International Journal of Computer Science and Applications,

Vol. 13, No. 1, pp. 1 – 16, 2016

TRAFFIC AND TRANSPORTATION SMART WITH CLOUD COMPUTING ON BIG DATA

XIAOXIA WANG

School of Traffic and Transportation, Beijing Jiaotong University, No.3 Shangyuancun, Haidian District, Beijing, 100044, China
xxwang@bjtu.edu.cn

ZHANQIANG LI

School of Traffic and Transportation, Beijing Jiaotong University, No.3 Shangyuancun, Haidian District, Beijing, 100044, China
15120960@bjtu.edu.cn

Now it's the trend of using cloud computing capacities for the provision and support of ubiquitous connectivity and real-time applications and services for smart cities' needs. This paper presents the route map of big data relying on cloud computing to make urban traffic and transportation smarter by mining and pattern visualization with literature review and case studies. Although most of these technologies already commercialized, to be or not to be cloud is still a problem for organizations because of the top issues like security and privacy. However, a simple structured framework might help prevent some typical traps. For e-government and politics, data open and transparency to enhance decision-making are possible when not only enterprises carried out business intelligence but also general people are empowered with smart devices.

Keywords: Traffic and Transportation; Cloud Computing; Big Data; Visualization.

1. Introduction

The era of big data is upon us. Hidden inside a flood of heterogeneous raw data is the knowledge. With the advent of cloud computing, resizable infrastructure for data analysis is now available to everyone via an on-demand maybe free model. In order to unlock the potential of big data, there are a significant number of research challenges need to overcome including: managing diverse sources of unstructured data with no common schema, removing the complexity of writing auto-scaling algorithms (e.g. Amazon Elastic Compute Cloud), real-time analytics, suitable visualization techniques for petabyte scale data sets etc. Big data hold great promises for discovering subtle population patterns and heterogeneities that are not possible with small-scale data, which should care[Villars *et al.* (2011)].

The global laboratory of IBM defines the future of cities as smarter cities involving following features: (1) using big data and analytics for deeper insights; (2) cloud for collaboration among disparate agencies; (3) mobile to gather data and address problems directly at the source; (4) social technologies for better engagement with citizens. Into and around the city, people and goods are always moving. Intelligent transportation systems(ITS) as fundamental infrastructure services make a city “livable” by improving capacity, enhancing travel experiences and making moving anything safer, more efficient and more secure. For instance, Singapore already set up an excellent stylish example based on Internet of Things(IoT) to connect, collect and comprehend. The Land Transport Authority plans routes and establishes minimum service standards for bus lines managed by the Singapore Bus Service and Singapore Mass Rapid Transit.

To deal with dynamic traffic environments and assist users and city officials better understand traffic problems in large cities, the current trend for powerful intelligent smart city applications is managing data and processing in cloud-enabled large scale sensor networks [Cuzzocrea *et al.* (2013)], for actively and autonomously adaptation and smart provision of services and content, which increasingly converge with cloud computing environments.

This paper organized as following: (1)present smart urban transport conceptual framework based on cloud and IoT by literature review; (2)bring up the route map for deploy big traffic data on the cloud; (3)put forward the question of to be or not to be cloud with some key issues and introduced a structured method to frame out the procedure of introducing cloud computing into business; (4) when business enterprises already carried out cloud and the people empowered with smart devices, transparency government are urged to open data and provide tailor-made policy choices thanks to big data mining.

2. Big Data Based on Cloud Computing and IoT

Smart cities rely not only on sensors within the city infrastructure, but also on a large number of devices that sense and integrate their data into technology platforms used for introspection into the habits and situations of individuals and city-large communities. Fig. 1 presents the conceptual framework of urban smart transportation based on cloud and IoT. Contemporary smart urban transport systems: (1)employ secured IoT to generate big data, which comprise billions of devices that sense, communicate, compute and potentially actuate, massive connected via GIS-T[Weisbrod (2011)], widely available real time communication network(e.g. 4G, Wi-Fi, Bluetooth); (2)generate big data [Chen *et al.* (2014)], [Sagiroglu and Sinanc (2013)] with “4Vs” features, as raw material facing great challenges[Fan *et al.* (2014)]; (3)rely on resilient cloud computing[Marston *et al.* (2011)] to store, manage, mine and create values for insight as the solution to many of our society’s traffic and transportation problems [Richards and King (2013)]; (4)in an era of data abundance, there is a clear need for visualization tools to provide insight into how coordinated systems should be expected to operate under different parameter settings and to document coordinated system behavior.

Cloud computing is utilized to meet the requirements on infrastructure for big data. [Suciu *et al.* (2013)] presented a framework for data procured from highly distributed, heterogeneous, decentralized, real and virtual devices that can be automatically managed, analyzed and controlled by distributed cloud-based services. [Zaslavsky *et al.* (2012)] discussed emerging IoT architecture, large scale sensor network applications, federating sensor networks, sensor data and related context capturing techniques, challenges in cloud-based management, storing, archiving and processing of sensor data. [Zheng *et al.* (2013)] provided an overview of service-generated big data and big data-as-a-service, employed to provide common big data related services (e.g. accessing service-generated big data and data analytics results) to users to enhance efficiency and reduce cost. [Zimmermann *et al.* (2013)] proposed an integration model for service-oriented enterprise architecture based on enterprise services architecture reference cube and the systematic development, diagnostics and optimization of architecture artifacts of service-oriented cloud-based enterprise systems for big data applications. [Li *et al.* (2011)] proposed urban traffic management systems using intelligent transportation cloud, which generate, store, manage, test, optimize, and use mobile traffic strategy agents to maximize advantages of cloud computing and agent technology to effectively control and manage urban traffic systems. [He *et al.* (2014)] presented a multilayered IoT-based vehicular data cloud platform with an intelligent parking cloud service and a vehicular data mining cloud service.

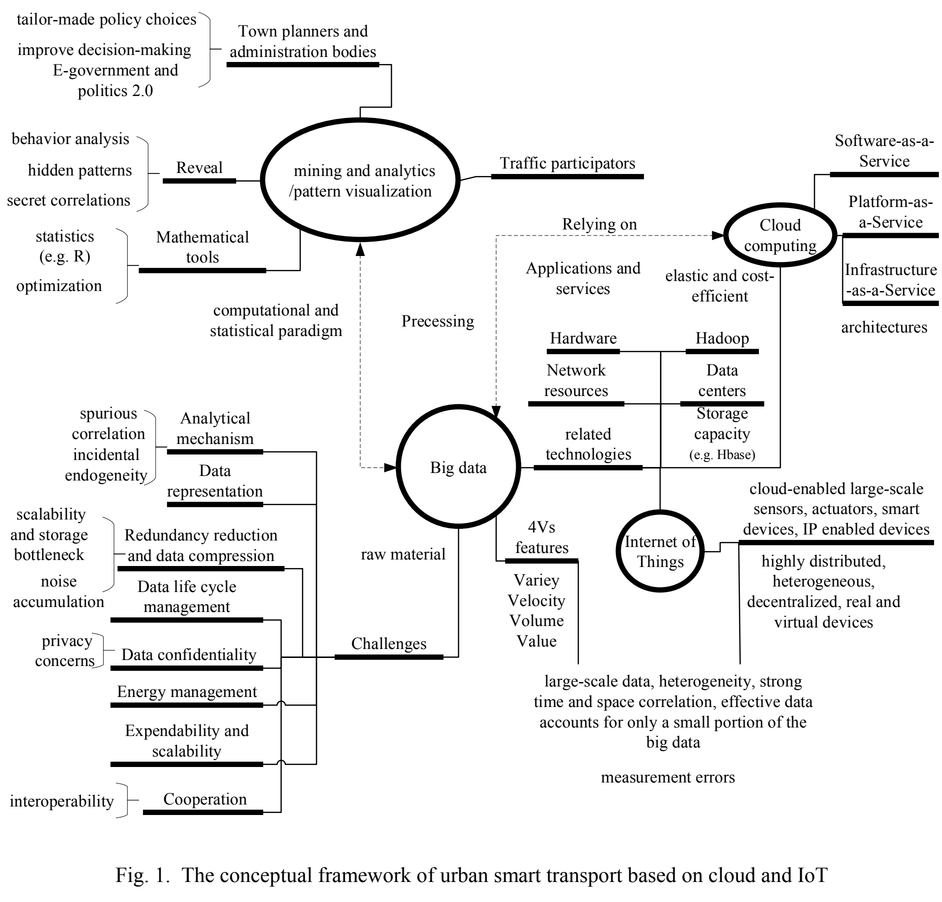
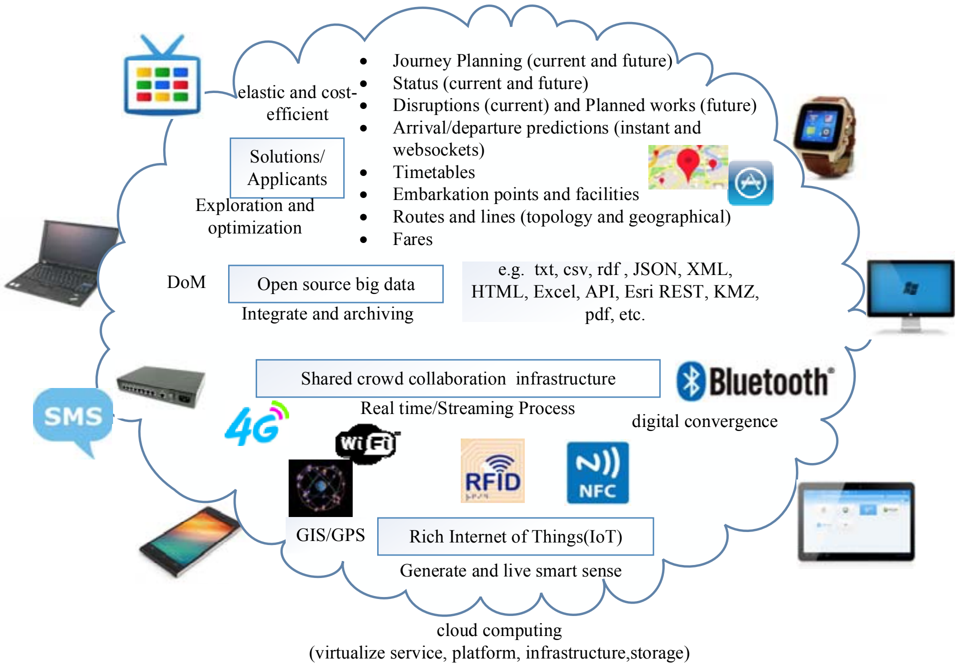


Fig. 1. The conceptual framework of urban smart transport based on cloud and IoT

3. Route Map for Deploy Big Traffic Data on the Cloud

With the tendency of asset-light, cloud computing with certain service-level agreement is under consideration. The term “moving to cloud” also refers to an organization moving away from a traditional capital expenditure model (buy the dedicated hardware and depreciate it over a period of time) to the operating expense model (use a shared cloud infrastructure and pay as one uses it).

Fig. 2 tap big open traffic data on cloud with IoT infrastructure, illustrate with the available options at each level. As a smart city begins with rich IoT involved with big data and cloud computing. When generate, big data transferred real time with streaming process through communication network(such as Bluetooth, Wi-Fi, 4G) to store(e.g. MongoDB), which usually combined with ETL(Extract, Transform and Load) Process. Standardized structure format, combined with open source software to develop planning tools; provide a platform to gather public involved more efficiently create to APPs and maps as well through crowdsourcing[(Lantz *et al.* (2015)]. As Harry Strasser’s vision, a connected world digital(technological) convergence where everything in people’s life will have computing power, wireless connectivity and a lot of smart sensors.



Notes: DoM(Document Object Model); Esri REST(Representational State Transfer); KMZ(Google Earth Placemark File)

Fig. 2. Cloud framework for tapping open traffic data based on multilayered IoT

3.1. *Physical layer: rich smart IoT incorporate with social media*

There are large scale highly distributed, heterogeneous, decentralized Geo-reference or GPS-enabled infrastructure networks for live data feeds, which from traditional (1) roadside sensor/detectors data, e.g. closed circuit television; (2) equipment/devices status, e.g. vehicular, parking lot; (3) traffic events; to (4) online and mobile social media information sources and weather services, which facilitate developing areas to catch up-to-date without too much heavy fixed investment. Traditionally, for instance, London where the legendary control system split cycle offset optimization technique in operation, Macroscopic Fundamental Diagrams (MFD, e.g. volume vs. density) present a macroscopic description of urban traffic. [Pascale *et al.* (2015)] designed a spatiotemporal clustering method to detect homogeneous areas over the network and overtime, which including a segmentation method to divide the time intervals into segments optimizing the clustering results within each of them. Recently developing countries(like Kenya, Turkey, Bangladesh) without properly widely deployed Geo-devices(or sensors), mobile device and Wi-Fi facilities, [(Lantz *et al.* (2015)] investigated Nairobi, Istanbul and Dhaka the potential of integration of text-based data from online media including Twitter, Facebook, FourSquare, and weather reporting services and location based data from mobile devices to support transit operation.

3.2. *Open resources: big data*

Not reduced the richness of the service-oriented big data, the challenges are the integration and archiving of highly dispersed and non-uniform even unstructured big data. [Philip *et al.* (2014)] discussed several underlying methodologies to handle the data deluge. The location-based data, from general matrix to ORNL's LandScan^{TM} (the community standard for global population distribution), sometimes go with user's profile and characteristics, as well as the environment parameter, should interpret with some standardized structure datasets format for improving transparency as well as further drill. For example, OneBusAway, the open-source platform for real-time transit info, includes a robust, secure, scalable back end, which accepts, stores, archives and interprets real time vehicle location data in combination with transit schedules and other related data and offers a suite of Application Programming Interfaces (APIs). Another example, the General Transit Feed Specification (GTFS), the industry standard introduced in 2007[Zeng *et al.* (2014)] making public transit data universally accessible, defines a common format for public transportation schedules and associated geographic information. GTFS "feeds" allow public transit agencies to publish their transit data and developers to write applications which consume that data in an interoperable way like CTA (Chicago Transit Authority). In 2013, with real-time train arrival data becoming widely available through its GTFS Realtime feed, New York City Transit capable of on-the-fly performance visualization and reporting.[Suchkov *et al.* (2014)] There are other open-source projects, such as TransiTime to make useful GPS based data readily available for both letting passengers know the status of their vehicles and helping agencies more effectively manage their systems.

3.3. *Big data visualizing exploration*

When live data became non real-time historical data, visualizing tools to extract traffic patterns and travel demand to demonstrate both historical data and forecast data. Big data hold great promises for discovering subtle population patterns and heterogeneities that are not possible with small-scale data, which we should care[Villars *et al.* (2011)]. Generally speaking, spatiotemporal data visualization usually based on statistical distributions and probabilistic analysis approaches, which divided into three categories: (1) the direct depiction; (2) the data summaries - aggregates; (3) the computationally extracted patterns.[Bahbouh and Morency (2014)] As early as [Hughes (2004)] summarized the practices of visualization in transportation at that time and indicated some future directions. Within the overall notion of context-sensitive design and public involvement, visualization cannot be only about how something looks but also about how it works. Analysis would be considerably simplified if one could visualize the data graphically.

Fig. 3 illustrate solutions that candidate to manage, move, analyze and visualize complex distributed data sets. [Batty (2013)] described how the growth of big data is shifting the emphasis from longer term strategic planning to short-term thinking about how cities function and can be managed, e.g. 6 months of smart travel card data of individual trips on Greater London’s public transport systems. [Dobre and Xhafa (2014)] presented a concrete implementation of an ITS designed on top of context-aware platform using integrated mobile services, which integrates services designed to collect context data (location, user’s profile and characteristics, as well as the environment). At macro level, GIS-based decision support system[Carder and Christie (2007)] linked to a GIS and can be used to visualize relief planning scenarios. To understand macro long-term demand, [Brooks *et al.* (2012)] detailed picture of potential mineral revenues and train freight volumes; [Bahbouh and Morency (2014)] visualized an OD matrix and to identify major corridors with clustering tools. GIS assists in improving project cost estimation [Ashur and Crockett (1997)]. Time-space diagrams were developed when data is scarcity, now with the trajectories of vehicles GPS data in time-space cube, calculate and visualize the headways and separation with simple geometric methods[(Anwar *et al*. (2014)]. More conventional is Fundamental Diagram (FD, e.g. the flow-density relationship) or fundamental-diagram-like relationship between the network-aggregated vehicle occupancy and traffic flow regimes: under-saturated, saturated, over-saturated; capacity, spillbacks and gridlocks; congestion degree[Lin *et al*. (2013)]. However, the shape of MFDs is influenced by the structure of the traffic network, the heterogeneity of the networks, the traffic demands, and the control strategies. Five-minute interval detectors’ data to plot the traffic volume and occupancy in selected areas; such as time-of-day schedule change time, observed cycle length, green time and split time, coordinated phase actuation, early return to green, arrivals over advance detection relative to green indication, progression quality characteristics related to offset, adjacent signal synchronization, coordinated phase operation in rest, plan time changes, preemption, impact of queuing, and longitudinal analysis of splits.[Brennan *et al*. (2012)]

[Suchkov *et al.* (2014)] present near-real-time train movements in string line (time-distance chart) format.

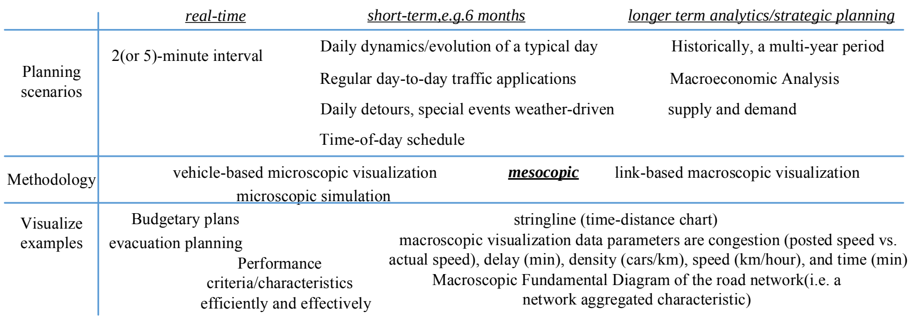


Fig. 3. Data visualization spectrum architecture[refer Lu *et al.* (2015)]

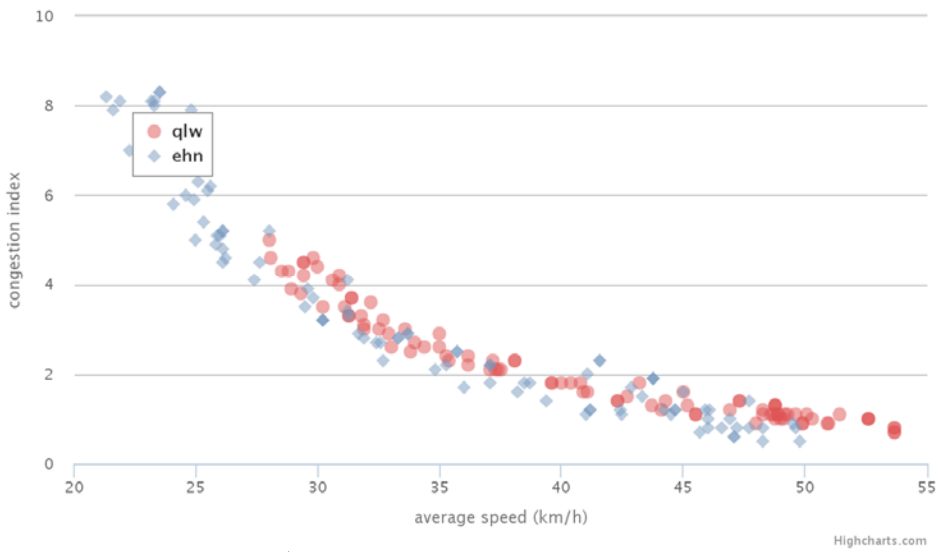
Cloud computing reveals hidden patters and secret correlations. According to the IDC iView "Extracting Value from Chaos"[(Gantz and Reinsel (2011)], data-intensive scientific discovery, also known as big data problems, converge into the big data explorations. [Fan *et al.* (2014)] emphasized on the viability of the sparsest solution in high-confidence set and point out that exogenous assumptions in most statistical methods for big data cannot be validated due to incidental endogeneity, which can lead to wrong statistical inferences and consequently wrong scientific conclusions. And there are more and more easily accessed visualize tools, which lower the barrier of processing data and focus on the underlying meaning. Traditionally, Tableau, as a business intelligence solution, explore and visualize data fast and easy with an intuitive drag and drop to tell a story. Table 1 lists some top amazing open source data visualization tools on GitHub. All of them facilitate average person concentrate on employing or analyzing data with multi-dimension presentation flexible and dynamic.

Table 1. Data visualization tools for the web list by GitHub

Tools	Features	Elements
D3	a JavaScript library for manipulating documents based on data	HTML, SVG and CSS(web standards); DOM manipulation
Chart.js	Simple HTML5 Charts using the <canvas> tag	HTML5; canvas
Leaflet	an open source JavaScript library for mobile-friendly interactive maps	Openstreetmap; GeoJSON
ECharts	Canvas(ZRender), Javascript	JS

Source: https://github.com/showcases/data-visualization

Here visualize dynamic(interactive) with Highcharts as Fig. 4 and 5, which is a charting library written in pure JavaScript, offering an easy way of adding interactive charts to website or web application. The procedure is: (1) crawl data every 15 minutes from Beijing Municipal Commission of Transport(BMCT); (2) save and organize data as excel format file; (3) clean, sort, organize source data for different purposes; (4) visualize organized data, which contribute to make all kinds of travel decision. For instance, comparing between days, months, years or regions by data mining software such as Rattle of R, identify temporal and spatial mutation and features, such as Monday and Friday are always congested.



Notes: ehn(erhuannei, within the 2nd ring road); qlw(quanluwang, the whole road network)

Fig. 4. Scatter of average speed and congestion index of qlw and ehn, Sept. 4th, 2015, Beijing

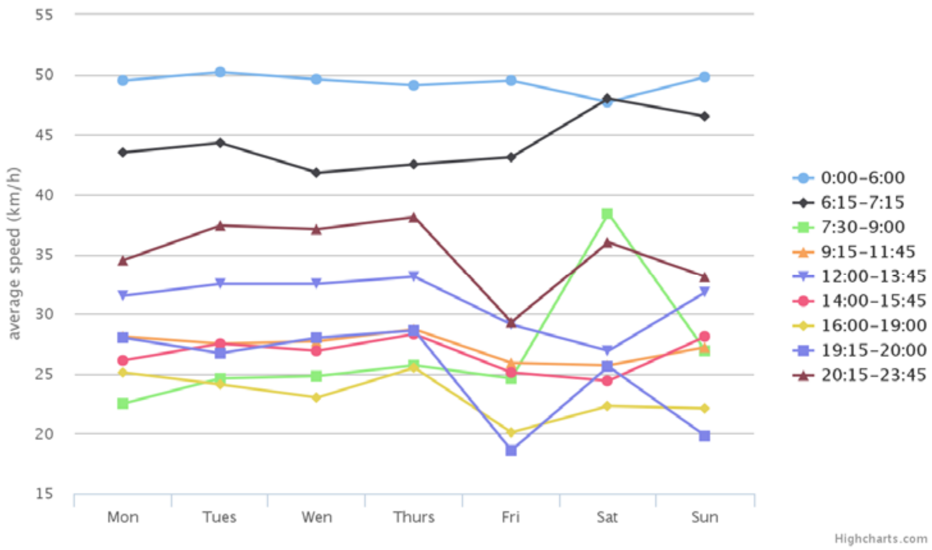
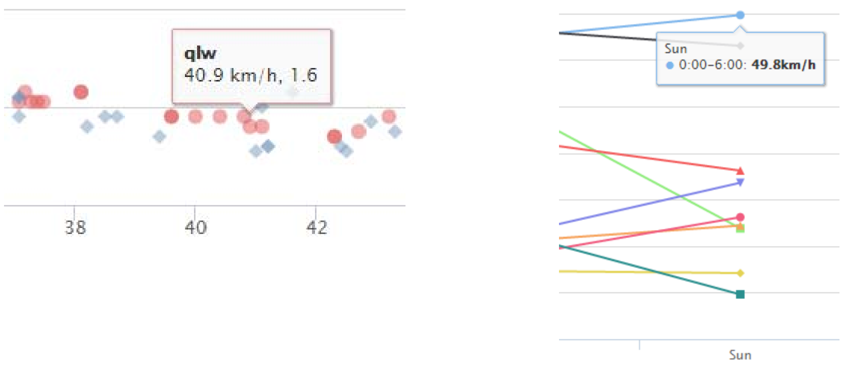


Fig. 5. Weekly average speed of Sept. in 2015 by time window, Beijing

Interactive visualization as Fig. 6 (a) and (b) by moving the mouse over every point of Fig. 4 and 5 to know the exact speed and congestion index.



(a)example of mouse over for Fig.4 (b) example of mouse over for Fig.5

Fig 6. Interactive and dynamic visualization

Another integration example is an interactive web Apps with Shiny by RStudio. As a web application framework for R(a free software environment for statistical computing and graphics), a Shiny APP is a web page (UI, User Interface) connected to a computer running a live R analyzing session (Server) and no HTML, CSS, or JavaScript knowledge required. Users manipulate the UI, which will result in the server to update the UI's displays (by running R code). Fig. 7 open source data from BMCT in the browser.

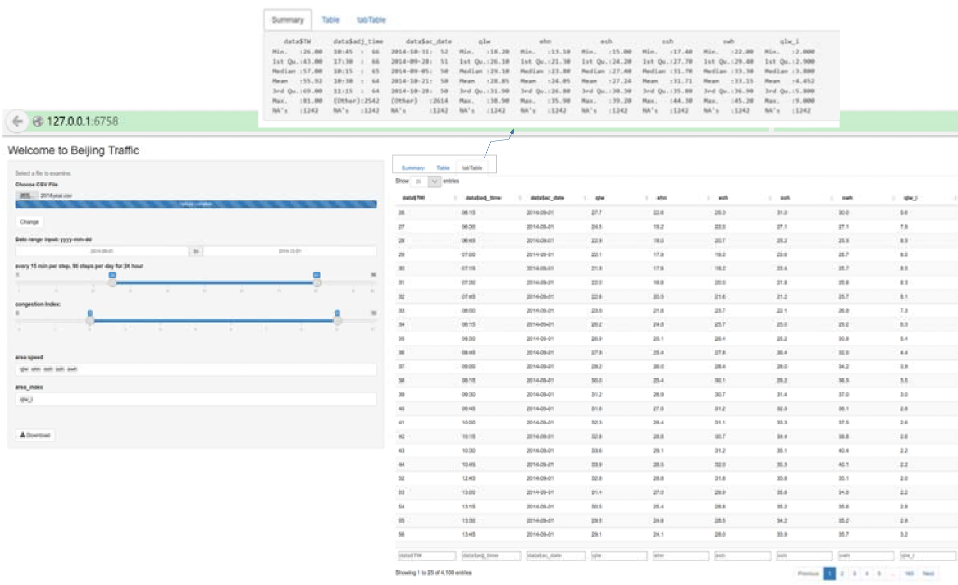


Fig. 7. Input, filter, output data with tab in the browser

Fig. 8 output the images of selected areas. Each area with 365 days as the X-axis, daily 24 hours as the Y-axis, each point is speed with color degree [Wang (2015)].

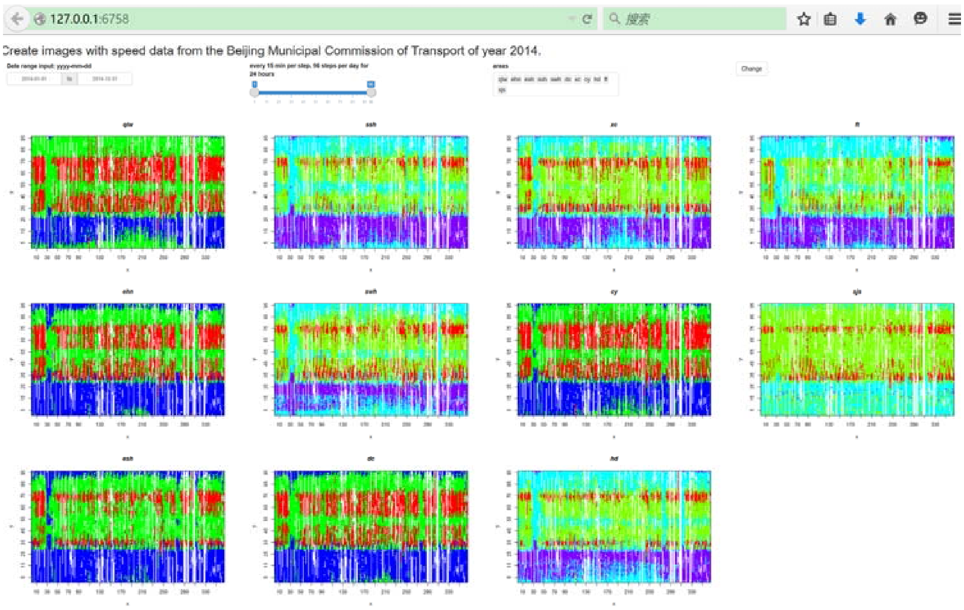


Fig. 8. Images output of selected areas

4. To be or not to be Cloud: Challenges for Organizations

Cloud-based big data analysis offers a convenient mean for providing an elastic and cost-efficient exploration of voluminous data sets. Gaode LBS(location based service) open platform, providing not only traditional location, maps, navigation but also guiding the application development from "LBS + O2O(Online to Offline)", "LBS + travel" to "LBS + intelligent devices" with free maps tools, open source code, demo and API(Application Programming Interface)/SDK(Software Development Kit) with Key. The circumstances that affect a firm's intention to adopt cloud computing technologies in support of its supply chain operations are investigated by considering the tenets of classical diffusion theory as framed within the context of the information processing view. [Cegielski (2012)] posit that various aspects of an organization and its respective environment represent both information processing requirements and capacity, which influence the firm's desire to adopt certain information technology innovations. The results suggest that business process complexity, entrepreneurial culture and the degree to which existing information systems embody compatibility and application functionality significantly affect a firm's propensity to adopt cloud computing technologies.

From the user perspective, provide a guideline for researchers in order to help them in making decisions regarding what might be a best fit for their goals. Highlighting the practical issues with interoperability, moving data to the cloud, and portability show opportunities which exist for future development of cloud computing and point to areas of concern that researchers face up to when making decisions about cloud computing. Particularly, security is of the utmost importance to researchers who wish to keep their sensitive data private. [(Branch *et al.* (2014)]

4.1. *Security and privacy*

[Zissis and Lekkas (2012)] proposed introducing a Trusted Third Party, tasked with assuring specific security characteristics within a cloud environment. The proposed solution calls upon cryptography, specifically public key infrastructure operating in concert with a single sign on and lightweight directory access protocol, to ensure the authentication, integrity and confidentiality of involved data and communications. The solution, present a horizontal level of service, obtainable to all implicated entities, that realizes a security mesh, within which essential trust is maintained.

In order to craft a balance between beneficial uses of data and in individual privacy, policymakers must address some of the most fundamental concepts of privacy law, including the definition of "personally identifiable information", the role of individual control, and the principles of data minimization and purpose limitation. [Tene and Polonetsky (2013)] emphasized the importance of providing individuals with access to their data in usable format. This will let people share the wealth created by their information and incentivize developers to offer user-side features and applications harnessing the value of big data. Where exclusive access to data is impracticable, data are likely to be de-identified to an extent sufficient to diminish privacy concerns. In addition, organizations should be required to disclose their decision criteria, since in a big data world it is often not the data but rather the inferences drawn from them that bring concern.

4.2. *Dilemma of cloud: traps to avoid*

In cloud-computing-based systems, allocated resources are encapsulated into cloud services and centrally managed, which allows high automation, flexibility, fast provision, and ease of integration at low cost. The integration between physical resources and cloud services can be enhanced by combining IoT technology and SaaS technology. In cloud manufacturing, distributed resources are encapsulated into cloud services and managed in a centralized way such as IT, pay-as-you-go business models, production scaling up and down per demand, and flexibility in deploying and customizing solutions. Clients utilize cloud services to meet their requirements. Cloud users request services varying from product design, manufacturing, testing, management, and all further stages of a product life cycle.[Xu (2012)] [Chen *et al.* (2014)] proposed an approach for developing cloud-based manufacturing systems based on a four-layer SaaS model and present a framework of logistic cloud. [Chen *et al.* (2012)] exploit the use of a cloud services platform for information exchange in combination with the contemporary popular community sites, while developing the site architecture based on the socio-technical system's theory as an indicator.

Following such a trend, industry leaders as Amazon, Google and IBM deploy various big data systems on their cloud platforms, aiming to occupy the huge market around the globe. Using the typical MapReduce framework as a case study, [Yuan *et al.* (2013)] found that its pipeline-based design intergrades the computational-intensive operations (such as mapping/reducing) together with the I/O-intensive operations (such as shuffling). Such computational-intensive and I/O-intensive operations will seriously influence the performance of each other and largely reduces the system efficiency especially on the low-end virtual machines (VMs). To address this problem, [Yuan *et al.* (2013)] re-model the resource provisioning problem in the cloud-based big data systems and present an interference-aware solution that smartly allocates the MapReduce jobs to different VMs.

However due to the cost of cloud based on resources allocation and harness, if employees or information department not properly gauge traffic and space, will result in two kinds of traps. One is not sufficient for business practices; the other is the upheaval of IT cost. Fortunately [Klems *et al.* (2009)] provided a step-by-step guide to determine the benefits from cloud computing, from describing a business scenario to comparing cloud computing services with a reference IT solution. They identify key components: business domain, objectives, demand behavior and technical requirements. Built on business objectives and technical requirements, the costs of a cloud computing service, as well as the costs of a reference it solution, can be calculated and compared if properly follow quality of service related service level agreements.

5. Transparency and Open Government

With national data transparency efforts like *President Obama's Open Data Initiative*(2009-1-21) and municipal projects like New York City's Big Apps or San Francisco's DataSF, government agencies across the country have been opening their raw data sets, some more reluctantly than others. Since fall 2010, as part of a contract with the Los Angeles Metropolitan Transportation Authority, researchers at the University of Southern California's Integrated Media Systems Center have been given access to high-resolution spatiotemporal transportation data from the LA County road network. [Jagadish *et al.* (2014)] Data.gov, the home of the U.S. Government's open data, datasets intended for public access and use. United States Census Bureau not only provides traditional data with tables, pdf and visualizations (like Infographics), Maps, Geographic data, even interactive data mapper, but also data tools and APIs for developers. OGP(Open Government Partnership) was launched in 2011 to provide an international platform for domestic reformers committed to making their governments more open, accountable, and responsive to citizens(or transparency, participation, accountability. Since then, OGP has increased from 8 to the 69 participating members. According to *The Notice of Action on Promoting the Development of Big Data* issued by the Chinese State Council, by the end of 2017 form a pattern of inter-department sharing data resources; by the end of 2018 united open platform of national government data and open to the social reasonable and moderate, leading the public carry out welfare development and innovative applications to bring incremental value for big data.

With the debut of City-Go-Round™ and media coverage generated about transit data transparency, various transit operators have taken steps to release their schedule and route information to third party developers, who in turn use the data to develop an array of applications to improve rider experience. Recognized the potential for new service management tools, TfL(Transport for London) pay extreme attention to every journey. To make sure their customers' journeys run as smoothly as possible, TfL not only invest in the latest technology to make plan and journeys easier, but also make all of their travel data freely available to the software developers to encourage them use all these public TfL data (or "open data") in their own software and services to present customer travel information in innovative ways - providing they adhere to the transport data terms and conditions. TfL facilitates them with (1) Data feeds, which spans a large spectrum of quality, accuracy and data formats. TfL simplifies the access to this data with their new front-end unified API, which supports all the data requirements of the TfL website. (2) Widgets for website or blog.

6. Conclusions

As grapple with increasingly large data sets, scientists uncork new bottlenecks.[Marx (2013)] Organizations are confronted with the choice of maintaining their own expensive storage devices or utilizing the cloud for their storage needs. Town planners and administrative bodies just need the right tools at their fingertips to consume all the data points that a town or city generates and is able to turn that into actions that improve peoples' lives. Publish a variety of useful data and encourage interested citizens to explore and take advantage of it, for instance, develop web and mobile applications, design data visualizations. E-government and politics 2.0 are on the way especially when business intelligence and analytics are performed by enterprises[Chen *et al.* 2012)]. Exploration of data-based knowledge to improve progress is not automatic and requires tailor-made policy choices that help to foster this emerging paradigm.[Hilbert (2013)] Thanks to quickly visualized data tools, identify correlations and conceive of innovative, unanticipated uses for existing information became easier. Cloud computing can transform the traditional government services model, help the country to align services innovation with administration strategy, and make intelligent executive networks that encourage effective collaboration.

Acknowledgments

This work was supported by NSFC (Project: 71303018), the MOE Key Laboratory for Transportation Complex Systems Theory and Technology of Beijing Jiaotong University, Center of Cooperative Innovation for Beijing Metropolitan Transportation.

References

Anwar, A., Zeng, W., Arisona, S. (2014). Time-Space Diagram Revisited. *Transportation Research Record: Journal of the Transportation Research Board*, *2442*(2442), 1–7.

Ashur, S. A., Crockett, B. (1997). Geographic Information Systems as a Support Tool in Construction Cost. *Transportation Research Record*, *1575*, 112–115.

Bahbouh, K., Morency, C. (2014). Encapsulating and Visualizing Disaggregated Origin-Destination Desire Lines to Identify Demand Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, *2430*, 162–169.

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*(3), 274–279.

Branch, R., Tjeerdsma, H., Wilson, C., Hurley, R., Mcconnell, S. (2014). Cloud Computing and Big Data : A Review of Current Service Models and Hardware Perspectives. *Journal of Software Engineering and Applications*, *7*(July), 686–693.

Brennan, T. M., Day, C. M., Sturdevant, J. R., Bullock, D. M. (2012). Visual Education Tools to Illustrate Coordinated System Operation. *Transportation Research Record: Journal of the Transportation Research Board*, *2259*, 59–72.

Brooks, C., Kourous-Harrigan, H., Billmire, M., Metz, P., Keefauver, D. E., Shuchman, R., … Taylor, M. (2012). Expanding Alaska-Canada Rail. *Transportation Research Record: Journal of the Transportation Research Board*, *2261*, 95–105.

Carder, D., Christie, C. (2007). Testing a Flexible GIS-Based Network Flow Model for Routing Hurricane Disaster Relief Goods. *Transportation Research Record*, (October 2006), 1–14.

Cegielski, C. G. (2012). Adoption of cloud computing technologies in supply chains: An organizational information processing theory approach. *The International Journal of Logistics Management*, *23*, 184–211.

Chen, C.-Y., Chang, C.-J., Lin, C.-H. (2012). On dynamic access control in web 2.0 and cloud interactive information hub: trends and theories. *Journal of Vibration and Control*, *20*, 548–560.

Chen, H., Chiang, R. H. L., Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data To Big Impact. *Mis Quarterly*, *36*(4), 1165–1188.

Chen, M., Mao, S., Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(January), 171–209.

Chen, S. L., Chen, Y. Y., Hsu, C. (2014). A new approach to integrate internet-of-things and software-as-a-service model for logistic systems: A case study. *Sensors (Switzerland)*, *14*(1), 6144–6164.

Cuzzocrea, A., Fortino, G., Rana, O. (2013). Managing data and processes in cloud-enabled large-scale sensor networks: State-of-the-art and future research directions. *Proceedings - 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, CCGrid 2013*, 583–588.

Dobre, C., Xhafa, F. (2014). Intelligent services for Big data science. *Future Generation Computer Systems*, *37*, 267–281.

Fan, J., Han, F., Liu, H. (2014). Challenges of Big Data analysis. *National Science Review*, (August 2013), 1–38.

Gantz, B. J., Reinsel, D. (2011). Extracting Value from Chaos State of the Universe : An Executive Summary. *IDC iView*, (June), 1–12. Retrieved from http://idcdocserv.com/1142

He, W., Yan, G., Xu, L. Da. (2014). Developing vehicular data cloud services in the IoT environment. *IEEE Transactions on Industrial Informatics*, *10*(2), 1587–1595.

Hilbert, M. (2013). Big data for development: From information-to knowledge societies. *Ssrn Electronic Journal*, 1–39.

Hughes, R. (2004). Visualization in Transportation: Current Practice and Future Directions. *Transportation Research Record*, *1899*(1), 167–174.

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, *57*, 86–94.

Klems, M., Nimis, J., Tai, S. (2009). Do Clouds Compute? A Framework for Estimating the Value of Cloud Computing. *Lecture Notes in Business Information Processing*, *22*, 110–123. Retrieved from http://www.guodao.cn:802/hdbsm/f.aspx?name=HCCOMP&gui=comp1109204027903

Lantz, K., Khan, S., Ngo, L. B., Chowdhury, M., Donaher, S., Amy Apon. (2015). POTENTIALS OF ONLINE MEDIA AND LOCATION-BASED BIG DATA FOR URBAN TRANSIT NETWORKS IN DEVELOPING COUNTRIES. In *94th Annual Meeting of the Transportation Research Board*.

Li, Z., Chen, C., Wang, K. (2011). Cloud computing for agent-based urban transportation systems. *IEEE Intelligent Systems*, *26*, 73–79.

Lin, S., Zhou, Z., Xi, Y. (2013). Analysis on Performance Criteria for Model-based Traffic Congestion Control in Urban Road Networks. *TRB 2013 Annual Meeting*, 1–18.

Lu, W., Liu, C., Thomas, N., Bhaduri, B. L., Han, L. D. (2015). Global System for Transportation Simulation and Visualization in Emergency Evacuation Scenarios. *Transportation Research Record: Journal of the Transportation Research Board*, *2529*, 46–55.

Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., Ghalsasi, A. (2011). Cloud computing — The business perspective. *Decision Support Systems*, *51*(1), 176–189.

Marx, V. (2013). Biology: The big challenges of big data. *Nature*, *498*, 255–260.

Pascale, A., Mavroeidis, D., Lam, H. T. (2015). Spatiotemporal Clustering of Urban Networks. *Transportation Research Record: Journal of the Transportation Research Board*, *2491*, 81–89.

Philip Chen, C. L., Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314–347.

Richards, N. M., King, J. H. (2013). Three Paradoxes of Big Data. *Stanford Law Review Online*, *66*, 41–46. Retrieved from http://www.stanfordlawreview.org/sites/default/files/online/topics/66_StanLRevOnline_41_RichardsKing.pdf

Sagiroglu, S., Sinanc, D. (2013). Big data: A review. *International Conference on Collaboration Technologies and Systems (CTS)*, 42–47.

Suchkov, B., Mikhail, B., Reddy, A. (2014). Development of a New , Lightweight GTFS Real Time Stringlines Tool to Visualize Subway Operations and Manage Service at New York City Transit. In *TRB* (pp. 1–19).

Suciu, G., Vulpe, A., Halunga, S., Fratu, O., Todoran, G., Suciu, V. (2013). Smart cities built on resilient cloud computing and secure internet of things. *Proceedings - 19th International Conference on Control Systems and Computer Science, CSCS 2013*, 513–518.

Tene, O., Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property Volume*, *11*, 240–273. Retrieved from http://heinonlinebackup.com/hol-cgibin/get_pdf.cgi?handle=hein.journals/nwteintp11&section=20\nhttp://ssrn.com/abstract=2149364

Villars, R. L., Olofson, C. W., Eastwood, M. (2011). Big Data: What It Is and Why You Should Care. *IDC*, *6*, 1–14.

Wang, X. (2015). Calibration of Big Traffic Data for a Transport Smart City. In *CICTP* (pp. 387–397).

Weisbrod, R. (2011). The Geography of Transport Systems. *Journal of Urban Technology*, *18*, 99–101.

Xu, X. (2012). From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, *28*(1), 75–86.

Yuan, Y., Wang, H., Wang, D., Liu, J. (2013). On interference-aware provisioning for cloud-based big data processing. *IEEE International Workshop on Quality of Service, IWQoS*, 201–206.

Zaslavsky, A., Perera, C., Georgakopoulos, D. (2012). Sensing as a Service and Big Data. *Eprint arXiv*.

Zeng, Q., Reddy, A., Lu, A., Levine, B. (2014). Develop New York City Surface Transit Boarding and Alighting Ridership Daily Production Application Using Big Data. *Draft for Trb 15*, 1–25.

Zheng, Z., Zhu, J., Lyu, M. R. (2013). Service-generated big data and big data-as-a-service: An overview. *Proceedings - 2013 IEEE International Congress on Big Data, BigData 2013*, 403–410.

Zimmermann, A., Pretz, M., Zimmermann, G., Firesmith, D. G., Petrov, I., El-Sheikh, E. (2013). Towards service-oriented enterprise architectures for big data applications in the cloud. *Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC*, 130–135.

Zissis, D., Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation Computer Systems*, *28*(3), 583–592.

http://bigapps.nyc/p/

http://datasf.org/

http://dev.socrata.com/

http://lbs.amap.com/?spm=0.0.0.0.Fxjn2L

http://onebusaway.org/

http://sf.streetsblog.org/2010/01/05/how-google-and-portlands-trimet-set-the-standard-for-open-transit-data/, accessed 2015-9-12

http://www.bjjtw.gov.cn/bmfw/jtzs/

http://www.census.gov/

http://www.citygoround.org/

http://www.digitalconvergence.eu/activities.html, accessed,2015-9-9

http://www.opengovpartnership.org/

http://www.research.ibm.com/index.shtml

http://www.transitchicago.com/apps/

http://www.transitchicago.com/data/

http://www.transitime.org/

https://data.cityofchicago.org/

https://tfl.gov.uk