Machine-to-Machine Communication Architecture as an Enabling Paradigm of Embedded Internet Evolution

ZORAN BOJKOVIC, BOJAN BAKMAZ, MIODRAG BAKMAZ
University of Belgrade
SERBIA
z.bojkovic@yahoo.com http://www.zoranbojkovic.com

Abstract: In addition to human-to-human (H2H) and machine-to-human (M2H) communications, an emerging technology enabling full mechanical automation that may change living styles is being developed. Communications among machine-type devices are known as machine-to-machine (M2M) communications. Current market penetration and recent predictions confirm that M2M system deployments are increasing exponentially. This is driven by the needs of industries to automate their real-time monitoring and control processes as well as the increasing popularity of smart applications to improve the living style. Owing to low power, cost efficiency and low human intervention, M2M communication has become a huge force for a number of wide variety of real-time applications such as environmental monitoring, smart grid, remote healthcare, building automation, etc. This work represents comprehensive survey on M2M communication architecture and considers some open research issues and challenges in reliability and security domains.

Key-Words: IoT, M2M communications, reliability, security, smart grid, wireless networks.

1 Introduction

Embedded Internet, also known as Internet of Things (IoT) paradigm, represents a dynamic global network infrastructure with self-configuring capabilities based on standard and communication protocols [1]. A promising technology for the development of IoT communications platforms with high potential to enable a wide range of applications in different domains is machine-to-machine (M2M) communication. This type of communication is characterized by involving a large number of low power intelligent objects sharing information and making collaborative decisions without human intervention [2]. Achieving better cost efficiency, M2M communications has become a market-changing force for a wide variety of real-time applications, such as environmental monitoring, remote healthcare, smart systems, industrial automation, etc. The advanced metering infrastructure of the smart grid presents the major growth potential in the M2M market today [3]. Also, with the penetration of embedded devices, M2M communications will become a dominant in home networks. Moreover, M2M communications are declared as one of the five disruptive technology directions for the fifth generation (5G) mobile networks [4].

This work presents a comprehensive survey on M2M architecture as an enabling paradigm in evolution of embedded Internet. Fundamental challenges in end-to-end architecture for M2M communications are identified. Then we focus on M2M communications in smart grid, considering networking for home energy management system. Finally, some challenging issues concerning reliability and security of M2M architecture are discussed.

2 End-to-End Architecture for M2M Communications

The characteristics of M2M communications are quite different from those of conventional networks. M2M networks are composed of large numbers of nodes, since the main subject participating in communication is a machine or object. Because most machines are battery operated, energy efficiency is the most important issue. As for the machine senses, itself or its surrounding physical environment, the traffic per object is very small. However, data are generated from a large number of objects, and because the data generation period, amount, and format are all different, a large quantity of data is generated. While M2M communication can occur without human intervention, operational stability and sustainability are also required.

In 2009, the European Telecommunications Standards Institute (ETSI) has established the M2M Technical Committee with the purpose to develop
an end-to-end architecture for M2M communications. According to ETSI, an M2M system is composed of the five key elements with functions as follows:

- The M2M component, embedded in a smart electrical device, transmits data or replies to requests.
- The M2M gateway enables connectivity between the M2M components and the communication network.
- The M2M server works as a middleware layer to pass data through various application services.
- The M2M area network provides connectivity between M2M components and M2M gateways.
- The M2M communication network provides connection between M2M gateways and M2M servers.

These key elements constitute the general M2M communication architecture in the three interlinked domains, i.e., the M2M device domain, network domain and application domain as shown in Fig. 1.

In the M2M domain, a potentially large number of nodes and M2M gateways (GW) are integrated to enable automated and diverse services. Each embedded node as a flexible and smart device should be equipped with various functions, such as data acquisition, data preprocessing, data storage, distinctive address, wireless transceiver, power supply, etc. They can make intelligent decision and transmit the sensory data packets to the GW in single-hop or multihop manner. The M2M GW is an integrated device. After collecting the packets from embedded nodes, it is able to intelligently manage the packets and provide efficient paths for forwarding these packets to the remote back-end server via wired/wireless networks.

In the network domain, a large number of heterogeneous points of attachment (PoA) potentially coexist. Here, convergence of heterogeneous networks (e.g., xDSL, LTE, WiMax, WiFi, etc.) in an optimal way provides cost-effective and reliable channels for sensing data packet transmission from M2M to the application domain.

Finally, in the application domain, various real-time services for remote management monitoring are provided and can be classified into several categories, such as traffic, logistic, business, home, etc. [5]. Back-end server is the key component for the whole M2M communication system. It forms the integration point for all collected data from M2M device domain.

Hierarchical network architectures offer effective solutions when increase in M2M devices poses a network capacity concern. Cost-effective solution, in this case, lies in hierarchical network architectures with both multiple tiers and multiple radios [6]. Large cells provide coverage to M2M devices and support high mobility in the multitier hierarchy. On the other side, smaller network elements such as relays and pico or femto PoA bring connectivity closer to the devices, increasing link reliability and system capacity. The lower cost of smaller PoA makes them an attractive method of adding capacity [7]. Multiple access networks can be integrated and managed as part of a single hierarchical network. The additional spectrum and connectivity available across heterogeneous networks can be used to further improve system capacity and quality of service (QoS). The costs associated with the additional capacity are very low, because of the fact that the alternate spectrum could be free (e.g., unlicensed spectrum). Network elements, such as the integrated femtocell PoA [6], can implement tighter coupling across wireless technologies, utilizing efficiently the available spectrum.

The past few years have seen increasing research interest in the design of home M2M networks in order to achieve a highly connected, efficient and reliable environment. A robust architecture for
home M2M network was proposed in [8] and [9]. This network architecture can be subdivided into three complementary M2M structures: home networking, healthcare, and home energy management. In particular, the home M2M network is essentially a heterogeneous network that consists of a backbone network and multiple sub-networks. In the backbone network, there is an intelligent home GW that manages the entire network and connects the home network to the Internet. The network-related functionalities are implemented in the home GW, including access control, multimedia encoding/format conversion, security management, and QoS provisioning. Each sub-network operating in a self-organized manner can be specially designed for an on-demand application, and has a sub-GW as an endpoint that connects the sub-network to the home GW and the backbone network. Home GW and sub-GW are logical entities, and their functionalities can be physically implemented in a single device.

Following this framework, an improved home M2M network architecture that incorporates the benefits of inter-cloud computing, was proposed in [5]. Since M2M communications will generate vast amounts of data and have many resource constrained devices, it can be envisaged that cloud computing will play a key role in M2M development by offering desirable features such as mass data storage, data offloading, processing, and virtualized infrastructure [10]. The optimal distribution of devices and cloud intelligence can provide another way to solve the problem of seamless connectivity for M2M communications.

3 Network Architecture for Home Energy Management System

Smart grid system enables the next-generation electrical power grid with the capability of adaptive and optimal power generation, distribution, and consumption [11]. It is one of the strongest driving forces for M2M communications. As M2M communications will connect a number of devices and systems together, the optimal network architecture design is an important issue in order to minimize cost of communications (e.g., hardware, maintenance, and radio resource management, etc.) while meeting QoS requirements of the traffic and applications.

Home energy management system (HEMS) focuses on the power consumer side in a smart grid [12]. Home appliances with smart meters can be monitored and controlled by a control center to optimize the power supply and consumption. M2M communications play a major role in a HEMS since information about home appliances has to be transferred to the control center for analysis and optimization. Because of this fact, HEMS networking is a representative example of M2M communication architecture (Fig. 2).

![Fig. 2. An example of M2M network architecture for the smart grid.](image)

The status and demand data are transferred from the smart meter of each house to the traffic concentrator/GW, which forward it to the wide area network (WAN) base station (BS, e.g., LTE) deployed for a particular service area. The deployment of a concentrator has to minimize the cost due to QoS degradation from packet delay, loss, error ratio, etc. The BS forwards HEMS traffic to the control center for data processing and storage. Home appliances (connected to a smart meter) are power consumption devices in the smart grid. A home area network (HAN) can be established among home appliances and a smart meter (e.g., over ZigBee). A neighborhood area network (NAN) is established among smart meters of the houses in an area to support the HEMS. A NAN has a concentrator to collect HEMS traffic from smart meters using WiFi as a short-range communication technology. The received packets are stored in one of the concentrator’s buffers depending on QoS requirements. The WAN transceiver of the concentrator retrieves a head-of-queue packet from the buffer and transmits it to the BS. After that, they are forwarded over the Internet backhaul to the
control center. HEMS data are received for processing and storage in control center, and can be used to optimize the electrical power generation and distribution.

When the smart meter reports the power consumption status, the control center periodically establishes a contract with the generators, for example every hour to buy power supply. For the given estimated aggregated power demand of all home appliances in a service area, the power price is determined by the power generator for a periodic contract. Estimated power demand can be incomplete and outdated due to the QoS degradation of data transmitted from home appliances. In this case, actual demand can be different from estimated demand. If estimated demand is larger than actual demand over-supply is occurred. On the other side, if actual demand is larger than the estimated demand, additional power supply is required. It is important to notice, that the QoS degradation can result in over- and under-supply situations that incur costs for consumers. Thus, the QoS requirements for HEMS traffic would be an important issue for HEMS network architecture designing.

4 Reliability of M2M Architecture in Traffic Domain

For achieving energy efficiency in M2M communications, since not all M2M nodes are expected to simultaneously be active, reliability is a challenging issue. In order to improve the reliability of M2M communication architecture, exploiting redundancy technologies, including information redundancy, spatial redundancy, and temporal redundancy, can be an efficient approach [2]. Here, three major reliability domains are identified:

- Reliability in sensing and processing. M2M node may not be sufficient to accurately sense and process monitoring data due to component failure.
- Reliability in transmission. In green communications, not all nodes are active, which may result in unreliable transmission in the M2M architecture.
- Reliability at the back-end server. When the number of arrival packets dramatically increases, the server cannot deal with the challenging situation, causing reliability and QoS degradation.

In order to achieve the reliability issue at the back-end, a pair of servers, the primary and secondary one, can be deployed at the application domain. This can be represented as complex queuing model, as shown in Fig. 3. When the number of arrival packets is small, only the primary server is active. On the other side, when the number of arrival packets is large, the secondary server will be active accordingly. Here \( \lambda \) denotes the packets arrival rate at the primary back-end server where they are served with intensity \( \mu_p \) or forwarded to the finite queue (memory). Packets which are not accepted by primary server are offered to the secondary server with arrival rate \( \alpha \lambda \) and serving intensity \( \mu_s \) or forwarded to the finite (or infinite) queue. For adequate analysis of this model application of simulating techniques is required.

![Fig. 3. Complex traffic model for back-end server reliability.](image)

The simplified model for primary and secondary servers is presented in [2], as M/M/1 queuing systems, where arrival rates are \((1 - \alpha)\lambda\) and \(\alpha \lambda\), respectively (Fig. 4).

![Fig. 4. M/M/1 queuing model for back-end server reliability.](image)

If \( \lambda \) is small, all packets will be served by the primary server considering energy saving. However, when \( \lambda \) increases, a fraction \( \alpha \) (\( 0 \leq \alpha < 1 \)), of the packets will be served by the secondary server, and the rest of packets \((1 - \alpha)\), will still be served by the primary server for guaranteeing the QoS in terms of average service delay. By calculating total average delay derivative per \( \alpha \), the relation

\[
\lambda > \mu_p - \sqrt{\mu_p \mu_s}
\]

is obtained, and it indicates condition for secondary server activation. Here, the overflow property of secondary arrival rate is neglected.

An approach where instead Poisson traffic, overflow traffic is considered is also applicable in
the case of the systems with losses. More details about systems with overflow traffic and changed serving intensities including traffic parameters determination could be found in [13, 14].

5 Security Requirements for M2M Communication Architecture

Research in security for M2M communications is still in its infancy [5]. M2M security mainly addresses the identification of potential attacks, threats, and vulnerabilities. As for attacks in M2M, they can be passive or active. A passive attack does not disrupt the operations of an M2M communication system, but attempts to learn information about M2M communications by eavesdropping. On the other hand, an active attack is an attempt to deliberately modify sensory and decision data in the M2M and network domains, or gain authentication to access the back-end server. Active attacks can be subdivided into external and internal attacks. An external attack is launched by attackers who are not equipped with key materials in an M2M communications system, while an internal attack is one from compromised M2M nodes that hold the key materials. Internal attacks cause more serious damage to the M2M systems compared to the external attacks.

In order to establish a secure M2M communication architecture, security mechanisms should achieve requirements such as [1, 2]: confidentiality, authentication, non-repudiation, access control, availability, and privacy. Confidentiality ensures that only authorized entities can read M2M sensing data. Integrity must be ensured so that illegal alteration of the sensing data (e.g., delaying, modifying, deleting, etc.) can be detected. Authentication allows the back-end server to certify the sensory data of the M2M nodes. Non-repudiation guarantees that M2M nodes, once sending data, cannot deny the transmission. Access control represents the ability to restrict and control access to the application domain, i.e., it allows only authorized M2M application systems to gain access to the back-end server. Availability ensures that whenever M2M application systems access the back-end server, it is always available. Privacy is very important in the case of privacy-sensitive M2M communication systems. Generally speaking, the security requirements in M2M communications can be achieved by cryptographic techniques, e.g., symmetric or asymmetric encryption primitives can be employed to achieve confidentiality, while digital signature and message authentication code techniques can achieve others. It should be noted that most security mechanisms only efficiently defend against external attacks. Once M2M nodes are compromised and launch some internal attacks, more sophisticated security mechanisms are needed. In [2], two mechanisms adapted to the M2M security domain are introduced, including early detection of a compromised node with bandwidth-efficient cooperative authentication to filter false data.

6 Conclusions

While mobile Internet is evolving towards embedded concept, M2M presents both challenging and opportunities to the science and industry. There exists significant economic motivation for network operators and equipment manufacturers to invest in future M2M communication architecture and new services.

Two major things needed for the realization of IoT vision are the development of new technologies that scale with the growth of M2M markets and a broad standardization effort in system interfaces, network architecture, and implementation platforms. M2M communications will play an important role in data exchange of a pervasive computing, and can be adopted in many applications (e.g., environmental monitoring, energy management, remote healthcare, etc.) with objectives to improve efficiency and reduce cost. The smart grid is one of the strongest driving forces for the advance of M2M communications.

Having in mind that M2M communications will be an exciting research area for many years to come, some future research directions regarding architecture design can be identified as:

- Achieving energy and cost efficiency of connections between internal nodes and external networks.
- Gateway connecting non-IP based and IP based components with ensured consistency and persistency of addressing.
- Supporting large-scale terminal devices by group control.
- Providing QoS enhancements by cross-layer joint design.
- Necessity for comprehensive analysis of M2M traffic features.

References:


