SCIENCE CHINA
Information Sciences

• **PROGRESS** •

October 2015, Vol. 58 108101:1–108101:12
doi: 10.1007/s11432-015-5396-5

Big data in smart cities

LI DeRen[1,2], CAO JianJun[1,3] & YAO Yuan[1,2,4*]

[1] *State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China;*
[2] *Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China;*
[3] *Geomatics Center of Gansu Province, Lanzhou 730000, China;*
[4] *National Engineering Research Center for Multimedia Software, Computer School, Wuhan University, Wuhan 430072, China*

Received January 20, 2015; accepted June 10, 2015; published online August 26, 2015

Abstract In this paper, we discuss the concept of the smart city, summarize its development, analyze the motivation and goals of building smart cities in China, and illustrate the supporting technologies of the smart city. Next, we propose a smart city infrastructure that is based on the concept of the digital city and integrates the Internet of Things and cloud computing technologies, thereby achieving automatic control and intelligence services for people and logistics in physical cities. We analyze the big data of smart cities according to ubiquitous sensor networks. Facing the problems and challenges caused by big data of smart cities, we propose a strategy to handle big data that focuses on cloud computing and data mining. In addition, we present a basic framework for cloud computing platforms and propose the establishment of an operating center for smart cities. In closing, we discuss the bright prospects of smart cities.

Keywords smart city, big data, digital city, Internet of Things, cloud computing, data mining, intelligence service

Citation Li D R, Cao J J, Yao Y. Big data in smart cities. Sci China Inf Sci, 2015, 58: 108101(12), doi: 10.1007/s11432-015-5396-5

1 Concept of the smart city

1.1 Concept and connotations of the smart city

The digital city exists in cyberspace. The digital city and the physical city map to each other; accordingly, the digital city is a digital reproduction of the city where we live in an online world [1]. The smart city is built on the basis of the digital city. Using a ubiquitous network of sensors, the smart city is associated with the physical city. A cloud computing platform handles the massive data storage, computation, analysis, and decision-making process, and conducts automated control based on the results of those analyses and decisions [2]. At the level of the smart city, the digital and physical cities can be linked by the Internet of Things, thereby forming an integrated cyber-physical space. In this space, the state and changes in the real world of both people and logistics will be sensed automatically in real time. Cloud computing handles the massive, complex computation and control. This structure will provide intelligence

* Corresponding author (email: whyaoyuan@whu.edu.cn)

services to facilitate human thriving, economic development, and social interactions to establish a low-carbon, green, and sustainable urban structure.

With a simple formula, we can denote the smart city as follows:

Smart City = digital city + Internet of Things + cloud computing.

1.2 Development of the smart city

The development of the smart city can be divided into three stages: an information stage, a digitalization stage, and an intelligence stage. The key events of each stage are as follows.

(1) In September 1993, the United States (U.S.) initiated an "information highway" program. In 1995, China began to promote an "eight gold" project that brought informatization nationwide, marking the beginning of urban informatization in China.

(2) In 1998, U.S. Vice President Al Gore proposed the concept of the "Digital Globe" [3]. The construction of "digitally comfortable communities" marks urban informatization's entry into a new stage of the digital city. More than 300 Chinese cities have built preliminary digital urban infrastructure frameworks. An Internet-based digital map, issued by the National Mapping Geographic Information Bureau, has become a carrier for digital China and digital cities and is used by hundreds of millions of Internet users.

(3) In 2006, next-generation information technologies such as the Internet of Things and cloud computing allowed for the comprehensive integration of information systems and applications within urban information system. In 2008, International Business Machines Corporation (IBM) proposed the new concept of the smart city. In 2009, Sam Palmisano, the Chief Executive Officer (CEO) of IBM, made a proposal to President Barack Obama to promote the establishment of intelligent basic infrastructures. This proposal aimed to overcome the financial crisis and to bring a new driving force to social and economic development. It represents a new era of cities moving from a digital stage to an intelligent stage.

The development of the smart city continues from the construction of the city's early information infrastructure to a digital city infrastructure. However, a smart city pays more attention to the integration of information resources, information sharing, and the integration of other infrastructures and services. It places a greater emphasis on the planning and coordination of municipal management and has timing requirements that are more stringent. It represents an advanced stage of urban informatization and digitalization and is marked by real-time, interactive, and intelligent services. It is also a sign of a high degree of integration between industrialization and information technology.

1.3 Objectives and motivation of building smart cities in China

In China, the goal of building smart cities is to promote low-carbon, green, harmonious, and sustainable development, thereby fulfilling the Chinese dream of the 1.3 billion people of China. China's urbanization and industrialization are still behind those of developed countries, but in terms of informatization, China is comparable to developed countries. Because urbanization and industrialization have become mature and stable in Western countries compared to China, these countries do not have huge and urgent support—in terms of demand, motivation, and capital—for the promotion of smart cities. China currently enjoys an era of great developments in urbanization, agricultural modernization, industrialization, and information technology. More than 270 cites and regions in China has proposed smart city construction planning and entered pilot list by Ministry of Housing and Urban-Rural Development of China. These developments collaboratively and efficiently promote each other. Thus, China has the advantage not only of having started at a late and more established stage but also of having a strong development motivation. National ministries, local governments, and related industries and businesses, all actively involved in building smart cities, have gradually developed from an exploration phase to a designing phase, and finally, to a construction phase.

Table 1 Urban functions and smart city applications

Urban function	Intelligent applications
Reproduction	intelligent public safety\intelligent environment\intelligent energy\intelligent urban management\intelligent retirement\intelligent healthcare\intelligent community\intelligent household...
Economic development	intelligent manufacturing\intelligent industry\intelligent logistics\intelligent city planning...
Social interactions	intelligent public transportation\intelligent shopping\intelligent general social management...
Culture enjoyment	intelligent education\intelligent tourism\intelligent outdoor stream media...

2 Overall architecture and supporting technologies of the smart city

Cities provide functions for human reproduction, economic development, social interaction, and cultural enjoyment. A smart city will provide a variety of intelligent applications and services in these four areas, thereby promoting more coordinated developments between man and society and between man and the nature. The intelligent applications provided by a smart city within the dimensions of reproduction, economic development, social interaction, and cultural enjoyment are shown in Table 1.

A smart city relies on the infrastructure of a digital city. It organizes the locations of people and objects according to geographic location and acquires and transports data and information through the Internet of Things. Massive real-time computation is conducted by cloud computing, and provides a feedback to the control unit using the Internet of Things. The control unit conducts intelligent and automatic control, thereby making the city smart. The overall architecture of a smart city consists of a perception layer for obtaining data, a network layer for transmitting information, a service layer for mass data storage, real-time analysis and processing, and an application layer for end users, as shown in Figure 1. The following sections introduce the three supporting technologies of a smart city, including digital city support, the Internet of Things, and cloud computing.

2.1 Digital city technology

The digital city is an information model of seamless coverage of the entire city. It organizes the city's scattered information of various types based on geographical locations. It not only reflects the inherent links of the city's various types of information (natural, cultural, social, etc.) but also allows for easy retrieval and use by geographical coordinates [4]. Based on geographic location, the digital city consolidates basic geographic data, ortho-photo data, street view image data, panoramic image data, three-dimensional model data, thematic data, and other types of data in the digital city. Through service-oriented architecture, it provides the users with the various types of spatial and attribute data through the network [5]. Different types of users share their information through registration on the network and publish that information in the form of services on the geospatial framework platforms of a digital city. Government, industry, the public, and other users can easily obtain relevant information about matters such as transportation, tourism, health care, education, emergency measures, and other related services [6].

2.2 Internet of Things

Using radio frequency identification (RFID), infrared sensors, global positioning systems (GPS), laser scanners, and other information-sensing devices, the Internet of Things connects things and the Internet based on protocols and conducts information exchange and communication to achieve intelligent identification, positioning, tracking, monitoring, and management. Specifically, the Internet of Things is formed by connecting sensors embedded into electrical networks, railways, bridges, tunnels, roads, buildings, water systems, dams, and oil and gas pipelines. The Internet of Things achieves connectivity between people, between people and machines, and between machines [7]. The World Wireless Research Forum predicts that the future world is a world of ubiquitous networks and that by 2017, there will be 7 trillion

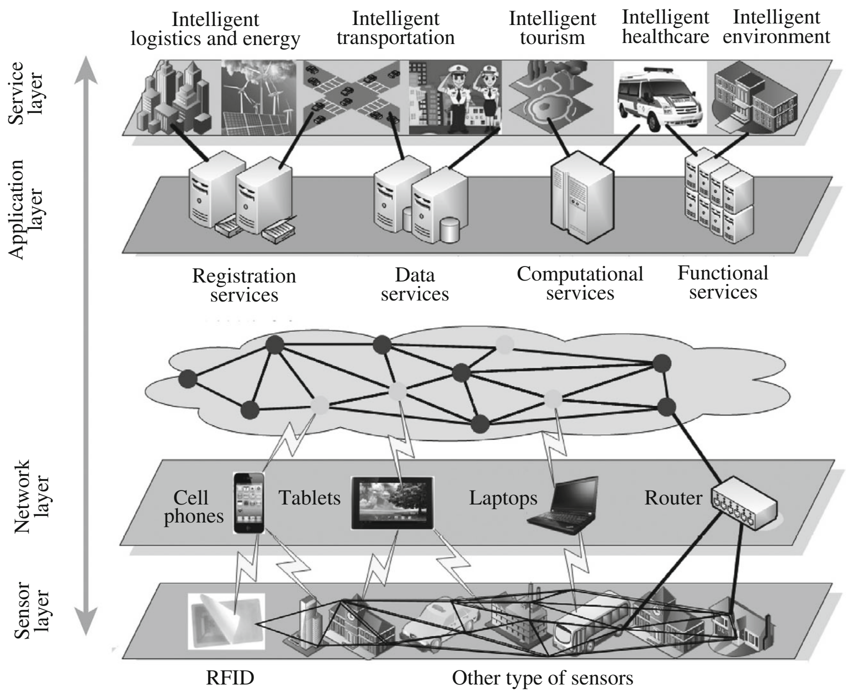

Figure 1 Architecture of the smart city.

sensors to serve the 7 billion people on the planet [8]. These sensors, using various wired and wireless networks, will provide users with fixed-spot, nomadic, and mobile applications and services.

2.3 Cloud computing technology

Cloud computing is an Internet-based computing model for public participation. Its computing resources (including computing power, storage capacity, interactivity, etc.) are dynamic, scalable, virtualized, and provided as a service. Cloud computing is a model of Internet-based computing, an extension of distributed computing and grid computing [9]. Cloud computing supports a cloud computing center which is socialized, integrated, and specialized for information services, to optimize service processes and reconstruction through the reuse and flexible reorganization of software, thereby improving utilization. Cloud computing promotes resource integration, information sharing, and collaboration among software; it also carries out service-oriented computing. Cloud computing can rapidly process the huge amounts of data produced around the world and simultaneously service millions of users [10].

3 Big data in smart cities

In September 2008, Nature published a special issue titled “Big Data” [11]. In February 2011, *Science* published a special issue titled “Dealing with Data” [12]. These issues show that the era of big data has arrived. In March 2012, the Obama administration officially announced and launched a “Big Data Research and Development Initiative”. This initiative considers big data to be the “oil” of the future world. The meaning of this initiative is comparable to 20th century’s information highway plan. Scientists and politicians have come to realize that big data will be a treasure for mining information and knowledge. Alongside the gradual construction and implementation of smart cities, human beings and various types of sensors will produce more and more data. There will be an increasing amount of data deposited, and the scale of data will gradually increase from the current Gigabyte (GB) and Terabyte (TB) magnitude to Petabyte (PB) or even Exabyte (EB). If we can thoroughly analyze these data, which have complex structures and are large in size, and integrate the analysis through cloud computing, we will be able to quickly convert the data into valuable information and to extract the laws of changes in nature and society. We are confronted by an exciting era of big data. An era to construct predictive models from valid and large-scale data, to visualize data, and to discover new laws is about to begin.

3.1 The increasing amount of big data

The smart city links the city with a digital city through the ubiquitous Internet of Things. Around the world, an increasing amount of big data is being produced, the total amount of which is beyond imagination. At all times, people exchange ideas, data, and information on the Web. Within a minute, Google receives two million search queries, Facebook has 680000 new posts, and more than 200 million e-mails are sent. Every day, Baidu processes 6 billion search queries, expands by 10 TB, processes more than 100 PB of data, and produces 1 TB of logs. Currently, the total number of Internet pages is nearly 1 trillion, and the amount of data is close to 1000 PB. Tencent QQ has more than 800 million monthly active users, more than 500 million Wechat users, and more than 100 billion online relationships in a chain. Even after compression, the total amount of data is 100 PB. Every day, the Internet receives hundreds of billions of time queries and adds more than 200 TB of data. Two- and three-dimensional data in the space of digital Earth grow rapidly and will reach the TB to PB level. A single high-definition (HD) webcam produces 3.6 GB of data per hour. Because the number of webcams throughout China is more than 20 million, the amount of data produced will reach a PB to an EB level. Aircrafts are equipped with a large number of sensors. Each engine generates 20 TB of data per flight hour. A flight from London to New York City will generate 640 TB of data. These engine status data are submitted by satellites to engine manufacturers for monitoring. According to a research report of Symantec Inc in 2012, the total enterprise data storage around the world reached 2.2 Zettabyte (ZB), with an annual growth rate of 67%. In Beijing, the public transportation card system receives approximately 40 million uses on a daily basis; and the subway system receives 1 million uses. The data of the Beijing traffic control center increases by 30 GB every day, and the total amount is approximately 20 TB. The China National Grid generates approximately 510 TB data on average each year (excluding video). The cumulative data generated is approximately 5 PB. Computed tomography (CT) images of a single patient are often up to two thousand images. The amount of data reaches dozens of GB. Currently, hospitals in large Chinese cities receive tens of thousands of outpatients every day. In a year, the total outpatient number in China is approximately several billion, with 200 million people being accepted into hospital wards. According to the relevant healthcare system regulations, patient data often must be kept for more than 50 years. Big medical data will also reach the EB level.

3.2 Challenges of big data in smart cities

In the smart city, the Internet of Things will continue to collect a vast amount of data using a vast number of sensors. Big data must undergo storage, processing, post queries, and analysis to be used for all types of intelligent applications and services. Moreover, there is an increasing need to conduct real-time storage, processing, query, and analysis for big data. This will bring a series of problems and challenges.

(1) The high cost of big data storage. The rate at which the data storage costs decrease, brought about by the development of storage technology, falls short of the growth rate of the data. Currently, cost is a challenge to all who store big data in accordance with the ideal standard for the storage and preservation of large data. For example, the Tianjin municipal security system will require a budget of 50 billion yuan to construct a storage capacity of 4.6 EB, which is ideal. This cost is equivalent to Tibet's annual GDP in 2012. The rapid growth in data size and the corresponding high cost have become important factors that restrict the development of urban security and other systems. At present, most of China's cities opt for a shorter data retention time and reduced quality of data storage to reduce construction costs. However, this would severely reduce the value of video data in terms of traceability and identification.

(2) Low degrees of automation in quick queries and the retrieval of big data. Traditional information systems only conduct simple data acquisition and storage and lack effective automatic retrieval and an analysis of key semantic information, such as behavior. In the era of big data, the rapid expansion

of large-scale data, such as spatial information and video further exposes the limitations of traditional methods. Different types of remote sensing observation satellite send PBs of data to the ground every day. Municipal video capture systems collect EB-level data every day. Currently, there remain difficulties in conducting a complex automatic semantic analysis and interpretation. In addition, municipal big data contains data related to important abnormal behaviors, events, and characteristics, such as those of falling, scuffles, climbing walls, hovering, vehicle collisions, and retrograde. Big data can be used for identification and warning. For example, major robberies are often accompanied by prior wanderings.

By efficiently using automation technologies for object and behavior identification, it is possible to retrieve semantic information from municipal big data, conduct prior warnings, and effectively deter criminal activities. This will allow for preventive actions, information retrieval during the course of an action, and post-event measurements. Ultimately, this will ensure a full range of protection of people's life, and property, as well as daily production during incidents.

(3) Mining for knowledge from big data is very difficult. Big data contains not only data and information but also a wealth of rules and knowledge. Such rules and knowledge are not given directly. Instead, to obtain them one must conduct in-depth data mining and analysis. However, the major characteristics of big data make it difficult for effective integration and management as wellas automatically process and analyze. It is particularly difficult to conduct data mining on a dataset containing spatial information. To mine rules and knowledge from big data, in addition to solving issues related to data heterogeneity and retrieval, in-depth data mining also must address a series of issues such as data selection, semantic description, semantic interpretation, uncertainty, knowledge representation, etc. [13]. Currently effective and feasible techniques for data selection, semantic description, and semantic interpretation cannot be directly applied to big data. Accordingly, the rules and knowledge related to big data cannot be fully utilized. The nine issues for sustainable human development, including human health, energy, weather forecasting, climate change, disaster emergency forecasting, water resources, sustainable agriculture, the environment, and biodiversity, have not been effectively addressed.

4 Cloud platform-based big data services

There are a large number and various types of big data in a smart city, especially related to spatial information, video, or other types of unstructured data. We should actively address the challenges of big data. We must take advantage of cloud computing and focus on data mining theory to effectively store and manage big data, to quickly retrieve and process data information, to mine information and knowledge from big data, and to make full use of its value.

With the support of cloud computing and data-mining technology, intelligence services will switch a model of service from one in which anyone, at any time, anywhere can obtain any information (4A) to a service that at the right time, in the right place, the right information is delivered to the right person (4R). Consequently, a city can achieve truly smart status.

4.1 Big data storage

Cloud-based storage can fully virtualize storage. All devices are completely transparent to the users of the cloud. Any authorized user can connect to cloud storage via the network, thereby allowing a user to have storage equivalent to the entire storage capacity of the cloud. This feature overcomes the performance and capacity bottlenecks of traditional storage and achieves linear expansion of performance and capacity. Intelligent compression methods are designed to achieve the redundancy of big data at an integrated level and extend the coding layer from the traditional signal and feature coding to a semantic coding layer. For big video data, the system will retrieve and express information about forecasts, backgrounds, global objects, and motion estimation from highly efficient video, as estimated based on global motion in the space of the video objects. The system will also retrieve characteristics of the global redundancy object, conduct semantic mapping, and encode on the semantic layer. This will also provide good support for later search and analysis. Combining cloud storage with intelligent compression, storage costs can be

Figure 2 Automatically extracting actions such as climbing the wall and running.

significantly reduced. Accordingly, it will be possible to effectively store all types of big data belonging to smart cities.

4.2 Retrieval and processing of big data

After solving the basic problem regarding storage, problems regarding application orientation, efficiency, and data retrieval speed must be addressed. After retrieving the effective datasets, the datasets will need to be processed based on the need for different types of intelligent applications. The following illustrate these issues using retrieval, remote sensing image processing, and location services as examples.

(1) The retrieval cloud. Supported by elastic computing power, a retrieval cloud service can not only automatically extract features from an image or a video but also extract and index activities such as climbing walls, running, trailing, gathering, wandering, and other behaviors, as shown in Figure 2.

For video retrieval, the process first divides the video into various frames, then extracts features from each frame to obtain a feature space that reflects the content of the view. This feature space serves as the basis for video clustering and retrieval. Feature extraction from key frames includes extracting visual features such as the color, texture, shape, and motion features of a frame. Extracting the motion feature is accomplished by first analyzing motions in a frame and then indexing the frame's motion features and static features of key frames. An end user simply selects the behavior and location of interest. For example, with respect to running behavior in a community, one can quickly retrieve location, key frames and video information relevant to all running objects.

(2) The remote sensing cloud. Taking advantage of the scalability of the cloud computing platform, the remote sensing cloud is a sharing and automatic control mechanism of computing resources, network resources, and storage resources. It effectively organizes the resources distributed among different geographic locations, including geographical space-land sensor resources, spatial data resources, processing algorithms, software resources, geoscience knowledge resources, and modular workflow resources. These resources are released through a registration service center. Additionally, a remote sensing cloud provides geo-spatial information for the entire community through a variety of network visualization services. MapWorld, a national geographic information public service platform constructed by the National Administration of Surveying, Mapping and Geoinformation, provides one-stop data resource services through the portal and presents basic geospatial data such as an electronic map, image, and terrain of all of China's provinces and cities. The data can be extended to three-dimensional city data, rainfall data, and other types of data. The OpenRS-Cloud visualizes the registered data and algorithms into a Web desktop under the support of a cloud computing service. Accordingly, users do not need to build a professional platform. Once the users have selected remote sensing data and algorithms, the cloud automatically allocates computing resources and quickly conducts relevant processing and analysis. An integrated space-sky-ground observation sensor network for ground observation is under construction. It

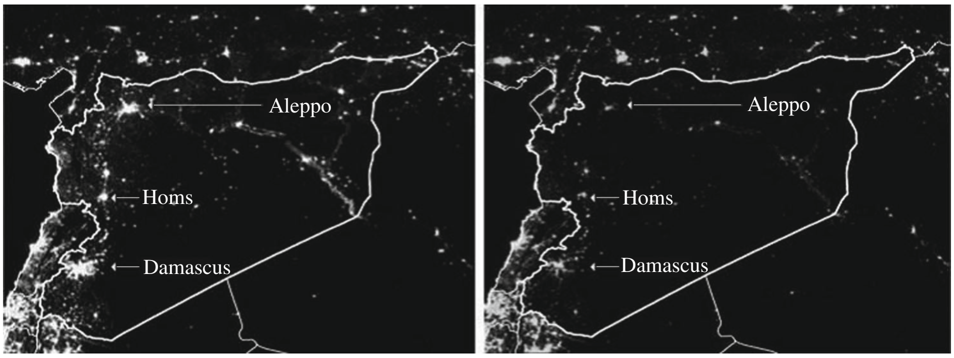

Figure 3 The night-time light monthly composites. (a) March 2011; (b) February 2014.

will allow global, all-weather, all-time, comprehensive, spatial data to be obtained. This system will serve as a basis in the remote sensing cloud for obtaining geospatial data, delivering rapid responses, warning of various types of major issues such as disasters and resource security, and handling the important problems related to global sustainable development. The International Organization for Earth Observation considers network-based satellite constellation observation as a core plan for the next 10 years. This observation played significant roles in monitoring major disasters such as floods in Africa in 2007 and in Myanmar in 2008. Night-time light images from satellite could be used for evaluating the ongoing crisis in Syria since reliable witness reports are hard to gather in a war zone [14]. Contrast light images of the war region in different time are shown in Figure 3.

(3) The location cloud. Due to various errors, the positioning accuracy of the global satellite navigation system cannot fulfill the requirements of many professional fields. To improve positioning accuracy, a Continuously Operating Reference Station System (CORS) has been developed. Now, users can transmit the satellite positioning information to a cloud service center, and the cloud services can calculate the location of the object on a sub-meter scale within 1 s. Through the enhanced services of the terrestrial base station system, it is possible to achieve high precision, meter-scale navigation by satellite positioning systems such as Beidou. In out-of-range areas not covered by the satellite signal, such as indoor and underground space, it is possible to use accelerometers, gyroscopes, electronic compasses, cameras, sensors and Wireless Fidelity (WiFi), wireless communication networks, Bluetooth, and other wireless signals to conduct high-precision indoor and outdoor continuous positioning. These methods will fulfill all types of needs by government sectors, industries, and the public, such as those involved in surveillance; exploration and surveys of the environment, forests, and other areas; municipal management; and public security. Figure 4 shows a multi-sensor multi-network location service.

4.3 Data mining for big data

Gray proposes four paradigms of scientific development [15]. Thousands of years ago, science used an empirical method. A few hundred years ago, science was theoretical field, a process of creating and testing a hypothesis. In recent decades, science has relied on calculation, conducting simulation and verification by computation. Currently, science involves the exploration and mining of scientific data and using data mining to unify theory, simulation, and experimental verification. Data mining has become an important means of conducting scientific discovery and exploration from big data and an effective way to solve major issues and to meet real needs. By mining the big data of a smart city, it is possible to explore and discover laws of natural and societal changes, including people's lives, behavior, and preferences; social trends; and trends in thinking and public opinion. This development will allow an inference of the market's responses to products and services and even to policies and various other aspects. The following sections provide a detailed analysis of the big-data mining process and data mining related to space and attributes.

(1) The process of data mining. Big-data mining needs to extend the entire process to the front and back ends, in accordance with the characteristics of big data. This particular process includes the

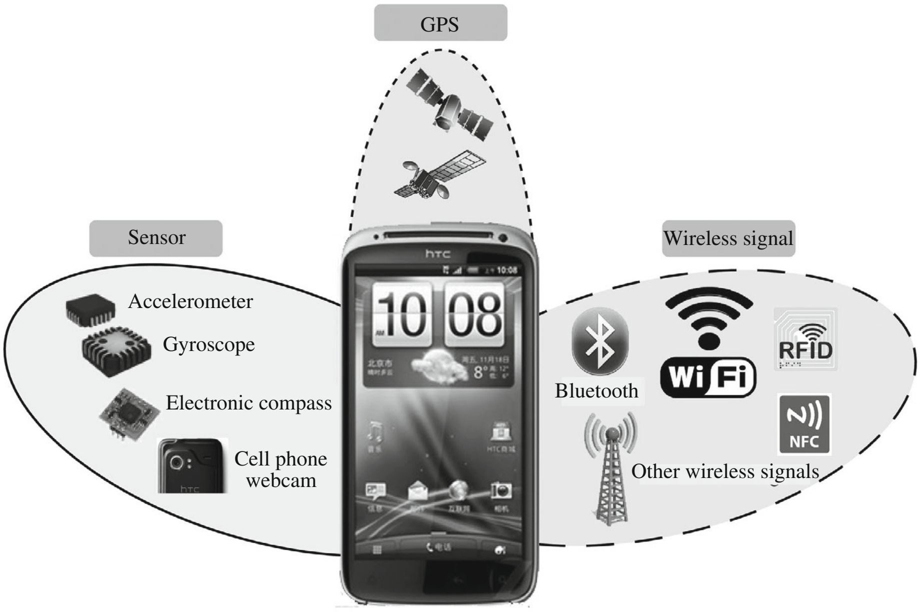

Figure 4 Location service based on multi-sensor and multi-network.

processing and analysis of mass and multisource data, automatically discovering and extracting implicit, nonobvious patterns, rules, and knowledge and visualizing and integrating them into a presentation that is easily understandable by humans. The specific procedure is shown in Figure 5. Data mining, data retrieval, processing, and information extraction are relatively more difficult to conduct. These processes need the support of theory and technology relevant to knowledge reasoning based on big data and the knowledge warehouse.

The data-mining process first obtains and stores data. In accordance with the demand for data mining, data are collected, retrieved, and integrated into big data, and data selection is performed, which includes de-noising, sampling, filtering, merging, and standardizing. By removing redundant and unnecessary data, a data set waiting to be processed is established. Next, the data are processed and analyzed, which includes linear analysis, nonlinear analysis, factor analysis, sequence analysis, linear regression, variable curves, bivariate statistics, etc. Based on the specific method used to classify the data, the relationships between data and between data types are analyzed. The intrinsic link between data is revealed by artificial neural networks, decision trees, genetic algorithms, and other methods. This will allow for the identification of in-depth patterns, rules, and knowledge. Based on the relationships between attributes, the data mining process presents an analysis of relationships between attributes from the discovery of these patterns, rules, and knowledge through means that are visually understandable by humans. The data-mining process can also fuse different types of content that have a certain association to achieve a more intuitive display for human analysis and utilization.

(2) Mining of data with integrated space and attributes. Spatial data is relevant to the object's geographic and spatial distribution. It is the record that reflects real-world phenomena and changes. Compared to general data, spatial data have spatial; temporal; multidimensional; massive temporal; multidimensional massive; and complex characteristics, along with spatial uncertainty [16]. Feature attribute data are non-spatial attribute data of an object. They may have either discrete or continuous values. Data mining based on spatial data and attribute data refers to a process that extracts spatial patterns and features. Relationships exist among spatial and non-spatial data and other implicit features, rules, knowledge of the user's interests from the spatial and feature data, and relevant information. Data mining based on spatial data and attribute data is highly relevant to the data-mining method and the type of knowledge being mined. The chosen spatial data-mining theory and method will directly affect the quality of the discovered knowledge. The main theories and methods of spatial data mining include probability theory, evidence theory, spatial statistics, rule induction, cluster analysis, spatial

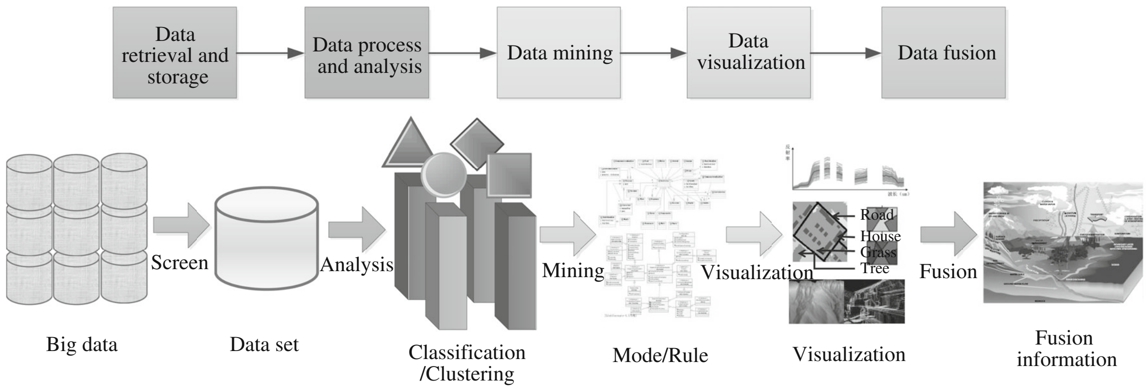

Figure 5 Data mining process of big data.

analysis, fuzzy sets, cloud models, data fields, rough sets, neural networks, genetic algorithms, decision trees, etc. These methods are not individually isolated. To locate certain types of knowledge, it is often required to integrate these methods and choose data mining theory, methods, and tools based on specific needs. By taking advantage of automatic learning techniques such as machine learning and artificial intelligence techniques, the degree of automation is increased and the degree of human-computer interaction is reduced.

5 Smart City Operation Center

5.1 A smart city needs an operation center

After establishing the Internet of Things, cloud computing platforms, the digital city geospatial framework, and other related infrastructure in a city, for that infrastructure to better play its full role in the city's operation, there is a need for a Smart City Operation Center (SCOC). The SCOC is an institute that integrates all types of real-time data, information, and services. As the "heart" of a smart city, the SCOC collects and monitors a full range of data and information, generated by the city's operation. It then provides customized services to the government, businesses, and individuals.

From the perspectives of planning and design, in building the SCOC, a city should first conduct a top-level smart-city design. This includes the development of relevant policies, regulations, and standards; the standardization of goals, frameworks, tasks, operation, and management of information technology development in various sectors; and unified norms and standards during the top-level design. In terms of urban operation, a city should integrate and share information resources relevant to the city, conduct real-time monitoring of operations, and achieve multisectoral coordination and command. In terms of servicing the entire society, a city should fully promote access to and the utilization of big data by the entire society, along with the formation of a service and trading system. SCOC building will achieve the goal of forming a complete application and service system that includes public service and commercial deep-processing based on big data.

5.2 The Composition of the SCOC

An SCOC generally consists of four parts: a big data center; a municipal center of operations, monitoring, and command; a municipal center of IT infrastructure maintenance and a center of intelligence services. The big data center will become a hub between the resource pool of city operation and the Internet of Things. It will allow for comprehensive and real-time awareness of data relevant to city operations. The center of operation, monitoring, and command will be based on comprehensive and real-time awareness and carries out cross-sectoral, cross-regional, and cross-system collaboration, along with efficient emergency responses. The center of IT infrastructure maintenance is responsible for maintaining the updating of the SCOC's infrastructures and ensuring the secure and stable operation of the SCOC at all times.

The center of intelligence services provides services not only to the government but also to all types of enterprises and to the public.

The traditional government information technology (IT) infrastructure will gradually be replaced by the "cloud" and the interactive smart city, which will reduce the costs of establishing and maintaining the city's IT, minimize administrative costs, and enhance operational efficiency. Through the SCOC, the mode of governance will be truly switched from city management to urban operations and services.

5.3 Functions of the SCOC

Through the SCOC, a city can provide operations and services. The city can also use big data as the basis for governance and major decisions. Based on data and empirical facts, the influence of subjective individual opinions and various business interest groups can be avoided.

City operation will achieve visibility, controllability, intelligence, predictability, quantifiable assessment, and continuous optimization. Consequently, the government will become more open, accountable, and efficient, thus minimizing the administrative risks. A study by McKinsey shows that after European public administrations adopt big data applications, the applications' potential value each year reaches $250 billion [17]. By adopting big data, enterprises can also reorganize production resources and optimize business models, thereby obtaining greater profit. For example, when a bank chooses a branch location, it considers not only the location, population distribution, consumption level and structure, aggregation effect, traffic, and other factors but also the feature data of its current operational status. By conducting sufficient data mining into spatial data and attributing data in big data, a bank can plan and locate branches of different levels based on its locations. Various types of intelligence services for the public will cover the major events of a lifetime, including birth, health, education, employment, marriage, child care, pension, and funeral, thereby improving urban residents' well-being.

In the US, 5 major cities, including the New York City, have opened their municipal databases and massive amounts of related information to the public. All businesses, organizations, and individuals can perform data mining from publicly available big data for information and knowledge relevant to their needs. Such cities also provide a wide range of intelligent services. These measures improve not only their competitiveness but also the city's overall competitiveness. The ultimate goals of releasing big data is to attract more investors and tourists and to provide more friendly services, thereby better promoting the city's development and prosperity.

6 Conclusion and prospects

The smart city is based on an integration of the digital city, the Internet of Things, and cloud computing. It realizes perception, control, and intelligence services to humans and has substance. The smart city has wide potential for economic restructuring and development, urban intelligence management, and intelligent services for the general public, making people and nature develop in a manner that is more coordinated.

Constructing a smart city is a systematic project, which is conducted step by step: each city performs top-level design according to its own characteristics and then conducts unified planning. To realize a smart city, we must build a more complete information infrastructure and a technical support infrastructure that includes the SCOC to ensure that the various smart city applications are both functional and affordable. The big data problem arising from the construction of a smart city is not only a cutting-edge problem for next-generation scientific research but also a motive that drives the development of a smart city. It brings new opportunities and challenges and urges us to expedite technology innovation and research related to big data, thereby promoting and accelerating the development of the intelligence service industry, better implementing the various intelligence applications in a smart city, allowing for more and better intelligent applications to serve the entire society, and making city operations more scientific, efficient, low-carbon and safe.

Acknowledgements
This work was supported by National Natural Science Foundation of China (Grant No. 71503099), Fundamental Research Funds for the Central Universities, Natural Science Foundation of Hubei Province of China (Grant No. 2015CFB325), and Geographic Conditions Monitoring Project of China.

References

1 Li D R, Yao Y, Shao Z F, et al. From digital Earth to smart Earth. Chin Sci Bull, 2014, 59: 722–733
2 Li D R, Shan J, Shao Z F, et al. Geomatics for smart cities-concept, key techniques, and applications. Geosp Inf Sci, 2013, 16: 13–24
3 Gore A. The digital Earth: understanding our planet in the 21st century. Photogramm Eng Remote Sens, 1999, 65: 528
4 Shao Z F, Li D R. Image city sharing platform and its typical applications. Sci China Inf Sci, 2011, 54: 1738–1746
5 Li D R, Shao Z F. The new era for geo-information. Sci China Ser-F: Inf Sci, 2009, 52: 1233–1242
6 Li D R, Yao Y, Shao Z F. The new mission for surveying, mapping and geographic information science in the smart Earth era. Sci Surv Mapp, 2012, 37: 5–8
7 ITU. ITU Internet Reports 2005: the Internet of things. ITU, Tunis, 2005
8 Uusitalo M A. Global visions for the future wireless world from the WWRF. IEEE Veh Technol Mag, 2006, 1: 4–8
9 Li D R. Cloud computing supports sociality, intensiveness and specialization of information service. J Chongqing Univ Posts Telecommun Nat Sci Ed, 2010, 22: 698–702
10 Barroso L A, Dean J, Hölzle U. Web search for a planet: the Google cluster architecture. IEEE Micro, 2003, 23: 22–28
11 Howe D, Costanzo M, Fey P, et al. Big data: the future of biocuration. Nature, 2008, 455: 47–50
12 Reichman O, Jones M, Schildhauer M. Challenges and opportunities of open data in ecology. Science, 2011, 331: 703–705
13 Li D R, Wang S L, Shi W Z, et al. On spatial data mining and knowledge discovery. Geomat Inf Sci Wuhan Univ, 2001, 26: 491–499
14 Li X, Li D R. Can night-time light images play a role in evaluating the Syrian Crisis? Int J Remote Sens, 2014, 35: 6648–6661
15 Hey T, Tansley S, Tolle K. The Fourth Paradigm: Data-intensive Scientific Discovery. Washington: Microsoft Research, 2009
16 Li D R, Wang S, Li D Y. Spatial Data Mining Theories and Applications. Beijing: Science Press, 2006
17 Manyika J, Chui M, Brown B, et al. Big data: the next frontier for innovation, competition, and productivity. The McKinsey Global Institute, 2011