http://www.urt.cn/

CrossMark

Public Transit Customer Satisfaction Dimensions Discovery from Online Reviews

Lu Gao[1] · Yao Yu[2] · Wuling Liang[1]

Abstract Online user-generated content provides a valuable source for identifying dimensions of services. This study proposes a framework for extracting the dimensions of consumer satisfaction of public transportation services using unsupervised latent Dirichlet allocation model. A pilot study was performed on 17,747 online user reviews collected from 1452 public transportation agencies (including streetcar, light rail, heavy rail, boat, and aerial tram) in the United States over 8 years. The proposed approach is able to identify a few dimensions that were not discussed in the previous literature. This research also provides an alternative method to collectively gather users' feedback and efficiently preprocess textual data related to transit customer satisfaction.

Keywords Public Transportation · User Comments · Text Mining · LDA Model · Customer Satisfaction

✉ Yao Yu
  yao.yu@udc.edu

  Lu Gao
  lgao5@central.uh.edu

  Wuling Liang
  wliang3@uh.edu

[1] Department of Construction Management, University of Houston, 111C Technology Annex Building, Houston, Texas 77204-4021, USA

[2] Department of Civil Engineering, University of the District of Columbia, 4200 Connecticut Ave NW 21, Washington, DC 20008, USA

1 Introduction

Public transportation has been the subject of increasing interest in recent years, chiefly due to its potential to alleviate congestion, reduce emissions and protect the environment, provide critical support during emergencies and disasters, and enhance mobility in small urban and rural areas. To increase the ridership, public transport service needs to be more market-oriented, which can help maintain consumer loyalty and improve the long-term financial performance of public transit companies [1].

Satisfaction is considered as the main driver of consumer loyalty and behavior [2]. Coffel [3] found that the satisfaction level of public transit customers has a significant influence on whether they choose public transit as their primary commute method. Declining satisfaction levels among transit users lead to significant decrease in their customer loyalty regarding using transit again or recommending transit to a friend or relative. Lai and Chen [1] revealed the vital role of customer satisfactory in understanding the behavioral intention of public transit users. The authors found that passenger behavioral intentions significantly rely on passenger satisfaction.

Customer satisfaction also reflects the performance of a transit system regarding meeting customers' needs [4]. Customer satisfaction measurement has been translated into service quality measures in the existing literature. For example, Eboli and Mazzulla [5] developed a customer satisfaction index to evaluate transit service quality. This index enables service quality monitoring, dissatisfaction identification, and future strategy definition. Nathanail [6] proposed a multi-criteria evaluation method to provide railway operator with a quality control toolbox. Results based on the multi-criteria evaluation provide transit

planners and practitioners with valuable information for effective decision-making and marketing strategies.

For the reasons discussed above, researches on user satisfaction toward public transit service allow a better understanding of their behavior and provide directions for future planning and improvement strategies. However, most of the existing work in this area has relied on the use of "customer satisfaction surveys," where participants express their point of view about services by filling out sample surveys. Two major concerns about the questionnaire-based studies are (1) the low response rate and (2) the potential lack of comprehensiveness due to the design of the questionnaire. To overcome these limitations, we propose a text mining framework utilizing online customer reviews to investigate customer satisfaction toward public transport services. In the following sections, previous studies on public transit customer satisfaction evaluation are reviewed and the latent Dirichlet allocation (LDA) topic model is discussed. Case study results and comparison with questionnaire surveys are presented at the end of the paper.

2 Related Work

As public transit becomes a more promising mode to serve all travel purposes, dedicated efforts were made to improve the existing service from various perspectives including accessibility, pricing, comfort, etc. The dimensions of customer satisfaction addressed by previous studies are summarized and discussed in the following sections.

2.1 Fare

Researchers have found that fare price has a great influence on the ridership of public transportation. For example, Cervero and Wachs [7] found that annual U.S. transit ridership declined by about 6 %, while average fares increased by 35 % between 1984 and 1987. According to the authors, customer dissatisfaction with fare price is the main reason for ridership decline. Goodwin [8] used fare elasticity index as an indicator to study customer satisfaction in public transportation. The author confirmed the significance of fare price in transit customer satisfaction. Coffel [3] also found that "service received for the fare paid" is one of the top satisfaction discriminators between "somewhat satisfied" and "very satisfied." Wallin [9] identified "price level" as one of the nine service attributes that are believed to impact customer satisfaction. The author stated that price becomes an important factor when the offered service is considered to be of low quality. Perone and Volinski [10] found that while a fare-free policy is appropriate for smaller transit systems, it does not have the same effect for larger transit systems in

major urban areas. Eboli and Mazzulla [5] used an index of customer satisfaction to evaluate transit service quality. The weight of ticket price was estimated to be 9.12 (scale from 1 to 10), which indicates that ticket price is crucial to customer satisfaction.

2.2 Wait and Travel Time

Wait and travel time are usually considered as critical measurements of transit customer satisfaction. For example, Cervero [11] found that transit riders are more sensitive to schedule reliability than almost any other service attributes. The author also found that riders are especially sensitive to out-of-vehicle travel time. Wall and McDonald [12] reported that in North West London, when the bus service frequency was changed from every 20 min to every 10 min, the estimated demand increased of around 20 %. More recently, Friman and Fellesson [13] found that there is a significant relationship between average public transportation speed and overall user satisfaction. Moreover, a survey report by Metropolitan Transportation Authority (MTA) in New York identified "how fast the public transit gets you where you want to go" as one of the highest satisfaction attributes concerned by the subway customers [14].

2.3 Cleanliness

Cleanliness is another popular topic in public transportation satisfaction surveys. Coffel [3] identified four top satisfaction discriminators between "somewhat satisfied" and "very satisfied." These discriminators include "cleanliness of light rail vehicle interior," "cleanliness of light rail vehicle exterior," "cleanliness of heavy rail vehicle interior," and "cleanliness of stations (waiting area)." Eboli and Mazzulla [5] studied cleanliness of both interior and exterior of transit vehicles and found that the weights of exterior and interior features are estimated to be 7.85 and 9.51, respectively (scale from 1 to 10). MTA [14] found that cleanliness had received a high attention from the subway user and the demand for improvement in cleanliness was overwhelming. The research concluded that cleanliness is one of the most important service attributes that transportation companies need to improve in the public transportation service.

2.4 Customer Service

Customer service is defined as the services provided by the employees of the public transportation agencies. It includes the behavior of the driver, conductor, in station customer service employees, etc. Wallin [9] developed a conceptual model to determine the relationships among customer preferences, customer satisfaction, and customer segments.

The author reported that information service such as schedule timetables and corresponding lines have a significant impact on customer satisfaction. Coffel [3] identified "courtesy of bus drivers," "courtesy and helpfulness of station staff (waiting area)," and "courtesy of the operator/conductor" as several staff behaviors that could attract more people to use public transit service.

2.5 Accessibility

The access to public transportation was another significant component of the overall transportation system. Coffel [3] identified "ease of making transfers from the station" as an important customer satisfaction dimension. Daganzo [15] investigated the structural effect of transit system on accessibility and proposed a combination of grid and hub-and-spoke network structure to improve the overall competitiveness of transit system over driving. Woldeamanuel and Cyganski [16] used a panel binomial probit model to analyze the parametric relationship between levels of traveler satisfaction and accessibility to public transport services. The results showed that travelers who tend to make frequent trips by public transportation demonstrate a higher probability of satisfaction with accessibility. According to MTA [14], "convenience of stops" is one of the most important dimensions among all public transit satisfaction measurements.

2.6 Safety

Turner [17] found that safety has great influence on commute experience of public transportation customers. Roberts et al. [18] reported that improving the security culture of public transportation will significantly improve the customer satisfaction as well as other aspects such as efficiency and employee morale. Safety is also among the nine service attributes identified by Wallin [9] that are believed to have impacts on customer satisfaction of public transportation. Coffel [3] identified the category of "safety from crime after getting off the bus" as an important factor to improve the customer satisfaction level from "somewhat satisfied" to "very satisfied."

2.7 Crowdedness

Many researchers found that crowdedness has considerable influence on public transit customer satisfaction. For example, Lundberg [19] found that the crowdedness condition of public transit contributes more to travelers' stress experience than trip duration. MTA [14] identified crowdedness as the most unsatisfied service attribute of

New York City Transit. More recently, researchers found that user satisfaction is lower when individuals lack space in transit vehicles and the space between transit passengers is found as one of the main qualities desired by users [20–22].

2.8 Comfortability

The comfortability of public transportation is related to conditions such as seat condition, temperature in the vehicles, and smoothness of the ride. Coffel [3] identified "smoothness of ride" and "seating comfort" as top attributes that can enhance the customer satisfaction level from "somewhat satisfied" to "very satisfied." It was also found that improving these attributes can increase loyalty and ridership among current and potential customers. MTA [14] also identified "comfort of temperature on vehicles" as an important service attribute.

2.9 Summary

The most common methods of data collection in the studies discussed above were interviews and surveys. However, these approaches are limited by the response rate and the variability and subjective nature of the response. Recently, a variety of new data sources and an expanding set of novel analysis methods open up new opportunities for studying transit user satisfaction. For example, Aranguren and Tonnelat [23] use transit users' facial expression to study their willingness to cope with the crowdedness in the Paris Metro. In this paper, we propose studying public transit customer satisfaction by analyzing online reviews and comments, which contain words expressing user sentiment or opinions about public transit service. In this research, we downloaded user comments from public transit review website and applied an unsupervised topic model to identify sets of satisfaction dimensions. A total of 17,747 reviews and comments were collected, and the extracted dimensions were compared with the findings reported in the previous studies.

3 Methodology

In this research, the LDA topic model was employed to extract opinions from user reviews. Topic models are usually used to analyze and summarize topics from large volume of textual documents. A topic is defined as a group of words that tend to occur together frequently and a document is defined as the mixture of different topics [24]. The LDA model is a generative probabilistic approach to analyze the collections of discrete data [25]. In this research, customer satisfaction attributes are considered as topics to be obtained from the documents (review

comments). The following gives a short overview of the mathematical basis of LDA.

Let M be the number of review comments, N_m be the number of words in the mth comments, V be the number of distinct words, and K be the number of topics. The number of topics, K, is a user-specified parameter that provides control over the level of details of the discovered topics. Also, let $w_{m,n}$ be the nth word in the mth comment, $z_{m,n}$ be the topic of topic of $w_{m,n}$, θ_m be the topic distribution for the mth document, ϕ_k be the word distribution for the kth topic, $\boldsymbol{\alpha}$ be the prior distribution for topics in a review, and $\boldsymbol{\beta}$ be the prior distribution for words in a topic. The words of the review comments are assumed to be generated in the following steps.

Step 1 The word distribution of the kth topic, ϕ_k, is generated from a Dirichlet distribution with parameter $\boldsymbol{\beta}$.

$$\phi_k \sim Dir(\boldsymbol{\beta}) \text{ for } 1 \leq k \leq K, \tag{1}$$

where Dir represents the Dirichlet distribution.

Step 2 The topic distribution of the mth review comment, θ_m, is generated from a Dirichlet distribution with parameter $\boldsymbol{\alpha}$.

$$\theta_m \sim Dir(\boldsymbol{\alpha}) \text{ for } 1 \leq m \leq M. \tag{2}$$

Step 3 The topic of the nth word in the mth review comment, $z_{m,n}$, is generated from the θ_m distribution as a discrete random variable.

$$z_{m,n} \sim Disc(\theta_m) \text{ for } 1 \leq m \leq M \text{ and } 1 \leq n \leq N_m. \tag{3}$$

Step 4 The nth word in the mth comment, $w_{m,n}$, is generated from the $\phi_{z_{m,n}}$ distribution as a discrete random variable.

$$w_{m,n} \sim Disc(\phi_{z_{m,n}}) \text{ for } 1 \leq m \leq M \text{ and } 1 \leq n \leq N_m. \tag{4}$$

Given the data generating process above, the joint probability of all the parameters is

$$p(w,z,\theta,\phi|\alpha,\beta) = p(w|\phi,z)p(\phi|\beta)p(z|\theta)p(\theta|\alpha), \tag{5}$$

where

$$p(w|\phi,z) = \prod_{m=1}^{M} \prod_{n=1}^{N_m} p\left(w_{m,n}|\phi_{z_{m,n}}\right);$$
$$p(\phi|\beta) = \prod_{k=1}^{K} p(\phi_k|\beta);$$
$$p(z|\theta) = \prod_{m=1}^{M} \prod_{n=1}^{N_m} p\left(z_{m,n}|\theta_m\right); \text{ and}$$
$$p(\theta|\alpha) = \prod_{m=1}^{M} p(\theta_m|\alpha).$$

By integrating θ and ϕ out, we have

$$p(w,z|\alpha,\beta) = \prod_{m=1}^{M} \left(\frac{\Gamma(A)}{\Gamma(A+N_m)} \prod_{k=1}^{K} \frac{\Gamma(\alpha_k+n_{m,k})}{\Gamma(\alpha_k)} \right) \\ \times \prod_{k=1}^{K} \left(\frac{\Gamma(B)}{\Gamma(B+n_k)} \prod_{v=1}^{V} \frac{\Gamma(\beta_v+n_k^v)}{\Gamma(\beta_v)} \right), \tag{6}$$

where

$n_{m,k}$ represents the number of words in the mth document that are assigned to the $k-$th topic,

n_k^v represents the number of v word assigned to the $k-$th topic,

$A = \sum_{k=1}^{K} \alpha_k$,

$B = \sum_{v=1}^{V} \beta_v$,

and n_k represents the number of words assigned to the $k-$topic.

In order to use Gibbs sampling to implement the LDA model, we need the following conditional probability:

$$p\left(z_{m,n}=k|z^{(-m,n)},w,\alpha,\beta\right) = \frac{p\left(z_{m,n}=k,z^{(-m,n)},w|\alpha,\beta\right)}{p(z^{(-m,n)},w|\alpha,\beta)} \\ \propto p\left(z_{m,n}=k,z^{(-m,n)},w|\alpha,\beta\right). \tag{7}$$

By plugging Eq. (6) into (7) and ignoring the terms that do not involve $z_{m,n}$, the conditional posterior of $z_{m,n}$ becomes as follows:

$$p\left(z_{m,n}=k|z^{(-m,n)},w,\alpha,\beta\right) \\ \propto \frac{\left(\alpha_k+n_{m,k}^{(-m,n)}\right) \times \left(\beta_v+n_k^{w_{m,n},(-m,n)}\right)}{\left(B+n_k^{(-m,n)}\right)}, \tag{8}$$

where

$z^{(-m,n)}$ represents all the topic assignments other than $z_{m,n}$;

$n_{m,k}^{(-m,n)}$ represents the number of words (excluding the nth word) in the mth review comment that have been assigned to the kth topic.

$n_k^{(-m,n)}$ represents the number of words (excluding the nth word in the mth review comment) assigned to the kth topic;

$n_k^{v,(-m,n)}$ represents the number of the vth word assigned to the kth topic (excluding the nth word in the mth document).

In this paper, we used the Gibbs sampling algorithm to draw random samples from the derived condition posterior distribution [Eq. (8)]. The idea behind Gibbs sampling is that we can obtain random samples from the joint posterior distribution by sequentially simulating individual parameters from the set of conditional distributions. Draws from this simulation algorithm will converge to the target posterior distribution.

4 Case Study

4.1 Data Description

In this study, the LDA topic model was applied to 17,747 review comments extracted from 1452 different public transportation agencies on the website of www.yelp.com.

Table 1 Descriptive statistics of collected comments

		Number of comments	Percent (%)
Transit methods	Street transit	4994	28.2
	Semi-rapid transit	162	0.9
	Rapid transit	11,062	62.3
	Others (boat, etc)	1529	8.6
	Total	17,747	100
Rating	0–1	66	0.4
	1–2	2104	11.9
	2–3	6203	34.9
	3–4	7859	44.3
	4–5	1515	8.5
	Total	17,747	100
Posted date	2005–2007	1730	9.7
	2008–2010	6718	37.9
	2011–2013	9299	52.4
	Total	17,747	100

Fig. 1 The word cloud of reviews of different transit methods

(a) Street Transit

(b) Semi-Transit

(c) Rapid Transit

(d) Others

These comments and reviews were posted between 2005 and 2013. Table 1 shows the descriptive statistics of the collected data. As shown in Table 1, most of the comments were collected from rapid transit services and only 0.9 % was collected from semi-rapid transit. A majority of the review ratings (five being the best) were between 2 and 4 (79.2 %). More than a half (52.4 %) of the comments was posted between 2011 and 2013, only 9.7 % of the comments were posted before 2008. Figure 1 shows the word cloud of reviews from four different transit methods (rapid transit, semi-rapid transit, street transit, and others). A word cloud is a visualization of the words frequency in a given text with words of higher frequency displayed in larger size. The word cloud using the online review in Fig. 1 clearly shows four themes of transit methods.

The 10 agencies/facilities with the most reviews are listed in Table 2. The Bay Area Rapid Transit (BART) received 755 reviews, which is the highest among the ten agencies/facilities. Moreover, seven of the 10 agencies/facilities are related to rapid transportation services, while two of the ten agencies/facilities are related to boat transportation service. Only one of them is related to street transportation service.

4.2 Topic Model Results

In this case study, the proposed unsupervised LDA topic model was used to summarize the top 10 customer satisfaction dimensions from the collected review comments. In the topic model, words with high associations were

Table 2 Agencies/facilities with the most reviews

Agencies	Number of reviews
BART—Bay Area Rapid Transit	755
Amtrak	633
Staten Island Ferry	417
Metrorail	345
Metro Transportation Authority	342
Union Station	338
Caltrain	336
Washington State Ferries	314
Chicago Transit Authority CTA	265
Port Authority Bus Terminal	255

grouped together. The topic model results are presented in Table 3. The dimensions most frequently identified include the following:

1. Waiting and travel time (the 1st topic of street transit, 1st and 2nd topics of semi-rapid transit, the 1st topic of rapid transit, the 7th and 9th topics of other public transits in Table 3).
2. Cleanliness of the vehicle (the 4th topic of street transit, the 9th topic of semi-rapid transit, the 10th topic of rapid transit, 2nd topic of other transit in Table 3). This is consistent with the existing studies, which have demonstrated the significance of the cleanliness in public transportation customer satisfaction.
3. Customer service (the 7th topic of street transit, the 4th topic of semi-rapid transit, the 3rd topic of rapid transit, the 5th topic of other transits in Table 3).

4. Transit price (the 9th topic of street transit, the 10th topic of semi-rapid transit, the 9th topic of rapid transit, the 8th topic of other transits in Table 3).
5. Accessibility (the 2nd topic of street transit, the 3rd topic of semi-rapid transit, the 4th topic of rapid transit, the 4th topic of other transits in Table 3).
6. Crowdedness (the 8th topic of street transit, the 2nd topic of rapid transit, the 8th topic of semi-rapid transit in Table 3).
7. Comfortability (the 3rd topic of street transit, the 6th topic of semi-rapid transit, the 6th topic of rapid transit, the 3rd and 6th topics of other public transits in Table 3).
8. Safety (the 10th topic of street transit, the 7th topic of semi-rapid transit, the 8th topic of rapid transit in Table 3).
9. Transfer service (the 6th topic in street transit, the 8th topic for semi-rapid transit, the 5th and 7th topics for rapid transit, the 10th topic for other public transit methods in Table 3).
10. Aesthetics (the 5th topic of street transit, the 1st topic for other public transits). This dimension was not found in the existing studies.

5 Conclusion

Transit customer satisfaction study helps understand customers' behavior intentions and lays foundation for toolbox development to monitor service quality, evaluate system performance, identify customers' dissatisfaction, and

Table 3 Extracted top 10 topics

No.	Street transit	Semi-transit	Rapid transit	Other public transit
1	Time, minutes, hour, late, hours, waiting	Waiting, between, times, half, driving	Train, time, minutes, hour, late, hours, wait	Ferry, deck, great, view, scenery
2	Stations, convenient, stops, walk, location	Time, ride, trolley, travel, between, pretty	Trains, crowd, time, passenger	Boat, clean, nice, pretty
3	Seat, walking, sitting, good, cold	Ride, hotel, parking, town	Stop, service, feel, notice, good, worst	Beer, ferry, cool, time, wind, years
4	Buses, clean, seat, pretty, nice	Station, driver, service, train	Stop, between, center, walk, downtown, parking	Dock, public, convenience, anywhere
5	Bus, view, good, great, experience	Trolley, lines, stops, best, time	Line, lines, blue, green, north, orange, anywhere, transfer	Service, ferries, quite, river
6	Station, stops, transfer, convenient	Trolley, seat, food, good	Train, seat, nice, great, sitting, love	Ferries, seat, food, relaxing
7	Service, good, worst,, customer, driver	Trolley, security, stops, station	Transfer to other public transit method	Waiting, ferry, tram, across, time
8	Bus, crowd, weekend, traffic	Trolley, stops, waiting, line	Passengers, door, safe, ride, driving	Ferry, cruise, money, time, best, trip
9	Riding, price, cheap, travel, good, options	Trolley, seat, clean, nice	Ticket, pass, money, price, fare	Trip, time, best, enough, quite
10	Security, bus, lady, public, accident	Trolley, free, money, trip, runs	Trains, pretty, station, nice, clean, seats	Dock, line, trips, terminal

develop improvement strategy. In this research, we applied a topic model to analyze the needs and expectations expressed by public transit customers. The proposed model identifies the most frequent customer satisfaction dimensions, which include waiting and travel time, cleanliness, customer service, price, accessibility, crowdedness, comfortability, safety, transfer service, and aesthetics. This research serves as a pilot study to test the feasibility and reliability of using online review comments to investigate transit users' satisfaction dimensions. With the research results confirming previous work on transit users' perception of service quality, the proposed method prove to be a reliable way to study various dimensions of customer satisfaction toward transit system. Since online review comments can be obtained with low cost and labor, transit agencies can use this method to collectively gather feedback from its users, which contrasts the expensive and low responsive survey/interview approaches. Moreover, when combining the text mining method with the traditional survey/interview approach, the joint investigation will ensure the comprehensiveness of the results. The future directions of this research include testing other text mining models and integrating data collected from both online reviews and questionnaire-based surveys. These approaches have the potential to enhance model efficiency and effectiveness and will facilitate future model selection and modification to serve the emerging needs of transit customer satisfaction analysis.

References

1. Lai W-T, Chen C-F (2011) Behavioral intentions of public transit passengers: the roles of service quality, perceived value, satisfaction and involvement. Transp Policy 18(2):318–325
2. Olsen SO (2007) Repurchase loyalty: the role of involvement and satisfaction. Psychol Mark 24(4):315–341
3. Coffel K (1995) Customer satisfaction index for the mass transit industry. TRANSIT-IDEA Program Project Final Report 10
4. Hill N, Brierley J, MacDougall R (2003) How to measure customer satisfaction. Gower Publishing, Ltd., Aldershot
5. Eboli L, Mazzulla G (2009) A new customer satisfaction index for evaluating transit service quality. J Public Transp 12(3):2
6. Nathanail E (2008) Measuring the quality of service for passengers on the hellenic railways. Transp Res Part A 42(1):48–66
7. Cervero R, Wachs M (1982) An answer to the transit crisis: the case for distance-based fares. J Contemp Stud 5(2):59–70
8. Goodwin PB (1992) A review of new demand elasticities with special reference to short and long run effects of price changes. J Transp Econ Policy 26:155–169
9. Andreassen TW (1995) (Dis) satisfaction with public services: the case of public transportation. J Serv Mark 9(5):30–41
10. Perone JS, Volinski JM (2003) Fare, free, or something in between? Free Fares
11. Cervero R (1990) Transit pricing research. Transportation 17(2):117–139
12. Wall G, McDonald M (2007) Improving bus service quality and information in winchester. Transp Policy 14(2):165–179
13. Friman M, Fellesson M (2009) Service supply and customer satisfaction in public transportation: the quality paradox. J Public Transp 12(4):4
14. Metropolitan Transportation Authority (MTA) (2014) Customer satisfaction survey subway report
15. Daganzo CF (2010) Structure of competitive transit networks. Transp Res Part B 44(4):434–446
16. Woldeamanuel MG, Cyganski R (2011) Factors affecting travellers's satisfaction with accessibility to public transportation. In Proceedings European transport conference
17. Turner BA (1994) Causes of disaster: sloppy management. Br J Manag 5(3):215–219
18. Roberts H, Retting R, Webb T, Colleary A, Turner B, Wang X, Toussaint R, Simpson G, White C (2015) Improving safety culture in public transportation
19. Lundberg ULF (1976) Urban commuting: crowdedness and catecholamine excretion. J Hum Stress 2(3):26–32
20. Litman T (2008) Valuing transit service quality improvements. J Public Transp 11(2):3
21. Cantwell M, Caulfield B, Mahony MO (2009) Examining the factors that impact public transport commuting satisfaction. J Public Transp 12(2):1
22. Olio LD, Ibeas I, Cecin P (2011) The quality of service desired by public transport users. Transp Policy 18(1):217–227
23. Aranguren M, Tonnelat S (2014) Emotional transactions in the paris subway: combining naturalistic videotaping, objective facial coding and sequential analysis in the study of nonverbal emotional behavior. J Nonverbal Behav 38(4):495–521
24. Charles E (2010) Text mining and topic models. Lecture notes
25. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022