Priority-based Data Collection Framework for Smart Cities

İzzet Fatih ŞENTÜRK, Siratigui Coulibaly

Bursa Technical University, Dept. Computer Engineering, 16310, Bursa, TURKEY

Abstract

Smart cities rely on continuous data collection from a vast number of sources including sensors and even their citizens via their mobile devices. Collected data enable monitoring urban phenomena and organizing infrastructure and services with improved utilization. For a better decision making, a unified platform that share the key information between different stakeholders can avoid fragmentation of data sources and enable data exchange for big data analysis. Sensor connectivity is crucial for data fidelity. However, sensors are typically conserved in terms of energy and employ low-rate wireless communication means. Consequently, network-wide connectivity is not always guaranteed. This papers assumes a smart city data platform that employs a mobile data collector to visit sensors that cannot deliver their data. To designate an efficient trajectory to collect data while avoiding data loss due to buffer overflow, it is important to consider different priorities and sampling rates for different sensors. This paper presents a conceptual framework to classify sensor priorities based on spatial features such as location and the nearby amenities for smart city applications.

Keywords: *Smart City, Data Collection, OpenStreetMap, Connectivity*

Akıllı Şehirler İçin Öncelik Temelli Veri Toplama Sistemi

Özet

Akıllı şehirler, algılayıcılar ve hatta akıllı cihazları sayesinde sakinlerinden topladıkları verilerle beslenirler. Toplanan veri, şehirde olup bitenlerin takibini ve altyapı ile hizmetlerin organize edilerek daha etkin kullanımını sağlar. Karar alma sürecini iyileştirmek için, kilit bilginin paydaşlar arasında paylaşımını sağlayan tümleşik bir platform, veri kaynaklarının parçalı yapısının önüne geçerek büyük veri analizi için veri paylaşımının önünü açar. Algılayıcıların bağlantı sürekliliği, veri doğruluğu için çok önemlidir. Fakat algılayıcılar genelde kısıtlı enerji kaynaklarına sahiptir ve dolayısıyla düşük güç tüketen kısa menzilli kablosuz iletişim yöntemlerini kullanırlar. Sonuç olarak, ağ çapında bağlantı her zaman mümkün olmayabilir. Bu bildiri, topladığı verileri aktaramayan algılayıcılardan veri toplayan hareketli veri taşıyıcılarının olduğu bir akıllı şehir veri platformu önermektedir. Veri toplama sırasında bellek taşması probleminin önüne geçecek etkin bir güzergâh belirlenebilmesi için farklı algılayıcılar için farklı önceliklerin ve veri üretme oranlarının dikkate alınması gerekmektedir. Bu bildiride, algılayıcıların, konum ve yakında bulunan tesislerin türü gibi mekânsal özelliklere bağlı olarak sınıflandırıldığı kavramsal bir çerçeve sunulmaktadır.

Anahtar Kelimeler: *Akıllı Şehir, Veri Toplama, OpenStreetMap, Bağlanabilirlik*

1 Introduction

The rapid urbanization of the world requires efficient use of natural resources in order to ensure a sustainable economic growth while providing a higher quality of life for residents. Disruptive digital technologies enable the transformation of urban areas into smart cities with the availability of data to manage assets and resources efficiently. Data is considered as one of the major drivers of smart cities. The prevalence of Internet of Things (IoT) sensors paves the way for collecting huge amount of data from multiple sources including mobile devices, video cameras, environmental sensors, etc [1]. Relying upon big data generated from cities, decision making process can be improved at the city level through data analysis [2]. Governing cities in a smarter way will increase asset utilization and decrease associated costs.

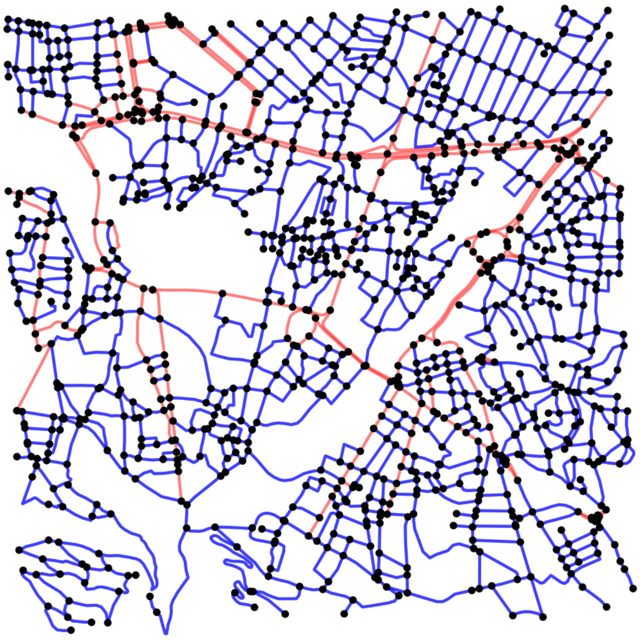

Figure 1. Partial road network for the city of Bursa, TURKEY. From the city center, the nodes within a bounding box of 1000 meters are included. Red color denotes one-way traffic. The figure is reproduced from [4].

Considering its major contribution to gas emissions and the waiting times in traffic, transportation is one of the key components of smart cities that has the potential to improve quality of life with safer, faster, and cleaner mobility with a reduced cost. For instance, smart traffic control can optimize the traffic flow by controlling traffic lights and other signals real-time [3]. Personalized transportation guidance can suggest the best means of transportation in terms of the cost, waiting time, etc. considering location-aware services. Smart parking can suggest the closest free parking space so that the time and resources to look for the parking space can be saved [3].

Smart cities rely on data collection from a vast number of data points across the city [4]. To ensure a certain degree of data fidelity, it is crucial to minimize data latency. On the other hand, a sensor can generate data more frequently than others depending on its location and the activity rate in the proximity. Therefore, it is very important to classify sensors in a smart city application based on their significance so that different data collection schemes can be applied for the individual sensor.

In this paper, we assume a smart city application and employ one of the volunteered geographic information (VGI) systems, OpenStreetMap (OSM) [5], to obtain spatial data. To model the physical world, OSM employs three basic components: node, way, and relation. Node represents a discrete point on the earth's surface associated with the respective latitude and longitude coordinates. Way, on the other hand, defines a polyline with an ordered list of nodes. OSM employs ways to represent roads. Similar to [4], we assume deployment of sensors on the respective node locations in the considered map obtained from OSM for the respective city. A sample road network with the defining nodes can be found in Fig. 1.

Sensors, typically, employ short-range communication technologies such as Zigbee and 6LoWPAN to conserve the limited energy and extend their lifetime. A base station (BS) is assumed to collect the data within the network through multi-hop routing and forward to the data analytics platform. However, network connectivity can be disrupted due to external damage and inhospitable surroundings [6]. Different recovery schemes can be applied to restore network connectivity. Similar to [7], this paper assumes a mobile data collector to visit nodes that cannot forward its data to the BS. Though, we assume that sensors can have different data generation rates depending on their location and the activities in the vicinity. Therefore, it is crucial to classify significance of nodes in the smart city application.

To classify the node significance, we have exploited various spatial features such as amenity types and road types in the vicinity and the speed limits for the respective road segments. In the experiments, we have considered metropolitan cities in Turkey and reported how the number of various importance types changes with the size of the application area.

The rest of the paper is organized as follows. Related work is summarized in Section II. Data platform is described in Section III. Experiments are discussed in Section IV. The paper is concluded in Section V.

2 Related Work

Availability of the mass amount of geotagged data contributed by non-professionals manifests itself in the growing interest in the volunteered geographic information systems. Despite data quality concerns, OSM is regarded as one of the most successful VGI applications. OSM has been studied extensively in the literature. While some studies focus on data quality and reliability [8-9], others considered various applications including parcel characterization [10], response to humanitarian events [11], remote damage assessment [12], etc.

Internet of Things (IoT) is the concept of connecting everyday dumb objects to the Internet so that the object will be digitally identifiable. IoT is regarded as a building block of the smart cities [13]. According to Gartner, IoT endpoints are expected to reach 25 billion by 2021 [14]. 6LoWPAN and LPWAN are two emerging low-power low-rate wireless communication technologies for IoT devices. 6LoWPAN provides short-range wireless IPv6 connectivity over IEEE 802.15.4 based networks. LPWAN, on the other hand, is a long-range non-cellular network protocol. The lack of IPv6 support and severe bandwidth and duty cycle constraints are the main drawbacks of LPWAN.

3 Data Platform

We obtained spatial data from OSM by exploiting OSMnx [15]. OSMnx models the obtained spatial data on a graph structure. In the graph, nodes represent OSM nodes. As discussed earlier, we assume sensor deployment on respective nodes. OSM describes useful and important facilities with the "amenity" keyword on the map. Different amenity values are defined for education, transportation, healthcare, etc. OSM assigns a road type to each road segment with the "highway" keyword. Available values for this keyword includes primary, secondary, residential, etc. The final tag that we consider in assessing the node significance is the maximum speed limit for the road segment. OSM identifies the speed limit with the "maxspeed" keyword. Recall that OSM defines ways with an ordered list of nodes. Since a node can be part of multiple node segments at intersections, we use one of the values for the road type and the speed limit.

4 Experiments

We changed the size of the application area in order to assess the impact of the network size. The size of the application area is controlled with the bounding box distance while collecting data. The bounding box distance is the distance from the city center to each direction (north, south, east, and west). In the experiments, we considered 30 metropolitan cities in Turkey [16] and reported the average result for significance.

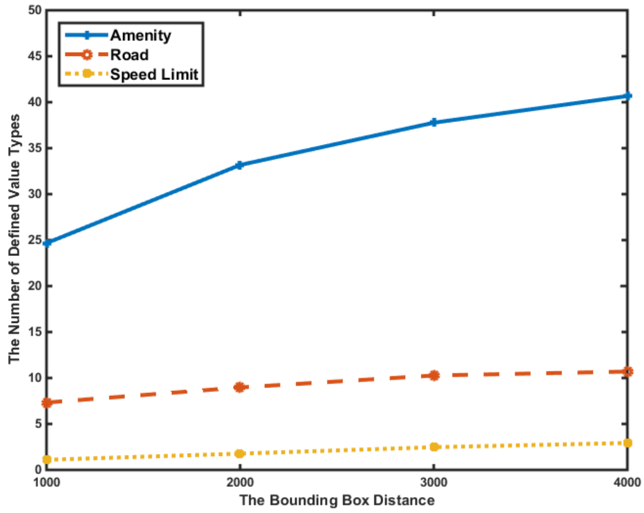

Figure 2. The number of defined values for different spatial features with respect to the bounding box distance.

Since OSM depends on volunteer work, missing or incorrect information is possible. Fig. 2. illustrates the number of defined values for different spatial features with respect to the bounding box distance. Amenity provides the most diversity. The number of defined road types is less than the amenity types. Speed limit is the least diverse spatial feature.

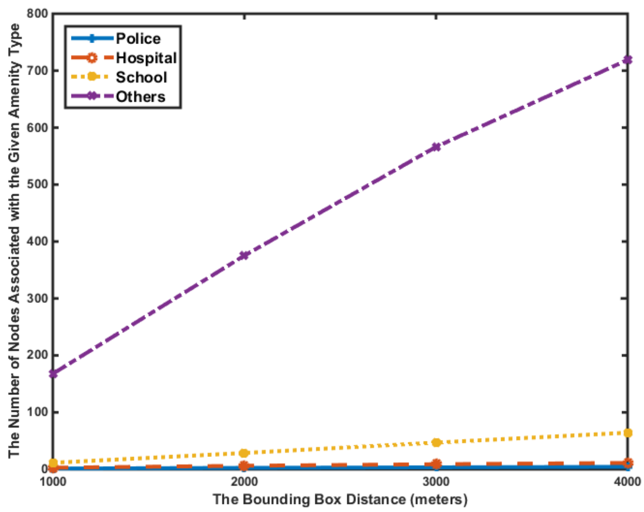

Figure 3. The number of nodes classified based on the amenity type. Available options are police station, hospital, school, and others.

Fig. 3. presents the number of nodes associated with different amenity types. This paper considers four different options while evaluating the node significance based on the amenity type. These options are police station, hospital, school, and others. It can be noticed from Fig. 3. that others represent the largest portion of the amenities. On the other hand, the number of schools is more than

the number of hospitals. The amenity type with the smallest set is police stations.

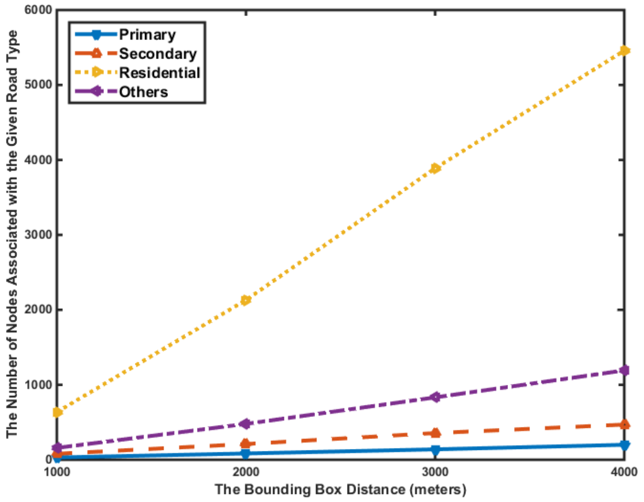

Figure 4. The number of nodes classified based on different road types. Available options are primary, secondary, residential, and others.

Fig. 4. denotes the number of nodes associated with the given road types. It can be noticed from Fig. 4. that most road segments are classified as residential. Primary roads represent the smallest group. The number of secondary roads is more than the number of primary roads.

5 Conclusion

In this paper, we proposed a smart city data platform framework to collect data based on the sensor priority. We employed one of the volunteered geographic information systems to obtain spatial data and assumed sensor deployment to certain locations. Considering the fact that the data generation rate can change based on the sensor location and the activity in its vicinity, we proposed to use three different spatial features to assess the importance of sensors. The idea is considering the nearby facility, road type, and the speed limit of the road segment. We have obtained data from 30 different metropolitan areas of Turkey and reported the number of nodes for different spatial features we have considered.

Acknowledgment

This work was supported by the Scientific and Technical Research Council of Turkey (TUBITAK) under Grant No. EEEAG-117E050. Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org.

References

[1] Yesner R, Ozdemir F. “Understanding Smart City Transformation with Best Practices”. IDC, 2017. (Accessed: 26.09.2019) https://www.isbak.istanbul/wp-content/uploads/2018/05/IDC-Smart-City-White-Paper-sponsored-by-ISBAK.pdf.

[2] Ferrer JR. "Barcelona’s Smart City vision: an opportunity for transformation", Field Actions Science Reports, Special Issue 16, 70-75, 2017.

[3] Deloitte. “Smart Cities, How rapid advances in technology are reshaping our economy and society”. November 2015. (Accessed: 26.09.2019). https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/public-sector/deloitte-nl-ps-smart-cities-report.pdf.

[4] Senturk IF, Kebe GY. “A New Approach to Simulating Node Deployment for Smart City Applications Using Geospatial Data”. The 2019 International Symposium on Networks, Computers and Communications (ISNCC’19), Istanbul, Turkey, 18-20 June 2019.

[5] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org. (Accessed: 26.09.2019). https://www.openstreetmap.org.

[6] Younis M, Senturk IF, Akkaya K, Lee S, Senel F. “Topology management techniques for tolerating node failures in wireless sensor networks: A survey”. Computer Networks, 58, 254-283, 2014.

[7] Senturk IF, Akkaya K. “Mobile data collector assignment and scheduling for minimizing data delay in partitioned wireless sensor networks”. 5th International Conference on Ad Hoc Networks, Barcelona, Spain, 16-17 October 2013.

[8] Brovelli MA, Minghini M, Molinari M, Mooney P. “Towards an Automated Comparison of OpenStreetMap with Authoritative Road Datasets”. Transactions in GIS, Volume: 21 issue: 2, pp. 191-206, 2017.

[9] Arsanjani JJ, Mooney P, Zipf A, Schauss A. “Quality Assessment of the Contributed Land Use Information from OpenStreetMap Versus Authoritative Datasets”. OpenStreetMap in GIScience, pp. 37-58, 2015.

[10] Liu X, Long Y. “Automated identification and characterization of parcels with OpenStreetMap and points of interest”. Environment and Planning B: Planning and Design, Volume: 43 issue: 2, pp. 341-360, 2015.

[11] Palen L, Soden R, Anderson TJ, Barrenechea M. “Success & Scale in a Data-Producing Organization: The Socio-Technical Evolution of OpenStreetMap in Response to Humanitarian Events”. 33rd Annual ACM Conference on Human Factors in Computing Systems, 18-23 April 2015, Seoul, Republic of Korea.

[12] Westrope C, Banick R, Levine M. “Groundtruthing openstreetmap building damage assessment”. Procedia Engineering, vol. 78, pp. 29–39, 2014.

[13] Park E, Del Pobil AP, Kwon SJ. “The role of internet of things (iot) in smart cities: Technology roadmap-oriented approaches”. Sustainability, vol. 10, no. 5, 2018.

[14] Gartner. “Forecast: Internet of Things Endpoints and Associated Services”, 2017. (Accessed: 26.09.2019). https://www.gartner.com/en/documents/3840665

[15] Boeing G. “Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks”. Computers, Environment and Urban Systems, vol. 65, pp. 126–139, 2017.

[16] Wikipedia. Metropolitan municipalities in Turkey. Last edited: 18 September 2019. (Accessed: 26.09.2019). https://en.wikipedia.org/wiki/Metropolitan_municipalities_in_Turkey.