Transportation Research Interdisciplinary Perspectives 3 (2019) 100062

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: https://www.journals.elsevier.com/transportation-research-interdisciplinary-perspectives

TravelBot: Utilising social media dialogue to provide journey disruption alerts

Paul Gault [a,1], Caitlin D. Cottrill [b,*], David Corsar [c,2], Peter Edwards [c], John D. Nelson [b,3], Milan Markovic [c], Mujtaba Mehdi [a,4], Somayajulu Sripada [c]

[a] Department of Computing Science, University of Aberdeen, UK
[b] Centre for Transport Research, University of Aberdeen, Aberdeen AB24 3UE, UK
[c] Department of Computing Science, University of Aberdeen, Meston Building, Aberdeen AB24 3UE, UK

ARTICLE INFO

Article history:
Received 28 December 2018
Received in revised form 5 October 2019
Accepted 23 October 2019
Available online 15 November 2019

Keywords:
Social media
Twitter
Travel disruption
Real-time passenger information
Passenger-operator dialogue

ABSTRACT

Use of social media in the public transport sector is rapidly increasing, driven by both passenger demand, and recognition by transport operators of the insights that social media enables. This paper explores the potential for utilising social media (specifically, the Twitter platform) to provide personalised information to public transport passengers, drawing from lessons learned from related studies. The *Tweeting Travel* study developed an understanding of the types of dialogues that can unfold on social media between passengers and a simulated travel advice system and then used this to shape development of the *TravelBot* system. This system provided users with real-time passenger information, including details of relevant travel disruptions that were automatically extracted from social media posts. A user evaluation of a *TravelBot* trial is presented, findings of which showed that participants highly valued the service and the information it provided, with most indicating a strong desire for the system to continue operation. These findings reveal the potential offered by social media for more personalised communication between public transport operators and their passengers, as well as indicating an efficient method by which this communication may be enabled.

1. Introduction

Recent years have seen rapid growth in the use of social media in the public transport sector. This has been driven, in part, by passenger desire to use social media as a channel for improved communication with transport providers; while transport operators have also recognized the value to be gained from insights about the passenger experience (Accenture, 2013). Conversations between passengers and operators are perhaps most frequently seen on Twitter (Passenger Focus and Abellio, 2012), an online social media platform that allows users to share short messages (originally 140 characters, though raised in November 2017 to 280 characters) and engage in dialogue via responses, re-Tweets, and shared hashtags.

With the use of social media on the rise – as of Q1 2018, Twitter alone has 336 million monthly active users (with 13.1 million users in the UK, or roughly 20% of the population (Aslam, 2018)) - and increasing expectations around the benefits that it can confer, it is unsurprising that many organisations are exploring ways in which they can more fully leverage the opportunities it offers. This paper presents the findings from research that has explored the potential for utilising social media (specifically, the Twitter platform) in the area of public transport and traveller information. The findings demonstrate how social media may be used to create an engaging dialogue between public transport information providers and passengers that draws upon contextual information and provides personalised, real-time information to facilitate an improved travel experience. The use and understanding of such dialogues is an important aspect of this research that has grounded the development of the experimental travel advice system presented here.

This paper reports on the implementation of a working prototype travel advice service (called *TravelBot*) which automates the collection of transport information and the provision of real-time passenger information (RTPI) through social media. A user evaluation of the prototype was conducted in order to better understand the user experience of a system that automates RTPI via social media. The research sought to assess how such a system could be used, the impact of the information it provides on users (e.g. impact on travel behaviour), and the challenges of detecting travel disruption from social media posts.

* Corresponding author.
E-mail addresses: paulg@young.scot, (P. Gault), c.cottrill@abdn.ac.uk, (C.D. Cottrill), d.corsar1@rgu.acuk, (D. Corsar), p.edwards@abdn.ac.uk, (P. Edwards), j.d.nelson@abdn.ac.uk, j.nelson@sydney.edu.au, (J.D. Nelson), milan.markovic@abdn.ac.uk, (M. Markovic), yaji.sripada@abdn.ac.uk. (S. Sripada).
[1] Current address: Young Scot, Rosebery House, 9 Haymarket Terrace, Edinburgh EH12 5EZ, UK.
[2] Current position: Lecturer, Robert Gordon University, UK.
[3] Current address: Chair in Public Transport, Institute of Transport and Logistics Studies (ITLS), University of Sydney, Australia.
[4] Current position: Kingston University, London, UK.

http://dx.doi.org/10.1016/j.trip.2019.100062

This paper presents the development, pilot implementation, and evaluation of the *TravelBot* system. Section 2 discusses previous research and related work that contributed to development of the underlying system requirements; Section 3 discusses the outcome of applied research to establish the design requirements for TravelBot; Section 4 describes the design of the *TravelBot* system and the *Tweeting Travel* study; Section 5 describes a user evaluation of *TravelBot*; Section 6 presents limitations of the study and future research needs and Section 7 presents conclusions and discussion.

2. Background and related work

A considerable literature is emerging on the role of social media in enabling the implementation of measures to promote smart mobility. Much of this is associated with the debate around smart cities and the critical role of information in enabling smarter travel choices. Nam and Pardo (2011) have stated that, 'The concept of smart city is not novel, but in the recent years it has taken on a new dimension of using Information and Communication Technologies (ICTs) to build and integrate critical infrastructures and services of a city. The initiatives of making a city smart have recently emerged as a model to mitigate and remedy current urban problems and make cities better as places to live (p. 283).' A core component in facilitating the smart city is smart or intelligent mobility, as identified in Chourabi et al. (2012) and Giffinger and Gudrun (2010), which has been defined as a system that uses innovative approaches for, 'Promoting more efficient and intelligent transportation systems—effectively leveraging networks to ensure more efficient movement of vehicles, people, and goods, thus reducing gridlock; and promoting new 'social' attitudes such as car sharing, carpooling, and car-bike combinations (Lim et al., 2016).' Mulley et al. (2019) similarly emphasise that a key characteristic of intelligent mobility is the appropriate use of new and emerging technologies linked to the wider societal objective of enabling the smarter, greener and more efficient movement of people and goods (Mulley et al., 2019). They also note that intelligent mobility application areas can be argued to encompass everything from connected and autonomous vehicles, to the provision of systems to support the planning and execution of seamless multi-modal journeys and the supporting intelligent infrastructure and services required to achieve this.

Data and information are key enablers of these smart mobility solutions; Intelligent Transport Systems (ITS) draw upon a wide range of data resources as inputs into mobility networks (Zhang et al., 2011), and new methods of information sharing between transport systems, operators, and passengers are evolving at a rapid rate (Lim et al., 2016; Papangelis et al., 2016). The role of high quality information as a key enabler of successful transport service provision and smart technology plays a critical role in ensuring that real-time journey planning, pre-trip and en route, is increasingly prevalent (Nelson, 2018). Social media is particularly relevant, as platforms such as Facebook, Twitter, and Whatsapp provide cost-efficient, timely, and public means for transport information to be shared in a multi-directional context (between transport operators and agencies, passengers, and other relevant parties) (Camacho et al., 2012; Gal-Tzur et al., 2014). The data generated through such engagement is being increasingly evaluated for inclusion into the transport ecosystem, with an increasing number of studies, for example, exploring the use of crowdsourced data to identify events or disruptions that may impact upon the transport network. Studies include those such as Gu et al. (2016), which mined Twitter data to identify traffic incidents; Markou et al. (2019), which predicted taxi demand hotspots during events using internet search queries of event listing sites such as Facebook and Ticketmaster; to Abbasi et al. (2015), which explored the use of crowdsourced Twitter data for contributing to travel demand models. These and other studies demonstrate the range of applications available for crowdsourced data; however a note of caution must also be sounded. As noted in Imran et al., 2015, processing social media data is not a straightforward task, but rather requires extensive system training, careful parsing of messages for relevance, and the ability to handle massive amounts of information. Additional challenges noted by Ganti et al. (2011) and Mehmood et al. (2017) include issues of data privacy, security and integrity.

Despite these challenges, however, IBM (n.d.) argues that there are several key characteristics of social media that make it especially relevant and appropriate for contributing to the development of smart cities and regions which they summarise as:

- Engaging: Promotes citizen involvement, and creates a sense of ownership and collaboration;
- Transparent: Provides a view into how the city works, with open dialogue and rapid feedback;
- Nimble: Services are delivered in real-time, and are able to adjust depending on needs;
- Secure: Respectful of privacy and leveraging technology to enhance citizen security.

While these characteristics may not be true in all cases (the security of social media data, for example, has been called into question by the practices used by Cambridge Analytica with respect to Facebook (ICO, 2018), amongst others), it is evident that social media does provide potential to facilitate convenient, flexible dialogue between multiple actors. The potential for social media to enhance smart mobility is perhaps in no place more prevalent than in the provision of public transport information. A number of recent studies have examined how public attitudes towards public transport are reflected in Twitter posts. Collins et al. (2013), for example, found that evaluation of short social media messages to determine rider satisfaction could be beneficial to the public transport agencies insofar as:

- The cost of data collection is minimal;
- Data can be collected in real-time;
- User-specific needs can be assessed;
- The data collected can provide meaningful insight as to why a particular sentiment is felt (Collins et al., 2013).

Such benefits are demonstrative of the emerging capabilities in terms of social media use for enhancing the analysis of public transport provision. Additional studies (such as Gault et al., 2014; Schweitzer, 2014; and Cottrill et al., 2017) have further expanded upon these findings by demonstrating that it is not only the *medium* but also the *method* of engagement that may have positive or negative impacts. Schweitzer (2014) states that, 'The evidence suggests that engagement online via interaction with individual commenters may pay off much more than blasting information, and those benefits can accrue both in content related to the agency's reputation and to the planning dialog' (Schweitzer, 2014). In other words, utilising the ability of social media to allow for engagement on a personal level may confer more benefits to the public transport agency than simply using it as a platform to share generic, widespread information. This finding was echoed in Gault et al. (2014), where they indicated that factors related to *persistent conversations* (their emphasis), the provision of real-time information, and identity management are all key to developing effective social media strategies for public transport operators.

In addition to evaluating social media itself, there has also been parallel interest in how social media may be linked with established journey planners or otherwise used to provide dynamic transport information. Public transport information has long been a robust area for smartphone application development and online services, with services ranging from bespoke apps (such as the UK-based First Bus App,[5] which integrates bus-based journey planning, ticket purchase, and service information) to more general tools (such as Google Maps[6]), to 'Mobility as a Service' (MaaS) apps, such

[5] https://www.firstgroup.com/tech-bus/first-bus-app.
[6] https://www.google.com/maps.

as Whim,[7] which aim to integrate multi-modal journey planning, payment, and information. There has also been exploration of more personalised services through, for example, Google Assistant,[8] which is working to integrate more personalised information on public transport services by integrating real-time data and walking directions (Hariramani, 2019); however, the potential provision of personalised information based on demonstrated or expected travel habits has been less widely investigated. Similarly, the integration of data 'crowdsourced' (gathered online from a wide variety of data suppliers) from social media with structured data from transport agencies is an approach that has been explored in areas such as campus transport (Qiao et al., 2011) and traffic prediction (Pereira et al., 2014). It has also found its way into transport information systems such as Waze (Xu and Dodds, 2015) and Moovit (Faber and Matthes, 2016). Heiskala et al. (2016) provide a useful summary of case study attributes of Moovit, Waze, and TrafficSense, indicating the incentives and disincentives of crowdsensing in the transport market, with particular attention to the need for widespread adoption for services to be beneficial. Often, however, data presented is considered to be timely and accurate, but not particularly personal (i.e. responsive to an individual's information needs) unless it is a direct response to a question or comment. In many cases, the broadcast of information from 'official' sources (such as a public transport operator) on social media platforms is generally useful, but perhaps not always particularly relevant. However, given that such feeds often serve as trusted sources of information (Gault et al., 2015), the potential for developing a system that builds upon crowdsourced data and institutional knowledge to provide information tailored to the individual is a topic of emerging interest.

These findings suggest that there is potential for more personal interactions between public transport operators and their passengers via social media, an observation that forms the basis of the work presented in this paper. By using the familiar, standard platform of Twitter, rather than developing a new stand-alone application, it is possible to both leverage the trust implicit in the brand (Pentina et al., 2013), and reduce the burden for participants (both passengers and operators).

3. Understanding requirements for the design of TravelBot

In this research a series of linked preparatory studies were conducted to develop an understanding of the role of social media in improving communication between public transport operators and passengers. The question of how social media is currently used by public transport operators was addressed first. The objective was to understand the existing professional practice of social media usage from a public transport operator's perspective, including both the internal processes they use to manage real-time travel information and the ways in which they present this externally through such a channel. An initial study was conducted in collaboration with FirstGroup plc, a prominent public transport operator in the UK and North America with headquarters in the study area (Scotland), which reviewed three of their Scottish subsidiaries to provide a comparison of their social media practices. Ethnographic fieldwork (Jones et al., 2013; Salvador et al., 1999), including observations and interviews at the different locations, was used to understand the internal working practices of the staff responsible for managing the delivery of their outwardly facing social media feeds. To supplement the ethnographic field work data, the Tweets flowing through the operator's accounts were collected over a four-week period in January and February 2014, which were then classified and a schema generated based on their content.

Following a content analysis (Sommer and Sommer, 1991) of the combined Tweets and the field work data, a set of themes emerged which helped inform an understanding of the internal practice for maintaining social media accounts across the subsidiaries. The themes are outlined here; for a more detailed discussion see Gault et al. (2014):

- *Identity management strategies*: Carefully maintained Twitter personas were used by operators as a means to relate to passengers.
- *Persistent conversation*: The persona(s) provide a consistent identity (or identities) that customers can engage with on a regular basis. Such conversations extend the operator-customer relationship beyond the time that the customer is travelling on an operator's vehicle(s).
- *Provision of real-time passenger information*: The approach to providing RTPI varied between the subsidiaries, primarily due to differences in site characteristics and the internal availability of relevant information. These factors dictated whether the operator was reactive or proactive about publishing updates on Twitter concerning disruptions to their services.

Building on these findings, the next activity looked at how public transport passengers engage with social media. An ethnographic study was conducted to explore the social media experience from the perspective of passengers who commute by bus. Nine participants were recruited through advertisement on a University email listserv that predominantly addresses postgraduate students, and staff, and they took part in activities designed to assess how they integrate social media technology into their daily travel routines. Participants were required, first, to regularly commute by bus; and, second, to be active users of online social media. Five participants made local journeys within the city of Aberdeen, UK, two made regional journeys, and two a combination of local and regional trips. All participants used Facebook and six used Twitter, with seven using at least two platforms. This study used the Cultural Probes (Gaver et al., 1999) method, which required participants to fill out a series of bespoke, incomplete artefacts as a mechanism to reflect on their commuting and social media experience (an example is shown below in Fig. 1). A number of factors were considered in the probe design in order to facilitate a meaningful dialogue with participants, including openness, boundedness, materiality, a sense of pace and challenge (Wallace et al., 2013). In addition, the nature of a cultural probes study, which is design-led and largely qualitative (Hutchinson et al., 2003), lends itself to a smaller sample size to allow for more robust evaluation of returned materials.

A total of 98 probe items were completed and evaluated by the project team, with the following resulting themes taken forward to help inform the next stage of system development:

- *Kinship networks*, which recognized other commuters in the extended social network of passengers as potential sources of relevant information.
- An *awareness of social media habits and usage patterns* was helpful to identify particular times of day where commuters are likely to engage with information made available through social media channels.
- *Lack of control* was found to have both positive (passengers reporting travel time as time used for other activities such as reading) and negative (reported frustration with unexplained delays and lack of ability to proactively respond) associations.

The operator and passenger perspectives explored in these studies helped shape and inform the subsequent *TravelBot* activities, revealing an opportunity to exploit social media to strengthen the relationships between operators and their passengers. This required an understanding of the types of dialogue that could unfold between these two groups, and the role that technology might play in supporting this activity.

4. Designing the TravelBot system

To further explore *what kinds of dialogue social media can facilitate between transport operators and passengers*, prototyping concepts to enable operators to exploit social media to strengthen ties with their passengers were developed, largely building upon the concept of dialogues. This exploration led to the further question of *how can such dialogues be utilised as a source of information*, which explored how relevant information could be extracted from any unfolding operator-passenger dialogues, and how such information could be directed to other passengers to help inform their journey planning.

[7] https://whimapp.com.
[8] https://assistant.google.com/.

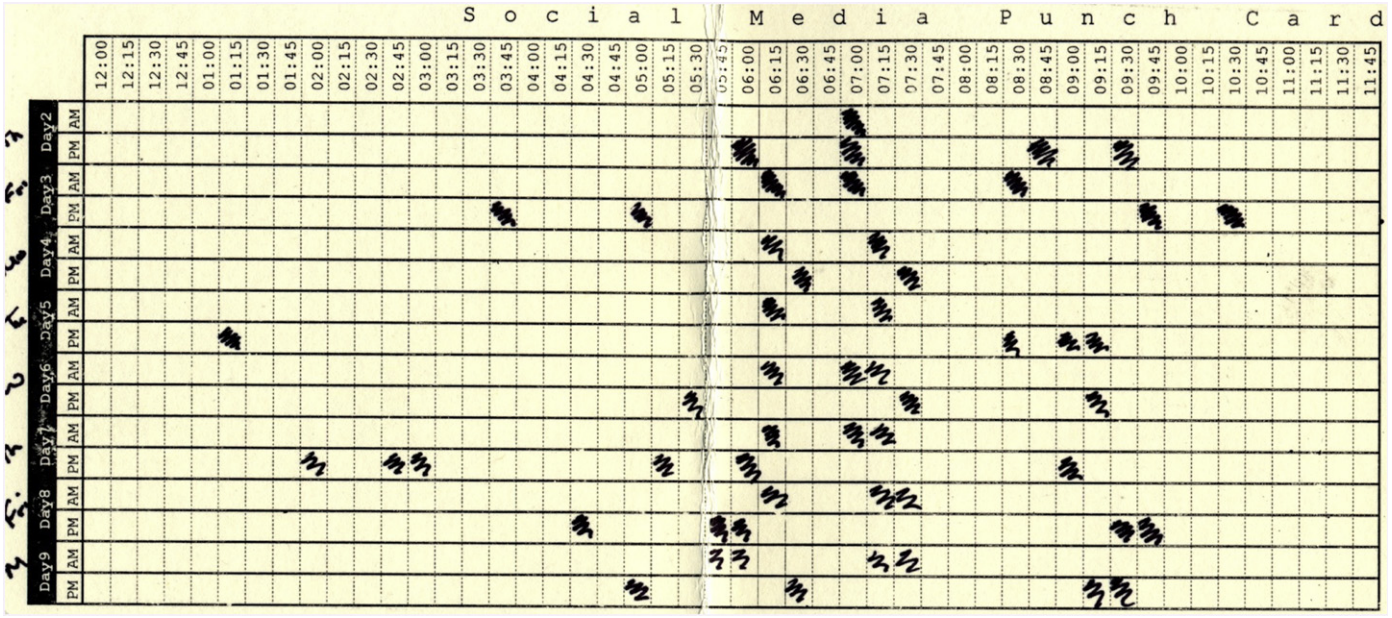

Fig. 1. Social Media Punch Card. Participants were asked to indicate the times at which they engaged with their social media accounts, and later reflect on what these patterns revealed about their social media habits.

4.1. Understanding dialogue between passengers and an online system

The findings presented above shaped the *Tweeting Travel* study, which aimed to understand the types of dialogues that would unfold on social media between passengers and an 'automated' travel advice system. The *Tweeting Travel* study utilised the Wizard of Oz (Dow et al., 2005) technique, a software prototyping method in which participants interact with a computer system that appears to be automated, but is actually operated by humans. In the *Tweeting Travel* study this involved two research team members simulating the role of a social media-based travel advice system by acting as "travel assistants". This allowed us to gain insights into the information and technology requirements for supporting such dialogues, and to avoid focusing on a specific system design based on an as yet incomplete group of assumptions regarding expected user interactions.

Seven commuters (four female, three male, all Twitter users) were recruited to participate in the study in Aberdeen, Scotland, again using the University email listserv. Participants were selected based on the similarity of their bus usage patterns in terms of services used, time, and places of travel. This similarity was required in order to allow for potential sharing of relevant information on the travel routes. Due to the intensive nature of the Wizard of Oz method, which required that operators be able to respond in real-time to participants who were travelling at similar times, participant numbers were kept small and this was considered adequate to test the concept. Participants were required to send at least one Tweet during each journey, which was added to a dataset that was analysed to determine the characteristics of these messages. Follow-up interviews with participants enabled further understanding of the user experience of the system.

At the beginning of the study, participants were asked to register information about their regular commute (including departure bus stop, destination bus stop, time of journey, journey frequency and amount of time before the journey they wished to be sent travel updates) using a registration website (see Fig. 2). This allowed the travel assistants to limit their correspondence to journey-appropriate messages. Participants were given an initial training session to show them how to register their journeys along with a step-by-step instruction booklet so they could do this independently.

The travel assistants made use of multiple sources of information in developing relevant messages to send to participants. A *journey monitoring interface* built by the developers displayed all of the current and upcoming registered journeys, along with the Google Maps website, which was used to both track the routes of participants' journeys and show any points where registered journeys might overlap. Tweetdeck,[9] a website which provides a dashboard for organising feeds of Tweets and Direct (private) Messages between two parties, was used to monitor messages from study participants, the bus operator and local media (radio) outlets for any travel updates (such as details of collisions, roadworks, etc.). Finally, a web site providing real-time bus arrival details was used as a source of information for sending updates to participants about their forthcoming journeys. These were used to provide participants with expected arrival times and other information pertaining to their journey, such as notifications of any disruptions that they might encounter.

The study took place over a period of two weeks in February 2015 with 24 journey itineraries registered with the system. During this time, participants were sent 208 Direct Messages, and contributed 102 Tweets and 113 Direct Messages. The balance between (public) Tweets and (private) Direct Messages sent by the participants indicated their willingness to share information about their journeys. The travel assistants communicated with participants using Direct Messages to protect their privacy by not publicly publishing details of their journeys and to ensure participants only received messages that were directly relevant to them.

Participant interviews (Kuniavsky, 2003) were then conducted to collect first-hand descriptive accounts of the opinions, attitudes, perceptions and experiences following the two-week period of *Tweeting Travel*. All interviews (which lasted 30 to 45 min each) were semi-structured, and so could evolve naturally based on where the conversation progressed.

Following analysis of the data collected during the study, including all of the Tweets sent/received and the follow-up interview transcripts, three key characteristics regarding the user experience of such a service were observed. These included: message tone, the potential for conversations to lead to useful information; and the certainty of language used.

Participants noted that the *tone of messages* they received differed between being 'impersonal or robotic' and 'friendly or conversational'. *Impersonal or robotic* messages were characterised by fitting as much information into the then-designated 140-character limit without necessarily attempting to offer a willingness to engage in further chat. *Friendly or*

[9] http://tweetdeck.twitter.com/.

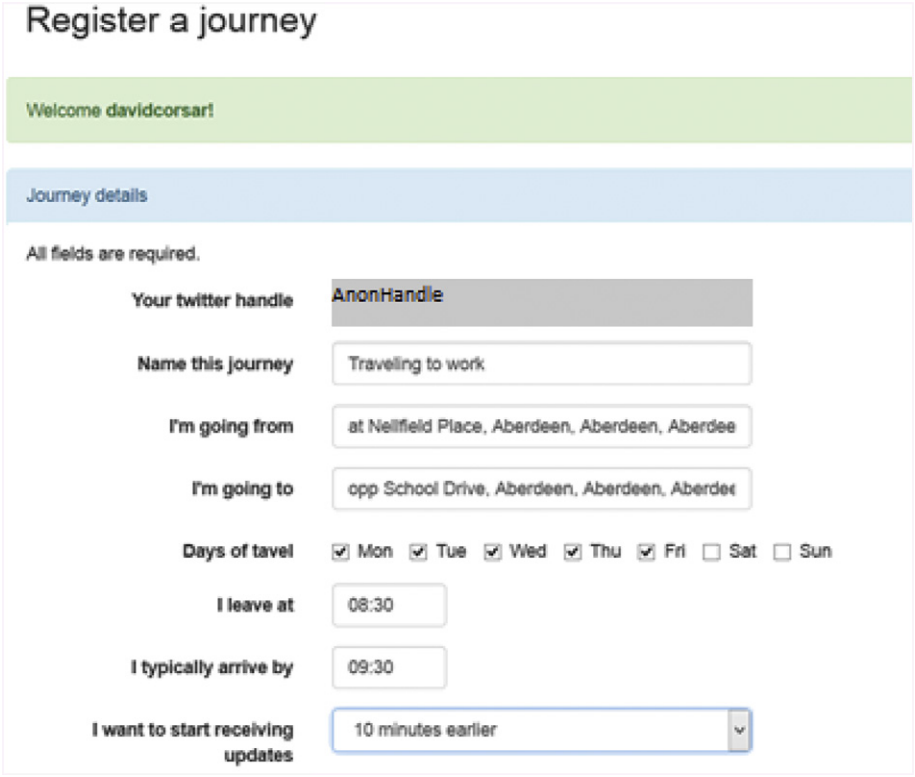

Fig. 2. The *Tweeting Travel* and *TravelBot* user journey registration form. (Note: The Twitter handle here has been di-identifiEd.)

conversational messages were more outgoing, so offered a greeting and responded to messages from participants in a more conversational way (even verging towards use of humour at some stages). The difference in tone was largely related to the different persons acting as travel assistants, though the findings related to tone were useful for the further study.

The *potential of conversations to lead to useful information* links to the previous finding, whereby if the travel assistant engaged in a conversation with the participant this could lead to the participant providing additional information about their journey. In one example, a participant Tweeted about missing the bus she was due to catch, leading to a conversation with the travel assistant about things to do while waiting for the next bus. Having caught the next bus, the participant went on to report a delay on a particular section of the route which could then be shared with other participants on the same route. This experience indicated that maintaining an open, friendly dialogue is important, as it increases the likelihood of further useful information being reported.

The *certainty of the language used* related to the delivery of real-time information about an upcoming journey, and the desire for the travel assistant to alter their choice of words based on the amount of time until a registered journey was due to take place. For example, 20 min before a journey was due to occur, the phrase “Scheduled for…” could be used, whereas 5 min before a journey, the phrase “Due in…” would be more appropriate. This subtle shift in the use of language was an important factor in influencing how a travel advice system might compose any automated messages.

4.2. Requirements summary

The studies discussed here revealed a complementary set of findings which informed the requirements for an automated travel advice service that uses social media:

- the system should be capable of supporting the transport operator with maintaining a persistent identity and automating the delivery of relevant, timely RTPI to customers;
- the system should complement the identity management strategies and collapsing of roles being exhibited by public transport operators in relation to their social media feed, allowing them to maintain a coherent identity while continuing to offer high levels of customer service;
- from the passenger perspective, the system should interpret knowledge of each user's social media and travel habits to determine time windows when any RTPI provision is likely to be beneficial;
- the system should seek to exploit passenger's positive perceptions of lack of control as an opportunity to obtain information from them, while playing a role in alleviating the negative perceptions of control loss by warning of/explaining journey delays and other disruptions;
- the system should be capable of using social media posts from passengers, their kinship network, public transport operators, media outlets, and other relevant parties, as potential sources of information relevant to the planned journeys of *TravelBot* users;
- the system should attempt to identify key terms and phrases within a message, and attempt to infer a machine processable representation of the information the message conveys. For example, if a user Tweets “My bus is running 10 minutes late” the system should infer a 10-minute delay to a trip and contextualise this by linking to details of the user's journey to identify the specific bus service to which the Tweet refers;
- Natural Language Generation (NLG) (Reiter and Robert, 2000) techniques should be used to produce messages that are personalised to the recipient and aim to engage them in some form of dialogue, recognizing the potential of conversations to lead to useful information. A human operator should also be able to send messages in tandem with the system to support this process. The certainty of the language in the generated messages should be affected by timing of the message delivery.

These requirements were then translated into a usable system (known as *TravelBot*), as described in detail in (Corsar et al., 2015), developed as a proof-of-concept for an automated approach to monitor Twitter for messages describing events that might impact upon a user's journey. If such messages were found, sending personalised messages to users to inform them of potential delays or disruptions (as shown in Fig. 3 below) to their registered journeys. In addition to Twitter, *TravelBot* utilised an open source

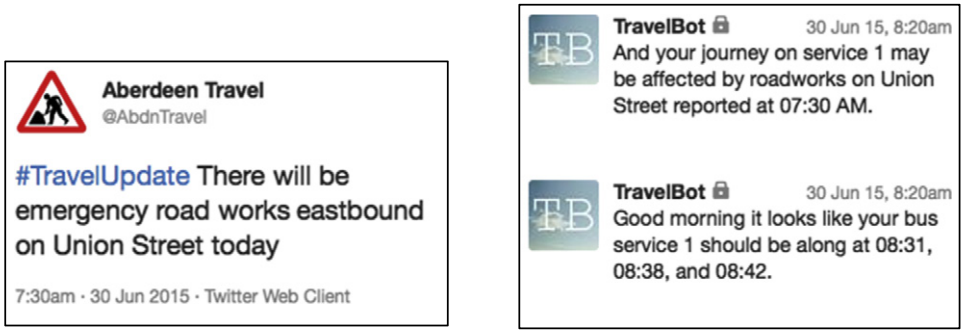

Fig. 3. (a) An example Tweet providing information about road works received by *TravelBot*; (b) sample messages sent by *TravelBot* (ordered by time bottom-to-top) informing a user of upcoming bus arrival times and warning of potential disruption due to road works described in an earlier Tweet.

web service that provides estimated real-time bus arrivals (the NextBus API), and a federated knowledge base of travel and transport-related information derived from various open datasets to ensure the quality of information provided to users. This knowledge base included details of public transport schedules (derived from the Traveline National Dataset), public transport access points (derived from the National Public Transport Access Nodes dataset), public transport localities (derived from the National Public Transport Gazetteer dataset), and the road network (derived from OpenStreetMap).

5. User evaluation

An evaluation of *TravelBot* was conducted to address the question of *What is the user experience of a system that automates RTPI via social media?*

5.1. Participants

Thirteen commuters (nine female, four male) were recruited as participants in the study. Again, while small, these numbers were deemed adequate for the purposes of the study. The recruitment criteria were that participants had to use Twitter and travel regularly on First Aberdeen services 1, 2, 3 and X40 during the three-week period of the study in August and September 2015. The participants were recruited by posting a message on the operator's Twitter feed and also strategically placing flyers at bus stops along the specific routes of interest. Two of the participants had been involved in the previous *Tweeting Travel* study (described in Section 4.1).

5.2. Procedure

The procedure for the evaluation study was as follows: The first stage verified that the participants met the recruitment criteria and had completed the ethics consent forms before any further participation. The second stage involved an entry survey with questions related to the participants' existing use and perceptions of travel information and whether or not it impacted upon their travel. The third stage of the study involved participants entering details of their planned bus journeys into the system. The fourth stage involved regular travel on the routes they had specified for the three weeks that followed. The final stage was an exit survey to gauge the response of the participants to information provided by the system; this included: repetition of questions from the entry survey to observe any shift in the participants' answers following use of the system; questions focusing on the experience of registering journeys, their consumption of the Direct Messages they received from the system (including frequency and volume of messages received), usefulness and reliability of the provided RTPI and any other information *TravelBot* provided (relating to disruptions or delays to their bus service). All thirteen participants fully completed both the entry and exit surveys.

5.3. Data collection

There were 36 journey itineraries (4 leisure, 32 commuting) registered with the system over the study period. During this time, participants were sent 615 Direct Messages while they contributed 56 Tweets and 47 Direct Messages. Although the *TravelBot* study lasted a week longer than the previous Tweeting Travel study, the amount of information contributed by participants was much less. They were encouraged to share information about any disruptions or delays that they had encountered but the emphasis in this study, rather than collecting a large body of user-generated Tweets, was on providing participants with relevant travel information automatically generated by the system.

5.4. Data journey examples

The *TravelBot* system is designed to support the sharing of information about disruptive events from "official sources" (such as the bus operator, media outlets, and local authorities) to users, between users, and from users to official sources. During the study, the typical flow was from official sources (particularly the bus operator and radio stations) to participants; however, the following example illustrates the sharing of information between users and the *TravelBot* system:

On /11 September 2015 at 13:13, participant "P1" sent a Direct Message to *TravelBot* stating: "The 1223 Auchinyell from Errol St is now running 25 min late. The bus has been diverted down Holburn St instead." From this message, *TravelBot* inferred a service diversion for bus services 1 and 2 (based on the location of the diversion), and, when sending messages to participants later that day included the information shown in Table 1.

Table 1
Information related to a service diversion sent to users during the *TravelBot* study.

Time	Recipient	Message extract
16:35	P2	"… service 1 is diverted at Holburn."
16:35	P2	"… service 2 Holburn diversion in place from 12:00 11 September to 16:59 11 September."
16:51	P3	"… service 2 Holburn diversion in place from 12:00 11 September to 16:59 11 September."
17:01	P4	"… service 2 diversion starts from 12:00 11 September to 16:59 11 September - Holburn."
17:01	P5	"… service 2 diversion starts from 12:00 11 September to 16:59 11 September - Holburn."
17:34	P1	"… service 2 diversion starts from 12:00 11 September to 16:59 11 September - Holburn."
17:34	P6	"… service 1 diversion starts from 12:00 11 September to 16:59 11 September - Holburn."
17:51	P7	"… service 2 Holburn diversion starts from 12:00 11 September to 16:59 11 September."
17:51	P7	"… service 1 Holburn diversion in place from 12:00 11 September to 16:59 11 September."

A further example of information flow from the transport operator (First Aberdeen) to *TravelBot* is the Tweet "Woodside Rd (BOD) closed until 14th Sept - Service X40 diverted. Details…" published by First Aberdeen on 09/09/2015. From this, *TravelBot* extracted details of the disruptive event (a service diversion) which was used to inform P2 at 16:36 on the same day that "…service X40 is diverted at Woodside." and at 17:15: "… service X40 Woodside diversion starts from 12:00 10 September to 16:59 10 September".

5.5. Survey results

Both the entry and exit surveys were designed to gather a mix of quantitative and qualitative feedback from the study participants. In the entry survey, participants were asked about their current experience with public transport information sources, with most responding that they use a number of different sources, as shown in Table 2. Most (8/13) also report that they 'Always' or 'Frequently' seek out information before making a regular bus journey, though only three participants responded that they will 'Always' or 'Frequently' change their travel plans based on RTPI. Views were mixed on the reliability of existing sources of RTPI, with eight of the 13 reporting that it is 'Very' or 'Somewhat' unreliable.

In the Exit survey, participants were asked to report on their overall experience of using the *TravelBot* system, along with recording any changes in attitudes from the Entry survey. Questions repeated from the Entry survey were mixed in consistency, with fewer (6 of 13) reporting that they 'Always' or 'Regularly' seek out information before taking regular bus journeys; however, while the same number of respondents reported that they 'Rarely' change their travel plans in response to RTPI, four respondents moved from 'Occasionally' to 'Always' or 'Frequently', with the highest number reporting that they may adjust their schedule or change their travel mode. In contrast to perceptions of reliability of existing sources of RTPI reported above, participants reported generally favourable perceptions of reliability of the *TravelBot* system, with 11 of 13 reporting that they 'Agree' or 'Strongly agree' both that the information provided was reliable and that it was useful. Although the number of respondents is relatively low, these initial results indicate that the perceived reliability of information obtained from the *TravelBot* system impacted positively upon the likelihood that participants would adjust their travel plans to a greater extent than similar sources of such information that they had used previously.

In the Exit survey, participants were also asked the multiple-choice question, "How often did you read the Twitter direct messages from *TravelBot*?". Eight participants responded "always" and five responded "frequently" with none responding either "occasionally", "rarely" or "never", indicating good user engagement with the *TravelBot* system. The value of this information was illustrated by a follow-up comment from one participant who stated, "…It isn't a big deal to check Google maps, but I disproportionately appreciated the info coming to me." This indicates that personalisation of travel alerts in the form of Twitter Direct Messages was found to be valuable for those making use of the system. The value of the information provided by *TravelBot* was further indicated in the following user response to the question "Did you receive any useful real-time information from *TravelBot* that helped with a journey?" (a question to which six participants responded affirmatively): "(*TravelBot*) Explained why and where a diversion was happening to the route 2 bus…*TravelBot* reassured me it was planned and gave me the information that cleared any worry I may have of it impacting my journey home." This illustrates the role of *TravelBot* and the information it provides in reducing the stress and negative perceptions of control experienced by passengers during such scenarios.

Table 2
Reported sources of public transport information (entry survey).

Reported sources of public transport information	
Paper timetable	2
Bus stop timetable	12
Bus stop real-time information display	7
Operator's Twitter feed (such as @FirstAberdeen)	5
Operator's web site	7
Operator's smartphone app (such as First Bus)	8
General travel information web site (such as Google Maps or Traveline)	1
General travel information smartphone app (such as Google Maps)	3

In addition to receiving information from the *TravelBot* system, eight participants also reported providing information on travel disruptions back to the system. The use of direct messages rather than public Tweets was viewed to be useful in this regard, as it allowed participants to communicate with the system without publicly reporting their travel behaviours. The issue of privacy was also considered in the study, with most participants reporting that they were comfortable sharing their regular bus stops and journey information with the system. The overall positive response to the system was evident, as when participants were asked: "Would you like the *TravelBot* service to continue to be available?", twelve responded "yes", with only one answering, "no". This response indicates there is potential for operators to offer a system providing the capabilities of *TravelBot*. While there were some issues reported with the system (in particular, difficulties in setting stop locations, timing of messages, and lack of information for non-routine trips), the overall experience demonstrates the potential for public transport operators to expand their use of social media services in communication with their passengers.

6. Limitations of the study and future needs

Due to the nature of the study and participant requirements (for example, being routine riders of specific bus services with regular travel times that overlapped with other participants), and for reasons discussed earlier, participant numbers were relatively small. While a larger scale study, in terms of participants and routes, would increase the potential for participants reporting disruptions or facilitating other information exchanges, this would also have increased the processing workload of the experimental *TravelBot* system (which relied on research team members simulating the role of a travel advice system by acting as "travel assistants"), which could have adversely affected response times. As this study was intended to demonstrate the proof-of-concept, the number of participants used was selected to maintain an acceptable system response time. There were also few transport disruptions that occurred during the trial, thus further reducing both the need for interaction and full exploration of the benefits of using the *TravelBot* system over other regular information sources; this would be a useful avenue for further research. The small numbers of participants, journey itineraries and messages generated also indicate that caution should be used in drawing any definitive conclusions from the study, though the general trends demonstrated offer promise for future exploration.

While the *TravelBot* system implemented the core functions of an automated travel advice service based on the findings reported in Sections 3 and 4, several options remain for attempting to increase user engagement with the system. Considering the difference in user engagement between the *Tweeting Travel* and *TravelBot* studies (in terms of number of messages contributed), this could be due to *TravelBot* being less conversational than the human operators, thus not encouraging further engagement from the participants. There is potential here to build on recent advances in automated "chatbot" technology to support a more conversational style dialogue between users and *TravelBot*, similar to that of the Tweeting Travel study which may lead to more participants providing useful information.

The information gathering approach adopted by *TravelBot* (from Tweets authored by participants, the bus operator, and media sources) utilised the Twitter network as an effective platform to connect with travellers; however, the kinship networks of travellers were less leveraged; future expansion of *TravelBot* could benefit from enabling users to provide information about others from whom they consider they may receive useful travel information. For example, when registering a journey, users could be prompted to indicate a subset of the Twitter accounts they follow from which they may receive (useful) travel information. *TravelBot* could then also monitor these accounts, rather than relying solely on the user forwarding relevant messages to *TravelBot*. The system also adopted a passive approach –

waiting for users to send it information about their journey. In an alternative approach, the system could utilise its knowledge of the user's social media habits (for example, that social media is often used during the commute, as seen in Section 3) and travel plans to actively message the user asking if they have experienced any disruptions. However, this would have to be performed carefully to avoid such messages being viewed as spam, thus potentially reducing future user engagement.

Another consideration is that while several of the studies discussed in the literature review (Section 2) have focused on the collection and analysis of large quantities of social media data in the study of transport disruptions, in this study we were primarily concerned with the creation of personalised dialogues between information providers and travellers. As such, the focus was not on the volume of data that could be collected, nor on the potential for further use of such data beyond the time-sensitive communication activity, though users were invited to share information about any disruptions they encountered (and they did so). However, as the *TravelBot* system develops further, it may be possible to create an aligned system to take more robust advantage of the volume of data available through the Twitter platform, and collected as part of the incident detection activity.

Finally, the scalability of the study needs further attention in order to determine if such a system is possible on a more extensive network; however, system response times were found to be acceptable, and should be capable of scaling to larger numbers of users and routes. Overall, it is considered that as a proof-of-concept, participant numbers were sufficient to demonstrate the viability and usefulness of such a system.

7. Discussion & conclusions

This paper has investigated the ways in which public transport operators and passengers can utilise social media to improve the overall travel experience. The overarching objective has been to demonstrate how social media may be used by operators to create an engaging dialogue with passengers that draws upon contextual information and provides personalised, real-time information to facilitate an improved travel experience.

The preparatory studies found that public transport commuters often use multiple social media platforms and regularly access them during and prior to their commutes, providing timely opportunities for public transport operators to use these channels to engage commuters with relevant travel information. The study also found that the social network of a commuter extends beyond their immediate friends and family to include fellow commuters, who are potential sources of information about their journeys, particularly when disruption occurs. An opportunity was also identified to exploit the mobile-based interfaces commonly used for accessing social media to improve the passenger experience, particularly by informing recipients of potential disruptions to their journey and so allowing them to maintain a sense of control over their journey.

The low cost, real-time nature of social media allows transport operators and passengers to engage in dialogues, which can feature the exchange of information of benefit to one or both parties. For example, operators can provide individuals with personalised warnings of potential disruptions, and passengers can report on aspects of the journey experience such as crowdedness, delays or disruptions that were unknown to the operator. Two types of dialogue were identified in this work: short, often one-way dialogues characterised by information-dense messages sent by either the operator or passengers offering little potential for further discussion; and more prolonged dialogues characterised by more conversational messages which do not necessarily include the exchange of travel related information. The tone of messages, particularly from the operator, plays a key role in encouraging (or not) the passenger to start/continue a dialogue, with impersonal or robotic messages that do not indicate a willingness to engage further tending to result in short, one-way dialogues.

From the perspective of passengers and public transport operators, passengers can utilise such dialogues to obtain information personalised to them and their journey directly from the operator. Similarly, operators can utilise information provided from passengers to complement and extend the knowledge available to them regarding the operational status of their bus network. From a technical perspective, the main type of information extracted from such conversations relates to details of events that may disrupt the transport network and services. The extraction process involves analysing each message to identify if a disruptive event is mentioned, and if so, attempting to determine its geolocation, time of occurrence, and any bus service(s) it affects. This can then be matched to details of passengers' (future) journeys and, if an overlap between the event and a journey is identified, the passenger can be sent an appropriate message. Enabling this requires the system to have access to data about the road network, public transport services, disruptions, and user journeys, which can be acquired from open data sources or the users directly.

The *TravelBot* evaluation study found that a system can be developed that utilises interactions between passengers and public transport operators as a source of information about the transport network. Regarding usage, the smaller number of messages sent by users to *TravelBot* (when compared to the number they received and the numbers sent during the earlier Tweeting Travel study) may indicate that participants' usage is influenced by a more information-consuming experience. However, the survey results found that participants valued the service and the information it provided, with all indicating a strong desire for the system to continue operating.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research described here is supported by the award made by the RCUK Digital Economy programme to the dot.rural Digital Economy Research Hub; award reference: EP/G066051/1. We extend our grateful thanks to the participants who have contributed to the studies throughout, and to the industry partner FirstGroup plc for their support.

References

Abbasi, A., Rashidi, T.H., Maghrebi, M., Waller, S.T., 2015. Utilising location based social media in travel survey methods: bringing Twitter data into the play. Proceedings of the 8th ACM SIGSPATIAL International Workshop on Location-based Social Networks. ACM.

Accenture, 2013. Public Transportation Users Predict Big Increases in the Use of Smartphones, Paperless Travel and Social Media, New Accenture Survey Reveals. http://newsroom.accenture.com/news/public-transportation-users-predict-big-increases-in-the-use-of-smartphones-paperless-travel-and-social-media-new-accenture-survey-reveals.htm.

Aslam, S., 2018, January 1. Twitter by the Numbers: Stats, Demographics & Fun Facts. Omnicore Retrieved August 14, 2018, from. https://www.omnicoreagency.com/twitter-statistics/.

Camacho, T., Foth, M., Rakotonirainy, A., 2012. Pervasive technology and public transport: opportunities beyond telematics. IEEE Pervasive Computing 12 (1), 18–25.

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J.R., Mellouli, S., Nahon, K., Pardo, T.A., Scholl, H.J., 2012. Understanding smart cities: an integrative framework. System Science (HICSS), 2012 45th Hawaii International Conference. IEEE, Wailea, Hawaii, pp. 2289–2297.

Collins, C., H., S., Ukkusuri, S.V., 2013. A novel transit rider satisfaction metric: rider sentiments measured from online social media data. J. Public Transp. 16 (2).

Corsar, D., Markovic, M., Gault, P., Mehdi, M., Edwards, P., Nelson, J.D., Cottrill, C., Sripada, S., 2015. TravelBot: journey disruption alerts utilising social media and linked data. Proceedings of the ISWC 2015 Posters & Demonstrations Track Co-located With the 14th International Semantic Web Conference (ISWC-2015). 1486. CEUR Workshop Proceedings.

Cottrill, C.D., Gault, P., Yeboah, G., Nelson, J.D., Anable, J., Budd, T., 2017. Tweeting transit: an examination of social media strategies for transport information management during a large event. Transportation Research Part C: Emerging Technologies 77, 421–432.

Dow, S., MacIntyre, B., Lee, J., Oezbek, C., Bolter, J.D., Gandy, M., 2005. Wizard of Oz support throughout an iterative design process. Pervasive Computing 4 (4), 18–26.

Faber, A., Matthes, F., 2016. Crowdsourcing and crowdinnovation. Digital Mobility Platforms and Ecosystems. TUM, pp. 36–47.

Gal-Tzur, A., Grant-Muller, S.M., Kuflik, T., Minkov, E., Nocera, S., Shoor, I., 2014. The potential of social media in delivering transport policy goals. Transp. Policy 32, 115–123.

Ganti, R.K., Ye, F., Lei, a.H., 2011. Mobile crowdsensing: current state and future challenges. IEEE Commun. Mag. 49 (11), 32–39.

Gault, P., Corsar, D., Edwards, P., Nelson, J.D., Cottrill, C., 2014. You'll never ride alone: the role of social media in supporting the bus passenger experience. Ethnographic Praxis in Industry Conference Proceedings 2014 (1), 199–212.

Gault, P., Cottrill, C., Yeboah, G., Nelson, J., Anable, J., 2015. Tweeting Public Transport: Examining the Use of Social Media During Large Events. IATBR, Windsor.
Gaver, B., Dunne, T., Pacenti, E., 1999. Design: cultural probes. Interactions 6 (1), 21–29.
Giffinger, R., Gudrun, H., 2010. Smart cities ranking: an effective instrument for the positioning of the cities? ACE: Architecture, City and Environment 4 (12), 7–26.
Gu, Y., Qian, Z., Chen, F., 2016. From Twitter to detector: real-time traffic incident detection using social media data. Transportation Research Part C: Emerging Technologies 67, 321–342.
Hariramani, P., 2019, May 23. Your commute just got easier with Google Pay and the Assistant. Retrieved September 2019, from. https://www.blog.google/products/google-pay/easier-commute-google-pay-assistant/.
Heiskala, M., Jokinen, J.P., Tinnilä, M., 2016. Crowdsensing-based transportation services—an analysis from business model and sustainability viewpoints. Res. Transp. Bus. Manag. 18, 38–48.
Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., 2003. Technology probes: inspiring design for and with families. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 17–24.
IBM. (n.d.). Social Media and the City. Retrieved February 2016, from IBM.com: https://www.ibm.com/smarterplanet/global/files/uk_uk_en_none_file1_488kb.pdf
Imran, M., Castillo, C., Diaz, F., Vieweg, S., 2015. Processing social media messages in mass emergency: a survey. ACM Computing Surveys (CSUR) 47 (4), 67.
Information Commissioner's Office, 2018, July 10. Findings, recommendations and actions from ICO investigation into data analytics in political campaigns. Retrieved August 14, 2018, from. Information Commissioner's Officehttps://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/07/findings-recommendations-and-actions-from-ico-investigation-into-data-analytics-in-political-campaigns/.
Jones, C.H., Cohn, S., Ogilvie, D., 2013. Making sense of a new transport system: an ethnographic study of the Cambridgeshire Guided Busway. PLoS One 8 (7).
Kuniavsky, M., 2003. Observing the User Experience: A Practitioner's Guide to User Research. Morgan kaufmann.
Lim, W.L., Lum, J.T., Yeo, I.J., Keoh, S.L., 2016. A Crowd-assisted Real-time Public Transport Information Service: No More Endless Wait. AwareCities'16 (Sydney, Australia).
Markou, I., Kaiser, K., Pereira, F.C., 2019. Predicting taxi demand hotspots using automated Internet search queries. Transportation Research Part C: Emerging Technologies 102, 73–86.
Mehmood, R., Meriton, R., Graham, G., Hennelly, P., Kumar, M., 2017. Exploring the influence of big data on city transport operations: a Markovian approach. Int. J. Oper. Prod. Manag. 37 (1), 75–104.
Mulley, C., Nelson, J.D., Hensher, D., 2019. Inteligent mobility and mobility as a service. In: Stanley, J., Hensher, D.A. (Eds.), A Research Agenda for Transport Policy. Edward Elgar Publishing, pp. 187–195.
Nam, T., Pardo, T., 2011. Conceptualizing smart city with dimensions of technology, people, and institutions. Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times. ACM, College Park, MD, USA, pp. 282–291.
Nelson, J.D., 2018. Perfect information and ITS. In: COwie, J., Ison, S. (Eds.), The Routledge Handbook of Transport Economics. Routledge, pp. 62–73.
Papangelis, K., Velaga, N.R., Ashmore, F., Sripada, S., Nelson, J.D., Beecroft, M., 2016. Exploring the rural passenger experience, information needs and decision making during public transport disruption. Res. Transp. Bus. Manag. 18, 57–69.
Passenger Focus and Abellio, 2012. Short and Tweet. How Passengers Want Social Media during Disruption. http://www.passengerfocus.org.uk/research/publications/short-and-tweet-how-passengers-want-social-media-during-disruption.
Pentina, I., Zhang, L., Basmanova, O., 2013. Antecedents and consequences of trust in a social media brand: a cross-cultural study of Twitter. Comput. Hum. Behav. 29 (4), 1546–1555.
Pereira, F.C., Bazzan, A.L., Ben-Akiva, M., 2014. The role of context in transport prediction. IEEE Intell. Syst. 29 (1), 76–80.
Qiao, F., Zhu, Q., Yu, L., 2011. Social media applications to publish dynamic transportation information on campus. ICCTP 2011, pp. 4318–4329.
Reiter, E., Robert, D., 2000. Building Natural Language Generation Systems. Cambridge University Press.
Salvador, T., Bell, G., Anderson, K., 1999. Design ethnography. Design Management Journal (Former Series) 10 (4), 35–41.
Schweitzer, L., 2014. Planning and social media: a case study of public transit and stigma on Twitter. J. Am. Plan. Assoc. 80 (3), 218–238.
Sommer, B., Sommer, R., 1991. A Practical Guide to Behavioral Research: Tools and Techniques. Oxford University Press.
Wallace, J., McCarthy, J., Wright, P.C., Olivier, P., 2013. Making design probes work. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 3441–3450.
Xu, Q., Dodds, E., 2015. Enhancing transport through open geospatial data and crowdsourcing. ICT for Transport: Opportunities and Threats. Elgar, pp. 140–160.
Zhang, J., Wang, F.Y., Wang, K., Lin, W.H., Xu, X., Chen, C., 2011. Data-driven intelligent transportation systems: a survey. IEEE Trans. Intell. Transp. Syst. 12 (4), 1624–1639.