

Modeling Aggregate Input Load of Interoperable Smart City Services

Aleksandar Antonić
University of Zagreb, Faculty of
Electrical Engineering and
Computing
Unska 3
Zagreb HR-10000, Croatia
aleksandar.antoniac@fer.hr

Martina Marjanović
University of Zagreb, Faculty of
Electrical Engineering and
Computing
Unska 3
Zagreb HR-10000, Croatia
martina.marjanovic@fer.hr

Ivana Podnar Žarko
University of Zagreb, Faculty of
Electrical Engineering and
Computing
Unska 3
Zagreb HR-10000, Croatia
ivana.podnar@fer.hr

ABSTRACT

The Internet of Things (IoT) is expanding and reaching the maturity level beyond initial deployments. An integrative and interoperable IoT platform proves to be a suitable execution environment for Smart City services because users simultaneously use multiple services, while an IoT platform enables cross-service data sharing. A large number of various IoT and mobile devices as well as the corresponding services can generate tremendous input load on an underlying IoT platform. Thus, it is crucial to analyze the overall input rate on Smart City services to ensure predefined quality of service (e.g., low latency required by some IoT services). An aggregate input rate which characterizes a real world deployment can be used to check if a platform is able to adequately support multiple services running in parallel and to evaluate its overall performance.

In this paper we review IoT-based Smart City services to identify key applications characterizing the domain, e.g., smart mobility, smart utilities, and citizen-driven mobile crowd sensing services. Next, we analyze the potential load which such applications pose on IoT services that continuously process the generated data streams. The analysis is used to create a model estimating an aggregate load generated by Smart City applications. We simulate a number of characteristic application compositions to provide insight about the aggregate input load and its potential impact on the performance of Smart City services. The proposed model is a first step towards predicting the processing load of Smart City services to facilitate the assessment and planning of required resources for continuous processing of sensor data in the context of Smart City services.

CCS CONCEPTS

•Computing methodologies → Model development and analysis; •Computer systems organization → Sensor networks; •Mathematics of computing → Distribution functions;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

DEBS '17, Barcelona, Spain

© 2017 ACM. 978-1-4503-5065-5/17/06...\$15.00

DOI: 10.1145/3093742.3093928

KEYWORDS

Internet of Things, Smart City services, input load, model

ACM Reference format:

Aleksandar Antonić, Martina Marjanović, and Ivana Podnar Žarko. 2017. Modeling Aggregate Input Load of Interoperable Smart City Services. In *Proceedings of DEBS '17, Barcelona, Spain, June 19-23, 2017*, 10 pages. DOI: 10.1145/3093742.3093928

1 INTRODUCTION

The Internet of Things (IoT) is reaching the peak of expectations according to Gartner [19], and we are witnessing a consolidation of the developed technologies and paradigms beyond initial trials and prototype solutions. IoT platforms are nowadays increasingly deployed to support and connect a large number of heterogeneous IoT devices, as well as to store and continuously process the generated data streams. Thus, the digitalization of our everyday environment results in a huge number of novel IoT services addressing the needs of citizens (e.g., monitoring of personal pollution exposure or live traffic data). In parallel, various IoT platforms which follow the requirements of domain-specific applications have emerged to create the so called IoT verticals, individual vertically integrated systems focusing on a single domain. The consolidation of IoT technologies has started by addressing one of the most pressing limitations of the fragmented IoT universe—the lack of interoperability. The H2020 project symbIoTe is developing an interoperability and mediation framework that enables cooperation and interaction between IoT verticals to create an environment for cross-domain IoT services [30]. To achieve an interoperable ecosystem, it is necessary to study and analyze individual performance of a service, but also to determine aggregate performance factors impacting all services running in parallel, i.e., to analyze individual performance requirements posed on underlying IoT platforms and the cumulative requirements related to all services.

Smart City is an ideal example of an interconnected ecosystem which serves as a driver for interoperable IoT deployments, especially in terms of connecting various domains and creating different IoT services. IoT platforms serving the Smart City domain collect, store and process all data generated in urban environments regardless of their source: these can be fixed sensors, mobile sensors mounted on public transportation, or smartphones with built-in sensors carried by citizens who wish to improve the quality of life in their city (mobile crowd sensing, MCS). The Smart Santander

project [29] is a large testbed and an example of a real-world Smart City deployment. Such platforms *connect* data sources with various services, and thus create an environment for deploying novel context-aware services which are particularly useful to end users (e.g., smart mobility services with alerts and notifications).

The publish/subscribe paradigm has been identified as a suitable communication solution enabling ad hoc and non-blocking component interactions in highly distributed environments, such as IoT platforms in the Smart City domain [21]. The paradigm is designed to send data only to parties that are interested in particular data objects, and offers the means to filter unnecessary data close to a production site, so that the system is not congested with irrelevant data [2]. Despite of its favorable properties, when designing publish/subscribe solutions for IoT platforms, we need to take into account the behavior of communicating parties to be able to validate the overall system performance. In the context of a Smart City, data consumer is a Smart City platform that has to store and process the incoming data, while all sensors and smartphone applications serve as data sources that constantly feed the platform with new data. To analyze the overall performance of such integrative platforms processing data streams from a multitude of sources, we first need to assess the load generated by those data sources.

In this paper we classify Smart City services with regard to their input load posed to an underlying platform assuming a normal workload. Additionally, we provide an estimation of an aggregate input load when multiple services are deployed in parallel. We focus on the load generated at the platform input point by all data sources (sensors and user applications involved in MCS tasks) during the process of data acquisition, but we do not take into account load that is generated by specific user requests, either one-time or continuous user queries. Although user requests also have significant impact on an aggregate input load, we do not investigate them further in this paper due to the lack of information about their characteristics (the frequency of requests, notification triggers and amount) in literature.

To summarize, our main contributions are the following: 1. analysis of Smart City services for an integrative IoT platforms; 2. classification and characterization of services in terms of the generated input load, and 3. simulation that estimates the cumulative input load for aggregated Smart City services. Evaluation of an IoT solution usually focuses on performance tests to identify service limitations by means of a synthetic input rate, and typically does not assess system performance under realistic input load. This paper provides insight about the input rates which can be expected in real-world service deployments, so that IoT solutions can be evaluated with regard to their performance in real-world environments.

The rest of the paper is structured in the following way: Section 2 provides an overview of related work, while Section 3 introduces an overview of deployments and classification of Smart City services according to their application domains. Section 4 provides insights about the characteristics of a service input load and analyzes an aggregate input load in cases when multiple services are running in parallel, while Section 5 concludes the paper and gives directions for future work.

2 RELATED WORK

Smart city. Early works have recognized that the IoT will drive a significant change in habitation of urban areas. Gluhak et al. studied multiple deployments of IoT testbeds, and have evaluated prototypical services in the Smart City environment [15]. The authors discuss requirements and challenges that need to be addressed to enable proper experimentation with IoT platforms. Although their work is published in early phase of the IoT, they already stressed the importance of enabling the concurrency in service execution, handling mobility of entities and impact of human users to the overall system performance and acceptance. Jalali et al. present enabling technologies and an architecture for the Smart City environment, and point to aggregation of data during its transfer from the source to the core network where an IoT platform will store data for future use [20]. The paper also presents applications that will drive the development of Smart City architectures. A more detailed analysis of Smart City services and application is available in [7] which focuses on positive synergy of a novel concept called the Cloud of Things, which interconnects the areas of Cloud Computing with IoT. In addition to the example usage of cloud-driven IoT applications, Botta et al. also identify several challenges, which include the performance of such platforms. The authors stress that the main challenge is to obtain stable and acceptable network performance to reach the Cloud where data is stored, because the broadband increase in capacity did not follow the storage and computation evolution [7]. In [34] Yin et al. present a literature review that analyzes the Smart City domain from the four different perspectives: technical, application, system integration and data processing. The authors consider some of non-technical issues important for further proliferation of Smart City services, such as city planning, citizen behavior and city traffic, which can significantly influence the overall performance of a Smart City environment. Neirotti et al. study Smart Cities from a socio-economical point of view, using statistical parameters of urban environments, such as population, size, economical development, to analyze adoption of different Smart City initiatives (i.e. applications) [27].

Input load/arrival rate. Modeling of input load or arrival rate is very important in different domains, not only in the area of computer networks. Literature contains various techniques to model and assess the input load or arrival rate of customer requests for almost all purposes where queueing theory [16] can be applied. For example, the arrival rate influences customer waiting time in a bank [33] or distribution of the input load is used as parameter to model the behavior of road traffic flow [31]. The area of telecommunications utilizes modeling of input load to evaluate the performance of call centers, and to optimize their operations. The authors of [28] are experimenting with the well-accepted arrival rate model for the call center to model and evaluate the impact of arrival rate uncertainty on the call center performance. The authors stress that the performance is highly sensitive to the arrival rate estimation. This points to the fact that such analysis should be done for other domains as well, e.g., for performance evaluation of web applications or distributed systems implemented using microservices. Zink et al. study network traffic generated within a campus network focusing on the YouTube video service [36]. The authors recorded traffic traces and modeled the number of requests posed

to the service in order to gain insight about the characteristics of the traffic, such as request distribution, frequency or clip popularity, which were used to create synthetic traffic traces that can be used in further experiments. Their work is somewhat similar to ours, with a difference that we model the IoT domain and we do not have access to real input load traces (e.g. traffic traces), rather we use input load distributions as reported in relevant literature. Some findings characterizing the IoT network traffic and corresponding models can be found in literature. Huang et al. report a model for congestion control in IoT in which they used the queueing theory to analyze the performance of the model [18]. The authors built their work on steady state probability distribution and they assume the exponential distribution for arrival rate of events. Similar work is reported by Awan et al. who study the Quality of Service for delay sensitive IoT applications, and also assume an exponential distribution of the overall input load [4]. In this paper, we provide a more thorough analysis of distributions characterizing input load generated by real-world IoT applications.

Performance evaluation. In our previous work we developed the CloUd-based PUBLISH/SUBSCRIBE middleware (CUPUS) used as an underlying communication solution for Mobile Crowd Sensing [2]. The CUPUS middleware was compared to the well-known protocols used in IoT, namely the Message Queue Telemetry Transport (MQTT) protocol and Constrained Application Protocol (CoAP) and, in addition, its performance was evaluated using a real-world data set [2]. Although, we used a data set obtained during a real-world trial of a mobile crowd sensing service in the evaluation, but the input rate was synthetically created using the acquired data set. The synthetic input rate was used to test the limitations of CUPUS, rather than to investigate its performance under a realistic load. In contrast to [2], this paper investigates input rates of real-world IoT service deployments, so that researchers can perform system evaluations, both in terms of performance limitations and expected performance in a real-world deployment. Similarly, Vandikas and Tsiatsis performed performance evaluation of IoT-Framework, a framework built on open source components used to disseminate the generated data streams in an IoT environment [32]. The authors evaluate their system with regard to maximum throughput without experimenting with the distribution of input rate and only focus on the total number of data producers (i.e. an overall input rate). In addition to the experimental evaluation and evaluation using the queueing theory principles, the literature reports on evaluation using analytical models developed for a specific group of solutions. For example, Mühl et al. analyze publish-subscribe systems by modeling the interrelationship between messages in the system and develop a novel general model that describes the system behavior in details [25] as opposed to typical queueing theory models and basic metrics. A similar approach is used in this paper. We try to dissect the aggregate input rate of IoT services into its basic components by analyzing individual services generating the load, instead of using a single distribution as a parameter to represent cumulative input rate.

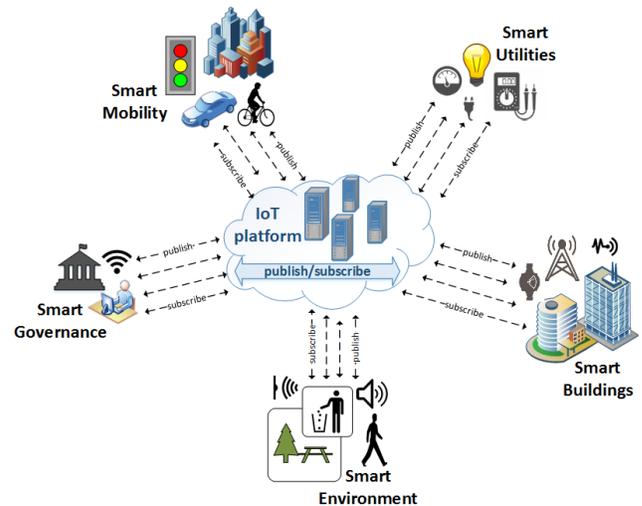


Figure 1: Smart City Environment

3 SMART CITY SERVICES

In recent years, the Smart City concept has attract a lot of interest. Although there is no single definition of a Smart City in the literature, all definitions point out that a Smart City can be defined as a system that uses Information and Communication Technologies (ICT) to meet the citizens' needs and improve the efficiency of city services. More specifically, the Smart City refers to safe, secure, environmental and efficient urban center with advanced infrastructure which integrates various public services, such as lighting, traffic or energy production, and thus increases their efficiency, reduces costs and power consumption, improves communication among the sub-systems and stimulates sustainable economic growth and a high quality of life [5, 9].

Furthermore, Pike Research¹ is forecasting that *the number of people living in cities will almost double - from 3.6 to 6.3 billion by 2050* which will require the adjustment of city authorities and services to enable the desired quality of life to their citizens. This can be achieved by using smart services which enable real-time monitoring and automated control of city infrastructure with less or even without human intervention [12]. Smart City services are usually categorized across multiple domains, including Smart Governance, Smart Mobility, Smart Utilities, Smart Buildings, and Smart Environment [35], which are recognized as key factors that express urban growth and development. Typically, smart services use numerous sensors deployed in an urban area (either heterogeneous or multiple instances of the same sensor type) which communicate with a remote IoT platform located in the cloud. Figure 1 shows a highly distributed architecture of a Smart City environment with multiple sensor instances which use publish/subscribe paradigm to communicate. Those sensors are either static (e.g., sensors deployed on traffic lights, within buildings, etc.) or mobile (e.g., sensors deployed on vehicles, carried by citizens, etc.), and create vast amounts of

¹<http://smartcitiescouncil.com/article/smart-cities-technology-market-top-20-billion-2020>

Table 1: Overview of services deployed in different Smart City testbeds

Service	Santander [29]	Padova [35]	Glasgow [8]	Cambridge [26]	Friedrichshafen [6]	Sophia Antipolis	Antwerp [23]
citizen services	✓				✓		
healthcare			✓		✓		
parking	✓	✓				✓	✓
traffic	✓	✓	✓	✓	✓		✓
smart metering			✓		✓		
smart lightning		✓	✓	✓			
smart building		✓	✓				
air quality (mobile)	✓						✓
air quality (fixed)	✓	✓	✓	✓		✓	
weather data	✓		✓	✓		✓	✓
noise detection		✓	✓				
waste management	✓	✓	✓				

data either periodically or on demand. Since each service has different requirements, they differ both in terms of generated input load and expected behavior. In the rest of this section, we briefly describe Smart City services in various domains and provide a short overview of services that have been implemented in different Smart City testbeds across Europe.

3.1 Smart Governance

Smart governance incorporates all public services which enable city authorities to efficiently communicate with citizens and to offer information in a secure and easily accessible way. Such services aim to address a number of challenges facing public sector organizations through citizen engagement platforms, such as e-Government. For example, the government can collect and analyze citizens' data to provide more efficient services for community management. Another example is a smart medical and healthcare system which serves to maintain all patient health records, can reduce cost and enhance the efficiency and quality of healthcare systems.

3.2 Smart Mobility

Smart mobility, i.e., efficient transportation has a significant role in the Smart City concept. Nowadays, great emphasis is put on the use of smart technologies to establish a smart traffic management system which includes monitoring of road conditions, free parking spots, automatic control of traffic lights, etc. Typically, such services exploit different sensors deployed on vehicles and public infrastructure, or involve individuals who continuously contribute traffic-related data to the application servers to estimate current road conditions. This information is of great importance both for citizens to adjust their routes while moving through the city, and for city authorities to plan the road infrastructure and devise adequate measures when needed. We distinguish between two types of services, those which use static sensors deployed on traffic lights or road segments that periodically send data, and those which exploit users who opportunistically collect data while moving through the city.

3.3 Smart Utilities

Smart utilities comprise different services deployed in homes to achieve intelligent control of various smart appliances (e.g., TV, refrigerator, washer, thermostat, etc.), lighting system, security cameras, gas sensors, or household energy consumption. For instance, by using a smart thermostat it is possible to remotely control house temperature and adjust heating/cooling to enhance the level of comfort before the owner enters the house, while intelligent control of security cameras and alarm systems enables real-time intrusion detections and appropriate reactions. Smart lighting service can be used to adjust the light illumination based on human movements, user preferences and ambient conditions. Similar, smart lighting service can be adopted on street lighting systems to reduce energy consumption since according to International Energy Agency report 19% of energy usage in the world is used for lighting [3]. Smart lighting service enables remote control of street lights to optimize the lamp intensity according to weather conditions and daylight availability. All those services require continuous monitoring and periodical data transmissions to a central IoT platform.

3.4 Smart Buildings

In recent years, a lot of attention is put on the intelligent design of buildings to enable advanced sensing, remote control and automation, as well as energy transmission and consumption monitoring. One example of a smart building service is continuous maintenance of its structural health which includes vibration monitoring, location of damages and predictions of its remaining lifetime. Such service typically uses different sensors deployed in buildings and surrounding areas which periodically communicate with a remote IoT platform. An important aspect of the smart building infrastructure is energy consumption monitoring which can be achieved with smart meters. Smart metering services can collect information from different devices, capture energy consumption in (near) real-time, as well as remotely control and adjust electrical power usage. Although smart meter typically refers to an electricity meter, smart buildings can also be equipped with smart devices measuring natural gas and water consumption. Such devices enable end-consumers

Table 2: Communication behavior of Smart City services

Service	Individual behavior	Number of deployments	Input load distribution	Distribution parameters
parking	publish every 10 mins	10-100	degenerate	$E(X) = 1 - 10pub/min$
traffic	publish every 10 mins	10-100	degenerate	$E(X) = 1 - 10pub/min$
smart metering	publish every 1-60 mins (mostly 15 mins)	100-1000	Poisson	$E(X) = \lambda = 1 - 100pub/min$
smart lightning	publish every 10-60 mins during nighttime	10-100	degenerate	$E(X) = 0.1 - 10pub/min$
smart building	publish every 10 mins	10-100	degenerate	$E(X) = 1 - 10pub/min$
air quality (mobile)	publish every 20 sec - 5 mins	10-100	Poisson	$E(X) = \lambda = 1 - 100pub/min$
air quality (fixed)	publish every 30 mins	1-10	degenerate	$E(X) = 0.01 - 0.1pub/min$
weather data	publish every 30 mins	10-100	degenerate	$E(X) = 0.1 - 10pub/min$
noise detection	publish every 10 mins	10-100	degenerate	$E(X) = 1 - 10pub/min$
waste management	publish every 60 mins	10-1000	degenerate	$E(X) = 0.1 - 10pub/min$

to adapt their energy, natural gas and water usage to different prices throughout the day to save money by reducing their consumption in higher price periods.

3.5 Smart Environment

Over the past few years, scientists are investigating the impact of environmental pollution on human health. It has been shown that exposure to traffic-related air pollution can cause different respiratory problems [17], while prolonged noise exposure can lead to sleep disturbance, cardiovascular diseases, hearing loss or mental health problems [11]. Therefore, city authorities aim to promptly identify contaminated areas and devise appropriate actions by using both static, as well as, mobile pollution sensors to densely monitor noise and air quality in big cities. Another important segment of smart environment is waste management control service which requires intelligent waste containers that are able to detect waste levels and improve the quality of recycling as well as garbage collection cycles.

Some of the aforementioned services have already been implemented in different real-world Smart City deployments. In Table 1 we give a brief overview of currently available services in various Smart City testbeds across Europe which will be used in the following section to estimate their data distributions and generated input load.

4 INPUT LOAD OF SMART CITY SERVICES

In this section we analyze the input load of a service in relation to its characteristics, namely: *individual behavior* which describes behavior of a single instance of a service, *number of deployments or installations* where we try to assess order of magnitude of running instances in a single Smart City, *input load* which represents probability distribution of input load that is created by a service, and *parameters* used in later analysis that describe the identified distribution. Additionally, we study the input load of service composition, which provides insight into the aggregate input load and its characteristics.

Table 2 summarizes the description of analyzed services regarding its load posed onto an underlying platform. In comparison to

Table 1, we removed the smart governance services, namely citizen services and healthcare, since those two services are usually focused on a single citizen, i.e., most of data is personal and confidential and the data does not have real value for anyone else except the current user, so it is not widely shared within a community. Such services are often centralized and literature does not report on usage patterns, so they are excluded from further analysis. We distinguish three types of sensors used across Smart City services: 1) fixed sensors that are mounted on a physical object and do not change location (e.g., sensors for monitoring building's structural health), 2) nomadic sensors that can change their location while they are offline (e.g., sensors mounted on a waste bin) and 3) mobile sensors which are mobile during their operation (e.g., wearable sensors for air quality monitoring). We do not specifically distinguish the services based on the type of used sensors, and in further analysis we consider only the mobile air quality service as fully mobile, and we do not make a distinction between fixed and nomadic sensors. Individual behavior of a sensor installation is taken from literature, and the number of deployments indicates only the order of magnitude, without the intention to give a real number, because it is hard to assess it correctly, since Smart City deployments grow continuously. The input load distribution is derived from the individual behavior, with an assumption that fixed sensors are not synchronized in their sensing cycles (i.e., we assume uniform distribution of the sensing cycles start time). A mobile air quality service depends on citizens who start them and later on in this section we report our findings regarding the input load distribution of such a service. Distribution parameters were derived from the first two columns, and the goal is to give an order of magnitude of the distribution parameters, not the exact values.

We identified the two different probability distributions of input load for Smart City services. One is the degenerate distribution, a distribution in which a random variable can have only a single value, i.e. a distribution that gives a constant value for all outcomes. The second identified distribution is the Poisson distribution, which is widely used for modeling the probability of an event occurring over a certain interval. The Poisson distribution is used in queueing theory to model the input load of a system and it has only one parameter which can be obtained empirically (e.g., it is used to

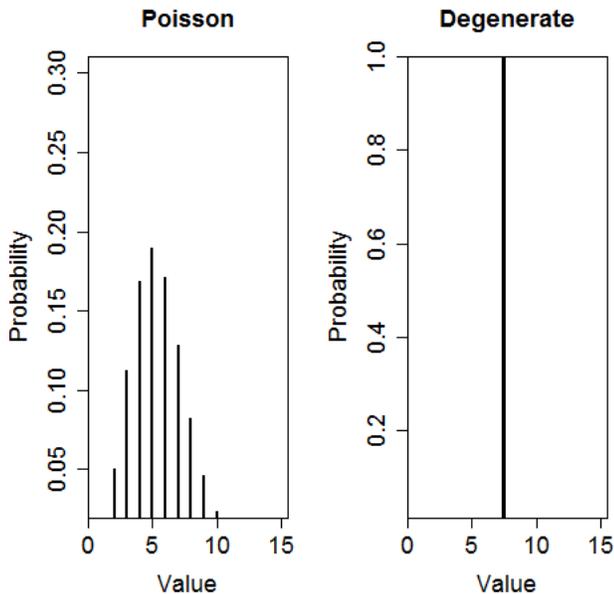


Figure 2: Probability mass function of identified distributions

model the probability of expected number of calls to a call center in a single time interval when the expected (or average) number of calls for that time interval is known). Figure 2 presents the probability mass function of both distributions.

For example a traffic congestion service developed in the Padova Smart City project which sends one data packet every 10 minutes per each deployed device, where the number of devices is constant in time [35] and sensing intervals are uniformly distributed in time, is described with a degenerate distribution modeling the input load. If 10 such sensors are deployed in a Smart City, the expected value of the input load is 1 publication/minute. Another example of a fixed deployment of sensors is the smart metering service in which devices periodically collect information every 10 to 60 minutes, depending on country regulations [10], and due to uncertainty of inter-arrival times which are modeled by the exponential distribution, the input load of such service can be modeled using the Poisson distribution [14].

The air quality and noise monitoring services in the city of Padova use static sensors which periodically send data to the application servers, while the 'Sense the Zagreb Air' project [1] uses mobile users to opportunistically collect air quality data with mobile phones and wearable sensors. The setup with fixed stations produces data with a constant rate, while the input load of mobile service is not easily predictable. We have analyzed the data acquired by real users during the 'Sense the Zagreb Air' project to determine the distribution of the input load generated by such service. The project organized a measurement campaign in July 2014 in Zagreb, Croatia, with volunteers that were collecting data while they were being mobile. We analyzed the data and obtained two graphs that characterize the input load of the mobile air quality service. Figure

3 shows the distribution of input load of the service for two different periods. The campaign was divided in two parts, during one part volunteers were carrying a sensor on their own, and during the second part a guided tour was organized when all volunteers received exact directions where to perform air quality measurements. Figure 3a represents freelance sensing between 5 PM and 6 PM every day of the campaign and Figure 3b represents a guided tour on the first day of the campaign (between 11 AM and 2 PM). We modeled input load with the Poisson distribution with good results for both scenarios. For the first scenario, the MAE parameter was 0.0025, while for the second scenario the MAE parameter was 0.0122. The analysis also shows that the distribution parameter (i.e., λ which represents the expected value) changes depending on the time of day, daily migrations and user incentives. It is interesting to observe that the guided measurement tour involved all 20 volunteers with all sensors adjusted to generated measurements periodically every 20 seconds, so the expected number of measurements (i.e. input load) would be close to 60. However, the analysis showed that the Poisson distribution with the $\lambda = 30$ shows the best fit. Further investigation of this phenomenon can be made, but it is beyond the scope of this paper.

4.1 Aggregate input load of multiple services

This subsection presents the analysis of aggregate input load when multiple services are running in parallel. The goal is to present the probability mass function that describes the aggregate load generated by services. Such distribution can be used to generate synthetic input load for testing the performance of a system with real-world parameters of generated data. First, we present a generic formula to calculate probability mass function and later we demonstrate it using the above mentioned services.

To obtain distribution (i.e., the probability mass function) of the aggregate input load of two services, it is necessary to calculate convolution of the two probability distributions, where each service input load is represented by its distribution. More formally, we form a new independent random variable Z which is defined as $Z = X + Y$, where X represents an independent random variable of input load of the first service and Y represents an independent random variable of input load of the second service. The probability mass function is calculated as follows:

$$P(Z = z) = \sum_{i=-\infty}^{\infty} P(X = i) \cdot P(Y = z - i), \quad (1)$$

where $P(X)$ and $P(Y)$ represent the probability mass functions of input load of the first and second service, respectively.

To demonstrate the aggregate input load with different distributions, we calculate the probability mass functions for three mixture of services with various combinations of the distributions.

To calculate the aggregate input load of the traffic and parking sensors, we calculate convolution of two degenerate distributions, with different distribution parameters. The probability mass function of the degenerate distribution is defined as follows:

$$P_{deg}(X = x; c) = \begin{cases} 1, & x = c \\ 0, & x \neq c \end{cases} \quad (2)$$

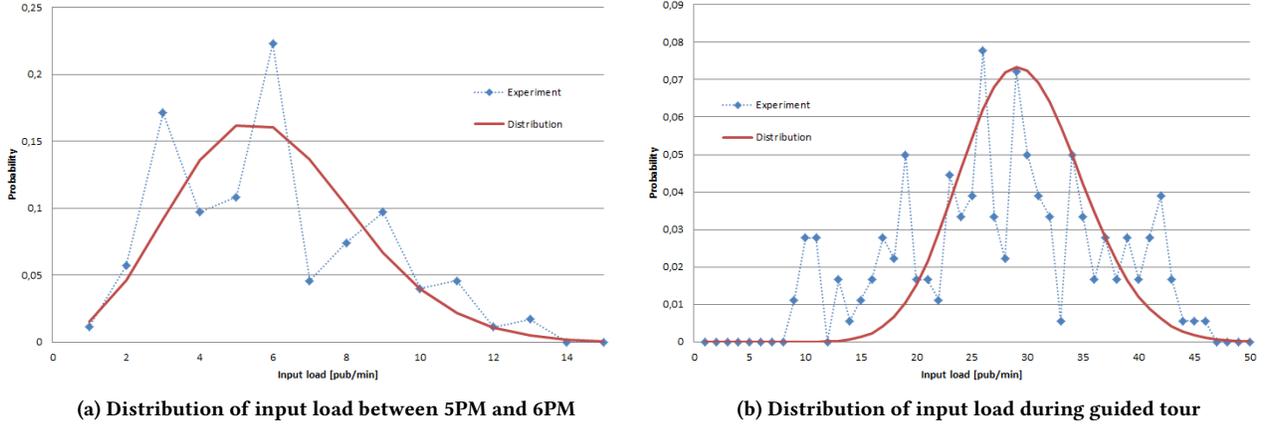


Figure 3: Distribution of input load of the mobile air quality service

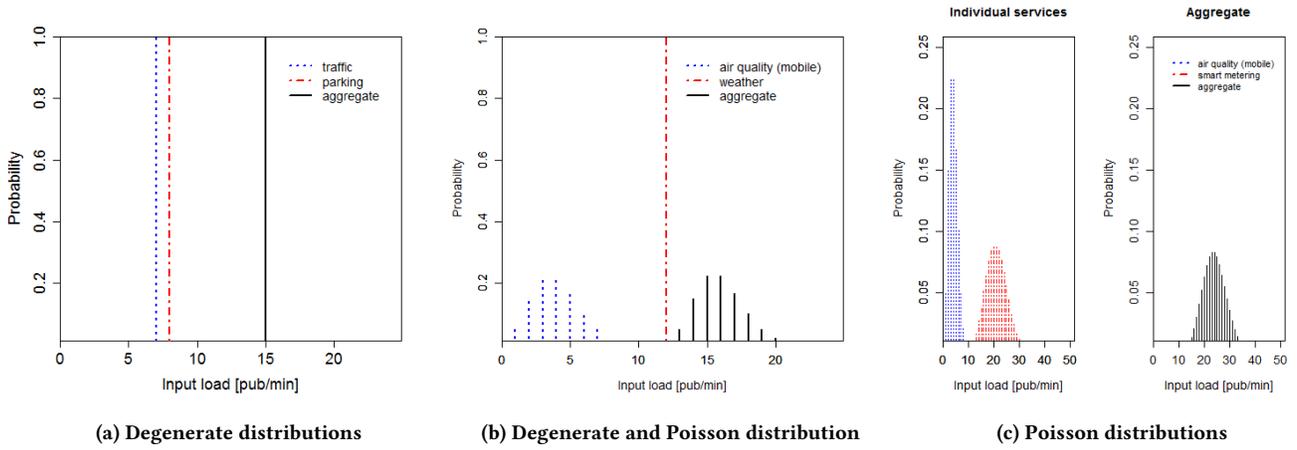


Figure 4: Probability mass functions of individual and aggregate service

To calculate the probability mass function of the aggregate input load for the traffic and parking service we calculate the convolution using Equation 1 as follows:

$$P(Z = z) = \sum_{i=0}^z P_{deg}(X = i; c_{traf}) \cdot P_{deg}(Y = z - i; c_{park})$$

$$P(Z = c_{traf} + c_{park}) = P_{deg}(X = c_{traf}) \cdot P_{deg}(Y = c_{park}) = 1$$

where X and Y represent independent random variables of input load for the traffic and parking service, respectively. The only point where the product of aggregate probability mass functions is equal to 1 is when aggregate independent variable $Z = c_{traf} + c_{park}$. All three aggregate probability mass functions are shown in Figure 4a. The \sum is limited between 0 and z since individual distributions do not have defined value for non-positive arguments.

The same approach can be used to calculate the aggregate input load of two services which have different distributions, i.e. the degenerate and Poisson distribution. The probability mass function of the Poisson distribution with parameter λ , which also defines the expected value of the distribution, is defined as follows:

$$P_{Pois}(X = x; \lambda) = \frac{\lambda^x \cdot e^{-\lambda}}{x!} \quad (3)$$

The aggregate probability mass function is calculated as follows:

$$P(Z = z) = \sum_{i=0}^z P_{deg}(X = i; c_{traf}) \cdot P_{Pois}(Y = z - i; \lambda_{mob-air})$$

$$= P_{Pois}(Y = z - c_{traf}; \lambda_{mob-air})$$

where $P_{deg}(X = i; c_{traf})$ represents the degenerate distribution of the traffic service and $P_{Pois}(Y = z - i; \lambda_{mob-air})$ represents the input load distribution of the mobile air quality service. Such service can be used for example to discover a correlation between traffic congestion and level of air pollutants. The aggregate distribution mass function is in fact the shifted Poisson distribution for the value of degenerate distribution. The expected value of the aggregate input load is $E(Z) = c_{traf} + \lambda_{mob-air}$, while the variance is the same as for the Poisson distribution of the mobile air quality service $Var(Z) = \lambda_{mob-air}$. The aggregate probability mass function of the degenerate and Poisson distribution is shown in Figure 4b.

The convolution of the two Poisson distributions, i.e., the input load of mobile air quality and smart metering service is resulting also in the Poisson distribution with the parameter which is the sum of the two parameters from individual services, as shown in Figure 4c. The aggregate probability mass function is calculated as follows:

$$\begin{aligned} P(Z = z) &= \sum_{i=0}^z P_{Pois}(X = i; \lambda_{meter}) \cdot P_{Pois}(Y = z - i; \lambda_{mob-air}) \\ &= P_{Pois}(Z = z; \lambda_{meter} + \lambda_{mob-air}) \end{aligned}$$

The expected value and the variance is defined by the Poisson distribution parameter $E(Z) = Var(Z) = \lambda_{meter} + \lambda_{mob-air}$.

Except the two distributions reported in literature, also the power-law probability distribution can be interesting, because it is used to model geographical distribution of mobile users [24] for a single time interval. We omitted it from the analysis in this paper, because further investigation is necessary to demonstrate if a geo-aware service (i.e., a service which utilizes current location of a user) would also produce an input load that follows the power-law distribution.

So far, we analyzed the aggregate input load for a case when multiple services are running in parallel, but services can be mutually exclusive where execution of the first service stops execution of the second service. In such a case, to calculate the aggregate input load of two services, it is necessary to calculate mixture distribution, where each service input load is represented by its probability mass function (P_i) and its weight (i.e. the occurrence probability) w_i . The probability mass function for mixture distribution is calculated as follows:

$$\begin{aligned} P(Z = z) &= \sum_{i=1}^n w_i \cdot P_i(z); \\ \sum_i w_i &= 1, w_i > 0. \end{aligned} \quad (4)$$

Mixture distribution is used to model an overall input load when users are migrating from one service to another, and weights represent the service share. Additionally, it is used to model an overall input load of a service that has multiple modes of operation, where each mode is represented with its own distribution. We presented that the mobile air quality service has different distribution parameters during the day and related to the user involvement, so to model overall input load of such service we calculate mixture distribution:

$$\begin{aligned} P(Z = z) &= w_1 \cdot P_{Pois}(X = z; \lambda_{mob-air1}) \\ &+ w_2 \cdot P_{Pois}(Y = z; \lambda_{mob-air2}), \end{aligned}$$

where $P_{Pois}(z; \lambda_{mob-air1/2})$ represent the distribution of sensing modes, and weights $w_{1/2}$ represent the share of users that are involved in the one or another mode. The probability mass function with different shares of user (i.e. mixtures) is shown in Figure 5. The expected value is calculated as a weighted sum of individual expected values: $E(Z) = w_1 \cdot \lambda_{mob-air1} + w_2 \cdot \lambda_{mob-air2}$

The output of the mixture distribution can be used as an input to convolution of distributions and vice versa, so with these two

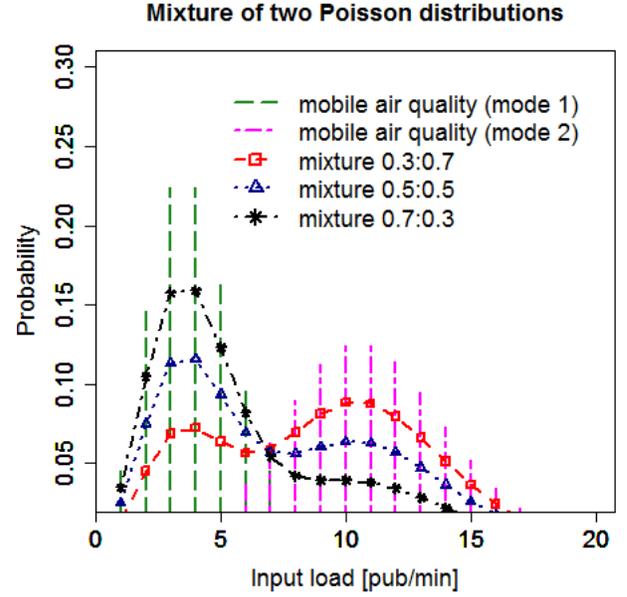


Figure 5: Probability mass function of the Poisson distribution mixture

approaches it is possible to determine a distribution of an overall input load for any combination of Smart City services.

4.2 Implications on performance evaluation

To assess the total aggregate input load of all Smart City services to an underlying IoT platform, we combine all distributions identified in Table 2. The aggregate input load consists of the degenerate and Poisson distribution, where the expected value is $E(all - services) = 270.1pub/min$. Since we did not present an exact number of deployments, but rather only identified an order of magnitude, we can conclude that a Smart City IoT platform should support 1000 publications per minute to be able to process all data in (near) real-time. Studies regarding platform evaluation show that IoT platforms can process input load of that size and even 10 times higher load [2, 32]. Although IoT platforms possess suitable techniques to process the identified load, continuous improvement is very important, because the number of deployed devices and involved users in IoT is constantly increasing (with expectations up to 50 billion by 2020 [13]). If we analyze a share of individual services in the total aggregate input load, we can observe that a mobile crowd sensing paradigm (e.g. the mobile air quality monitoring) is already responsible for a large part of the total input load. Services such as mobile air quality monitoring do not represent full potential of the MCS paradigm because they require an adequate equipment to be operable (i.e. a wearable sensor), but utilization of MCS paradigm with services that do not require anything except a smartphone (e.g., noise monitoring/detection) can generate massive amounts of data.

Evaluation of an IoT platform is often done experimentally, by executing performance tests with various input loads to test limitations of the platform and to get performance parameters. To analyze the performance with real-world services, evaluation should be made with real world data and input load which correctly represents a test case. If a distribution of input load of individual services is known, a probability mass function can be calculated and used to generate aggregate input load to a platform. Input load is generated by using the probability mass function (or the cumulative function) together with a random number generator (uniformly generated).

In addition, IoT platforms are evaluated using the analytical models based on the queueing theory principles [22]. The most common queueing theory model used in the analysis is the M/M/1 model, which represents the model where an input load is modeled by the Poisson distribution and service time is modeled by exponential distribution. The queueing theory also includes the models which use general distribution of input rate, which offer expressions to calculate parameters of analyzed system and do not follow neither the degenerate nor Poisson distributions.

5 CONCLUSIONS

The rapid expansion of the Internet of Things has opened new perspectives for deployment of different smart services. However, the lack of interoperability among platforms and services prevents IoT to reach its full potential which is particularly visible in the Smart City domain. This has led to a need for an integrative and interoperable IoT platform which supports a multitude of services with different requirements and provides prerequisites for the Smart City deployment. To achieve an interoperable ecosystem, it is necessary to analyze both the individual performance requirements posed on underlying IoT platform, as well as the cumulative requirements that represent all platform services.

In this paper we review IoT-based Smart City services with regard to their input load posed to an underlying platform during normal workload. The analysis takes into account the individual service behavior, deployment size, the probability distribution of input load created by observed service, and parameters that describe the identified distribution. Additionally, we provide an estimation of an aggregate input load generated at the platform input point when multiple services are deployed in parallel. We have identified two types of Smart City services regarding the probability distribution of input load, those which generate data following the degenerate distribution, such as parking or waste management services, and those whose input load follows the Poisson distribution, such as smart metering service or mobile air quality monitoring with wearable sensors. We have shown that the total aggregate input load of all Smart City services to an underlying IoT platform can be expressed as convolution of the degenerate and Poisson distribution with the expected value of $E(\text{all} - \text{services}) = 270.1 \text{ pub/min}$. The aggregate probability mass function can be used to generate an overall input load necessary for the platform performance evaluation.

As future work we plan to investigate the characteristics of input load of citizen-based services which do not require additional equipment and thus generate huge amounts of data (e.g., noise monitoring with smartphones). Since such services are geo-aware, we

plan to more thoroughly investigate whether their input load corresponds to the power-law distribution. Another possible direction for future work is creation of a test suite with some pre-defined input loads that represent IoT-based Smart City services which would be a step forward to the standardization of evaluation process for the IoT platforms.

ACKNOWLEDGMENTS

This work is supported in part by the H2020 symbIoTe project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688156. This work has been supported in part by the Croatian Science Foundation under the project number 8065 (Human-centric Communications in Smart Networks).

REFERENCES

- [1] A. Antonić, V. Bilas, M. Marjanović, M. Matijašević, D. Oletić, M. Pavelić, I. P. Žarko, K. Pripuzić, and L. Skorin-Kapov. 2014. Urban crowd sensing demonstrator: Sense the Zagreb Air. In *2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. 423–424. DOI: <http://dx.doi.org/10.1109/SOFTCOM.2014.7039132>
- [2] Aleksandar Antonić, Martina Marjanović, Krešimir Pripuzić, and Ivana Podnar Žarko. 2016. A Mobile Crowd Sensing Ecosystem Enabled by CUPUS. *Future Gener. Comput. Syst.* 56, C (March 2016), 607–622. DOI: <http://dx.doi.org/10.1016/j.future.2015.08.005>
- [3] Vijay Madiseti Arshdeep Bahga. 2014. *Internet of Things (A Hands-on-Approach)*.
- [4] Irfan Awan, Muhammad Younas, and Wajia Naveed. 2014. Modelling QoS in IoT Applications. In *17th International Conference on Network-Based Information Systems, NBIS 2014, Salerno, Italy, September 10–12, 2014*. 99–105. DOI: <http://dx.doi.org/10.1109/NBIS.2014.97>
- [5] Tuba Bakıcı, Esteve Almirall, and Jonathan Wareham. 2013. A Smart City Initiative: the Case of Barcelona. *Journal of the Knowledge Economy* 4, 2 (2013), 135–148.
- [6] Karsten Bolz. 2015. T-City Friedrichshafen. *Responsible-Industry* (2015).
- [7] Alessio Botta, Walter de Donato, Valerio Persico, and Antonio Pescap. 2016. Integration of Cloud computing and Internet of Things: A survey. *Future Generation Computer Systems* 56 (2016), 684 – 700. DOI: <http://dx.doi.org/10.1016/j.future.2015.09.021>
- [8] Nick Taylor Buck and Aidan While. 2015. Competitive urbanism and the limits to smart city innovation: the UK Future Cities initiative. *Urban Studies* 54, 2 (August 2015), 501–519.
- [9] Andrea Caragliu, Chiara Del Bo, and Peter Nijkamp. 2009. *Smart cities in Europe*. Serie Research Memoranda 0048. VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics.
- [10] European Commission. 2014. Cost-benefit analyses and state of play of smart metering deployment in the EU-27. (2014).
- [11] European Commission. 2002. Environmental Noise Directive 2002/49/EC. (2002).
- [12] Mischa Dohler, Ignasi Vilajosana, Xavi Vilajosana, and Jordi Llosa. 2011. Smart Cities: An action plan. 1–6.
- [13] D. Evans. 2011. The Internet of Things: How the Next Evolution of the Internet Is Changing Everything. White Paper. (2011).
- [14] A. Ghasempour. 2016. Optimum Packet Service and Arrival Rates in Advanced Metering Infrastructure Architecture of Smart Grid. In *2016 IEEE Green Technologies Conference (GreenTech)*. 1–5. DOI: <http://dx.doi.org/10.1109/GreenTech.2016.8>
- [15] A. Gluhak, S. Krco, M. Nati, D. Pfisterer, N. Mitton, and T. Razafindralambo. 2011. A survey on facilities for experimental internet of things research. *IEEE Communications Magazine* 49, 11 (November 2011), 58–67. DOI: <http://dx.doi.org/10.1109/MCOM.2011.6069710>
- [16] Donald Gross, John F. Shortle, James M. Thompson, and Carl M. Harris. 2008. *Fundamentals of Queueing Theory* (4th ed.). Wiley-Interscience, New York, NY, USA.
- [17] HEI Panel on the Health Effects of Traffic-Related Air Pollution. 2010. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. *HEI Special Report* 17 (2010).
- [18] J. Huang, D. Du, Q. Duan, Y. Sun, Y. Yin, T. Zhou, and Y. Zhang. 2014. Modeling and analysis on congestion control in the Internet of Things. In *2014 IEEE International Conference on Communications (ICC)*. 434–439. DOI: <http://dx.doi.org/10.1109/ICC.2014.6883357>
- [19] Gartner Inc. 2016. Gartner's 2016 Hype Cycle for Emerging Technologies. (Aug 2016). <http://www.gartner.com/newsroom/id/3412017>

- [20] R. Jalali, K. El-khatib, and C. McGregor. 2015. Smart city architecture for community level services through the internet of things. In *2015 18th International Conference on Intelligence in Next Generation Networks*. 108–113. DOI : <http://dx.doi.org/10.1109/ICIN.2015.7073815>
- [21] V. Karagiannis, P. Chatzimisios, F. Vázquez-Gallego, and J. Alonso-Zarate. 2015. A Survey on Application Layer Protocols for the Internet of Things. *Transaction on IoT and Cloud Computing (TICC)*, 2015 1, 1 (Jan. 2015).
- [22] Leonard Kleinrock. 1975. *Theory, Volume 1, Queueing Systems*. Wiley-Interscience.
- [23] Steven Latre, Philip Leroux, Tanguy Coenen, Bart Braem, Pieter Ballon, and Piet Demeester. 2016. City of things: An integrated and multi-technology testbed for IoT smart city experiments. In *2016 IEEE International Smart Cities Conference (ISC2)*. 1–8.
- [24] Martina Marjanović, Lea Skorin-Kapov, Krešimir Pripuzić, Aleksandar Antonić, and Ivana Podnar Žarko. 2016. Energy-aware and quality-driven sensor management for green mobile crowd sensing. *Journal of Network and Computer Applications* 59 (2016), 95 – 108. DOI : <http://dx.doi.org/10.1016/j.jnca.2015.06.023>
- [25] Gero Mühl, Helge Parzyjegl, and Matthias Prellwitz. 2015. Analyzing Content-based Publish/Subscribe Systems. In *Proceedings of the 9th ACM International Conference on Distributed Event-Based Systems (DEBS '15)*. ACM, New York, NY, USA, 128–139.
- [26] Rohan Narayana Murty, Geoffrey Mainland, Ian Rose, Atanu Roy Chowdhury, Abhimanyu Gosaint, Josh Berst, , and Matt Welsh. 2008. CitySense: An Urban-Scale Wireless Sensor Network and Testbed. In *2008 IEEE Conference on Technologies for Homeland Security*. 583–588.
- [27] Paolo Neirotti, Alberto De Marco, Anna Corinna Cagliano, Giulio Mangano, and Francesco Scorrano. 2014. Current trends in Smart City initiatives: Some stylised facts. *Cities* 38 (2014), 25 – 36. DOI : <http://dx.doi.org/10.1016/j.cities.2013.12.010>
- [28] T. R. Robbins, D. J. Medeiros, and P. Dum. 2006. Evaluating Arrival Rate Uncertainty in Call Centers. In *Proceedings of the 2006 Winter Simulation Conference*. 2180–2187.
- [29] L. Snchez, V. Gutierrez, J. A. Galache, P. Sotres, J. R. Santana, J. Casanueva, and L. Muoz. 2013. SmartSantander: Experimentation and service provision in the smart city. In *2013 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. 1–6.
- [30] S. Soursos, I. P. arko, P. Zwickl, I. Gojmerac, G. Bianchi, and G. Carrozzo. 2016. Towards the cross-domain interoperability of IoT platforms. In *2016 European Conference on Networks and Communications (EuCNC)*. 398–402. DOI : <http://dx.doi.org/10.1109/EuCNC.2016.7561070>
- [31] Nico Vandaele, Tom Van Woensel, and Aviel Verbruggen. 2000. A queueing based traffic flow model. *Transportation Research Part D: Transport and Environment* 5, 2 (2000), 121 – 135.
- [32] K. Vandikas and V. Tsiatsis. 2014. Performance Evaluation of an IoT Platform. In *2014 Eighth International Conference on Next Generation Mobile Apps, Services and Technologies*. 141–146. DOI : <http://dx.doi.org/10.1109/NGMAST.2014.66>
- [33] H. Xiao and G. Zhang. 2010. The queueing theory application in bank service optimization. In *2010 International Conference on Logistics Systems and Intelligent Management (ICLSIM)*, Vol. 2. 1097–1100.
- [34] ChuanTao Yin, Zhang Xiong, Hui Chen, JingYuan Wang, Daven Cooper, and Bertrand David. 2015. A literature survey on smart cities. *Science China Information Sciences* 58, 10 (2015), 1–18. DOI : <http://dx.doi.org/10.1007/s11432-015-5397-4>
- [35] Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi. 2014. Internet of Things for Smart Cities. *IEEE Internet of Things Journal* 1, 1 (Feb 2014), 22–32.
- [36] Michael Zink, Kyoungwon Suh, Yu Gu, and Jim Kurose. 2009. Characteristics of YouTube network traffic at a campus network fi?! Measurements, models, and implications. *Computer Networks* 53, 4 (2009), 501 – 514. Content Distribution Infrastructures for Community Networks.