Building Smarter City through Big Data - Best Practices in Seoul Metropolitan Gov.

Ki-Byoung Kim[1]

Abstract: Since 2013, Seoul Metropolitan Government (SMG) has introduced big data initiatively in administration and put into practices in transportation, safety, welfare in order to overcome limited resources and conflicting interests. For establishing a new midnight bus service, SMG prepared optimized midnight bus routes by analyzing big data from mobile phone Call Data Record (CDR) through collaboration with a telecommunication company. Despite of limited budget and resources, newly identified routes can cover over 42% of the citizen with 9 routes and less than 1% of buses compare with day time operation. In addition to solve transportation problem, SMG utilizes big data to resolve location selection problem for choosing new facility locations such as life double cropping centers and senior citizen leisure centers. As results, SMG demonstrates big data as a good tool to make policies and to build smarter city by overcome space-time limitation of resources, mediation of conflicts, and maximizes benefit of the citizen.

Keywords: Smart city, Night bus route optimization, Location optimization, Big data modeling, Big data administration

I. Introduction

A. Background

Smart city has two important characteristics such as utilization of IT infrastructure and overcome of resource limitation [1][2]. SMG recently has utilized big data as one of the means to build smarter city.

Due to a small twitter request from a university student, SMG started using big data to establish night bus service in 2013. Based on a success of route optimization by big data, SMG expands big data administration into 9 area including taxi analysis, location analysis of welfare facilities, and traffic accident analysis, etc. in 2014.

II. Route Optimization of Night Bus

A. Problem definition

SMG planned to provide midnight bus (Owl bus) service in early of 2013, however, there were several problems SMG had faced such that there were very small experience in operation of night bus in Seoul at that time. Among them, the most important problems were to find transportation demand such as

1) where the passengers were, and
2) where they wanted to move in mid-night

in addition to limited resources and budget problems.

B. Modeling administration problems into data problems

While these problems could be handled with traditional approaches by transportation experts, due to increased value of big data and data-smart governance, it was a good chance for SMG to adopt big data to solve administration problems [3]. SMG decided to adopt big data to sort out these two transportation questions. To make problems more manageable and find optimized bus routes, these administration problems were transformed into data problems at the first stage. These steps are

1) Administration problems were transformed into big data problems
2) Demand/supply questions of public transportation in midnight were modeled as data questions
3) To solve data problems, big data analysis was adopted.

Transportation demand based on locations and origin-destination (OD) was modeled. Transportation demand based on a location was identified by Call Data Record (CDR) provided by a telecommunication company. In fact, a count of CDR was not exactly matched with de facto population of a certain location in the city. So, the number of count are normalized by eliminating double counts and resident population counts. In case of OD demand, the problem was more complex, so, it needed more data.

In order to find OD demand, billing addresses of each CDR were used. Contract subscribers had their own billing addresses, and, fortunately, a billing address can be identified as home or work, and, during the midnight, it could be assumed that the most of the passengers not in their home needed to return home located in their home address. Then, OD data could be easily created by associating CDR and billing addresses as home.

Based on this model, 3 billion call data during 1 month and associated billing addresses in Seoul metropolitan area were analyzed and figured out transportation demand and direction in midnight clearly.

C. Simplifying data problems

[1] Chief Data Officer, a director of data & statistics division, Seoul metropolitan gov., Seoul city hall Jung-gu Seoul 04515, kbkim@seoul.go.kr, pskbkim@gmail.com

In order to find optimized routes of Seoul owl bus, big data analysis was tried. As big data was unstructured and very large in volume, it took 2 weeks to analyze one time for the first trial. So, simplified model was needed to reduce analysis time and put more time to find optimize routes rather than analysis itself.

Total area of Seoul metropolitan is about 605 km^2. According to a survey to the citizen, we paid attention to the citizen's behavior to take public buses. They were willing to walk about 500 m to take public transportation. In order to simplify the data model of de facto population, whole area of Seoul had been remodeled into 500m radius circles, and, in order to remove duplicated area between circles, circles were remodeled as hexagons.

After changing the data model from whole unstructured Seoul area into 500m radius hexagons, total number of hexagons became 1,252. All transportation demand were put into these hexagons. Then, number of analysis became dramatically reduced, and, the analysis team could speed up transportation demand analysis of each locations and OD in a day from 1 time per 2 weeks.

D. Results

A big data practice of Seoul Owl bus demonstrates big data analysis one of the effective approaches to solve administration problems innovatively as a leading case.

Owl bus service plan was considered from the early year of 2013, so, budget and resources were very limited because the planning had been finished the previous year. However, big data analysis found optimized 9 routes based on transportation demand in midnight, Owl bus could start running on well optimized routes even with very limited buses. Compared with day services, Seoul Owl bus runs less than 0.5% of buses, but, covers more than 42% of midnight transportation demand from the citizen.

From the perspective of operation efficiency, Owl bus on optimized routes can carry 10% more passengers at the most than the night bus on non-optimized routes.

III. Additional Practices

A. Additional Problems

Location selection is one of the frequent questions in smart cities as well as traditional cities. SMG applied big data analysis to select the best location of welfare facilities such as senior citizen leisure facilities and second life double cropping centers. Due to limited budget and resources, there were more demand than the capability to build centers every year.

In order to identify best locations of each facilities, SMG utilized de facto population from big data analysis as well as residential population in each time of a day in addition to traditional demographic statistics of the citizen. Using additional big data analysis, SMG could choose the best location of second center for second life double cropping center in Jongro-gu (district) against 3 candidate locations for maximizing benefit to the citizen.

Safety problems were also put in practices by analyzing operational data of public transportations in Seoul, and, policies for the safety of children and senior citizen safety has been developed and run since 2015.

B. Modeling more problems into data problems

In order to utilize floating population as well as residential population, Seoul metropolitan area has been modeled as 50m by 50m square for multipurpose, so for the most of the administration problems, floating population can be used based on this model. In case of transportation data, node & link model has been adapted. Node & link model was originally developed to show traffic information along with road and cross in navigational system [4]. And, SMG modified the length of links into 150m each which are suitable to represent vehicle information on the road in at least 10 seconds.

Safety policies were also put into practices by analyzing big data produced by public transportations in Seoul, and, policies for the safety of children and senior citizen safety has been developed and run since 2015.

C. Results

Location selection based on big data could choose the best location based on demand from the citizen represented by big data. Traffic accident analysis could find weak spots of minorities such as children or senior citizen who could not explain their own problems strongly to the metropolitan government. Big data analysis has been a good tool to find weak spots of minorities who has small or no voices in policy making.

In a different perspective, data could be a good tool to mediate different demands and conflict requests. Without it, arbitration may be concluded based on stronger requests or louder voices.

IV. Summary

Seoul Owl bus demonstrates big data analysis as one of the good approaches to solve administration problems. SMG models data into simplified one to make problems simpler. Area of SMG and traffic data has been modeled with simplified one without losing important insights. We found and verify that the transformation of administration problems into data problems make the original problems more manageable and simplified model of big data make analysis more practical.

References

[1] Hyojin Hong, Junghoon Sim, "Future Cities through Smart City", IT & Future Strategy, Vol. 2010, NIA, No. 2010-13, 2010.
[2] Jongsung Hwang, Jaepil Cha, "Fever of Smart City in Overseas and Implication", IT & Future Strategy, NIA, No. 2013-11, 2013.
[3] Stephen Goldsmith, Susan Crawford, The Responsive City Engaging Communities through Data-Smart Governance, Jossey-Bass, 2014.
[4] Harvey J. Miller, S. Shaw, Geographic Information Systems for Transportation, Oxford University Press, 2001.