Investigating the Role of Twitter in E-Governance by Extracting Information on Citizen Complaints and Grievances Reports

Swati Agarwal[1] and Ashish Sureka[2(✉)]

[1] Birla Institute of Technology and Science, Pilani, Goa, India
swatia@goa.bits-pilani.ac.in
[2] Ashoka University, Sonepat, Haryana, India
ashish.sureka@ashoka.edu.in

Abstract. Open Source Social Media Intelligence (OSSMInt) is a field that focuses on extracting useful information and actionable insights from publicly available and overt sources of data on social media platforms. There are several applications that can be built by applying OSSMInt techniques on this human-sensor data. In this paper, we present some of the use-cases of OSSMInt that are useful for the public sector agencies for e-governance. E-governance on social media include the identification of complaints and grievances reported online by the public citizens for the government authorities and facilitate public agencies to response those complaints, provide better services and improve their connections with public citizens. We present the basic Natural Language Processing and Machine Learning based framework, tools and techniques within the context of OSSMInt and E-governance. The focus of this paper is on mining user-generated content on Twitter (the most popular social media and microblogging website) to identify public citizens complaints and grievances. In particular, we focus on two important applications: (1) complaints which are reported to spread awareness among other citizens and to bring government's attention to the issues reported in the complaint, and (2) complaints which seek for immediate action and response from the concerned authorities. In addition to the basic introduction and motivation, we will discuss the unique challenges to these applications, open research problems, important literature, proposed approach, experimental results, and future directions.

Keywords: Bad roads complaints · Complaints and grievances
Government applications · Information visualization
Lexical knowledgebase · Mining user generated content
Natural Language Processing · Social media analytics
Text classification · Twitter

1 Introduction

Recently, there has been a noticeable adoption of social media and upward trend in its usage by government agencies for not just disseminating information but

P. K. Reddy et al. (Eds.): BDA 2017, LNCS 10721, pp. 1–11, 2017.
https://doi.org/10.1007/978-3-319-72413-3_21

also acquiring information such as complaints and grievances from citizens (a phenomenon referred to as *citizensourcing*) [8, 12]. Specifically, social media websites like Twitter and Facebook are gaining popularity as social-media based citizen grievance management system or platforms on which people can lodge complaints. Evidence shows that Twitter is the most widely used micro-blogging platform on Internet. Due to the wide reachability and connectivity among its users, Twitter is being used by the National Government to reach out the public. For example, in India, ministry of the railway (@railminindia), state police (@delhipolice and @mumbaipolice), ministry of road and transport (@morthindia), traffic police (@dtptraffic) and income tax department (@incometaxindia) have some of the most active Twitter accounts. Unlike other countries, one of the primary objectives of Indian Government's Twitter accounts is to not only reach out to the public but to also address their complaints and grievances[1,2]. Based on our analysis of several Indian Government Twitter accounts, we find that an active Government Twitter handle receives an approximate of 5 tweets per minute. Based on our analysis of several Indian governments' Twitter handle data, we found that 50% of the tweets posted in an hour are complaints and grievances reported from various regions of India. The government bodies on Twitter forward these complaints and redirect authors to the concerned department for resolving their complaints efficiently.

We conduct a literature survey in the area of complaints and grievances identification on social media and divide the existing literature into two lines of research: (1) **usage of Twitter microblogging website to report and complaints and grievances** Heverin et al. [11] examine the use of Twitter by city police departments in large U.S. cities (cities with populations greater than 300,000) that have active Twitter accounts. Anderson et al. [5] present a study on Twitter adoption across American municipal police departments serving populations over $100,000$. Meijer et al. [15] present an empirical analysis of Twitter usage by the Dutch police and conduct an analysis of 982 Twitter handle. Edwards et al. [7] present a study on webcare, i.e. the act of engaging in online communication with citizens. Vanessa et al. [8] present a study for analyzing the behavioral similarities and differences of 3-1-1 phone service (formal) and Twitter (informal) channels for reporting issues in the community. (2) **mining public complaints and communication on Twitter for building prediciton models for situation awareness.** Kumar et al. [13] present an application of Twitter to atomically determine road hazards by building language models based on Twitter users' online communication. Gu et al. [10] propose a methodology to mine tweet texts to extract incident information on both highways and arterial roads. Fu et al. [9] describe an approach to extract and analyze real-time traffic related Twitter data for incident management purpose. Eleonora [6] present a real-time monitoring system for traffic event detection from Twitter stream analysis and conduct a case-study for the Italian road network. Mai et al. [14] demonstrate the use of data from public social interactions on Twitter as a

[1] http://bit.ly/2fR7R1s.
[2] http://bit.ly/2xLyEpN.

potential complement to traffic incident data. Schulz et al. [17] present a solution for a real-time identification of small-scale incidents using microblogs, thereby allowing to increase the situational awareness by harvesting additional information about incidents. Napong et al. [18] present a study on social-based traffic information extraction and classification consisting of mining Twitter for traffic congestion, incidents, and weather. Furthermore, in order to address the complaints and grievances of citizens, the Indian government also initiated several policies and organizations. The aim of these organizations (TwitterSeva, MociSeva, Cybercell, DOTSeva) is to mine online complaints specific to the concerned department and address them in a timely and efficient manner.

The research work presented in this paper is motivated by the need to develop a solution to automatically resolve the challenges of manual inspection. Twitter allows users to post a maximum of 140 characters and therefore involves usage of slang and abbreviations. Due to the presence of free-form text tweets do not have a defined structure or language format and hence are high likely to have grammar and spelling errors. Due to the presence of multilingual texts and scripts in tweets, it is challenging to identify the linguistic features for building NLP (Natural Language Processing) based applications. Text classification or categorization and information extraction from tweets is thus a technically challenging problem [2–4]. Further, filtering these complaint reports from non-complaint tweets is technically challenging due to the wide range of complaints. Our research aim is to build a text analysis based model to address the NLP challenges in microposts (very short text such as tweets). In particular, the research aim of the work presented in this paper is to investigate text classification based techniques for automatically identifying complaints tweets and assigning them to predefined labels based on the topic of the content. Our research aim is also to investigate information extraction and visualization to extract useful information and insights from the complaint reports. Furthermore, our aim is to create an annotated dataset and make it publicly available to the research community.

2 Data Collection and Pre-processing

We formulate the problem of automatic identification of citizens' complaints (reported to official Twitter handle of a public agency) as a one-class classification problem. We propose a text classification based approach consisting of various components performing several tasks: tweets extraction from public agencies' account, enrichment and enhancement of raw microposts (tweets), learning the features of non-complaint and complaint report tweets, developing a baseline classification approach, use of ensemble techniques to improve the baseline method, empirical analysis and performance evaluation. Based on our inspection on complaints and grievances reports on Twitter, we divide them into two case studies (as discussed in the literature survey. See literature survey in Sect. 1). Due to the page limit, we present the data collection and statistics in form of a table. Tables 1 and 2 demonstrate the statistics of dataset for Case Study 1 (4 weeks duration-11 April 2016 to 8 May 2016) and Case Study 2 (8 weeks-from

Table 1. Dataset for case study 1

Account	Original	Sampled
@RailMinIndia	36182	1500
@dtpTraffic	1524	1000
@DelhiPolice	1720	1000
@IncomeTaxIndia	383	200

Table 2. Dataset for case study 2

@nitin_gadkari, @MORTHIndia		
Collected	Original	Replied
81304	17511	11092
Retweets	Sampled English tweets	Users
52701	3302	2604

July 18, 2016 to September 13, 2016) respectively. Tables 1 and 2 shows the number of tweets collected, and sampled (after pre-processing and filtering) for our experiments. The statistics shown in the tables also shows the name of the public agencies' accounts for which we collected the data.

We preprocess the sampled datasets (for both case study 1 and 2) and address the challenge of noisy content in the tweets. We use the micropost enrichment algorithm proposed in our previous study Mittal et al. (focused on the problem of complaints and grievances identification) [16]. The proposed algorithm performs a syntactic enhancement of the tweets and consists of five phases primarily named as hashtag expansion (splitting the joint hashtags), @username mention expansion (replacing the @mentioned usernames with their actual profile names), spell error correction (using the application of Bing Search Engine to correct the spelling error in tweets), acronym and slang treatment (correcting the domain specific and generalized slang and abbreviations in tweets) and sentence segmentation (removing filler terms, consecutive special characters, correcting spaces and conjunctions).

3 Identifying Non-complaints Reports

The official public service agencies account on Twitter are open and anyone can mention them in their tweets. We observe that not all tweets posted on these accounts are complaint reports and rather are either off-topic or discussions not relevant to the complaint department. In order to classify a complaint report related to a government service (case studies 1 and 2), we identify various features that are a strong indicator of a tweet to be a complaint report. However, we also observe that there are several discriminatory features that are a strong indicator of a tweet certainly not to be complaint report. Based on our manual inspection on official Twitter handle and non-complaint tweets identification model proposed in our previous study [16], we divide such tweets into 4 categories (AISP): Appreciation posts, Information Sharing and Promotional tweets.

Appreciation tweets are the posts made by citizens for praising the government for their work or resolving their previous complaints. Information sharing tweets are the tweets posted by users to share daily news about the events or policies initiated by the government. Further, due to the presence of several official accounts on Twitter, we categorize tweets posted by a different official

Table 3. Examples of non-complaint tweets-classified into 4 categories: appreciation (APP), information sharing (IS) and promotional (PRL) tweets.

AISP	Tweet
APP	Big **big move** @mansukhmandviya ji. **Thank you** for personally monitoring this. **Thanks** to @MORTHIndia and @nitin_gadkari ji also for this
IS	Commuters, get ready for more sun-kissed rides on the waters `https://goo.gl/SlrCWE` @PiyushGoyal @nitin_gadkari
PRL	Income declaration scheme: Government assures complete confidentiality: `http://goo.gl/jVU5qK` @FinMinIndia **@IncomeTaxIndia**

account of same public agencies categorized as promotional tweets. Table 3 shows examples of appreciation, information sharing, and promotional tweets posted on official accounts of @nitin_gadkari and @MORTHIndia. We use the AISP tweet classifier method proposed in our previous study [16] and classify AISP tweets and filter the unknown posts that may or may not be a complaint report about killer roads. We use these unknown tweets for further feature extraction and classification and complaints reports classification.

4 Features Extraction

In this Section, we identify various linguistic and contextual features that can be used to classify a complaint tweet. Due to the page limit, we present all features in form of a table. Tables 4 and 5 shows the list of all features extracted for

Table 4. List of features extracted in case study 1. CDK = Closed Domain Keywords

Feature	Summary	Technique	Presentation
N-Grams	Character n-grams frequently occurring in complaint posts	WordNet	Grouped Triplets of n-grams similar to each other
CDK	Keywords specific to different departments of public agencies	Manual Inspection	Each word is a column in feature vector space
Events	Activities, events and substances reported in the complaints	IBM Watson Bluemix Concept and Relationship (ICR)	Each substance is a column in feature vector space
Location	Location of the incident or reported complaint	ICR, Google Geocoding API	Location is an attribute in the feature space
Media	Presence of multimedia files such as video or images	Twitter API	A boolean vector in the feature space denoting the presence or absence of media in a tweet

Table 5. List of features extracted in case study 2.

Feature	Summary	Technique	Presentation
Problem	The issue reported in the complaint for prompt addressal by concerned department	CoreNLP, POS Tagging, ConceptNet	Problem or issue reported in a tweet is a unique vector in feature space
Location	The region (or city) from where the complaint is reported	Indico API	Location is an attribute in our column-divided into city, state, town, region
Landmark	The exact pinpoint location (or landmark) of the problem	OpenStreetMap API	Landmark is a unique vector in our feature space

case study 1 and case study 2 respectively. Since this tutorial is compiled from our previous papers, in this paper, we only provide the summary of extracted features. We recommend our readers to read the full version of our previous studies Mittal et al. [16] and Agarwal et al. [1].

5 Classification

5.1 Case Study 1

In next step of the processing pipeline of our proposed solution method, we use an ensemble learning based Support Vector Machine (SVM) classifier. We divide our data (tweets categorized as unknown in AISP classifier) into training (25%) and testing dataset (75%). We use the feature vector model created in previous phase and learn our one-class classification model-a tweet either belongs to "complaint and grievances (C&G)" class or is identified as "unknown". Previous work indicates that the performance of an SVM classifier can be improved by modifying the kernels of the classifier. To investigate the performance of our proposed approach and evaluate the performance across various dimensions, we train our model by varying the kernel parameter of our SVM: linear, polynomial and RBF (Radial Basis Function) Kernels. Further, we use the application of ensemble methods to boost the performance of our baseline SVM classifier by combining kernels into cascaded and parallel manner.

In addition to the identification of complaint reports, we also identify the issue reported in the complaints for a quick addressal by concerned department. Due to the diversity in issues reported in complaints, we perform topic modeling on C&G reports. We use Alchemy Concept API by IBM Watson and identify the hidden topics in complaints. For example, in @RailMinIndia, delay in train, refund-issue, cleanliness, poor service and assistance in train coach and several more related complaints. Similarly, in @dtptraffic, complaints focus on the topics like traffic rules violation, illegal challans, riding motor-bikes without helmets and similar complaints with different issues can be there. We address the challenges of keyword spotting methods and use NLP based methods to find such words and label these complaints into the most likely topic and sub-topic defined in the taxonomy hierarchy. Examples include, riding without helmet or driving without a number plate comes under traffic violation related complaints. More examples: unhygienic food serving or low quality facilities to train passengers are tagged as poor assistance in coaches.

5.2 Case Study 2

Based on our inspection of complaints reported posted to @MORTHIndia and @nitin_gadkari, we divide tweets into 3 categories: Useful tweets, Nearly-Useful tweets, and Irrelevant tweets. We use Rule-based classifier trained on the features extracted in previous phase (problem, location and landmark).

Irrelevant Tweets (IRT): We define a tweet as irrelevant if the post is not about the poor conditions and irregularities of roads or highways causing life risks, discomfort, hazard or poor experience to the citizens. We observe that the tweet labeled as irrelevant is either off-topic or the authors discuss the problems related to road and transport; however the complaint is not about the poor conditions of roads or faulty and dysfunctional facilities.

Useful Tweets (UT): Useful tweets U_t are the posts which are a clear indicator of complaints and can be used to identify the low-level details of the issue faced by the citizens. Given a tweet t_i and a set of named entities $X = < T_c, T_l, T_p >$, we define t_i as a useful tweet- $t_i \in U_t$ if $t_i \in N$ (dataset) and $t_i = \{X, T_o \mid \forall x \in X : P(x)$ where $P(x) \neq \phi\}$. While, T_c denotes the name of city or region, T_l denotes exact geographical location or landmark, T_p denotes the problem or issue reported in the complaint and T_o denotes other words in the tweet.

Nearly-Useful Tweets (N-UT): Nearly-Useful tweets are the tweets posted for complaining a report but containing incomplete or insufficient information about the issue. For example, missing the exact location of the problem, ambiguity in reporting the issue, defining the problem on an abstract level and lacking the details. Given a tweet t_i, we define it as a nearly-useful tweet $t_i \in NU_t$ if $t_i \in N$ and $t_i = \{X, T_o \mid \exists x \in X : P(x)$ where $P(x) = \phi\}$. For simplicity of tweets, we removed the @username mentions from the tweets and used the corrected and enriched form of tweets (after applying micropost-enricment algorithm). We further apply the geographical location hierarchy model (bottom-up) and use graph backtracing method to identify the region or locality for a given pinpoint location in the tweet. For a given pinpoint location T_l, we use OpenStreetMap API and extract the detailed information associated with a geographical location.

6 Empirical Analysis and Results

In this Section, we discuss the empirical analysis performed on the tweets, and acquired experimental results performed on the complaint reports.

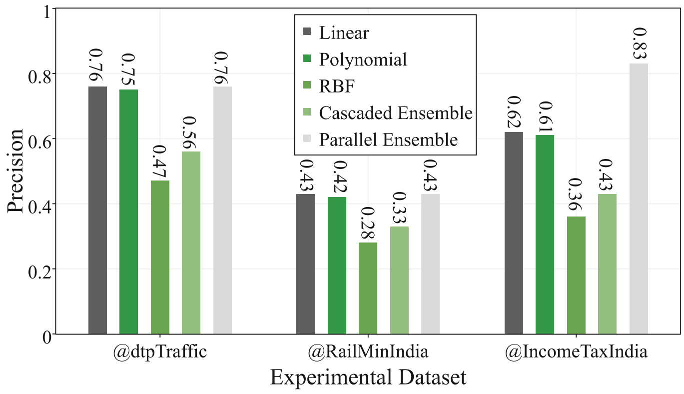

Fig. 1. Confusion matrix for C&G Tweets classifiers. Three different SVM kernel functions & Ensemble Classifiers. Column graphs illustrate - linear kernel outperforms other kernels. Ensembling all kernels in cascaded or parallel boost overall performance of every kernel

6.1 Case Study 1

Proposed AISP classifier identifies 47 (A:12, IS:27, P:8), 132 (A:20, IS: 99, P: 13), 121 (A:41, IS:76, P:4) and 35 (A:4, IS:30, P:1) tweets as AISP for @dtpTraffic, @DelhiPolice, @RailMinIndia and @IncomeTaxIndia respectively. According to our empirical results, we are able to correctly categorize 124, 34, 32 and 103 tweets for @DelhiPolice, @dtpTraffic, @IncomeTaxIndia and @RailMinIndia respectively. Figure 1 reveals that linear kernel in SVM outperforms RBF kernel with a reasonably high margin (variation from 20% to 30%). Figure 1 reveals that for @dtpTraffic (which is linear kernel), we are able to attain the maximum precision rate: 76% ($184/(184 + 58)$). However, for @IncomeTaxIndia & @RailMinIndia, we were able to attain a precision value of 62% ($31/(31+19)$) and 43% ($188/(188+252)$) respectively. Our result shows that using linear kernel in one-class SVM classifier, we are able to get an accuracy upto 60%. However, there is an overall misclassification of upto 10% in detecting complaint tweets as unknown. The column graphs in Fig. 1 reveals that linear and polynomial kernels gives similar results with a small difference 1% to 2% in precision value. Using SVM polynomial kernel, in @dtpTraffic experimental dataset, we were able to identify complaint tweets with a precision of 75% ($170/(170 + 56)$). While, for @RailMinIndia and @IncomeTaxIndia, we were able to identify complaints tweets with a precision rate of 42%($139/(139 + 189)$) and 61% ($22/(22 + 14)$) respectively. In order to compute the efficacy of our approach for correct classification, we record an overall misclassification of 12% (complaint tweets wrongly classified as unknown) for all accounts for polynomial kernel SVM classifier. Parallel ensemble SVM classifier outperforms cascaded ensemble classifier. Using a combination of linear, polynomial and RBF kernels in parallel manner, we were able to achieve a precision of 75%, 83% and 39% for @dtpTraffic, @IncomeTaxIndia and @RailMinIndia respectively. While, there is an overall misclassification of 18% in identifying complaint tweets as unknown. Figure 1 reveals that by arranging these kernels in cascaded order, it decreases the performance of overall classification from 10% to 20%. For example, for @dtpTraffic and @IncomeTaxIndia datasets, we achieve a precision of 56% and 43% respectively which are approximately 20% lesser than the individual precision of linear kernel SVM classifier. In comparison to cascaded ensembling, in parallel ensemble classification, we are able to boost the accuracy for @IncomeTaxIndia dataset by 21% whereas for @dtpTraffic, the performance is maintained with a precision of 76%.

6.2 Case Study 2

We conduct our experiments on $3,302$ random sampled tweets collected for @MORTHIndia and @nitin_gadkari and report the accuracy of AISP and complaint tweets classifiers. Proposed AISP classifier identifies a total of 20.5% (680 out of 3302) tweets as certainly non-complaint (AISP) reports. Among 680 AISP tweets, 417 tweets are identified as news or information sharing tweets. While, 166 and 97 tweets are identified as appreciation and promotional tweets respectively. Based on our AISP experimental results, we perform rule-based classifier on the remaining $2,622$ tweets and identify useful, nearly-useful and irrelevant complaint reports. Table 6 reveals that a very small percentage of tweets (6.4% - 170 tweets out of 2622 reports) are identified as complete reports that contains all three important components of a killer road complaint. Whereas, the largest chunk of reported tweets is classified as incomplete or nearly-useful tweets (1718 reports out of 2622 tweets). Our experimental results reveal that further only a very small percentage of tweets are convertible (N-UT-C) into complete or useful tweets (50 tweets out of 1718 nearly-useful tweets) while 97% (1668 out of 1718) of nearly-useful tweets have either landmark or concrete problem component missing from the tweets and hence not-convertible.

Table 6. Confusion matrix results for the rule-based classifier

		Predicted			Total
		N-UT	IRT	UT	
Actual	N-UT	1088	131	59	1278
	IRT	376	569	17	962
	UT	254	34	94	382
Total		1718	734	170	**2,622**

Based on the results acquired by our rule-based classifier, we classify bad road related complaints with an overall accuracy of 67%. In addition to measuring the accuracy of our classifier, we also measure the overall recall value of the classification. Based on our results and the Table 6, we record a recall of 65%. As discussed earlier, the complaints reported to public agencies' accounts are user-generated content and lacks a standard format or terminology for com-plaining a report. Further, the excessive use of metaphor and sarcasm while reporting a complaint generates false alarms and impacts the overall accuracy of the classification.

7 Conclusions

We present case studies on identification of complaints reported to public agencies' accounts on Twitter. We formulate our problem as a one class classification problem and conduct two case studies on "complaints seeking for an immediate response" and "complaints reported to bring the attention of the government (bad road conditions)". We identify various linguistic and contextual features for identifying complaints reports tweets. We also propose various features that are strong indicators of a tweet to certainly not to be a complaint report. Our results reveal that linear kernel one-class SVM outperforms RBF with a margin of 20% while polynomial and linear kernels produce the similar results with a difference of 1% to 2% of precision. Furthermore, parallel ensembling of kernels outperforms cascaded ensembling. In second case study, we apply rule based classifier on three important components of a killer road complaint; problem reported in the complaint, landmark or pinpoint location, city or location information. Our results shows that the proposed approach classifies these complaint reports with an accuracy of 67% and a recall of 65%. Our results reveal that maximum number of complaints are reported about the risky and accident prone roads while most of them are due to the poor condition of amenities.

References

1. Agarwal, S., Mittal, N., Sureka, A.: Potholes and bad road conditions-mining Twitter to extract information on killer roads. In: The ACM India Joint International Conference (CoDS-COMAD), India. ACM (2018, under-review)
2. Agarwal, S., Sureka, A.: Using KNN and SVM based one-class classifier for detecting online radicalization on Twitter. In: Natarajan, R., Barua, G., Patra, M.R. (eds.) ICDCIT 2015. LNCS, vol. 8956, pp. 431–442. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-14977-6_47
3. Agarwal, S., Sureka, A.: Investigating the potential of aggregated tweets as surrogate data for forecasting civil protests. In: Proceedings of the 3rd IKDD Conference on Data Science, p. 8. ACM (2016)
4. Agarwal, S., Sureka, A., Goyal, V.: Open source social media analytics for intelligence and security informatics applications. In: Kumar, N., Bhatnagar, V. (eds.) BDA 2015. LNCS, vol. 9498, pp. 21–37. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-27057-9_2
5. Anderson, M., Lewis, K., Dedehayir, O.: Diffusion of innovation in the public sector: Twitter adoption by municipal police departments in the US. In: Portland International Conference on Management of Engineering and Technology (2015)

6. D'Andrea, E.: Real-time detection of traffic from Twitter stream analysis. IEEE Trans. Intell. Transp. Syst. **16**(4), 2269–2283 (2015)
7. Edwards, A., de Kool, D.: Webcare in public services: deliver better with less? In: Nepal, S., Paris, C., Georgakopoulos, D. (eds.) Social Media for Government Services, pp. 151–166. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-27237-5_8
8. Frias-Martinez, V., Sae-Tang, A., Frias-Martinez, E.: To call, or to tweet? Understanding 3-1-1 citizen complaint behaviors. In: ASE (2014)
9. Fu, K., Nune, R., Tao, J.X.: Social media data analysis for traffic incident detection and management. In: Transportation Research Board 94th Annual Meeting (2015)
10. Gu, Y., Qian, Z.S., Chen, F.: From Twitter to detector: real-time traffic incident detection using social media data. Transp. Res. Part C Emerg. Technol. **67**, 321–342 (2016)
11. Heverin, T., Zach, L.: Twitter for city police department information sharing. Proc. Am. Soc. Inf. Sci. Technol. **47**, 1–7 (2010)
12. Khan, G.F., Swar, B., Lee, S.K.: Social media risks and benefits a public sector perspective. Soc. Sci. Comput. Rev. **32**(5), 606–627 (2014)
13. Kumar, A., Jiang, M., Fang, Y.: Where not to go? Detecting road hazards using Twitter. In: SIGIR. ACM (2014)
14. Mai, E., Hranac, R.: Twitter interactions as a data source for transportation incidents. In: Proceedings of Transportation Research Board 92nd Annual Meeting (2013)
15. Meijer, A.J., Torenvlied, R.: Social media and the new organization of government communications an empirical analysis of Twitter usage by the Dutch police. Am. Rev. Public Adm. (2014). http://journals.sagepub.com/doi/abs/10.1177/0275074014551381
16. Mittal, N., Agarwal, S., Sureka, A.: Got a complaint?- Keep calm and tweet it!. In: Li, J., Li, X., Wang, S., Li, J., Sheng, Q.Z. (eds.) ADMA 2016. LNCS (LNAI), vol. 10086, pp. 619–635. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49586-6_44
17. Schulz, A., Ristoski, P., Paulheim, H.: I see a car crash: real-time detection of small scale incidents in microblogs. In: Cimiano, P., Fernández, M., Lopez, V., Schlobach, S., Völker, J. (eds.) ESWC 2013. LNCS, vol. 7955, pp. 22–33. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41242-4_3
18. Wanichayapong, N.: Social-based traffic information extraction and classification. In: 11th International Conference on ITS Telecommunications (ITST). IEEE (2011)