Challenges and Opportunities of Internet of Things

Yen-Kuang Chen

Intel Labs
Intel Corporation
Santa Clara, CA, USA

E-mail: yen-kuang.chen@intel.com

Abstract - To date, most Internet applications focus on providing information, interaction, and entertainment for humans. However, with the widespread deployment of networked, intelligent sensor technologies, an Internet of Things (IoT) is steadily evolving, much like the Internet decades ago. In the future, hundreds of billions of smart sensors and devices will interact with one another without human intervention, on a Machine-to-Machine (M2M) basis. They will generate an enormous amount of data at an unprecedented scale and resolution, providing humans with information and control of events and objects even in remote physical environments. The scale of the M2M Internet will be several orders of magnitude larger than the existing Internet, posing serious research challenges. This paper will provide an overview of challenges and opportunities presented by this new paradigm.

I. Introduction

Technologies change our life. Out of many emerging technologies, Internet of Things (IoT), also known as machine-to-machine (M2M) (where smart devices that collect data, relay information to one another, process the information collaboratively, and take action automatically) is a new paradigm offering both challenges and opportunities.

Over last twenty years innovative information technologies have wrought significant change in human civilization. For example, imagine that you were to give a speech in another city which is normally one hour away from where you live. Unfortunately, there was a traffic jam on the highway and you were expected to be late. If this happened 20 years ago, there was literally no way you can communicate the news to your audience. Nowadays, if you get stuck on the highway, you can simply use your cell phone to tell your audience. This is because cell phones and telecommunication service are affordable and available to almost everyone. Question: can technologies do better?
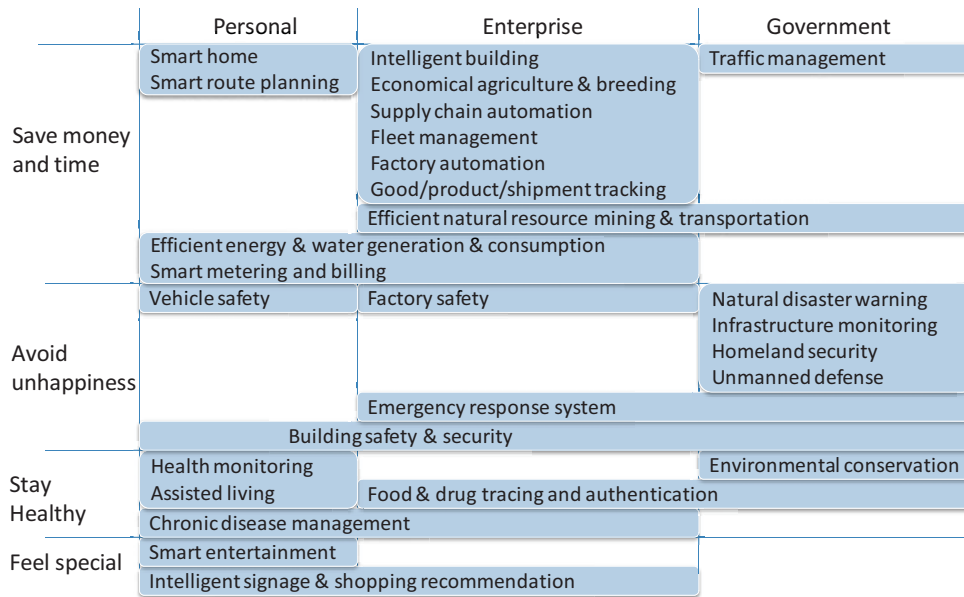
More game-changing capabilities are emerging. First, new types of sensors enhance our perceptual abilities by detecting information that humans cannot detect and collecting such information anytime and anywhere. Second, robots similarly enhance our ability to act, overcoming our physical limitations. They can be given greater-than-human strength and may go where humans cannot. After the recent earthquake and nuclear disaster in Japan, robots were used to explore the damaged nuclear plants due to radiation fears [5]. Third, wireless communication and broadband internet technologies enhance our communication capabilities, a power

that will only grow when 4G wireless [42] and greater internet bandwidth become available. Fourth, emerging cloud computing and machine intelligence will enhance our analytical capabilities. With massive computation capabilities and more mature machine learning techniques, we may begin analyzing information that could not be processed in the past.

Further, humans keep demanding technologies to improve their lives. Fundamentally, people want to be happier. How can people be happier? First, humans want more time, money, and quality of life. Technologies can help humans save money, enhance their appearance, eat better, and sleep better. Second, most people want to avoid being in threatening or troublesome situations. Technologies such as weather forecasting or fire warning systems help predict future events. Third, they want to be healthier. Technologies help humans to provide better and safer environments for healing, to take care of aging and sick people, and to avoid accident and injury. Fourth, most people desire companionship. Email, internet phones and social networks connect people. Finally, people want to feel special or be respected. Personalized services and products have become very popular.

Many of the above human desires can be further fulfilled by emerging M2M and related technologies [12][32][36]. For the same example given earlier, imagine that you were to give a speech in another city and you got stuck on the highway. With the improvement in M2M technology, your calendar and your car can communicate. If you are expected to be late, your smart phone will send a message to your audience automatically telling them approximately how much time they have to wait. Or even better, the calendar planner can look up the traffic condition in advance and suggests what time you should leave. Sensors can monitor the traffic conditions along the routes to your destination so that you are able to select the best route to get to the venue on time. There are many similar examples. The nexus of human needs and emerging computing and sensor technologies is bringing about a new digital revolution.

Section II will give an overview of the new paradigm. Section III will describe the challenges of the new paradigm. Section IV will discuss the opportunities created by these challenges.

	Personal	Enterprise	Government
Save money and time	Smart home Smart route planning	Intelligent building Economical agriculture & breeding Supply chain automation Fleet management Factory automation Good/product/shipment tracking	Traffic management
		Efficient natural resource mining & transportation	
	Efficient energy & water generation & consumption Smart metering and billing		
Avoid unhappiness	Vehicle safety	Factory safety	Natural disaster warning Infrastructure monitoring Homeland security Unmanned defense
		Emergency response system	
	Building safety & security		
Stay Healthy	Health monitoring Assisted living	Food & drug tracing and authentication	Environmental conservation
	Chronic disease management		
Feel special	Smart entertainment		
	Intelligent signage & shopping recommendation		

Figure 1: M2M application examples. Different rows reflect different human wishes. Different columns reflect who pay for the services/applications

II. New Paradigm: Connected Context Computing

Recently, sensor networks, cyber physical systems, and internet of things have become more common as sensing, communication, and analytics technologies have matured. In the future, digital sensing, communication, and processing capabilities will be ubiquitously embedded into everyday objects, turning them into the Internet of Things (IoT, or machine-to-machine, M2M). In this new paradigm, smart devices will collect data, relay the information or context to each another, and process the information collaboratively using cloud computing and similar technologies. Finally, either humans will be prompted to take action, or the machines themselves will act automatically.

The world is on the edge of revolutions in both quantity and quality. Today more things than humans are connected to the internet [9]. In the near future, the number of connected devices will be tens or hundreds of times larger than the number of connected people. It is estimated that 20 households in 2012 years will generate more internet traffic than the entire internet in 2008 [10]. Furthermore, the computation capabilities per device are also increasing. For example, today's smart phone is about 100,000 times faster than the ENIAC I computer, built in the 1940s [4].

There are many applications of this new paradigm [14][16][17], as shown in Figure 1. Some examples are highlighted below:

1. Smart home: At home, embedded sensors can understand the human activities and properly adjust the air temperatures or lighting to reduce our energy usage without sacrificing human comforts.
2. Economical agriculture: In a farm field, remote bug traps can detect the outbreak of pests and initiate spreading the right amount of pesticide. This will reduce the chance of overspreading and potential damage to the crops.
3. Vehicle safety: Sensors on a car can help drivers understand the potential risk of running into each other; in particular, sensors and inter-vehicle communication can help us see what we cannot see. With timely and proper warning, we can reduce the vehicle collision rate.
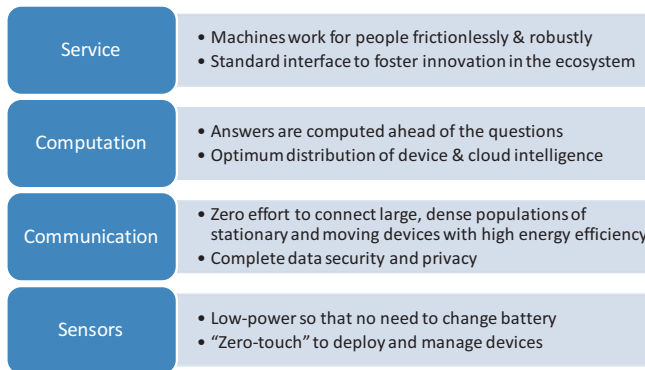4. Assisted living: The population is aging. There is an increasing need to take care of more elders. Sensors can help us monitor the health condition of elders and properly provide help (e.g., reminders of missing a dose, warning of high blood pressure, requesting medical emergency).

In short, connected embedded sensors help humans "hear/see" things that they could not hear or see in the past and do something that we could not do in the past. This paradigm shift creates numerous challenges and opportunities for engineering.

III. Challenges

M2M can be broken down into four major layers as shown in Figure 2. Sensors collect data, communication units relay the information collected, computing units analyze the information, and service layers take action.

In the future, enormous numbers of sensors will be deployed. The costs of servicing such sensors will be a major concern. Hence, one challenge is sensor technology that requires minimal or even zero effort to deploy and maintain. According to [3], many domestic ubiquitous computing projects have failed because the complexity of sensor deployment. Additionally, one important sensor service cost is battery replacement. It is often almost impossible to replace sensor batteries once they are in the field. Therefore, another challenge is low power sensor design, or designs which do not require a battery change over the lifetime of the sensor. For example, if a sensor is deployed on an animal for tracking purposes, the battery of the sensor

Figure 2: Major components in M2M and their challenges

Figure 3: Refinement process of transforming sea of raw data into information, knowledge, and finally useful wisdom
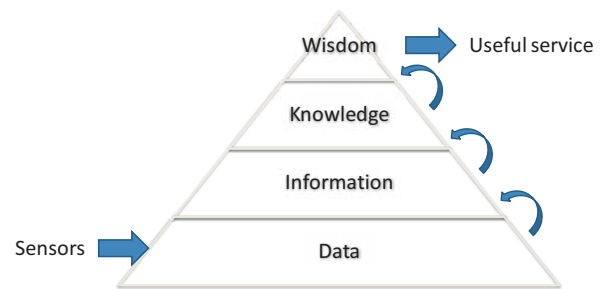
should outlive the animal.

After the sensors collect the data, the next step is to communicate the information collected. Even today the number of devices connected to the internet exceeds the number of humans; in the future this gap will only increase. Many of the sensors will be connected wirelessly through systems like Bluetooth, WiFi, or 3G/4G cellular networks. Connecting the growing number of devices is a huge challenge. Most base stations are designed to provide a certain quality of service up to a given number of users. When there are too many simultaneous users, some users will not receive service. Since the number of devices will be orders of magnitude larger than the number of human users, this problem will become even more serious. Too, as the number of devices connected to the internet grows, so do security and privacy issues [6].

Connected devices (sensors) can produce oceans of data. According to Cisco, the number of objects on the internet exceeded the number of humans in 2008 or 2009 [9], a trend that accelerates every year. Thus, in the future, the amount of data generated by machines will be orders of magnitude greater than that generated by humans. However, we need layers of intelligence to transform this data into wisdom (Figure 3). In this new computing era, analysis of data and its context will play a key role [2].

Sensor generated data are different from human generated data:

1. There are often real-time requirements for data processing, e.g., natural disaster warning. Stream data processing or mining is a critical component in the analysis.
2. There is often a huge amount of temporal and spatial redundancy in the information. It will be more efficient if the analytical algorithm can take advantage of such redundancy. However, synchronization of data from different sensors may be inaccurate.
3. Reliability or accuracy of the data may be unpredictable. Separating in reliable from unreliable signals will be a critical component of the analysis.
4. The goal for the sensor data processing is either machine or human action, meaning that the requirement for data processing is more stringent.

Generating accurate and timely answers is a tremendous challenge. Further, many analytic approaches assume that all of the data is present on the server. However, it takes both power and bandwidth to communicate the data to the server. As M2M devices become more powerful, intelligent computation must be distributed across both the devices and the cloud.

Finally, after understanding the contexts, machines should either take proper action or prompt humans for proper action. Ideally, machines should work for people. Today, when using search engine, a significant portion of the work is done by the users instead of the machines. The machines give us a list of possible results and ask users to refine the search results. As the users search through the list of results, the search engine uses the results to refine future search results. Users actually work for the machines in this case. Ultimately, we want the machines to work for users, not vice versa. Furthermore, to build a successful M2M ecosystem, it is crucial to have unified standards for everyone to follow. However, major standards are still under development [1][7][13][18]-[29][38][41] and many emerging applications are using their own standards. For example, EPC Global and Ubiquitous ID are two different, non-compatible ways of identifying objects [37]. Existing solutions are highly fragmented and technology is typically dedicated to a single application. Since the solutions are being designed from scratch, such projects take longer than conventional IT projects [11]. Diverse technical solutions and standards are slowing the development of the global M2M market [39].

IV. Opportunities

While there are many challenges, they are also many research and development opportunities. The following subsections will discuss these opportunities.

A. Low-power wireless sensors

One of the challenges is to design low-power sensors which do not need battery replacement over their lifetimes. This creates a demand for energy-efficient designs. A typical wireless sensor has 4 major components: a sensing unit, a processing unit, a transmitting/receiving unit, and a power unit, as shown in Figure 4.

One of the opportunities is to design low-power sensing unit. Highly accurate sensor modules often consume great amounts of power. One alternative is to use an array of low-accuracy modules with lower power consumption, and then use data fusion to create high-accuracy information.

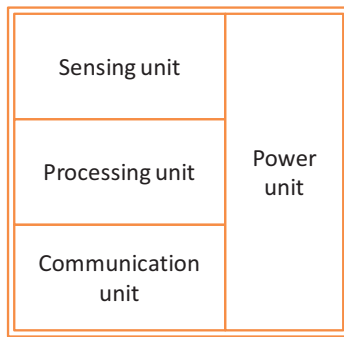Another opportunity is to design a new video processing

Figure 4: Schematic of a typical wireless sensor

and encoding algorithm. Traditionally, video encoders have higher complexity than video decoders. This is because encoders must analyze the redundancy in the video in order to compress the video efficiently. It is acceptable conventionally because usually the video is encoded only once while it is decoded multiple times, e.g., DVD or video on-demand. However, in the new paradigm, due to low power budgets, the redundancy cannot be analyzed completely at the video sensor node. A good video encoding algorithm must shift the computational complexity from encoders to decoders [8].

After the collected data is processed, the information must be transmitted to the gateway or even to the backend server. While the power of digital circuits scales reasonably with Moore's Law, the power of analog circuits does not scale well. Therefore, analog circuits in wireless communication will consume relatively larger amounts of power than digital circuits. For low-power sensor node, design of low-power transmitter circuits should also be considered. One solution is to use digitally intensive circuits to replace analog circuits in wireless communication.

To prolong the battery life of the sensor, we can harvest energy from the ambient environment from sources such as lights, heat, vibration, or radio frequency. The efficiency of today's RF energy harvesting solutions is around 16.3% [35]. More efficient solution are thus an urgent need. Further, the amount of energy that can be harvested from RF is at the uW level, still to low to power a wireless sensor. We also need a high-efficiency energy harvesting circuit adaptable to different sources.

B. Better connectivity

Another challenge in M2M is autonomous networking to connect large, dense populations of stationary and moving devices with efficient energy usage.

First, many existing wireless standards are optimized for human-to-human applications, but may not be able to support large numbers of devices in a limited spectrum. Moreover, in the future, the rate of increase in the available spectrum will be slower than the rate of increase in the number of wireless sensors. Fortunately, machine-to-machine communication has several unique characteristics: the data rate is often lower, the information from different sensors or at different time steps may have strong correlations, and some messages do not require real-time delivery. Therefore, one approach to these problems is to form clusters of machines. Instead of communicates with the base station directly, machines talk to nearby cluster head, which in turn pass on to the base station. This will reduce the machine transmission power demand and increase spatial reuse of the spectrum. Another possibility is to fuse information or remove redundancy of the information, e.g., via distributed coding, to further reduce bandwidth usage.

Second, many connected devices are mobile, such as sensors install on vehicles. These sensors not only need to communicate with other sensors via intra-vehicle networks, but also inter-vehicle networks. Existing radios most likely underperform in on-road wireless channels. To provide reliable vehicle-to-vehicle communication, we should first study the vehicle mobility model and then develop the optimal communication protocols based on the channel model.

Finally, wireless communication consumes large amounts of bandwidth. A recent study shows that up to 70% of the power used when a person is playing an on-line game on a mobile device goes to wireless communication [33]. As Moore's Law continues to scale down the power consumption of the computation circuits, the power consumption of the communication circuits will become even more dominant. Many connected devices are battery-powered, and current human-to-human communication designs do not consider energy efficiency as the first priority. Yet, energy-efficient machine-to-machine communication is crucial. In particular, signaling costs are high. Self-organizing hybrid distributed and centralized competition and cooperation framework may be one approach to reducing signaling overhead.

C. "Zero-touch" management & analysis

Today, the application development infrastructure of the M2M is highly fragmented and proprietary [30]. Lack of common tool kits for application developers actually hinders the growth of M2M. Development of a flexible, scalable, robust, and optimized applications development framework for data analytics, data security, and sensor management is also important.

Turning the sea of data into useful contexts or wisdom is extremely critical. Massive multi-modality, heterogeneous sensors create great complexity in analyzing data. Since data may be collected by hundreds or thousands of different sensors, the incoming data rate may be billions or trillions of bytes per day, with the information noisy and unreliable in nature. An API should be defined so that (1) common machine learning and data mining techniques can be easily used and (2) application specific problems can be effectively addressed. To increase the reliability of the information, the temporal and spatial redundancy should be leveraged to detect anomalies. To reduce the computation time of network inferences, fast hashing techniques may be used, as the majority (95%) of the data may not be useful.

Another great opportunity is to design a context-aware, decentralized analysis algorithm. Many analytic algorithms assume that the system has all the data on the server. However, it takes power and bandwidth to communicate the data to the server. Moreover, not all the data are important or useful. Architecture-wise, it is natural for sensor data to be processed in a hierarchical and distributed fashion. Data may be analyzed and fused in sensors or gateways before arriving at the data center to save energy and bandwidth. If the system can first understand what context is important, it

only needs to transmit the relevant information to the backend server or the cloud. As devices become more computationally capable, intelligent computation may easily be distributed between the sensors and the backend servers. Furthermore, it will help if the computation can easily migrate from one piece of hardware to another piece of hardware. This is because the computational capabilities of the devices and the bandwidth availability and energy consumption of the communication will improve over time. However, we need to explore programming techniques to improve software development processes for distributed heterogeneous computing. Furthermore, open challenges in this direction remain. These include (1) network protocols, and data formats are not yet compatible across different devices, applications, and servers; (2) sharing devices resources across multiple applications is not a common practice; and (3) infrastructure will become incompatible over time.

Data security and privacy is always a major concern, even more important in M2M, which touches many aspects of human life. Some low-cost devices have a limited budget to implement strong security or cryptography features. These lightweight devices can become the weakest links in the system. Conventional firewalls that provide network security by blocking malicious traffic can no longer work in M2M because of its decentralized nature. If the lightweight devices are not properly secured, the data they produce cannot be fully trusted. Attackers may provide false information that alters the behaviors of the system. Designing low-cost and scalable crypto algorithms and hardware accelerators is crucial.

A system-level security analytics and self-adaptive security policy framework are also needed. Security policies specify trust relationships between entities, information creation and protection rules, and access control rules to any target asset in the system. Security policies also define requirements of the actual security mechanism that enforces security policies. In the current common practice, the security policies are either administered by IT departments or statically configured with fairly naïve rules. However, often these policies do not satisfy the requirements of the applications. This leads to an incomplete system. Further, unique challenges are faced in creating suitable security policies for M2M. First, the expected behaviors of a M2M system may evolve over time due to environmental and/or policy changes. It is thus necessary to develop solutions for detecting system behaviors outside original specifications while adapting to changes yet be able to detect deviations from expected system communications in real-time. Finally, distributed peer-to-peer machine interactions and wireless communications demand attack-defense systems that collect and analyze incoming attack information in a distributed fashion. However, usually the amount of available data is limited and sometimes only a partial view of the entire system is possible.

Most current wireless sensor networks require great human effort to configure and deploy applications. Thus, human beings are often the bottleneck in large-scale deployment and long-term sustainability of systems in the field. One of the key components that will increase the adoption rate and breadth of M2M across multiple vertical markets is a configurable middleware that supports intelligent devices
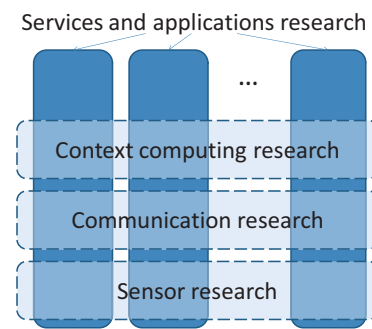
Figure 5: An inter-discipline research effort to quest for unified system architecture and standard interface, which can increase the rate and breadth of M2M adoption

management with self-configuration, self-optimization, self-healing, and self-protection capabilities. For example, gateway devices should be able to detect, diagnose, decide, and defuse faulty end nodes. Master devices should be able to auto-configure devices by deploying in-situ program updates based on context changes or user commands. In this case, the whole system can be reloaded and upgraded remotely. Further, if the middleware is designed well, it is not only easy for application developers to reprogram devices after deployment, but also easy to develop the applications/services in a distributed heterogeneous computing environment before deployment.

D. Smart service for People

Diverse and fragmented standards and interfaces between layers of the systems hinder the innovation capabilities of M2M application developers and service providers. If we want the M2M industry to grow as reliably and powerfully as Moore's Law, we will need to develop the same kind of interdependent, mutually advantageous industry structure that causes the PC software and hardware industry spiral to function. Standard interfaces are the key to the spiral. When each individual component provider has a fixed boundary condition, it is easier and faster to create innovative solutions. It is not easy for one single component or solution provider to define a generic system architecture and the interface. The best way is to bring community efforts together, as shown in Figure 5. While some works on across-the-broad sensor, communication, and context computing technologies, some should develop a few key vertical applications and interact deeply with the horizontal components. After understanding the requirements from multiple verticals, we may be able to create standardized platforms to better serve human beings. The broader the diverse set of applications is (use fixed and mobile sensors to monitor both human and non-human objects and perform energy saving, vehicle safety, health/wellness monitoring, and environmental monitoring functions), the easier we can define the generic architecture and standard interfaces. After that, we can focus on building innovative M2M services to server humans.

V. Summary

As growing numbers of devices are added to the internet,

M2M will transform the way we live, play, and work. An exciting area for innovation, it offers numerous challenges and opportunities, from scaling applications and services from billions to trillions of connected devices, and from tera to zeta bytes of data.

References

[1] 3GPP MTC, Machine-type communication

[2] M. Balazinska, A. Deshpande, M.J. Franklin, P.B. Gibbons, J. Gray, S. Nath, M. Hansen, M. Liebhold, A. Szalay, V. Tao, "Data Management in the Worldwide Sensor Web," IEEE Pervasive Computing, vol.6, no.2, pp.30-40, April-June 2007

[3] C. Beckmann, S. Consolvo, and A. LaMarca, "Some Assembly Required: Supporting End-User Sensor Installation in Domestic Ubiquitous Computing Environments," UbiComp 2004, LNCS 3205, pp. 107-124, 2004.

[4] M. Bellis, "The History of the ENIAC Computer," About.com Guide

[5] G. Brumfiel, "First Eyes Inside Nuclear Plant May Be A Robot's," NPR, March 23, 2011.

[6] I. Cha, Y. Shah, A. U. Schmidt, A. Leicher, M. V. Meyerstein, "Trust in M2M communication," *Vehicular Technology Magazine, IEEE* , vol.4, no.3, pp.69-75, Sept. 2009

[7] K. Chang, A. Soong, M. Tseng, and Z. Xiang, "Global Wireless Machine-to-Machine Standardization," IEEE Internet Computing, vol.15, no.2, pp.64-69, March-April 2011.

[8] C.-C. Chiu, S.-Y. Chien, C.-H. Lee, V. S. Somayazulu, and Y.-K. Chen, "Distributed video coding: a promising solution for distributed wireless video sensors or not?" in Proc. Visual Communications and Image Processing 2011, Nov. 2011.

[9] CISCO, "How the Internet of Things Will Change Everything," http://youtu.be/mf7HxU0ZR_Q

[10] CISCO, "Imagine the Possibilities," http://youtu.be/RGk7llP9Z8k

[11] T. Cory, "Building M2M services means overcoming new challenges", Telecom Engine, December 13, 2009

[12] D. Estrin, "Participatory sensing: applications and architecture [Internet Predictions]," Internet Computing, IEEE , vol.14, no.1, pp.12-42, Jan.-Feb. 2010

[13] ETSI TC M2M, European Telecommunications Standards Institute technical committee machine-to-machine

[14] European Commission, Directorate-General for the Information Society and Media, "Vision and challenges for realising the Internet of things," 2010.

[15] L. Greenemeier, "Can Mobile Phone Networks Be Improved to Better Cope with Emergencies?" Scientific American, August 2011.

[16] IEEE Communications Magazine, special issue on "Recent progress in machine-to-machine communications," April 2011

[17] IEEE Wireless Communications, special issue on "The Internet of things," December 2010.

[18] IoT-A, Internet of Things - Architecture, http://www.iot-a.eu

[19] IoT-i, Internet of Things Imitative, http://www.iot-i.eu

[20] IPSO Alliance, IP for Smart Objects

[21] ISO/IEC JTC 1/SC 6 Telecommunications and information exchange between systems

[22] ISO/IEC JTC 1/SC 31 Automatic identification and data capture techniques

[23] ITU-T IoT-GSI, Internet of Things Global Standards Initiative

[24] ITU-T JCA-IoT, Joint Coordination Activity on Internet of Things

[25] ITU-T Study Group 2 Operational aspects of service provision and telecommunications management

[26] ITU-T Study Group 11 Signaling requirements, protocols and test specifications

[27] ITU-T Study Group 13 Future networks including mobile and NGN

[28] ITU-T Study Group 16 Multimedia coding, systems and applications

[29] ITU-T Study Group 17 Security

[30] K. Johnson, "Machine-to-Machine: Reinventing Embedded Devices for Smart Services," http://www.intel.com/content/dam/doc/technology-brief/m2m-reinventing-devices-for-smart-services.pdf

[31] A. M. Kelly, "Telecom Milestone: More Cellphone-Only Than Landline-Only Households," Mediamark Research Inc., Sept 2007.

[32] G. Lawton, "Machine-to-machine technology gears up for growth," Computer , vol.37, no.9, pp. 12- 15, Sept. 2004

[33] T. Pering, Y. Agarwal, R. Gupta, and R. Want, "Coolspots: reducing the power consumption of wireless mobile devices with multiple radio interfaces," in Proc. of ACM Mobile Systems, Applications and Services, pp. 220–232, 2006.

[34] Public Polls from EE5158 students, National Taiwan University, Sept. 2011.

[35] A. Sample and J. R. Smith, "Experimental results with two wireless power transfer systems," Proceedings of the 4th international conference on Radio and wireless symposium, pp. 16-18, Jan. 2009.

[36] J. Santa, A.F. Gomez-Skarmeta, "Sharing Context-Aware Road and Safety Information," Pervasive Computing, IEEE , vol.8, no.3, pp.58-65, July-Sept. 2009

[37] J. Sen, "Internet of Things-A Standardization Perspective," July 2010, www.gisfi.org/wg_documents/GISFI_IoT_2010062.pdf

[38] SENSEI, Integrating the Physical with the Digital World of the Network of the Future, http://www.sensei-project.eu/

[39] J. O. Smith, "The Coming of Age of M2M Standards," Connected World Conference, Keynote presentation, 2010.

[40] R. Staszewski, T. Jung, R. B. Staszewski, K. Muhammad, D. Leipold, T. Murphy, S. Sabin, J. Wallberg, S. Larson, M. Entezari, J. Fresquez, S. Dondershine, and S. Syed, "Software Assisted Digital RF Processor for Single-Chip GSM Radio in 90 nm CMOS," IEEE Custom Integrated Circuits Conference, 2006.

[41] TIA TR-50, Telecommunications Industry Association Smart Device Communications Standards Committee

[42] uniteChannel, "The simple truth about LTE," http://youtu.be/asxYJw7wlHE