiency were used as traffic characteristics.

E. BROWSING BEHAVIOR PATTERN
Shafiq Muhammad Zubair one of the scientist characterized the geospatial dynamics of application usage in a 3G cellular data network. The authors in proposed and developed a scalable co-clustering methodology, Phantom, to group both users and browsing profiles simultaneously in 3G networks. They found that there existed distinct “behavior patterns” among mobile users, and the behavior of most users could be classified as either homogeneous or heterogeneous. Although these works provided some deep insights on certain aspects of mobile user behavior, further study on the user behavior defined by data usage, mobility pattern and application usage is expected.

IV. DATA COLLECTIONS
The collected traffic data come from a large Chinese 2G and 3G service provider. The high level view of a mobile network is shown in given figure. There are three major components in one mobile network, including mobile devices, radio access network and core network.

1. Mobile device is the terminal connecting to the mobile network.
2. 2G subscribers access the network via BTS (Base Transceiver Station) to BSC (Base Station Controller).
   In the case of 3G subscribers, request data is collected by node Bs, and send to RNC (Radio Network Controller).
3. Core Network is composed of SGSN (Serving GPRS Support Node) and GGSN (Gateway GPRS Support Node). Gn interface is between SGSN and GGSN. GGSN send the data to Internet through the Gi interface.

The data sets used in this study are collected by our Traffic Monitoring System (TMS) (this device has been placed in the production networks by several ISPs for traffic monitoring purposes), which is connected to the Gn interface. Mirrored packets with HTTP head have been collected from a large Chinese ISP that owns a large metropolitan area network in Southern China. This ISP has 4.5 million mobile subscribers in that area. We group the packets into different flows by their 5-tuples {IP source address, IP destination address, source port number, destination port number, transport protocol}, i.e., a 5-tuple flow is a sequence of packets that share the same 5-tuple during a certain period (e.g. 64s). For the security reason, user privacy information in packets is replaced by a hashed number, which could be used for identifying subscribers, without affecting the usefulness of our analysis.
With the popularity of smart phones and the development of mobile applications (social network, e-commerce, video streaming and etc.), mobile traffic is undergoing a significant change over the last few years and the compositions of the mobile traffic have changed. However, web applications are still the dominant service in Mobile Internet, and mobile video tend to generate more and more traffic data in the next few years. The analysis of HTTP-head traffic give us the opportunity to deeply understand all kinds of applications, include web browsing, web video, web music and so on.

V. DEVELOPMENT

The first access to the mobile Web was commercially offered in Finland at the end of 1996 on the Nokia 9000 Communicator phone via the Sonera and Radiolinja networks. This was access to the real internet. The first commercial launch of a mobile-specific browser-based web service was in 1999 in Japan when i-mode was launched by NTT DoCoMo.

The mobile Web primarily utilises lightweight pages like this one written in Extensible Hypertext Markup Language (XHTML) or Wireless Markup Language (WML) to deliver content to mobile devices. Many new mobile browsers are moving beyond these limits by supporting a wider range of Web formats, including variants of HTML commonly found on the desktop Web.

VI. CONCLUSIONS

In this paper, by using the real traffic data collected from mobile Internet in a large metropolitan area of China, we studied the mobile user behavior with detailed multi-dimension analysis by focusing on three features - data usage, mobility pattern and application usage.

REFERENCES