Many innovations are on the way for users of portable devices, especially smartphones. Edward Chang, vice president for research at hTC Corporation, outlined some of the frontiers of research in a plenary session at the 75th Anniversary ASIS&T Annual Meeting in Baltimore.

Chang, who used to direct research for Google in China, began with an overview of the past few years. Web 1.0 appeared 15 years ago, he said. Google developed page ranking. Then Web 2.0 arrived. In addition to document content, it involved people with the web, permitting them to develop their own content and enhancing connections between people. New search engines, still under development, can do people searching. We’re also experiencing a rapid increase in mobile access to the web. There are now more than four billion users. iOS and Android systems are everywhere. But, he said, these changes are more evolutionary than revolutionary.

Providers want to sense the user, to make the smartphone smarter. They want to change the information model from pulling to pushing, matching the users to things that are relevant to him or her.

New developments include human-computer interaction improvements as well as many more sensors on cellphones, sensors that report on users and their environments and enable providers such as Google to do context-centered computation. For example, on a mobile device, input methods such as voice or touch are preferred to keystrokes. And because there are output space limitations, a terse summary using natural language is preferred to a list of results.

Chang’s team built Confucius, a Google Q&A system, to provide high quality, timely answers. They showcased it at the International Conference on Very Large Data Bases in 2010. When it came to market, it faced six
competitors already. With future enhancements to Confucius, your smartphone receives a spoken query and delivers an answer using a voice interface. Then the software provides labels for the query semi-automatically. It generates answers on search results using natural language processing (NLP) techniques. Depending on the question asked, it can parse the web pages to zero in on your information request. A simple ranking can provide you with top-ranked answers. And if that can’t be done, you can resort to users on the Internet. But there, you need to evaluate user credentials and route questions to experts. Once the question has been answered, the phone can speak the answers back.

Current research topics include things like speech recognition. Questions can be answered using a model-based or a data-driven approach. Certain questions are about opinions. As Chang spoke, hurricane Sandy was approaching Baltimore; he noted that persons asking about the storm could be frustrated because the information they wanted was not yet available.

The reason that voice recognition works better than it used to, Chang said, is because the model-based approach is giving way to a data-driven approach. That’s what Google Translate uses. Google can look for translated documents in the United Nations document collection, and if they exist, the software can return the translated phrase. That’s the data-driven approach. If they can collect data about different data in different accents, they don’t need the model-based approach.

Google can achieve this level of quality in translation, he said, because of its massive computation power – Google has about 20-million CPUs worldwide. Translation may not be accurate in offline or airplane mode, because it must use the model-driven approach. But a translation request may be sent to the cloud, and the translation then becomes much more accurate. He quoted a 2001 article by Michele Banko and Eric Brill (available at http://acl.ldc.upenn.edu/P/P01/P01-1005.pdf) that discussed the advantages of large scale. Test accuracy increases with the size of the training corpus.

A second challenge for voice input, Chang said, is natural language understanding (NLU). Issues such as context awareness (for example, function, location) and dialog design (for failure recovery) come into play. Speech recognition is not just converting speech into text, but understanding the semantics. For example, if you ask about Japanese restaurants in a particular area, the software concludes you’re interested in booking something; the accuracy of results increases if that’s true.

With wi-fi and GPS sensors, Chang said, location-based services can be provided. Google is developing indoor and 3D positioning and navigation, because most commerce activities occur indoors. Your cellphone may know you’re interested in cat food and near a cat food store, and then offers you a coupon to redeem. But coupon providers generally aren’t on the highway – they’re in the mall. If sensors can achieve resolution to about five meters, coupon effectiveness will be greatly increased.

Existing technologies such as GPS, wi-fi and cell towers fall a bit short for indoor positioning. Also, the time to make the first fix can be more than 30 seconds – too long for many people. Wi-fi models use wave propagation based on computing the distance from a mobile device to a known access point (AP). But obstructions, wave deflection and noise can affect accuracy. Another wi-fi approach is the signal strength map (SSM) or RF fingerprint. It generates a heat map showing how close a device is to an access point. Chang said it’s promising but still suffers from noise problems. Its implementation can be laborious, too; it requires site surveys and re-surveys when the AP is broken. There are also privacy concerns.

Google’s solution is a patented technology called XINS (pronounced “SINS”) and what Chang called “killer apps.” Google started the XINS project a year ago. It involves inertial navigation systems (INS) such as accelerometers, gyroscopes and compasses. It computes a cellphone’s direction and speed of movement. Once a cellphone’s position has been
accurately plotted, a gyroscope can detect its roll, pitch and yaw. If the angular speed is known, the attitude of the cellphone can be determined too. If the acceleration is known, the velocity, and hence the position, of the phone can be inferred.

Of course, there are technical challenges, too. The INS devices must be inexpensive, and they can be prone to errors. The devices may suffer from precision bias, in which one full rotation does not equal 360 degrees. They can also be sensitive to temperature and noise. Errors are progressively multiplied, Chang pointed out, so even small deviations can result in huge errors. Still, with INS and a good indoor map, with 20 access points, it’s possible to achieve accuracy within five meters. Integration drift can be a problem, he said, because its effect is cumulative. The vibration energy model (VEM) is predicated upon noting the position of arms and legs as people walk. If Google can model that, Chang said, it can, with proper processing, obtain an accurate direction for a person’s movement.

A second problem involves error drift. Proper INS calibration helps with that. But, Chang said, they can’t ask users to do the calibration. The service providers must do that. Chang said they’ve developed a six-point calibration. If they can calibrate six parameters in three dimensions, they can do the calibration. It’s non-intrusive to the user. Once they’ve collected more than eight points, they can use a straightforward optimization method, which can form something like a sphere or elongated sphere around the user. Once we can convert it into a true sphere, they can get the needed calibration parameters.

Chang noted that biological and environmental sensors can be added as well. The proper sensors can allow services to predict someone’s transportation mode – walking, running, using an elevator, train, bike, car. A pressure sensor can detect vertical movement with great accuracy. Once service providers do ground truth, he said, they can determine someone’s elevation as well. In addition, there are health sensors that can be connected to a cellphone. For example, a message can be sent directly to a doctor when a sensor detects precursors to a heart attack. Sensors can monitor stress and serve as a fitness coach. Sensors can also do security alerts. Once a lock has been touched, a photo can be transmitted to your cellphone. Increasingly, sensors are everywhere.

In closing, Chang said mobile opportunities are tremendous. With increasing context-aware computing, recommendations on the information a person needs at a particular time and in a particular place can be made. We can make sure, he said, that people get the transportation information they need, or the music they like. To make it all happen, we’ll need cloud computing, infrastructure, big data, social networks and sensors for cellphones. The coming changes will be both evolutionary and revolutionary.

The question and answer session dealt with a number of topics, but four persons raised issues involving privacy. Big data can have a negative social side. Totalitarian regimes can use it for surveillance. Will applications be developed that will permit users to opt out of providing selected data or choose not to see certain ads? Even the emergency management applications require surrendering information about health. Chang acknowledged that everyone is concerned about privacy. It certainly needs to be preserved, he said. But he added he doesn’t have the answer to everyone’s questions. The issue is being addressed, he said.