User Behavior Pattern Analysis and Prediction Based on Mobile Phone Sensors

Jiqiang Song[1], Eugene Y. Tang[1], and Leibo Liu[2]

[1] Intel Labs China, Beijing, China
{jiqiang.song,eugene.y.tang}@intel.com
[2] Tsinghua University, Beijing, China
liulb@tsinghua.edu.cn

Abstract. More and more mobile phones are equipped with multiple sensors today. This creates a new opportunity to analyze users' daily behaviors and evolve mobile phones into truly intelligent personal devices, which provide accurate context-adaptive and individualized services. This paper proposed a MAST (Movement, Action, and Situation over Time) model to explore along this direction and identified key technologies required. The sensing results gathered from some mobile phone sensors were presented to demonstrate the feasibility. To enable always sensing while reducing power consumption for mobile phones, an independent sensor subsystem and a phone-cloud collaboration model were proposed. This paper also listed typical usage models powered by mobile phone sensor based user behavior prediction.

Keywords: Mobile computing, sensor, user behavior analysis, pattern prediction, MAST, context-adaptive, individualized.

1 Introduction

Sensors are essential components of mobile phones from the beginning. The major functionality of early mobile phones is voice call, for which the microphone sensor is the input device. Later on, the camera sensor is integrated into mobile phones as a must-have feature. Today's mobile phones are far different from a decade ago. More and more sensors are integrated into mobile phones to provide revolutionary user experience. For example, Apple's iPhone [1] uses multi-touch screen, gravity sensor, light sensor to enable a unique user interface, whereas Google's Nexus One [2] is equipped with dual microphones, one is on the front, the other is on the back, to do noise cancellation. With the fast development of highly-integrated MEMS sensors [3], mobile phones with multiple useful sensors will become affordable for mass consumers. This creates a new opportunity to analyze users' daily behaviors and evolve mobile phones into truly intelligent personal devices, which provide accurate context-adaptive and individualized services. This paper presents a preliminary work on this focus. Section 3 proposes a MAST (Movement, Action, and Situation over Time) model to

characterize human behaviors, and provides a probabilistic method to analyze and predict user behavior patterns. Key technologies required for this research are discussed in Section 4. Finally, we list three typical usage models based on user behavior pattern analysis and prediction using mobile phone sensors.

2   Related Work

User behavior analysis is not a new topic. Plenty of research work [4-6] has been done by analyzing applications and contents that users interact with. However, these analysis tools are either embedded in applications or in the service provider's servers, which are not taking into account the new and powerful sensing capability of next generation mobile phones. Some work on wearable sensors [7] utilized multiple sensors, but assuming sensors are worn on several positions of human body, which is very different from the usage model of mobile phones. On the other hand, existing research on sensors mainly focuses on the environmental infrastructure buildup and inter-device communications, such as ubiquitous computing [8], pervasive computing [9], and wireless sensor network [10]. Different from the above research, the work presented in this paper is mobile phone centric, and will be complementary with the advancement of above research and form synergy for mobile computing.

3   MAST Model

3.1   Concept

This section proposes a simple model for mobile phone user behavior analysis. The basis of user behavior analysis is to understand human activities and link them to other factors. In general, human daily activities can be classified into movements and actions according to their effects. First, we give their definitions in our context.

- A *movement* is a salient physical activity which causes location change or pose change of human body, e.g., sit down, walk, run, etc.
- An *action* is a subtle physical activity which generates impact to the rest of the world other than location change or pose change of human body, e.g., talk, press on a key, tap on a screen, etc.

There are two important factors related to human activities, time and situation. *Time* is the natural time perceivable by human, measured by seconds. *Situation* is the location where human activities happen and the environment that affects human activities. Note that the meaning of location in the situation defined here is not the same as the precise geographic location which is measured by longitude and latitude. It is what human regard as a "place", such as office, home, bus, etc.

Movements and actions have different relations with time and situation. Movements usually take measurable time, and it may bring human from one situation to another, while actions take variable time (very short for pressing a key or pretty long for watching a video), and it does not lead to situation changes. Figure 1 shows an example of recording a person's behavior over a short period of time, where S_i (i=1..3) denote different situations, M_i (i=1..4) denote different movements, and A_i (i=1..5) denote different actions.
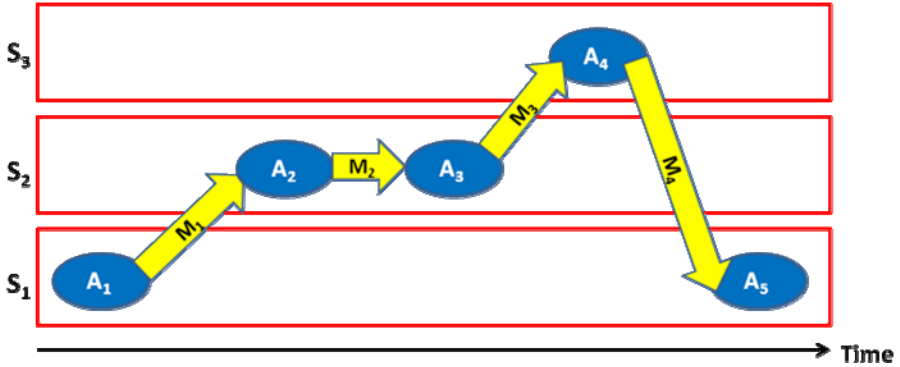

Fig. 1. Temporal record of a person's behavior over time

From Figure 1, one can clearly determine where the person is and what the person is doing at any particular time. We call this the MAST (Movement, Action, Situation over Time) model. $M(t)$, $A(t)$ and $S(t)$ denote the value of movement, action and situation at time t, respectively.

3.2 Probabilistic Model

If we use one day as the analyzing period for human behavior, $M(t)$, $A(t)$ and $S(t)$ may be different at the same t of different dates. Therefore, $M(t)$, $A(t)$ and $S(t)$ are not scalar values, but vectors containing probabilities of possible values. For example, suppose the possible movements of a person at a particular time are M_i (i=1..n), $M(t) = [P_1, P_2, .., P_n]$, where P_i is the probability of M_i happening at that time, and $\sum P_i = 1.0$. Same for $A(t)$ and $S(t)$. $M(t)$, $A(t)$ and $S(t)$ are denoted as column vectors.

$M(t)$, $A(t)$ and $S(t)$ can be independently observed over time and over days. Since a person's daily life does show certain patterns due to his/her occupation and habits, $M(t)$, $A(t)$ and $S(t)$ will converge after a certain period. They may converge at more than one peaks, e.g., corresponding to weekday pattern, weekend pattern, etc. This will be useful to predict a person's movement, action or situation by time. However, the prediction is sensitive to the possible time shift of human activity. On the contrary, the correlations among movements, actions and situations are less time sensitive, so that they are more robust to characterize human behavior patterns. Figure 2 shows the probabilistic relations among $M(t)$, $A(t)$ and $S(t)$. In short, when a user is observed in

state M(t), A(t) or S(t) at time t, how possible he/she will be in state M(t+ Δt), A(t+ Δt) or S(t+ Δt) after time Δt, where Δt is the time interval between observation and prediction.
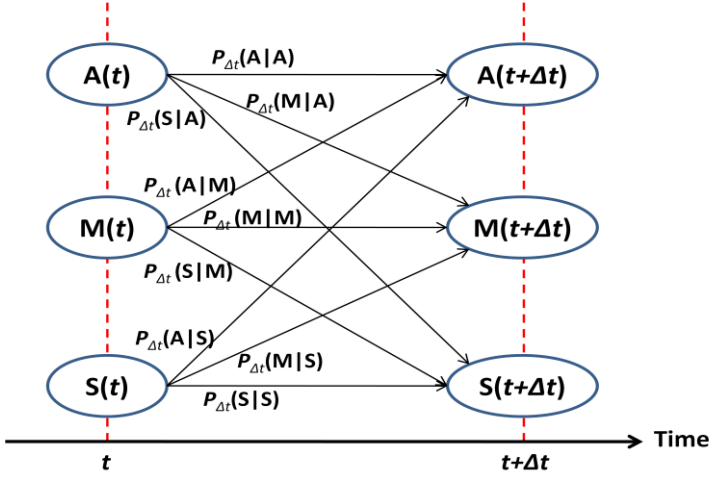

Fig. 2. Probabilistic relation model among M(t), A(t) and S(t)

Below are the physical meanings of probabilities defined in Figure 2, where i, j are indexes of possible actions, movements or situations.

- $P_{\Delta t}(A|A)$ indicates when A_i happens at t, the probability of A_j happening at $(t + \Delta t)$.
- $P_{\Delta t}(M|A)$ indicates when A_i happens at t, the probability of M_j happening at $(t + \Delta t)$.
- $P_{\Delta t}(S|A)$ indicates when A_i happens at t, the probability of being S_j at $(t + \Delta t)$.
- $P_{\Delta t}(M|M)$ indicates when M_i happens at t, the probability of M_j happening at $(t + \Delta t)$.
- $P_{\Delta t}(A|M)$ indicates when M_i happens at t, the probability of A_j happening at $(t + \Delta t)$.
- $P_{\Delta t}(S|M)$ indicates when M_i happens at t, the probability of being S_j at $(t + \Delta t)$.
- $P_{\Delta t}(S|S)$ indicates when being S_i at t, the probability of being S_j at $(t + \Delta t)$.
- $P_{\Delta t}(A|S)$ indicates when being S_i at t, the probability of A_j happening at $(t + \Delta t)$.
- $P_{\Delta t}(M|S)$ indicates when being S_i at t, the probability of M_j happening at $(t + \Delta t)$.

Therefore, for a specified time t, $P_{\Delta t}(Y|X)$ is a two-dimension probability matrix, which is trained by day-to-day observations of user activities. Its row dimension equals to the dimension of $X(t)$, and its column dimension equals that of $Y(t)$. For any given X_i detected at time t, $P_{\Delta t}(Y|X)$ stores the probability of each Y_j happening after Δt. The minor time shifts of user activities should be considered during training $P_{\Delta t}(Y|X)$. When $X(t)$ or $Y(t)$ is updated with new observation data, $P_{\Delta t}(Y|X)$ should be updated as well.

With the converged MAST probability model trained from user activities over many days, one can extract many interesting user behavior patterns by time-movement correlation, time-action correlation, time-situation correlation, movement-movement correlation, movement-action correlation, etc. It is also interesting to extract time-related correlations and time-unrelated correlations.

3.3 MAST-Based Behavior Prediction

The MAST probabilistic model described in the above section is designed for characterizing a user's daily behavior pattern. As long as the model has been trained to converge, it is very useful to predict the user's next activity based on his/her current activity being detected. For example, assuming time is always a known value, below is a basic flow for predicting action by movement.

    i.   Detect a human movement, denoted by M_i;
    ii.   Using the current time t as a starting point, search the neighborhood of $M(t)$ for a local probability peak for M_i;
    iii.   If the local peak is larger than a threshold, a strong prediction can be made; otherwise, a weak prediction is made;
    iv.   Get the time t' corresponding to the local peak;
    v.   Using t' as the starting point, search the neighborhood of $P_{\Delta t}(A|M)$ for a local probability peak for M_i;
    vi.   Take the A_j corresponding to the local peak as the prediction.

The prediction flows for other input/output requirements are similar. One can also further develop to use more than one input to predict more precisely.

4   Key Technology

There are many methods, such as Bayesian-based methods [11] and HMM-based methods [12], available for training the MAST probabilistic model, so that we do not repeat them in this paper. This section describes the other key technologies required for user behavior pattern analysis and prediction based on mobile phone sensors. They include user activity detection, independent sensor subsystem, and phone-cloud collaboration.

4.1 User Activity Detection

When using sensors embedded in a mobile phone to collect its user's activity data in a non-intrusive way, the detectable movements and situations are not as many as those detectable by surrounding sensors or wearable sensors. On the other hand, not all human movements and actions are concerned for user behavior analysis. Table 1 shows some concerned movements, actions and situations.

Table 1. Concerned movements, actions and situations

Movements	Actions	Situations
Sit down	Phone call	Home
Stand up	Read news	Office
Walk on flat	Web search	Car
Walk upstairs	Read mail	Bus
Walk downstairs	Write mail	Subway
Take lift up	Take photo	Street
Take lift down	Listen music	Canteen
Run	Watch video	Meeting room
Drive	Visit SNS website	Playfield

Among several sensors on a mobile phone (e.g., Nexus One), we identified that the 3-axis accelerometer and the 3-axis orientation sensor (a.k.a. digital compass) played important roles in detecting movements. Figure 3 shows the definition of 3D output of these two sensors, where $G(i)$ means gravity acceleration on axis i.
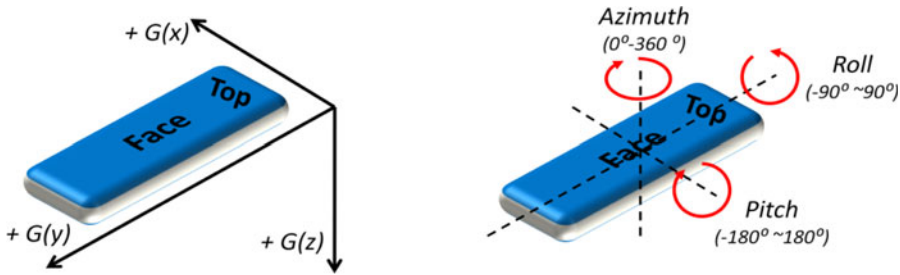

Fig. 3. 3D output definitions of accelerometer and orientation sensor on a phone

Figure 4 shows the accelerometer data graphs over time for some typical movements, which demonstrate salient features. In the figure, the horizontal axis is time, and the vertical axis is $G(i)$. Since the accelerometer data are 3-dimension vectors, 2-dimension and 3-dimension graphics pattern recognition method [13] is employed to detect movements.

Orientation sensor data are used to determine the orientation and pose of phone, which does not necessary indicate its user's orientation because there are different positions that the phone may be placed on human body, e.g., shirt pocket, coat pocket, trousers pocket or bag. When the user is not moving, it is impossible to determine the user's orientation from orientation sensor output. Fortunately, we observed that when the user is doing the same walking movement on flat ground, the accelerometer sensor
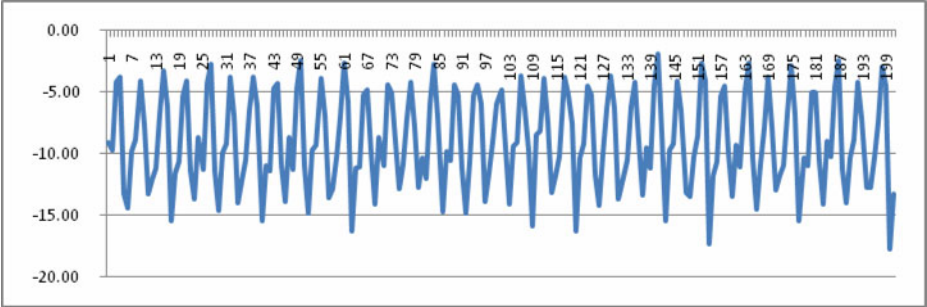
a. G(x) data for walking on flat
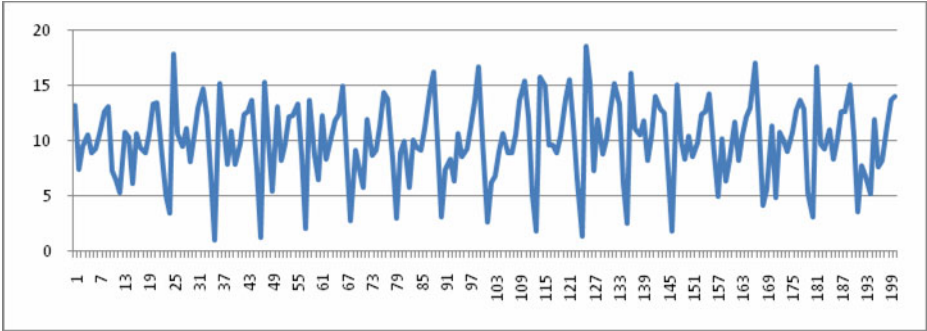
b. G(x) data for walking upstairs
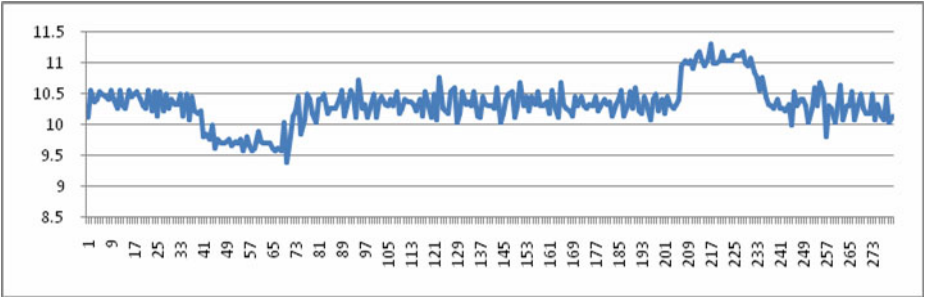
c. G(z) data for taking lift up

Fig. 4. Accelerometer data graphs for some typical movements

output has different features for different phone poses, so that one can derive the relative position of phone on its user's body. The user orientation detection algorithm may fuse the accelerometer output and orientation sensor output to determine user's current

orientation. Table 2 shows some results of user walking orientation detection using sensor output fusion. In the table, the accelerometer output, $G^*(i)$, is the averaged pace-wise sum of gravity acceleration values on each axis for walking, and the granularity of fused user orientation is 15^0 in the $0^0 \sim 360^0$ scope. From the result, we can see that when the phone moves with user at different poses, e.g., holding phone in hand to read messages so that the phone is heading front and facing up, the feature ($G^*(x)$, $G^*(y)$, $G^*(z)$) varies much, and the fused user orientation can be significantly different from the original "Azimuth" output.

Table 2. Sensor output fusion for user orientation detection

Phone pose		Accelerometer output			Orientation sensor output			Fused user orientation
Heading	Facing	$G^*(x)$	$G^*(y)$	$G^*(z)$	Azimuth	Pitch	Roll	
Front	Up	-0.8	101.1	301.4	287^0	-18^0	0^0	285^0
Up	Back	49.7	308.9	-39.3	231^0	-99^0	9^0	240^0
Down	Front	56.2	-299.9	58.6	43^0	75^0	10^0	75^0
Back	Down	-18.0	-60.1	-297.9	102^0	152^0	-4^0	255^0

All concerned actions listed in Table 1 are through a user's interaction with the applications installed in a mobile phone. These actions cannot be detected by sensors, but can be detected by software approach. We used an application manager which monitors the execution status of applications, such as start, active, inactive, and stop. Therefore, the user actions can be recorded by the virtual desktop.

For the situation detection, the microphone sensor is most useful. There are some pretty good methods [14] can be used to detect situations by audio only. The fusion with other sensors, such as accelerometer, will further clear the ambiguity between similar audio scenes, like canteen and playfield, since the body movements are quite different.

4.2 Independent Sensor Subsystem

Unlike other wearable sensor based user behavior analysis approach, we must consider the limitation imposed by mobile phone sensors. The most critical limitation is that mobile phone is a battery-powered small device. Detecting a user's movement requires always sensing in the background, which means sensors are working even the user does not use the phone or phone is in sleep mode. This will pose a severe power consumption issue, since sensor data need CPU processing in today's mobile phone SoC (System on Chip) architecture, so that CPU cannot sleep when sensors are working. Figure 5 shows a typical mobile phone SoC architecture.
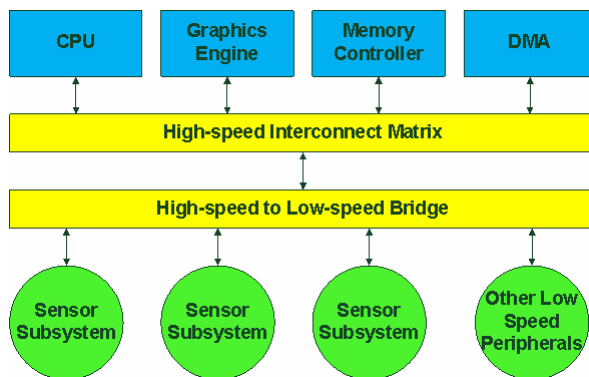
Fig. 5. Current mobile phone SoC architecture

In details, following reasons make the architecture in Figure 5 not suitable for always sensing.

i.   Sensors' always-on traffic characteristic is different from other low-speed peripherals;

ii. There are two levels of buses to separate high-speed and low-speed traffics to alleviate burden on main bus. However, sensor data need to travel upwards to be buffered and processed by CPU. It does not make sense either to keep the high-speed bus on just waiting for the low-speed sensor data;

iii. Furthermore, there is no direct communication between sensors, which prevents sensors from autonomous inter-control and collaboration.

To save the power consumption for always sensing, we propose to design an independent sensor subsystem in mobile phone SoC to resolve the above problems, as shown in Figure 6. This design isolates the core network from the sensor network. It not only allows CPU and main bus to sleep while one or more sensor subsystems are working, but also adds intelligence to sensor peripheral for inter-sensor communication, sensor data manipulation, and sensor data fusion.

Isolation of the core network from the sensor network mainly occurs in two different aspects – power and clocking. There are well-known techniques in IC (Integrated Circuit) design for creating different power islands and clock sub-networks, so they will not be repeated here. Adding intelligence to the sensor network is more complicated. On the one hand, we do not want to add a processing block to each sensor, as it creates significant overhead. On the other hand, it is not a good idea to have one processing block for all sensors, as this will cause a significant burden on the routing and processing capability. Instead, what we are proposing is a network of sensor subsystems, where each subsystem has its own processing unit. The size of each subsystem depends on the sensor processing requirements, size of chip and capability of the processing unit. We can employ power saving techniques to such a network of sensor subsystems by further dividing the sensor subsystems into sub-groups.
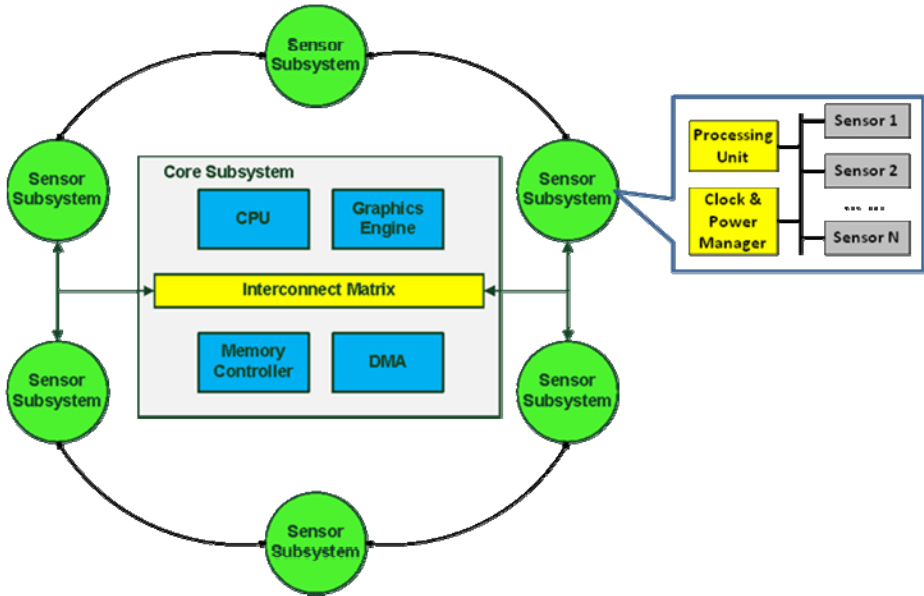
Fig. 6. Independent sensor subsystem

4.3   Phone-Cloud Collaboration

It's well known that probabilistic model training is computation intensive. Mobile phone itself is not an ideal platform to do that, unless there is specific hardware accelerator for that, which is not true nowadays. Therefore, how to leverage the network side capability is essential. With the fast development of wireless communication, always-connected mobile phones will become mainstream and the average available bandwidth for each user will exceed 1Mbps at the 4G era. In this circumstance, cloud computing and storage service [15, 16] is going to be the most easy-to-access and always-available computation and storage resource for mobile phones. We propose a phone-cloud collaboration model to fully exploit the advantages of both sides, as shown in Figure 7.

In this collaboration model, the phone acts like a sensory organ, while the cloud acts like a brain. The cloud must emphasize privacy management, so that its users feel like they are using a private computer. The MAST probabilistic model for a user is stored in the cloud side, so that switching phones does not impact behavior analysis. The cloud provides standard interface to interact with phones. The phone uses its sensors and/or software to detect its user's movements, actions and situations. The detection results are sent to the cloud for two purposes. One is to train the MAST probabilistic model for a specific user; the other is to serve as input to predict the user's next behavior. With this partition, the communication data rate is very low since only results are transmitted.
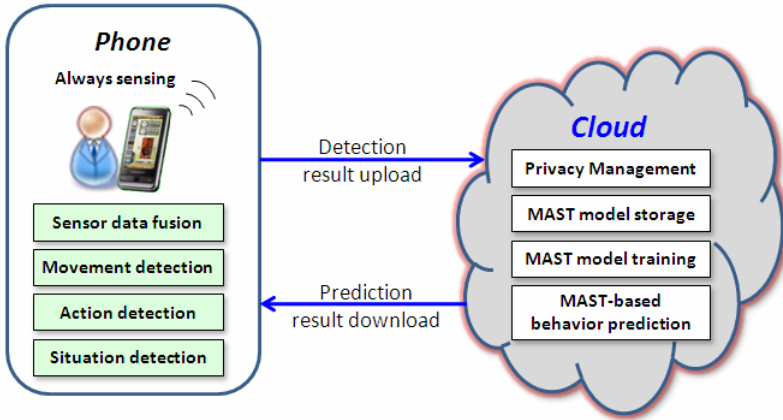
Fig. 7. Phone-cloud collaboration model

5   Usage Model

Mobile phone sensor based user behavior pattern analysis and prediction is the foundation of evolving mobile phones from a life-enriching device to an intelligent user behavior assistant. This section depicts three usage models powered by this technology.

5.1   Proactive UI Adaptation

This usage model is between user and phone. Today's mobile phone contains many applications that cannot be easily navigated on the small screen. Even though there are many UI innovations on mobile phones, but they still cannot fulfill individual's need. From the user behavior analysis, the user pattern about using what application at what time, what situation, and after what movement, what action is known. With the real-time user behavior prediction result, the phone's UI can be reorganized to adapt to the next action or situation, so that the user will always see the application he/she wants to use on the first screen.

5.2   Proactive Situation Adaptation

This usage model is between user and situation. There are two aspects: one is that the phone adapts to the situation, the other is that the site adapts to the user. For the former, when a user is in different situations (e.g., home, office, driving, playfield, etc.), he/she wants to use the phone in different modes, such as turning the ring volume louder or lower, declining all calls, etc. For the latter, the same site may need to change settings or services for different users. For example, a car shared by several users need to change seat position and radio channels when a user is coming. With the real-time user behavior prediction result, the proactive adaptation can be realized.

5.3 Proactive Service Adaptation

The above two usage models are beneficial to phone users only. This usage model is between user and service provider. If a user is willing to share his/her behavior pattern information to service providers, service providers can access the user behavior pattern analysis and prediction interface in the cloud to classify the user's profile and predict the user's activity, so that to push accurate commercial services to him/her. This is a win-win usage model. Users get timely individualized service, and service providers save costs on untargeted promotions.

6 Summary

This paper proposed a MAST probabilistic model for using mobile phone sensors to analyze and predict user behavior patterns. Some concerned new technologies, like user activity detection, independent sensor system and phone-cloud collaboration, are discussed. From those preliminary studies, we find that this research direction is promising while challenging. In the future, we will further investigate those key technologies and complete a prototype system.

References

[1] Apple's iPhone Information (April 6, 2010), http://www.apple.com/iphone/
[2] Google's Nexus One Information (April 6, 2010), http://www.google.com/phone
[3] Bryzek, J., et al.: Marvelous MEMS. IEEE Circuits and Devices Magazine 22, 8–28 (2006)
[4] Dou, Y., Yan, H., Lei, Z.: Broadband dial-up user behavior identification and analysis. In: Proceedings of the 2nd IEEE International Conference on Broadband Network & Multimedia Technology, pp. 316–322 (2009)
[5] Vilas, M., et al.: User behavior analysis of a video-on-demand service with a wide variety of subjects and lengths. In: Proceedings of the 31st EUROMICRO Conference on Software Engineering and Advanced Applications, pp. 330–337 (2005)
[6] Chung, Y.W., Chung, M.Y., Sung, D.K.: Modeling and Analysis of Mobile Terminal Power on/off-State Management Considering User Behavior. IEEE Transactions on Vehicular Technology 57, 3708–3722 (2008)
[7] Corbellini, S., Ferraris, F., Parvis, M.: A System for Monitoring Workers Safety in an Unhealthy Environment by means of Wearable Sensors. In: Proceedings of IEEE Instrumentation and Measurement Technology Conference, pp. 951–955 (2008)
[8] Cheng, J.: Testing and Debugging Persistent Computing Systems: A New Challenge in Ubiquitous. In: Proceedings of IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, pp. 408–414 (2008)
[9] Reddy, Y.V.: Pervasive Computing: Implications, Opportunities and Challenges for the Society. In: Proceedings of the 1st International Symposium on Pervasive Computing and Applications, pp. 5–5 (2006)
[10] Chong, C.-Y., Kumar, S.P.: Sensor Networks: Evolution, Opportunities, and Challenges. Proceedings of the IEEE 91, 1247–1256 (2003)
[11] Mirikitani, D.T., Nikolaev, N.: Recursive Bayesian Recurrent Neural Networks for Time-Series Modeling. IEEE Transactions on Neural Networks 21, 262–274 (2010)

[12] Wu, F., Chiu, I.-H., Lin, J.-R.: Prediction of the intention of purchase of the user surfing on the Web using hidden Markov model. In: Proceedings of ICSSSM 2005, International Conference on Services Systems and Services Management, pp. 387–390 (2005)

[13] Lu, T., et al.: A novel knowledge-based system for interpreting complex engineering drawings: theory, representation and implementation. IEEE Transactions on Pattern Analysis and Machine Intelligence 31, 1444–1457 (2009)

[14] Lu, L., Zhang, H., Jiang, H.: Content Analysis for Audio Classification and Segmentation. IEEE Transactions on Speech and Audio Processing 10, 504–515 (2002)

[15] Amazon Cloud Computing Service Information (April 6, 2010),
`http://aws.amazon.com/ec2/`

[16] Amazon Cloud Storage Service Information (April 6, 2010),
`http://aws.amazon.com/s3/`