Toward an Intelligent Traffic Management Based on Big Data for Smart City

Yassine Karouani(✉) and Ziyati Elhoussaine(✉)

RITM Laboratory, Computer Science and Networks Team,
ENSEM-ESTC-UH2C, Casablanca, Morocco
ykarouani@gmail.com, ziyati@gmail.com

Abstract. It is anticipated that the Smart City research initiative will create new breakthroughs to revolutionize transportation system operations, infrastructure design, construction and management, as Big Data progresses. This latter will focus on the modeling, analysis and optimization of data-intensive intelligent transport systems, which will allow for more efficient system-wide operations. The focus is on the use of non-traditional data generated by smart city initiatives and emerging mobile applications, including data from social media, smart phones and more generally all connected objects. Research on this subject allows us to have a global view on the studies carried out in this field not on the infrastructure side but control and management of road traffic, based on the main objectives according to the users of the road. These objectives are the elaboration of a shortest path between a source and a destination, as well as the time required to traverse this path. We study different existing solutions such as solution employed by: Google, Japan (VICS, PCS) trying to find the advantage, the weak points and the common points to better bring out a new model which gathers the maximum advantages of these methods.

Keywords: Smart city · Big Data · Mongo DB · Traffic road · Traffic congestion Vehicle routing

1 Introduction

Recently the entire world and in particular Morocco, everyone is looking for solution to improve traffic management and congestion. People for the sake of convenience use the automobile, and road traffic demand is on the increase. Traffic jam and an increase in traffic accidents have been the result, and the original objective of automobile transportation reaching one's destination with safety, comfort, confidence and speed has become more and more difficult to achieve. This problem is widely recognized. Nevertheless, traffic control is becoming more sophisticated thanks to new roads, road improvement, traffic signals and traffic control systems, and traffic safety education as administrative measure. The fact remains, however, that traffic continues to System increase, making road traffic more difficult and unpleasant. The objective of this document is to propose a solution and to highlight methods with the aim of improving the management of road traffic by respecting requirements for optimization and efficiency

M. Ben Ahmed and A. A. Boudhir (Eds.): SCAMS 2017, LNNS 37, pp. 502–514, 2018.
https://doi.org/10.1007/978-3-319-74500-8_47

in the functioning of social systems, including systems more sophisticated transportation, CO2 emissions, and modernization of aging infrastructure. Today we have a lot of tools that can help us to collect information on urban activities such as people and traffic flows.

For example, smartphones, smart cards, and cameras, without forgetting the Internet. While information on these flows has been provided in the past only in the form of general public statistics. Currently the availability of these data extends to the sectors that use it; the amount of data received and collected is growing exponentially and the leaps in storage and computational power within the last decade underpin the unique selling proposition of Big Data of being able to provide better insights into various business processes or everyday life in a way that was not possible in the past.

The role of Big Data here is not just for massive data storage, but the most important is to analyze large datasets and discover the relationships between structured and unstructured datasets.

This article touches on two themes:

- Foundations and principles, discussions on some solutions for the management of existing road traffic.
- Applications, providing a variety of use cases of data analysis in the field of traffic management based on "Outsourcing", and GPS for data collection and large data processing side, analysis.

2 Related Works

In paper [1], authors implement a distributed system for solving the problem of garbage collection. The proposed solution tries to create a decentralized system for transportation and collection of garbage. An improved Ant Colony Optimization was implemented to create the best shortest routes that visit all collections points such as each collection point is visited once.

Authors in paper [2] review the applications of big data to support smart cities. It discusses and compares different definitions of the smart city and big data and explores the opportunities, challenges and benefits of incorporating big data applications for smart cities. In addition it attempts to identify the requirements that support the implementation of big data applications for smart city services. The review reveals that several opportunities are available for utilizing big data in smart cities; however, there are still many issues and challenges to be addressed to achieve better utilization of this technology.

Work [21] compare three evaluation methods that have been applied in their field tests in the past few years. The shortcoming and advantages of each method are presented. An improvement of Single Vehicle Test method is proposed to increase the reliability of reference value without jeopardizing the advantages of Single Vehicle Test. The comparison of old and new Single Vehicle Test method in their field tests shows the improvement can effectively increase the reliability of reference value, therefore increase the reliability of evaluation result.

A Smart and intelligent transportation is proposed in [22] using the graph technology, which analyzes the real time transport data using the parallel environment of Hadoop ecosystem. Based on the analysis, the system performs real time decisions related to any transport problem. The proposed system consists of various layers starts from capturing layer that collects the vehicular network data from the external vehicular network or road sensors. The whole system is implemented using apache Spark on the top of the Hadoop ecosystem in order to perform real time analysis and decision-making. For graph processing, Giraph, with the capability of parallel processing, is used, which dramatically increase the performance of the system.

3 Advanced Models and Architectures

The safety and comfort of motorists is a priority. That is why all stakeholders (state, local authorities, engineering offices, etc.) are mobilizing to optimize the use of infrastructures, that's why exist different solution and proposals such us:

3.1 Google Traffic

Early versions of Google Maps provided information to users about how long it would take to travel a particular road in heavy traffic conditions. Traffic information was based on historical traffic data and was not particularly accurate.

In this time Google offering traffic data based on information gathered anonymously from cellular phone users.

Google Traffic works by analyzing the data coming from Crowdsourced traffic data that's include a raw data of GPS determined locations transmitted by a large number of mobile phone users. By calculating the speed of users along a length of road, Google is able to generate a live traffic map (Fig. 1).

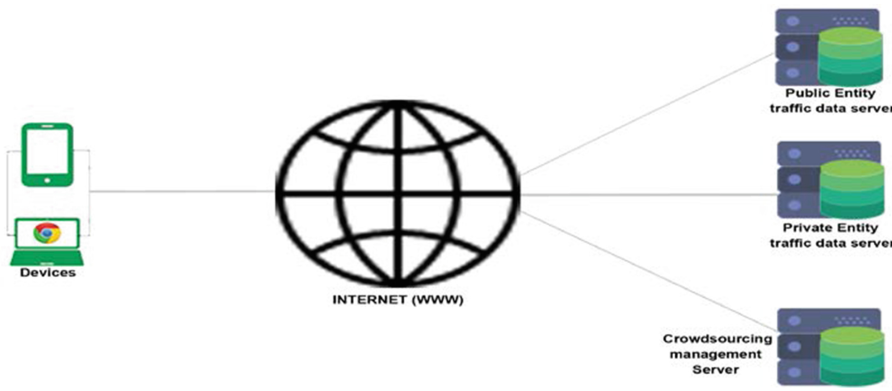

Fig. 1. Collect traffic information for Google.

Google stated: "When we combine your speed with the speed of other phones on the road, across thousands of phones moving around a city at any given time, we can get a pretty good picture of live traffic conditions" [23].

Current and predicted traffic information is provided from incident data, traffic flow data, and media related to traffic received from multiple sources. The Crowdsourced data may be provided passively by applications on remote mobile devices or actively by users operating the remote mobile devices. An application on a mobile device may receive the multiple data types, aggregate and validate the data, and provides traffic information for a user. The traffic information may relate to the current position and route of the user or a future route. The present technology may also provide driving efficiency information such as fuel consumption data, carbon footprint data, and a driving rating for a user associated with a vehicle.

Figure 1 is a block diagram of an exemplary system for analyzing traffic data. System includes devices, network (internet), public entity traffic data server, private entity traffic data server, and crowd sourcing management server. Mobile device may communicate with network and be operated by a user in a moving vehicle. Mobile device may include circuitry, logic, software, and other components for determining a position, speed, and acceleration of the mobile device. In some embodiments, mobile device may include a global positioning system (GPS) mechanism for determining the location of the device. Based on the location data, the speed and acceleration of the device may be determined.

Network may communicate with devices, entity public and private data servers and crowdsourcing management server. Network may include a private network, public network, local area network, wide area network, the Internet, an intranet, and a combination of these networks.

Public entity traffic data server may include one or more servers, including one or more network servers, web servers, applications servers and database servers that provide traffic data by a public sector organization. The public sector organization may be, for example, the Department of Transportation or some other public entity. Public entity traffic data server may provide traffic data such as incident data for a planned or unplanned traffic incident, traffic speed and flow data, or traffic camera image and video data. A planned traffic incident may include data for a highway closure or construction work. An unplanned traffic incident may include data for a disabled vehicle or a car crash. The traffic speed and flow data may be determined from radar outposts, tollbooth data collection, or other data. Public sector traffic cams located on roadways may collect the traffic camera image and video.

Private entity traffic data server may include one or more servers, including one or more network servers, web servers, applications servers and database servers that provide traffic data such as incident data, traffic speed and flow data and traffic camera and image data. Examples of private entity traffic data servers are those provided by companies such as Inrix, Traffic Cast, Clear Channel, and Traffic.com.

Crowdsourcing management server may include one or more servers, including one or more network servers, web servers, applications servers and database servers, that receive crowd source data from devices, aggregate and organize the data, and provide traffic data to a device application on any of mobile devices. Data is received from a plurality of remote mobile device applications regarding current traffic information. The traffic information is aggregated to create a unified set of data and broadcast to mobile devices to which the data is relevant.

3.2 VICS: Vehicle Information and Communication System

The VICS [3–7] is a telecommunication system that transmits information such as congestion and regulation of traffic by detecting car movements with sensor installed on the road, this information is edited and processed by Vehicle Information and Communication System Center, and shown on the navigation screen by text or graphical form such as car-navigation system installed in individual cars (Fig. 2).

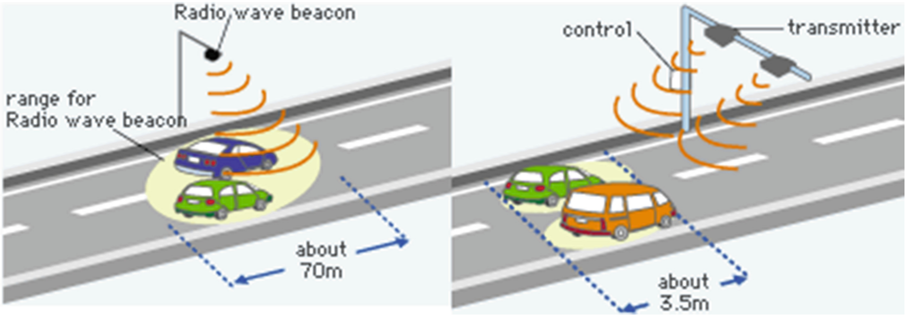

Fig. 2. Vehicle information and communication system.

First, a road traffic information collection system has already been built by the public sector. The information for the new information communications systems is already in existence and waiting for effective utilization.

Second, on board intelligence of vehicles is improved and popularized so that it can use the new information technologies. Already, approximately 400,000 vehicles have navigation systems installed, and this is a major factor in the successful building of a nation wide digital road map database.

Third, development of a mobile communications system, which provides information for such vehicle equipment, is near completion. Tests of the Advanced Mobile Traffic Information and Communication system (AMTICS), the Road and Automobile Communication System (RACS), and of FM multiplex broadcasting by the Telecommunications Technology Council indicate a promising future for these technologies.

3.3 PCS: Probe Car System

Acquisition of road traffic data is an important aspect of Road traffic Informatics systems. An innovative approach is utilizing the vehicles themselves as a source of real-time traffic data, functioning as roving traffic probes. This principle was probably first touched upon in the early 1970s in the course of a pilot with the Japanese Comprehensive Automobile Traffic Control System (CACS) [3–7], which was a government sponsored program aimed at the development of a route guidance and traffic information system. Besides inductive loops to transmit guidance information to CACS-equipped vehicles, also employment of the vehicles themselves to collect traffic data was propounded [8]. The traditional collection methods using roadside sensors are necessary but not sufficient because of their limited coverage and expensive costs of implementation and maintenance. To solve the limit of roadside sensors, the idea of collecting real time traffic data from in vehicle devices through mobile phones or GPS is quite popular known as probe car (floating car). Raw data of geolocalisation sent anonymously to a central processing center. After being collected and analyzed, useful information can be redistributed to the drivers on the road with real time (Fig. 3).

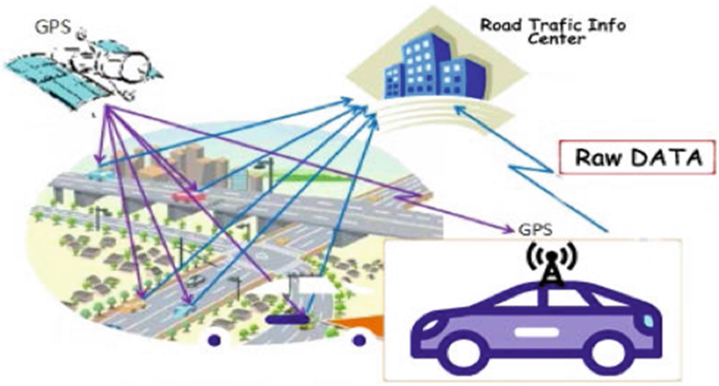

Fig. 3. Probe car system.

PCS [3–7] provides information on car movements and that on forecasting traffic congestion to individual drivers the same as VICS does. Different from VICS, this system collects traffic information from all cars, which are considered to be movable sensor units. Each car has a telecommunication unit and transmits several kinds of information such as position, velocity, and the status of the car to the central server.

4 An Advanced Model for Traffic Road Congestion in Smart City

Generally, the improvement of the roads is done by the installation of road equipment (signaling, radars, etc.) that can be expensive and long to set up [9].

Fortunately, the growth of intelligent transport systems (ITS) such as connected vehicles, driving assistants (GPS, radar warning devices, etc.) and smartphones has revolutionized traffic management.

4.1 System Model

The use of the Floating Car Data makes it possible to respond to the various issues of road safety and the reduction of congestion. In the Fig. 4 we explain our model for traffic congestion control in the smart city.

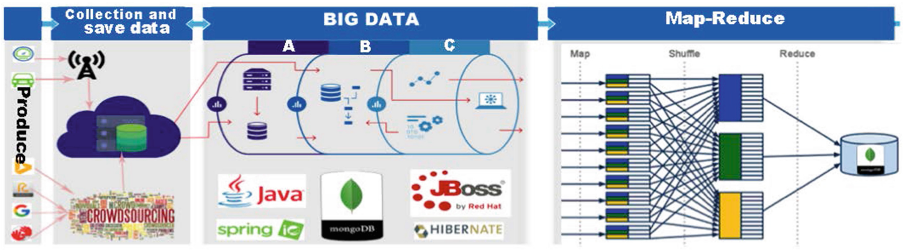

Fig. 4. Model for traffic congestion.

Our model is based on several layers that work in parallel to collect and manage all the data collected from outside and create several scenarios that will be used by all drivers.

Layer 1: Data Producers

This layer consists in producing the data that can be used in the system. It is the birth side of the ingredients of the final solution, for example: vehicles transmit the geo-location data generated by an installed GPS tag, pre-run data on the Web Services of Google, Bing, Tom-tom, Navcities…

Layer 2: Data Center

In this part it is a cloud where we store all the data received, and regenerated to a single format because they come in many different formats, that comes from sources installed and managed by us is encrypted and in hex format examples of resources used: Sensors on the roads, GPS devices on vehicles, applications on smartphones, and Data on the internet (Crowdsourcing) is come by the web services rest in general is in format JSON.

Layer 3: The Processing Center (Big DATA)

Big Data is the term that describes a set of non-structured and complex data, which come from our center data, it is used to store, manage, analyze and visualize information.

Big Data Analysis (BDA) is a way to innovate the marketing since it primarily generates a global revolution in the use of mass media and devices, given that treatment of data is going to create support processes which allow the integration of intelligent systems [10, 11], there upon BDA is implemented in order to deliver a better management of massive data in real time.

Nowadays, Big Data is enabling that great aspects of human life be studied as scientist and marketing area, given the volume of data generated daily and the analysis of its complexity [12]. Its challenge is in generated convert data in tools that allow to managers to solve problems acting rightly in making decision [13].

Our model depends the Big Data and these processing and analysis algorithms for improving the systems performance [14]. Some challenges and difficulties of Big Data's role in this model are shown below:

– **(A)** Data manipulation, storage availability, compatibility of systems, among others, attending to the autonomy and the adaptability of data for the development of applications [15]. Especially in a Smart Cites. Security and monitoring: To enhance user's reliability, sensible data should be protected and restricted to private access [16].

- **(B)** Technological development: It looks for reducing the heterogeneity between operative systems and to face the energetic barriers of devices [17].
- **(C)** Standardization: In order to avoid the wrong usage of critical information [18], the requirement is on the protection of sensible data and the use of it without authorization [16, 17] has block the manipulation of existent IoT's platforms. Privacy: People does not feel entrusted with the idea of sharing personal data with the world, they believe in technology but not in its managers [19, 20].

Layer 4: Map-Reduce

This layer for condensing large volumes of data into useful aggregated results. In this map-reduce operation, the map phase is applies to each input document (i.e. the documents in the collection that match the query condition). The map function emits key-value pairs. For those keys that have multiple values, and applies the reduce phase, which collects and condenses the aggregated data, then stores the results in a collection, the output of the reduce function may pass through a finalize function to further condense or process the results of the aggregation.

- **Map():** The Map worker parses key-value pairs out of the assigned split and passes each pair to the user-defined Map function. The intermediate key-value pairs produced by the Map function are buffered in memory. Periodically, the buffered pairs are written to local disk, partitioned into R regions for sharding purposes by the partitioning function that is given the key and the number of reducers R and returns the index of the desired reducer.
- **Shuffle:** the Map output to the Reduce processors: When ready, a reduce worker reads remotely the buffered data from the local disks of the map workers. When a reduce worker has read all intermediate data, it sorts the data by the intermediate keys so that all occurrences of the same key are grouped together. Typically many different keys map to the same reduce task.
- **Reduce():** The reduce worker iterates over the sorted intermediate data and for each unique intermediate key encountered, it passes the key and the corresponding set of intermediate values to the user's Reduce function.

Produce the final output: The final output is available in the R output files (one per reduce task).

4.2 Application

The block of code in Fig. 5 represent a function that helps us to display all the vehicles which find it at a distance given by parameter in the back office and the current position of the driver.

```
function predicte_state_off_all_viheculare_after_map_reduced() {
    var results = [];
    var all = db.device.find({}, {
        _id: 0,
        Identifier: 1,
        name: 1
    });
    var names = [];
    var identifiers = [];
    all.forEach(function(imei) {
        identifiers.push("position_"+imei['identifier']);
        names.push(imei['name']);
    });
    for (var i = 1; i < identifiers.length; i++) {
        scol = db.getCollection(identifiers[i]);
        scol.find().limit(1).sort({
            date_gps: 1
        }).forEach(function(data) {
            results.push({
                name: names[i],
                position: data
            });
        });
    }
    return results;
};
```

Fig. 5. Magic function for state all vehicles in current time.

In input a request that's select all existence car in the system who can manipulate data represented by the variable "**all**", by looping on this last we obtain the list of the frames received from these vehicles (Fig. 6).


```
{
    "name" : "Moto1 11",
    "position" : {
        "_id" : ObjectId("589f45ec5d08412bf59b3d4f"),
        "_class" : "com.karouani.ensem.mongodb.models.Position",
        "rmuid" : "522613",
        "position_id" : 93617,
        "latitude" : 35.7614,
        "longitude" : -5.83877,
        "altitude" : 77,
        "speed" : 4,
        "direction" : 146,
        ....
        "raw_data" : "4D4347500075F907000004E8DE1F0400EB00200102000800F06B
F48E00B7EE0100000000000003992000000050F8164FF3A62B803151E00007F0000
00020A0E3A0B1204E00723",
        "date_creation" : "2016-04-18 12:58:01.0",
        "date_gps" : "2016-04-18 08:57:59.0",
        "date_cell" : "2016-04-18 11:58:01.0"
    }
}....
```


Fig. 6. Result description as JSON format.

For check our solution we used this ingredient:

$$MacBook\ Pro\ I5 = CPU\ 2.4\,\mathrm{GHZ}\ and\ 8G\ of\ RAM \tag{1}$$

$$Distance = 8\,\mathrm{km}^2 \tag{2}$$

$$Number\ of\ Vehicle = 1200 \tag{3}$$

The result was given after 16 ms in the format Json in Fig. 6 which is composed of:

- **name:** The name represents the name of the vehicle tracker.
- **position:** the position object contains all the data required and received by the gps beacon installed in the vehicle, e.g. geolocation coordinates (latitude, longitude,

altitude), speed, date time to pick up and send the frame, reception frame And of course the direction where goes!

We make calculations to estimate the time consumed to generate the state of the traffic road in real time.

Suppose that Casablanca city contains 4 million vehicles [23] then mathematically speaking, if 1200 give the result in 16 ms then with 4 millions we get:

$$(4000000000 \,*\, 16)/1200 = 26666.67\,\text{ms} \tag{4}$$

$$26666.67/(1000 * 60) = 0.889\,\text{min} \tag{5}$$

In our solution for looking in the traffic real time you have just less than 1 min for got a result.

Figure 7 describes the front end in showing the result generated in maps using the Google services (Polylines, markers, and Heatmaps Layer), Jquery and Node JS.

Fig. 7. Traffic result after map-reduce layer.

The colored lines representing traffic conditions on major highways refer to the speed at which one can travel on that road.

The dreaded red lines mean highway traffic is moving at less than 40 km/h and could indicate an accident or congestion on that route. Orange lines on the map mean traffic is moving faster, from 40 to 80 km/h, while green lines mean traffic is zipping along at 80 km/h or more. If you see gray lines, that means there's no traffic information available at the time and a red-black line refers to extremely slow or stopped traffic.

If you are looking at traffic on city streets, where the speed limits are much lower than on the highways, the colors take on more of a relative meaning. Red (or red-black) lines mean a lot of slow going and general congestion. Orange is a little better but still not the best for city travel, and green means traffic conditions are good.

Vehicle has 3 state: if speed is greater than 5 km/h and the ignition is on that's mine the vehicle is moving the sign is icon with the speed and a cap direction else ignition off and the speed is less than 5 km/h that's mine it's in parking with blue icon otherwise is stopped with the icon "STOP" color red example it's stop at a red traffic light. The info showing all others information related to the vehicle like date creating the data and Signal GPS…

5 Conclusion

Smart cities and urban planning leave a major impact on the development of the nations. It increases the decision power of the societies by making an intelligent and effective decision at the appropriate time. In this paper, we presented a solution for traffic management in a smart city; a four layer system is proposed to help drivers minimize travel time. We exploit the power of Big Data in the implementation of the system. In this model without the Big Data technology it was not possible to generate value in the information, since the platforms did not have speed for Processing Data, coverage to contain a large volume of information and the ability to categorize data according to their variety.

Taking advantage of the automobiles good freedom and free movability, and allowing each driver to freely select routes in a natural way, Our Model will contribute to traffic safety and smoothness. This will lead to more effective utilization of road assets.

Our solution combine the tree best solution in the world related to management traffic congestion is "VICS" this method collected information of all vehicles crossed a road using sensor installed, the PCS where we installed the device GPS into a vehicle to collect a data of the geolocalisation (Latitude, Longitude, speed, altitude, ignition… etc.), those method send the data to the central storage data (Layer 2) to save it into a system storage based in big data (Mongo DB) without missing that's part has some listner scheduled as role is collecting the data from the web using a web services REST (Crowdsourcing) in a major platform offer this solution like Google, Tom Tom, Waze, Open Signal… etc.

6 Perspectives

For the next works, we aim to enhance our model across the multiples stages mentioned above, and propose a short path to the front-end user by implementing some related algorithms, which take in consideration the complexity and near real-time processing, change map-reduce by apache spark, and java (Spring, hibernate…) by NODEJS

References

1. Elgarej, M., Mansouri, K., Youssfi, M.: An improved swarm optimization algorithm for vehicle path planning problem. In: 4th IEEE International Colloquium on Information Science and Technology (CiSt) (2016)
2. Al Nuaimi, E., Al Neyadi, H., Mohamed, N., Al-Jaroodi, J.: Applications of big data to smart cities. J. Internet Serv. Appl. **6**, 25 (2015)

3. Wang, S., Djahel, S., Zhang, Z., McManis, J.: Next road rerouting: a multiagent system for mitigating unexpected urban traffic congestion. IEEE Trans. Intell. Transp. Syst. **17**(10), 2888–2899 (2016)
4. Koyama, A., Inoue, D., Shoji, S.: An implementation of visualization system for vehicles and pedestrians. In: The 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA) (2016)
5. Schmied, R., Moser, D., Waschl, H., del Re, L.: Scenario model predictive control for robust adaptive cruise control in multi-vehicle traffic situations. In: Intelligent Vehicles Symposium (IV). IEEE (2016)
6. Ma, D., Luo, X., Li, W., Jin, S., Guo, W., Wang, D.: Traffic demand estimation for lane groups at signal-controlled intersections using travel times from video-imaging detectors. IET Intell. Transp. Syst. **11**(4), 222–229 (2017)
7. El Hatri, C., Boumhidi, J.: Q-learning based intelligent multi-objective particle swarm optimization of light control for traffic urban congestion management. In: 4th IEEE International Colloquium on Information Science and Technology (CiSt) (2016)
8. Sánchez-Medina, J., Gálan-Moreno, M.J., Rubio-Royo, E.: Traffic signal optimization in "La Almozara" district in Saragossa under congestion conditions, using genetic algorithms, traffic microsimulation, and cluster computing. IEEE Trans. Intell. Transp. Syst. **11**(1), 132–141 (2010)
9. Ram, S., Wang, Y., Currim, F., Dong, F., Dantas, E., Sabóia, L.A.: SMARTBUS: a web application for smart urban mobility and transportation. In: 25th International Conference on World Wide Web Companion (2016)
10. Zhou, K., Fu, C., Yang, S.: Big Data driven smart energy management: from big data to big insights. Renew. Sustain. Energy Rev. **56**, 215–225 (2016)
11. Mohamed, N., Al-Jaroodi, J.: Real-time big data analytics: applications and challenges. In: 2014 International Conference on High Performance Computing and Simulation (HPCS), pp. 305–310 (2014)
12. Sharma, S.: Expanded cloud plumes hiding Big Data ecosystem. Future Gener. Comput. Syst. **59**, 63–92 (2016)
13. Xu, Z., Frankwick, G.L., Ramirez, E.: Effects of big data analytics and traditional marketing analytics on new product success: a knowledge fusion perspective. J. Bus. Res. **69**(5), 1562–1566 (2015). Glova, J., Sabol, T., Vajda, V.: Business models for the Internet of Things environment. Procedia Econ. Financ.
14. Kyriazis, D., Varvarigou, T.: Smart, autonomous and reliable Internet of Things. Procedia Comput. Sci. **21**, 442–448 (2013)
15. Henze, M., Hermerschmidt, L., Kerpen, D., Häußling, R., Rumpe, B., Wehrle, K.: A comprehensive approach to privacy in the cloud-based Internet of Things. Future Gener. Comput. Syst. **56**, 701–718 (2015)
16. Yan, Z., Zhang, P., Vasilakos, A.V.: A survey on trust management for Internet of Things. J. Netw. Comput. Appl. **42**, 120–134 (2014)
17. Botta, A., de Donato, W., Persico, V., Pescapé, A.: Integration of cloud computing and Internet of Things: a survey. Future Gener. Comput. Syst. **56**, 684–700 (2015)
18. Weber, R.H.: Internet of Things: privacy issues revisited. Comput. Law Secur. Rev. **31**(5), 618–627 (2015)
19. Lee, I., Lee, K.: The Internet of Things (IoT): applications, investments, and challenges for enterprises. Bus. Horiz. **58**(4), 431–440 (2015)
20. Liu, B., Li, L., Liu, K.: Study on the evaluation method of probe car system. In: IEEE Intelligent Vehicles Symposium (2010)

21. Rathore, M.M., Ahmad, A., Paul, A., Thikshaja, U.K.: Exploiting real-time big data to empower smart transportation using big graphs. In: 2016 IEEE Region 10 Symposium (TENSYMP), pp. 135–139 (2016)
22. Google Official Blog. https://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html
23. La nouvelle Tribune du maroc. https://lnt.ma/parc-automobile-au-maroc-compte-3-437-948-unites-fin-2014/