Big data logistics: a health-care transport capacity sharing model

Rashid Mehmooda*, Gary Grahamb

aKing Khalid University, Abha, Saudi Arabia
bUniversity of Leeds, Leeds, United Kingdom

Abstract

The growth of cities in the 21st century has put more pressure on resources and conditions of urban life. There are several reasons why the health-care industry is the focus of this investigation. For instance, in the UK various studies point to the lack of failure of basic quality control procedures and misalignment between customer needs and provider services and duplication of logistics practices. The development of smart cities and big data present unprecedented challenges and opportunities for operations managers; they need to develop new tools and techniques for network planning and control. Our paper aims to make a contribution to big data and city operations theory by exploring how big data can lead to improvements in transport capacity sharing. We explore using Markov models the integration of big data with future city (health-care) transport sharing. A mathematical model was designed to illustrate how sharing transport load (and capacity) in a smart city can improve efficiencies in meeting demand for city services. The results from our analysis of 13 different sharing/demand scenarios are presented. A key finding is that the probability for system failure and performance variance tends to be highest in a scenario of high demand/zero sharing.

Keywords: future city, Big data; transport operation management; healthcare information systems; integrated systems; shared resources

1. Introduction

The growth of e-commerce, home delivery, automated packaging stations and click & collect services are pushing the limits of existing city-network designs. Big data logistics will need to support “omni-channel” retail models, smaller store formats, increased intensity of deliveries, coordinate multiple trans-shipment points, engage a wider...
range of vehicle technologies – including electric and autonomous vehicles – and support complex inventory balancing and deployment strategies. Both large and small cities are proposing a new model called the “smart city” which represents highly technical, sustainable, comfortable and secure living environment.

Big data logistics can be defined as the modelling and analysis of (urban) transport and distribution systems through large data sets created by GPS, cell phone and transactional data of company operations, combined with human generated activity (e.g. social media, public transport)\(^3\). The logistic industry is undergoing a fundamental shift from “product-related” services to “information related” services. The demands and requirements are literally changing on a daily basis with the innovations in technologies with smart computing. Increasingly the real time tracking of vehicles could facilitate more accurate resource pooling and capacity sharing.

The introduction of “eco-vehicles”, particularly Fully Electric and Hybrid Vehicles (FEVs, PHEVs) for city logistics operation is becoming a viable option for local administrations and logistics service providers addressing sustainability policies. In most cases electric vehicles are vans and small trucks (up to 3.5t) but also other types of FEVs that started to be used for operating last mile and several forms of B2C services, like the cargo cycles used in the Petite Reine scheme in Rouen (FR) or Gnewt Cargo scheme in London. Besides last mile services, FEVs are also often used to support sustainable own-transport services (for shops, businesses and citizens) like van sharing schemes. Overall, the surveyed best practices operating FEVs have shown that electric vehicles bring clear benefits as regards the abatement of exhausted gases, CO2 and noise emissions. Not least, FEVs are accepted by the public and have an “image” which may be a helping factor for the introduction of new sustainable logistics services.

The logistic firms require more technical and technological supports to handle the three V’s of Big Data & analytics that is “Volume”, “Variety” and “Velocity”\(^4\). Health care providers have to adapt to the changing customer demands, while at the same time exploit the availability of new data sources and management frameworks. Our study aims to provide new understanding about load (capacity) sharing and optimization in a smart city context. The primary purpose of this investigation is to explore using Markov models the integration of big data/smart cities with future city (health-care) transport capacity sharing. There are several reasons why the health-care industry is the focus of this investigation. For instance, in the UK various studies point to the lack or failure of basic quality control procedures and misalignment among customer needs and provider services and duplication of logistics practices. A recent Audit Scotland report\(^1\) revealed wide variation in costs per emergency call and the way that resources were being deployed to meet demand. Examples were cited of ambulances queueing outside hospitals. Another concern raised was that of the high variance in ambulance performance (response lead time) and even factoring in for geographical distortion, this was above acceptable statistical process control limits. The report concluded that the ambulance services greatest future challenge is to improve efficiency in its resource base and be more effective in managing demand variation. Further that different hospitals (trusts) needed to work more closely together with each other to utilise capacity efficiencies and share best practice. This was even more pressing given the dramatic rise in emergency call volumes from 4.4m in 2000/01 to 7.9m in 2009/10.

The development of smart cities and big data present unprecedented challenges and opportunities for operations managers; they need to develop new tools and techniques for network planning and control. Our paper aims to make a contribution to big data and city operations theory by exploring how big data can lead to improvements in transport capacity sharing. We explore using Markov models the integration of big data with future city (health-care) transport sharing. This approach is justified as it allows the complexity of decision making at a city-system level to be investigated. Therefore a mathematical model was designed to illustrate how sharing transport load (and capacity) in a smart city can improve efficiencies in meeting demand for city services. The results from our analysis of 13 different sharing/demand scenarios are presented. A key finding is that the probability for system failure and performance variance tends to be highest in a scenario of high demand/zero sharing. Our work complements and extends that of Fosso-Wamba and Sertia & Patel who theoretically link together big data to building city absorptive capacity. These initial findings are part of a long run investigation into how load sharing could lead to efficiency improvements at a city-operational level. Such ideas are radical changes driven by, or leading to more decentralized rather than having centralized transport solutions. Finally, the limitations of our work are presented. It is clear that while Markov analysis begins to reveal the inherent complexity of health care provision there is a need for more empirical case data if we are to complete fully our decision modelling.
2. Literature – health care logistics

In many communities, under the influence of scarcity of transport, patients’ accessibility to gain sufficient medical treatment could be reduced. As a result, the best medical service in the world would be worthless if the intended recipient cannot get access to the service. There are many causes of transport problems, such as: (a) poor coordination of transport, that is, different facilities and authorities within a district control their vehicles separately, which leads to inefficient planning; (b) lack of vehicles; (c) abuse of government vehicles; and (d) poor maintenance and repair of vehicles (ibid., p. 140). The strategic approach to transport service provision in health-care has traditionally been modular and piecemeal, separating patient and non-patient services, and with regard to individual efficiencies rather than a more holistic effectiveness. Because provision is fragmented, management control is problematic to the extent that identification of the different aspects and their associated costs are not readily available. When more services are moved from traditional acute settings and into the community to provide care closer to home, the greater the demand on the provision of different transport models, and the greater the demand for a more integrated approach to transport provision between care providers in the district.

3. Integrative framework

A framework was developed by integrating together the literatures on smart city transport, big data logistics and capacity sharing. The resulting integrating framework is presented in Figure 1. Operationally the transport management system receives requests from the residential areas for transport, obtains the required information through GPS or the electronic health records (EHR) and in real time, synchronously organizes the relevant transport arrangements.

3.1. Capacity sharing

Formal processes of transport systems sharing (such as Uber, Zipcar and the many city bike schemes) are now available to subscribers. These use technology, principally GPS location sensors and Internet of Things connectors, to provide a joined-up service across large areas, supported by mapping, coordinating and payment infrastructure on a large scale. Uber (https://www.uber.com/) is like Airbnb in that it allows private individuals to use spare capacity: it brings together the drivers of private cars with time on their hands with those needing a ride, through a service that coordinates this pairing, handles payment and registers/vets. It is a paid taxi service, not a ride-sharing scheme. We are also observing the rise of bike sharing systems. By contrast, the provision of timeshare bikes as a city service in London is closer to the provision of buses, with a hop-on-and-off quality that cannot be booked, though also with digital technology as enabler. This is used similarly, to register, unlock, record duration and re-dock, but also to
monitor clumping and trigger vans to transport bikes between locations, so that docking points are neither full nor empty as peak flows move bikes unequally round the city. As we have noted, when distribution of resources is in municipal control, it is largely invisible as sharing per se, though here it is the provision, not the distribution, that Transport for London oversees. Orsi comments in her analysis of sharing and coordination, that this level of sharing requires formal infrastructure: ‘Whether this is publicly or privately managed, [t]aking sharing to the fourth degree can require getting government buy-in, mobilizing multiple players (legislators, investors, banks, developers, planners, etc.), or even restructuring our communities.’ It is, of course, much harder to move car gluts around than bikes, and generally you return the car to where you found it. Again, the emphasis is on access as needed, replacing ownership of dormant resources.

3.2. Smart city transport

Two items come across as central in most descriptions on smart cities: the transportation/logistics aspects (includes e.g. disaster and emergency management), predominately from a sustainability point of view; and new technology to facilitate the organizing of smart city activities, including the capturing of data and its analysis. As for technology, smart cities are still associated with either sensor or household data. Big data is seen as central to smart cities, in how large sets of data would inform about activities of different city actors. Big Data refers to the emerging technologies that are designed to extract value from data having four Vs characteristics; volume, variety, velocity and veracity. It has also been defined as large pools of unstructured data that can be stored, managed, and analyzed. It refers to the emerging technologies that are designed to extract value from data having four Vs characteristics; volume, variety, velocity and veracity. Integration of city systems is an important sub-theme of city government-led smart city visions and plans, which though seems to suggest that data is pooled rather than integrated, and that analyses need to establish interaction patterns rather than be based on how actors actually interact. The big data analyses are referred to as what bring meaning to the data through interlinking it, while it in its capturing is unstructured and unconnected. The logistics/transport aspects point to how transportation would need to be reorganized so as to deal with CO2 footprint. The logistic aspects could be seen as a move from individual firms optimizing their transportations, over collaborative, or system level analyses of flows, to redrawing the landscape and focus on local production, and thereby foremost short-distance transportations. Ideas are radical changes driven by, or leading to local production, thus rather going from centralized transport solutions to distributed, than reverse. Logistics firms would in a sense lose business, while companies utilizing their offerings would change interaction partners to more local alternatives. The changed interaction patterns would hence transform entire interaction systems, with geography increasing the impact on interaction decisions.

3.3. Big data logistics

The explosion of measurements and statistics produced by and available to cities - the emergence of big data - is providing new opportunities for citizen engagement and citizen-led innovation. City authorities and communities can also use ever-growing bodies of data to improve understanding of citizen behaviour and service usage and build transparency and accountability by opening up their records and statistics for public consumption - the growth of “open data”. With the growth of technology and datasets also come new piracy surveillance and data misuse challenges for future cities. Cities also face challenges around data quality, comprehensiveness, data collection and analysis particularly aligning data from data sources and managing the sheer volume of data which is produced. Big data need to be robust, accessible, and “interpret-able” if it is to provide cities and companies with meaningful opportunities and solutions. Big data analytics is the process of examining large amounts of unstructured data to uncover hidden patterns, unknown correlations and other useful information. Smart cities provide an ideal background for exploitation of big data and the interactions in the value chain can generate “exhaust data”. Indeed many big data applications are implemented far from the purposes for which the data was collected. For example, location information that cell companies gather (so that they can efficiently route calls) can be used to make predictions. The transport network applications of big data can be utilised to the key operational processes. For
example, big data is useful to define mobility strategies based on actual consumer patterns (e.g. location based data generated by mobile phones) rather than surveys and samples. Additionally transport planners can use big data algorithms (instead of small data samples) to fine tune mobility planning based on real-time in store and online sales\(^{10}\). In the next section we mathematically model using Markov chains the interplay of big data logistics, smart city transport and capacity sharing.

4. The Model

We extend our earlier work\(^{12}\) on modeling logistics of healthcare systems. The Markov model that we have developed in this paper focuses on matching the transport demands (of patients) with city health-care transport service provision. This model was designed to illustrate how sharing of transport capacity in a smart city can improve efficiencies in meeting patient demand for city health-care services. First we construct a model of healthcare in a future city in Northern England. Our mathematical model to acquire and calculate the traffic demand in a future city is derived as follows. We consider a future city healthcare mobility service system where there are three main types of nodes of interests. These are Residential Areas (RA), (Primary Healthcare) Medical Centers (MC), and Hospitals (H). We will see later that this model can be extended to contain additional types of nodes of interests (healthcare related or otherwise). The mobility management system we propose here exploits big data technologies and keeps static and real time data about the healthcare related supply and demand. The total demand for healthcare related mobility in the city can be calculated as follows:

\[
\lambda_T = \lambda_{ra,mc} + \lambda_{ra,h} + \lambda_{mc,h} + \lambda_{mc,ra} + \lambda_{h,ra} + \lambda_{h,mc}
\]

(1)

where \(\lambda_T\) is the total arrival rate for the healthcare related mobility demand in the city, \(\lambda_{ra,mc}\) is the arrival rate for the mobility demand from a residential area to a medical center, \(\lambda_{ra,h}\) is the arrival rate for the mobility demand from a residential area to a hospital, \(\lambda_{mc,h}\) is the arrival rate for the mobility demand from a medical center to a hospital, \(\lambda_{mc,ra}\) is the arrival rate for the mobility demand from a medical center to a residential area, and so on. The arrival rate for the healthcare related mobility demand in the city in each of these cases is represented as the number of kilometers demand per hour. The Medical centers and the hospitals are built in such a way to localize and minimize the traffic. However, there will still be traffic between RAs and remote MCs and hospitals due to reasons such as medical specialization, user/patient preferences, etc. Note that we use the word ‘user’ because current trends are towards preventive healthcare where people will try to stay proactively healthy rather than being reactive and going through treatment once became ill. The individual arrival rates are in km demand per hour for the mobility between two types of nodes of interest; these can be calculated as:

\[
\lambda_{ra,mc} = D_{ra,mc} \circ \xi_{ra,mc}
\]

(2)

where \(D_{ra,mc}\) is the distance between a residential area and a medical center, and \(\xi_{ra,mc}\) is the number of requests per hour received for mobility service required between a residential area and medical center. The Hadamard product (\(\circ\)) in Equation (2) will produce \(\xi_{ra,mc}\) where its element \((i,j)\) is the product of elements \((i,j)\) of the matrices \(D_{ra,mc}\) and \(\xi_{ra,mc}\). The arrival rate \(\lambda_{ra,h}\) (and the other arrival rates) included in Equation (1) can be calculated similarly as given by Equation (2). Note that the distance \(D_{ra,mc}\) can be a static, dynamic, or real-time value obtained through connecting to the application programming interface (API) of a geographical information system, web mapping software, etc (e.g. Google maps). A real-time value from an appropriate navigation software can allow options for intelligent decisions such as fastest or shortest route depending on the context, user or system preference, etc. Similarly, the number of requests \(\xi_{ra,mc}\) can also be a static, dynamic, or real-time value obtained through a sub-system that produces these values using historical and real-time data. The resources for sharing and the vehicle sharing related data can also come from applications mentioned in Section 3.1 such as Zipcar.

There are a total of \(M\) residential areas, \(N\) medical centers and \(P\) hospitals. Therefore, the terms, such as \(\lambda_{ra,mc}\), represent a matrix. That is, suppose there are 10 residential areas and 20 medical centers, \(\lambda_{ra,mc}\) can take 200 values...
(a 10 × 20 matrix), where \( \lambda_{1,1} \) represents the arrival rate of mobility demand between residential area 1 and medical center 1 in terms of km demand per hour. Similarly, \( \lambda_{10,20} \) is the arrival rate of mobility demand between residential area number 10 and medical center number 20. The terms \( D_{ra,mc} \) and \( \xi_{ra,mc} \) are also matrices with same dimension as for \( \lambda_{ra,mc} \). Note here that \( \lambda_T \) in Equation (1) cannot be obtained as a scalar because the terms on the right hand side of the equation are matrices of different dimensions. One possibility is to obtain \( \lambda_T \) by direct sum (operator \( \otimes \)) of the terms on the right hand side in Equation (1). Essentially, it will create a matrix containing demand values for all (source, destination) pairs. The resultant matrix will be a square matrix of dimension \( R \), where \( R = 2 \times M \times N \times P \). That is, we will have source to destination demands of all the nodes of interest in the mobility network. In this paper, for demonstration purposes we have calculated the total demand as a scalar value by taking average values of the distances between the nodes of interest and the number of requests for mobility between nodes of interest.

### Table 1 Data and calculations for the model

<table>
<thead>
<tr>
<th>( \lambda ) (demand: km/h)</th>
<th>( D_{ra,mc} ) (for all arrival nodes)</th>
<th>( \xi_T ) (ra, mc, h)</th>
<th>nodes of arrival</th>
<th>( \xi_{ra,mc} ) (per arrival node)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{ra,mc} )</td>
<td>2000</td>
<td>5</td>
<td>400</td>
<td>10</td>
</tr>
<tr>
<td>( \lambda_{ra,h} )</td>
<td>800</td>
<td>20</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>( \lambda_{mc,h} )</td>
<td>400</td>
<td>20</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>( \lambda_{mc,ra} )</td>
<td>1940</td>
<td>5</td>
<td>388</td>
<td>5</td>
</tr>
<tr>
<td>( \lambda_{h,ra} )</td>
<td>1040</td>
<td>20</td>
<td>52</td>
<td>5</td>
</tr>
<tr>
<td>( \lambda_{h,mc} )</td>
<td>160</td>
<td>20</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>( \lambda_T )</td>
<td>5540</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now consider that the city has a total of 5 residential areas, 10 medical centers, and 2 hospitals. The city planners calculated that on average the mobility service system will receive 40 requests per hour for transport from residential areas to each medical centre (see Row 2, Table 1). There are ten MCs and hence the total arrival rate of people is 400. The average distance between an RA and an MC is 5 KM, hence the total mobility demand from RA to MC (\( \lambda_{ra,mc} \)) is 2000 km per hour (reader should not confuse this demand with speed units). The total mobility demands for other nodes of interest and the relevant data are given in Rows 3 to 7 of Table 1, and the total healthcare related mobility demand (\( \lambda_T \)) for the whole network is given in the last row. The planners considered this value of \( \lambda_T \) and built a capacity for healthcare related mobility in the city for satisfying 5600 km (slightly above the calculated \( \lambda_T \)) demand each hour between the nodes of interest.

5. Results and Analysis

We now consider a range of scenarios where there are various levels of vehicle sharing and therefore the demand for vehicles varies accordingly. The planners considered a total of 13 such scenarios. The best case scenario is one of a high level of vehicle sharing and demand of 500 km transport per hour for the whole city. At the other extreme we have the worst case scenario where there is no vehicle sharing and we have the highest possible demand (5540 KM demand per hour). In this scenario we are very close to the total service capacity of the healthcare-related transport service. The planners considered 10 other scenarios in between these two extremes with total transport demands of 1000, 1500, ..., 5500 km per hour linked to the corresponding levels of vehicle sharing (high to low). Another scenario with demand of 1km was selected so that it is near zero demand to show that such a system will have high probability of being in the near idle states.
We use the model from the previous section and the above given data and build a simple Markov model to understand the long run (steady state) transport demand for various vehicle sharing scenarios considered by the planners in the future city. It should be noted that the future city were actively pursuing smart health policies to reduce the actual level of demand for its medical services (e.g. bicycle sharing schemes, gym sharing, healthy lifestyle campaigns). Figure 2 presents the results obtained by solving the Markov model (see \[13, 14\] for solution methods). The horizontal axis is used for the states of the system. State 1 means that the system is idle and there is no request in the system for transport. State 2 implies that there is a request for 1km transport and it is being processed by the system. The final state number is 5601, which implies that the system is operating on its full capacity with 5600 KM transport demand being processed by the system. The vertical axis gives the steady state probabilities for the system. The results for the 13 scenarios considered by the city planners are illustrated in the figure by the 13 separate plots. Each plot gives the probability of the system to be in each of the system states for a particular (demand/sharing) scenario. Consider the plot for the idle scenario with mobility demand of 1km per hour. The plot shows that the system will operate with a very low number of jobs with very high probability. The actual probability results show that such a system will be in the state number zero (no job in the system) with probability 0.9944, and in state number one with probability 0.00018. Note also the slope and location of this plot relative to the other plots. The best case scenario of highest sharing with the transport demand of 500km per hour shows that the system will operate with a very low number of jobs with very high probability (the probability of states numbered zero and 500 are 0.21267 and 0.00007, respectively). Note that the plots with a higher transport demand shift the peak of the curves (higher probability) towards the right of the figure: the probability curve for the worst case scenario (5540: no sharing) has the peak near the highest numbered states. The interpretation of the results in the figure is that the system will be in danger of crashing (or failing) if the per hour demand reaches the worst case scenario (of near full capacity demand and zero sharing).

**Implications for city healthcare logistics planners:** In respect to optimization the strategy of the city health-care transport planner would be to co-ordinate with other city services to work strategies to reduce demand on the system (e.g. by improving the health and well-being of the city population, car free cities and greater use of bike sharing) while at the same time encouraging vehicle sharing schemes. The potential for system failure and performance variance seems to be when there is high demand for city health care services but zero sharing.

6. Conclusions

The development of smart cities and big data presents unprecedented challenges and opportunities for operations managers: they need to develop new tools and techniques for network planning and control, and the increased transparency and convenience that can be derived from smart city infrastructure and services call for the
development of new operations models. The paper aims to begin to make a contribution to theory by presenting the potential of big data to facilitate a city-network perspective to capacity sharing decision making, which is more efficient than individual health care transport schemes taking independent decisions, which often leads to duplication and inefficiency with ambulance capacity failing to meet volatile and rapidly changing demand with resulting unacceptable levels of performance variance, in particular in dealing with emergencies. Our primary purpose was to build a framework and to provide initial Markovian results investigating the interplay of big data and smart cities with transport sharing and to assess how this could improve performance. To advance the framework and preliminary Markovian model we intend to extend our research investigation through intensive case studies of health care transport operations in the UK, US, France and the Middle East. We emphasise the importance of “big data” orientations and related management and operations issues to be analysed with Markovian theoretical framing as an area in which further research is urgently needed. Future operational performance is linked with these sharing orientations which can ensure unique service delivery competitive advantage and urban performance. Further case studies are therefore needed to explore load optimization in areas such as “bike sharing”, “manufacturing plant location/freight delivery” and “waste management” with preferably a time-series longitudinal dimension (where we can explore dynamic as well as static constructs over time). This will enable a series of cases to be investigated and emerging theoretical constructs to be identified, advanced and tested through a comparison of inter-case variance.

References