A New Smart Waste City Management

Dung D. Vu

Electrical Engineering Department,
École de Technologie Supérieure (ETS)
Montréal, Québec, Canada.
Email: do-dung.vu.1@ens.etmstl.ca

Georges Kaddoum

Electrical Engineering Department,
École de Technologie Supérieure (ETS)
Montréal, Québec, Canada.
Email: georges.kaddoum@etsmtl.ca

***Abstract*—This paper presents a new method of smart waste city management which makes the environment of the city cleanly with a low cost. In this approach, the sensor model detects, measures, and transmits waste volume data over the Internet. Since each trash bin has its geo-location and serial number, the collected data is processed by using regression, classification, and graph theory. Thenceforth a new method is proposed to dynamically and efficiently manage the waste collection by predicting waste status, classifying trash bin location, and monitoring the amount of waste. Then, it recommends the optimization of the route to manage the garbage truck efficiently. Finally, the simulation results are presented and estimated.**

1. Introduction

A smart city is an urban development vision to integrate information and communication technology (ICT) and Internet of things (IoT) technology which is the way of becoming the next technological revolution [1] securely to manage a city's assets. IoT is a framework in which all things have a representation and a presence on the Internet. More specifically, the IoT aims at offering new applications and services bridging the physical and virtual worlds, in which Machine-to-Machine(M2M) communication represents the baseline interface that enables the interactions between things and applications in the cloud such as environment monitoring [2] [3], object tracking [4], traffic management [5], health care [6], and smart home technology [7] [8]. Organizations can use IoT to drive considerable cost savings by improving asset utilization, enhancing process efficiency and boosting productivity. IoT combines the exponential growth of smart devices, the confluence of low-cost technologies (sensors, wireless networks, big data, and computing power), the pervasive connectivity, and the massive volumes of big data. Hence, IoT and big data are two sides of the same coin. An IoT device generates continuous information streams in a scalable way. Companies must handle the high volume of stream data and perform actions on that one. This actions can be event correlation, metric calculation, statistic preparation, and analytic which are applied to construct the smart city. In the smart city, the waste management system is a crucial point for the living environment, and its quality is considered seriously. A possible Smart waste city management system requires a way to cluster the location of trash bins, detect the status of waste in each bin, and process this collected data. The result of this work will be a valuable input data for the garbage truck management system which calculates the most optimal route to prevent the hazard of damage, pollution of waste, and resource consumptions.

To manage the waste of a smart city, the system incorporates a model for sharing data between truck drivers in the real time to perform garbage collection and the route optimization was proposed in [9]. A waste collection solution based on providing intelligence of trash cans including sensors and IoT prototype, which can read, collect, and transmit trash volume data via wireless network was proposed in [10] [11]. An adaptive large neighborhood search algorithm of finding optimal cost routes of garbage trucks is presented in [12], which processes the data of all emptied trash bins and the driven waste to disposal sites while respecting customer time windows. An improved dynamic route planning is discussed in [13] where the authors enhanced a guided variable neighborhood threshold meta-heuristic adapted to the problem of waste collection. On the other hand, the most important part of a waste management system is the smart bin [14] [15] [16] which collects the data of waste by using sensor and sends them over the internet to the server for monitoring the status of waste. While the network of ultrasonic sensors enabled smart bins to connect through the cellular network and generated a large amount of data which is further analyzed and visualized at real time to gain insights about the status of waste around the city [14], a Smart waste management with self-describing objects can detect the kind of waste based on its Radio-frequency identification (RFID) information [15]. However, they don't know the hazard of each bin such as explosion or flame from the bottle of perfume, batteries, and electronic wastes. In [16], a hazard detection method was proposed to detect and prevent these issues. Hence, a waste city management was conducted by an optimal garbage truck routes algorithm based on the status of the smart bin. We observed that citizens in the city tend to throw their waste into the bins without a particular time of a day. Therefore, a method for predicting the status of each bin will help the city's waste management system to operate more efficiently.

Contributions: This paper presents a new method of smart waste city management to provide a clean and hygienic environment to the city resident. In this approach, the collected data from the trash bin has been transferred over the Internet to the server which contains the status of each bin. The contributions of this paper are summarized as follows:

- The number of working clusters in each city are optimized and the location of new trash bins are classified automatically.
- A regression algorithm is applied to predict the situation of waste which can reduce the overload trash bin phenomenon while the garbage truck is coming.
- Finally, the priority weight of each bin is considered to improve the optimal garbage truck routes algorithm more efficiently.

This paper is organized as the followings: In section 2, a new method of smart waste city management (SWCM) is explained and the structure of the algorithm is proposed. Section 3 presents the SWCM simulation model. Finally, the related concluding remarks are discussed in section 4.

1.1. Data acquisition

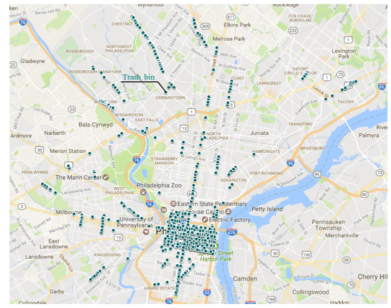
2. IoT-Based Smart waste city management model

In this section, we describe how to collect waste meta data associated with their status and locations. We evaluate our model with real big data in order to validate its output result.

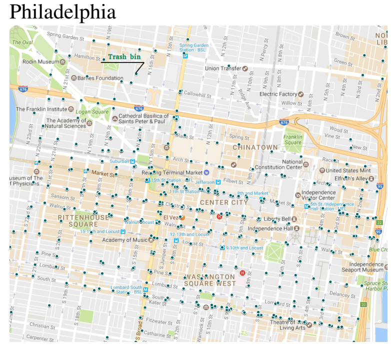
To obtain a set of waste data, we use an open source database, [1] which has a significant amount of geo-location information and status of each trash bin at the largest city in the Commonwealth of Pennsylvania in the United Stated named Philadelphia. Figure 1 illustrates the distribution of trash bin (green node) in Philadelphia City. We removed the content that was not useful with our scenario. For data analysis, we adopt the following five essential fields from this meta data:

- *sn*: The serial number of each trash bin
- *timestamp*: The milestone of recording data
- *level*: The amount of waste in each trash bin at the given *timestamp*
- *lat*: The latitude of each trash bin
- *lon* The longitude of each trash bin

1. https://www.opendataphilly.org/dataset



(a) The whole distribution of trash bin in Philadelphia



(b) A part of distribution trash bin in Philadelphia

Figure 1: Philadelphia trash bin

2.1. System description

2.1.1. The overview of functionality. The proposed system is based on the waste status of each trash bin in the city. The collected data is sent over the Internet to the server where it is stored and processed. At here, it is used for monitoring and predicting the status of each trash bin daily. Moreover, it will be utilized for calculating the optimal garbage truck routes, accordingly. Every day, the system sends the newly calculated routes to the workers based on the waste status, traffic congestion, the state of the forecast, and the balanced cost-efficiency functions. Since the limitation of our data, our optimal garbage truck routes algorithm is constructed by using the waste status of each bin. The prediction status of each bin can be analyzed based on the given training data before it occurs. Then, it will be considered to update the weight of the trash bin accordingly which is the most important input parameter of optimal garbage truck routes algorithm. The system overview is shown in the Figure 2.

2.1.2. Data processing and classifications. The status of each bin is not homogeneous and significantly differs from each other according to the state of each location. In this section, we introduce an algorithm that can be used to dynamically and efficiently manage waste collection strategies.

K-means [17] is an unsupervised machine learning algorithm that groups a dataset into a user-specified number (k) of clusters. We use it to make the working cluster of each garbage truck. However, this algorithm is somewhat naive, it clusters the data into k clusters, even if k is not the right

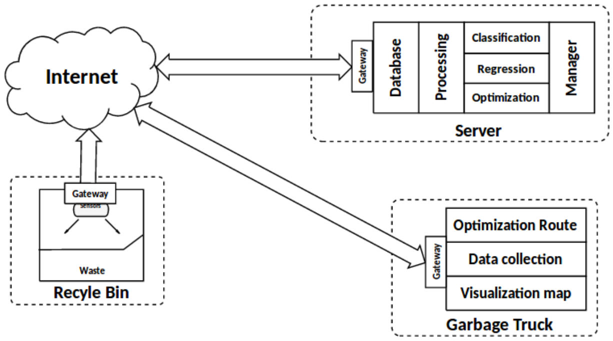


Figure 2: The smart waste management system overview

number of clusters to use. Therefore, when using k-means clustering, users need some way to determine whether they are using the right number of clusters. In this paper, we use Elbow method [18] to validate the number of clusters. The collecting routes are the travelling cycles containing a set of trash bins within a given cluster. The optimization of these cycles is a combinatorial optimization problem. When the objective function of this optimization is to minimize the driving distance. By considering a large number of routes, we use a Genetic Algorithm (GA) [19] which are relatively fast in providing near optimal solutions. Since the garbage truck needs time to collect every trash bin, it is very delightful if a status of trash bin can be predicted. Hence, after predicting, the system will recommend which one should be collected to prevent the overload phenomenon. In this paper, we use the Logistic regression algorithm [2] to predict the status of each trash bin based on its historical data.

The overall procedure is summarized in Algorithm 1. In this algorithm, the parameters are defined in Table 1. We suppose that the input dataset $\mathcal{D}$ is the database in 1.1. The parameter η, n, t and k denotes the threshold of amount waste in each bin, the number of bins, the time stamp of collecting data, and the number of working clusters, respectively. At here, when a status of the trash bin in a cluster s_{jh} is greater than the given threshold η, the weight of this one will be updated to 1, then the GA algorithm minimizes the driving distance for visiting the selected bins and returning to the headquarters. On the other hand, at the given time stamp t', a logistic regression algorithm is applied to predict the status of each bin. If the prediction status of bin is greater than the given threshold η, the weight is updated to 1, respectively. Therefore, we find a new optimal garbage truck routes which help us to prevent the overload trash bin case.

3. Analysis result and discussion

In this section, using the proposed Algorithm 1 in Section 2.1, we first analyze the number of working cluster

2. https://nlp.stanford.edu/manning/courses/ling289/logistic.pdf

TABLE 1: The definition of each variable

Variable	Description
$W = \{w_i \vert i \in (1, n)\}$	The weight of each bin
$T = \{t_{di} \vert d \in (1, m); i \in (1, n)\}$	The collected data milestone of each bin
$\mathcal{D} = \{S, L, W, T\}$	The input data
η	The threshold of amount waste in a bin
$\mathcal{M} = \{M_j \vert j \in (1, k)\}$	The optimal route of each garbage truck within each cluster
$\mathcal{DC} = \{DC_j \vert j \in (1, k);$ $DC_j = \{S_j, L_j, W_j, T_j\}\}$	The working cluster

Algorithm 1

Input: $\mathcal{D}, \eta$
Output: $\mathcal{M}, \mathcal{M}'$

```
Initialization:
k = Elbow (D →L)
DC = K-means (D →L, k)
At a given time stamp t in the original data.
for j = 1 to k do
    for h = 1 to size(DC_j) do
        if (s_jh ≥ η) then
            w_jh = 1
            M_j = Optimalroute (DC_j → L_j, DC_j → W_j)
        end if
    end for
    Plot (M_j)
end for
At a given time stamp t' in the near future.
for j = 1 to k do
    for h = 1 to size(DC_j) do
        s'_jh = LogisticRegression(DC_j → S_j, DC_j → T'_j)
        if (s'_jh ≥ η) then
            w'_jh = 1
            M'_j = Optimalroute (DC_j → L_j, DC_j → W'_j)
        end if
    end for
    Plot (M'_j)
end for
```

and then show the experimental result. We simply assume $\eta = 0.5$, which can also to be set another value to monitor the amount of waste.

3.1. Number of working cluster

One method of validating the number clusters for K-means algorithm is the Elbow method. It looks at the percentage of variance explained as a function of the number of clusters: One should choose a number of clusters so that adding another cluster doesnt give much better modeling of the data. In our work, we choose $k \in [1; 15]$. To estimate the optimal value of k, we consider the ratio $\frac{BSS}{TSS}$, where BSS and TSS are the between and total sum of squares, respectively. A good choice value of k is considered if this ratio tends to change slowly and remain less changing as

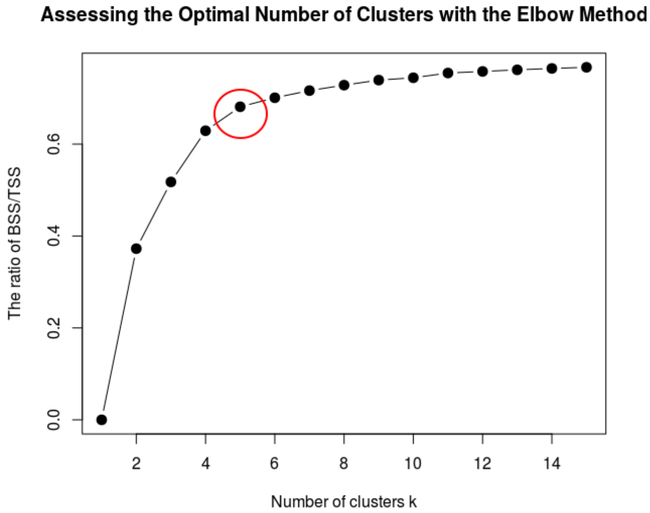


Figure 3: The optimal value of number working clusters

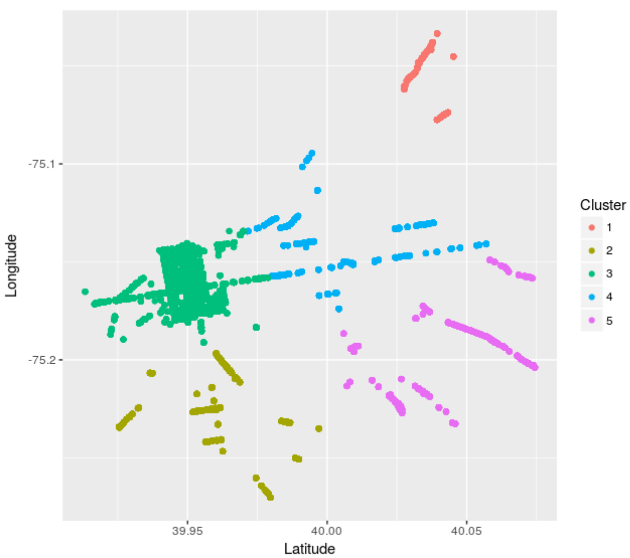


Figure 4: The working clusters in Philadelphia

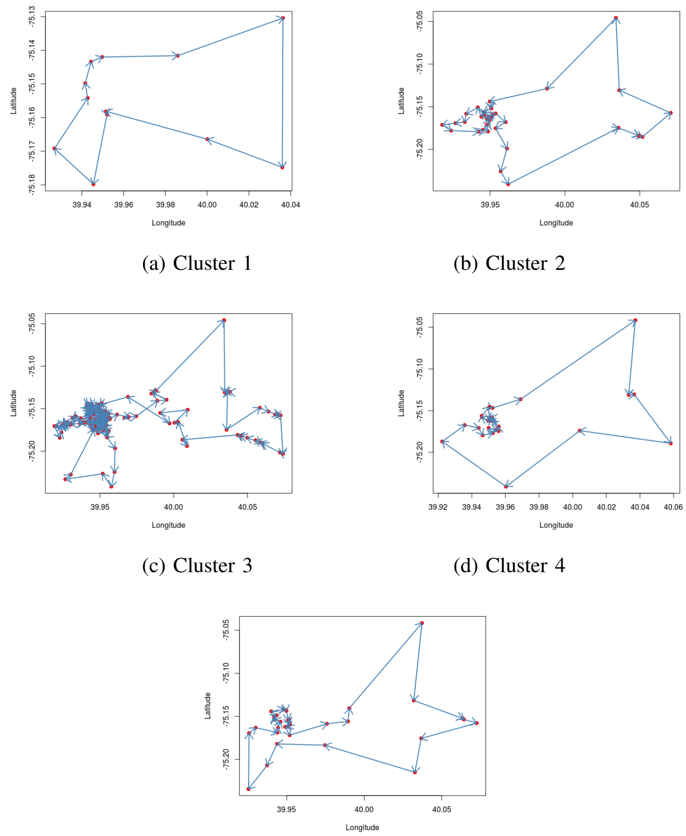
compared to other k's.[3] We see a pretty clear Elbow at $k = 5$ in Figure 3, which indicates that 5 is the best number of clusters.

Using the value of k above, we apply the K-means algorithm to make the working clusers for our system. The output is represented in the Figure 4. If the manager tend to add more trash bins into our system, the system will automatically classify them to the each cluster.

3.2. Optimal route of garbage truck in each cluster

In the each cluster dataset $\mathcal{DC}_j$, where $j \in [1 : k]$, we observed that the original system has three status levels of

3. https://www.r-bloggers.com/finding-optimal-number-of-clusters/



(a) Cluster 1

(b) Cluster 2

(c) Cluster 3

(d) Cluster 4

(e) Cluster 5

Figure 5: The optimal route of garbage truck at 2014/06/16

bin such as RED, YELLOW, and GREEN. Let we assume that when the level of trash bin is RED or YELLOW, its weight is 1, otherwise it is 0.

Since we need to update the optimal garbage truck routes daily for collecting the high weight trash bin, so we would like to pick up a milestone randomly which is "2014/06/16" in our dataset to make the simulation. By using the GA algorithm, the optimal garbage truck routes of each working cluster are constructed based on the weight and coordinate of trash bins. In the Figure 5, the red point is the high weight trash bin and the arrow is the route of garbage truck. It means that the garbage trucks collected all the trash bin whose weigh is 1 by using the optimal routes. Hence, the system will help the smart city to reduce the traffic congestion, fuel consumption, and pollution.

3.3. Prediction the status of trash bin

In fact, when the garbage trucks are running on the road, the status of trash bins can be modified. It is very delightful to predict the status of trash bin, then update the weight of this one. After that, the system will update the optimal garbage truck route. A Logistic regression algorithm is applied to work on this issue.

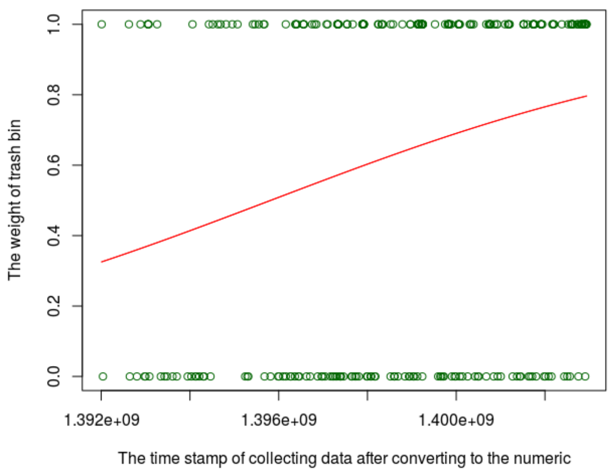


Figure 6: The logistic regression of the 171758^{th} trash bin

We pick up randomly one trash bin in one cluster whose serial number and location are 171758 and ($39.95204^o - 75.15911^o$), respectively to analyze the output result for convenience. Fistly, We convert all the date to the numeric data type[4] which is utilized for the Logistic regression algorithm. Figure 6 represents the relation between weight (green circle) and the time stamp of this trash bin. In this figure, the red line represents the trend of trash bin's weight. By considering a given time, if the red line approach the value 1, the weight of trash bin will be updated to 1, respectively. Hence, the optimal garbage truck route will be constructed based on the new weight.

4. Conclusion and future work

In this paper, we first clustered the working area by using the open source database of Philadelphia. Especially, we introduced an algorithm which automatically makes the working clusters and calculates the optimal garbage truck routes. Also, we use the Logistic regression to predict and update the weight of each trash bin. They will be used for creating the new optimal garbage route which reduces the pollution and fuel consumption more efficiently.

To further improve the algorithm, future work, the optimal garbage truck routes within each cluster need to be fitted with the street map data.

References

[1] F. C. Delicato, P. F. Pires, T. Batista, E. Cavalcante, B. Costa, and T. Barros, *Towards an IoT eco-system*, In the Proceedings of the 1st ACM International Workshop on Software Engineering for Systems-of-Systems, SESoS13, 2013, pp. 25–28.

[2] M. T. Lazarescu, *Design of a WSN platform for long-term environmental monitoring for IoT application*, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Volume 3, Number 1 (2013): 45–54.

[3] S. D. T. Kelly, N. K. Suryadevara, and S. C. Mukhopadhyay, *Towards the implementation of IoT for environmental condition monitoring in homes*, IEEE Sensors Journal, Volume 13, Number 10 (2013): 3846–3853.

[4] K. Gama, L. Touseau, and D. Donsez, *Combining heterogeneous service technologies for building an internet of things middle-ware*, Computer Communications, Volume 35, Number 4 (2012): 405–417.

[5] L. Foschini, T, Taleb, A. Corradi, and D. Bottazzi, *M2M-based metropolitan platform for IMS-enabled road traffic management in IoT*, IEEE Communications Magazine, Volume 49, Number 11 (2011): 50–57.

[6] A. J. Jara, A. Zamora, and A. F. G. Skarmeta, *An internet of things-based personal device for diabetes therapy management in ambient assited living (AAL)*, Personal and Ubiquitous Computing, Volume 15, Number 4 (2011): 431–440.

[7] S. Tozlu, M. Senel, W. Mao, and A. Keshavarzian, *Wi-Fi enabled sensors for internet of things: A pratical approach*, IEEE Communications Magazine, Volume 50, Number 6 (2012): 134–143.

[8] X. Li, R. Lu, X. Liang, X. Shen, J. Chen, and X. Lin, *Smart community: An internet of things appilication*, IEEE Communications Magazine, Volume 49, Number 11 (2011): 68–75.

[9] A. Medvedev, P. Fedchenkov, A. Zaslavsky, T. Anagnostopoulos, and S. Khoruzhnikov, *Waste management as an IoT enabled service in smart cites*, 15th ed, St. Petersburg, Russia: Internet of Things, Smart Spaces, and Next Generation Networks and Systems, 2015, pp 104–115.

[10] J. M. Gutierreza, M. Jensenb, M. Heniusa, and T. Riaze, *Smart waste collection system based on location intelligence*, Procedia Computer Science 61 (2015): 120–127.

[11] J. Hong, S. Park, B. Lee, J. Lee, D. Jeong, and S. Park, *IoT-based smart garbage system for efficient food waste management*, The Scientific World Journal (2014).

[12] K. Buhrkala, A. Larsena, and S. Ropkea, *The waste collection vehicle routing problem with time windows in a city logistics context*, Procedia Social and Behavioral Sciences 39 (2012): 241–254.

[13] T. Nuortio, J. Kytöjoki, H. Niska, and O. Bräysy, *Improved route planning and scheduling of waste collection and transport*, Expert Systems with Applications, Volume 30, Issue 2 (2006): 223–232.

[14] N. Sharma, N. Singha, and T. Dutta, *Smart bin implementation for smart cities*, International Journal of Scientific & Engineering Research (2015), Volume 6, Issue 9: 787–791.

[15] Y. Glouche and P. Couderc, *A smart waste management with self-describing objects*, The second International Conference on Smart Systems, Devices, and Technologies, 2013, pp 63–70.

[16] A. Sinhan and P. Couderc, *Smart bin for imcompatible waste items*, The ninth International Conference on Autonomic and Autonomous Systems, 2013, pp 40–45.

[17] T. Kanungo, D.M. Mount, N.S. Netanyahu, C.D. Piatko, R. Silverman, and A.Y. Wu, *An efficient k-means clustering algorithm: Analysis and implementation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 24, Issue 7 (2002): 881–892.

[18] T. M. Kodinariya and P. R. Makwana, *Review on determining number of cluster in K-means clustering*, International Journal of Advance Research in Computer Science and Management Studies, Volume 1, Issue 6 (2013): 90–95.

[19] J. M. Gutierrez, M. Imine, and O. B. Madsen, *Network planning using GA for regular topologies*, Proceedings of IEEE International Conference on Communications, 2008, pp: 5258–5262.

4. https://cran.r-project.org/web/packages/lubridate/lubridate.pdf