



## EWGT2013 – 16<sup>th</sup> Meeting of the EURO Working Group on Transportation The Impact of Social Media Usage on Transport Policy: Issues, Challenges and Recommendations

Ayelet Gal-Tzur<sup>(a)</sup>, Susan M. Grant-Muller<sup>(b)</sup>, Einat Minkov<sup>(c)</sup>, and Silvio Nocera<sup>(d)\*</sup>

<sup>a</sup>Technion - Israel Institute of Technology - Technion City - Haifa 32000, Israel

<sup>b</sup>University of Leeds - Woodhouse Ln, Leeds LS1 3HE, United Kingdom

<sup>c</sup>University of Haifa - Mount Carmel - Haifa 31905, Israel

<sup>d</sup>IUAV University of Venice – Santa Croce 191 – I-30135 Venice, Italy

### Abstract

Transport stakeholders are increasingly recognizing the value of social media in connecting with their customers in many forms. This study explores two sides to engagement with social media – firstly the potential uses of social media by transport service suppliers and secondly the potential value to policy development of shared transport related information by the public. Samples of stakeholders of different sizes were selected for the study, reflecting alternative modes and providing a variety of services and functions in the transport sector. The findings give insights into the practices of organizations of different size, function and longevity of social media use. The early results of a study to harvest freely available transport information from the public and transport system users are also presented, demonstrating that transport policy relevant information can be harvested from online social media sources.

© 2013 The Authors. Published by Elsevier Ltd.

Selection and/or peer-review under responsibility of Scientific Committee.

*Keywords: Social Media; Transport Planning; Transport Policy; Travel Behaviour*

### 1. Introduction

Rapid and recent technological development in social networks is providing a vision amongst transport suppliers, governments and academia of what can be called ‘next-generation’ information channels. Social media<sup>†</sup> in general has become an important source for unprecedented reach, speed and democratization of communication. It is having a dramatic influence on customers, news, stakeholders and policy makers: from trend identification in levels of support for politicians to gathering supporters for political protest and lobbying, social media plays a crucial role in modern life, including potentially for transport. Policy makers in many Western countries have high expectations of the potential effects of such information services on, for example, network efficiency (e.g. European Commission, 2001; Transport Research Board, 2004), with the view that strategic policy may reduce

\* Corresponding author. Tel.: +39-041-2571370; fax: +39-041-2571392.

E-mail address: [nocera@iuav.it](mailto:nocera@iuav.it)

<sup>†</sup> The term “Social media” is used in this research to refer to a group of web-based applications that encourage users to interact and share ‘content’, examples of which include blogs, social and professional networking sites, micro-blogging sites and media-sharing sites

transport impacts significantly (Nocera and Cavallaro, 2011, 2013). These expectations provide momentum for research and development efforts aimed at designing efficient policy initiatives, particularly by those in the field of personal mobility and information provision. These include the need for new understandings of customer perceptions at the outset of a new generation of travel information services and to align these with several dominant theories on traveller behaviour. Transport policy-makers might apply some of these findings in order to derive technical or functional requirements of optimal services or to help estimate the potential levels of use and effects of services that are being developed. Social media data could be harvested on-line directly from other transport users in a type of Advanced Traveller Information System, influencing important travel choices: departure time, route and mode choice. The information harvested might also have secondary influences on traveller destination choice and activity patterns, whilst the framework developed might assist policy makers and stakeholders in an understanding those choices. For these reasons, social media may represent a new and effective means for organisations to communicate with customers and citizens. Increased engagement with social media has happened from both bottom up (the public and users of the transport system) and top down (by the transport system stakeholders who operate, provide services and develop policies for the system). Public transport companies and stakeholders have started using social media as a strategic tool to reach out to customers, by using social media as a ‘help line’ for customer service or for real time information. This paper discusses the potential of this new phenomena through both a top-down approach, and a bottom-up approach. The former is demonstrated through an investigation of the manner in which different selected transport providers currently use social media. The latter is demonstrated by reporting the first results of an experiment conducted jointly by four research groups in which a methodology for harvesting useful information for transport stakeholders was developed. It draws on a multidisciplinary interface between transport science and text mining expertise, thereby establishing ‘what’ should be harvested and ‘how’ it can be harvested, illustrated here using a case study based on trips to and from a large sporting event.

## 2. Use of social media by transport stakeholders

This section provides a short overview of the evidence on how the transport stakeholders are currently using the social media, specifically:

- How the media is being used – the apparent function eg for publicity of the organisations’ new initiatives, for customer relations (CR), for up to date information notices etc
- Whether there is evidence on the extent of engagement with the public (number of followers, number of likes) and whether there was is any apparent trend in this – for example by the organisation size.

Following the approach of Bregman et al (2012), a selection of transport providers of different types (size, location, business type) were chosen and their current use (March 2013) of social media reviewed using publically available information. The sample chosen is only a small, preliminary snapshot and was determined by a number of factors and not intended to be either random or fully representative. The availability of a website with English language was one restriction and in general, medium to large sized organisations were chosen. The sample is shown in table 1, comprising a local authority, a national transport policy/decision maker, two airports, transnational ticket sales and schedule information companies (rail, ferry) and two airlines. The two types of social media observed were Facebook and Twitter<sup>‡</sup>. The nature of the business of most of the stakeholders involves dynamic interaction with the public – for example in providing timely information or sales. These two media inherently function in such a way as to facilitate this and bring the possibility of a two way exchange – the main advantage over most standard websites. However during the course of the research it did become clear that some were also using Youtube as a third source of social media particularly for one-direction communication of new company initiatives or campaigns. Different models of use of the media were observed: 1) The public were

---

<sup>‡</sup> [www.facebook.com](http://www.facebook.com) and [www.twitter.com](http://www.twitter.com)

strongly invited to engage with either Facebook, Twitter within a section of the company website and the use of the media was purposeful, 2) the public were invited to engage with tailored and welcoming messages, the purpose of use was general or multifunctional, 3) Links to social media were given on the main website page but were small in size and the invitation to engage was simple/generic (e.g. ‘follow us’) and 4) the logos were present on the main website there were links to a twitter stream, but there was no direct link to the Facebook page or no Facebook page was available.

Whilst most organisations had a Facebook page and twitter line that was focused around their core business, some webpages gave links to Facebook/Twitter that were shared between organisations within a larger group (e.g. the Gran Canaria airport and Leeds City Council. Gran Canaria is part of a larger group of airports, whilst Leeds city council has a single Facebook and twitter line for a number of departments (transport, city cleansing, housing etc). One advantage of a dedicated social media line or page would be an improved ability to conduct further analysis on public postings, e.g. picking up trends in sentiments or information requests. With sites that cover multifunctions or several organisations this may well become more difficult with a wider variety and relevance of messages. The main advantage of the shared social media function however is the reduced cost in maintenance and active CR input. It may well be the case that over a period of time, the use of the media for organisations currently on model 2-4 above may develop and step increasingly towards the more directed use such as model 1.

Table 1: Review of Transport Stakeholder Social Media use

Transport stakeholder		Approx. turnover	Facebook page/Likes	Bespoke Invitation to engage in social media	Twitter account Followers/Tracking start date
KLM	Airline (international travel)	2011: KLM Group carried 25.2M passengers, 484.100 tonnes of freight. Income 6.985M Euros.	<a href="http://www.facebook.com/KLM">http://www.facebook.com/KLM</a> 3,106,797 likes	Yes	@KLM 355,569 Followers Tracking: 3/11/2008
Ryanair	Low cost airline	2012: 75.8m passengers Income 4.324.9M Euros	Strategic decision on no facebook account	N/A	@ryanairnews 21,122 Followers Tracking: 4/12/10
RailEuropeUK	Rail ticket service provider at trans european level	Not readily available (four independent companies in distinct geographical areas)	<a href="https://www.facebook.com/RailEuropeUK#_=_">https://www.facebook.com/RailEuropeUK#_=_</a> 10,137 likes	No	@RailEuropeUK 2,551 Followers Tracking: 26/12/10
Direct Ferries	Ferry services provider at international level	2011: Turnover £46.5M	<a href="https://www.facebook.com/directferries?sk=app_4949752878">https://www.facebook.com/directferries?sk=app_4949752878</a> 21,303 likes	Yes	@directferriesuk 614 Followers Tracking: 5/12/11
Frankfurt airport	Air Transport node (3rd busiest airport)	2012: 57.52M passengers	<a href="https://www.facebook.com/FrankfurtAirport">https://www.facebook.com/FrankfurtAirport</a> 117,638 likes	Yes	@Airport_FRA 7,278 Followers Tracking: 27/7/10
Gran Canaria airport	Air Transport node (38th busiest airport – tourism)	2011: 10.5M passengers, 23,7 million tonnes of cargo	<a href="https://www.facebook.com/AenaAeropuertos">https://www.facebook.com/AenaAeropuertos</a> 4,522 likes	No	@aenaeropuertos 23,930 Followers Tracking: 4/12/10
Department for Transport (UK)	NGO – Transport strategy and planning (national)	N/A	No facebook page	No	@transportgovuk 23,260 Followers Tracking: 14/7/09
Leeds City Council (UK)	LGO - Transport policy and planning (local)	N/A	<a href="https://www.facebook.com/LeedsCouncil">https://www.facebook.com/LeedsCouncil</a> 1,192 likes	No	@leedsc 12,196 Followers Tracking: 19/6/09
NS (Netherlands transport service)	National transport supplier	2011: 396M passengers Income €272M	<a href="https://www.facebook.com/nederlandsespoorwegen">https://www.facebook.com/nederlandsespoorwegen</a> 31,518 likes	No	@NSonline 70,078 Followers Tracking: 29/3/10
National rail enquiries (UK)	National information service on rail times and disruptions	N/A	<a href="https://www.facebook.com/nationalrailenq?group_id=0&amp;filter=2">https://www.facebook.com/nationalrailenq?group_id=0&amp;filter=2</a> 26,086 likes	No	@nationalrailenq 98,700 Followers Tracking: 8/6/09

As can be seen from Table 1, the organisation with the largest number of Facebook likes and highest Twitter tracking was KLM. This was not the largest airline in the survey but was the first from this group to establish a twitter. KLM use model (1) above, with Facebook and Twitter prominently used for CR purposes. Links to both media were provided from their customer services webpage. The organisation with the lowest social media presence is the low cost carrier Ryan Air, which in 2012 carried 75.8m passengers, one of the world's top 10 largest carriers according to the International Air Transport Association criteria. Ryan Air has taken a strategic decision not to engage with social media. *'such accounts would result in too many customer queries and would require more resources from the airline. "A Facebook account would not be helpful to us, as we would have so many people looking for a response," he said adding that it would having to hire "two more people just to sit on Facebook all day". Kiely said that passengers can get in touch through Ryanair's customer care line (Kiely, 2013)* More than one Twitter line was found on searching and it wasn't entirely clear which was the 'official' line. It was noted by Bregman et al (2012) that some organisations in the USA deliberately use more than one twitter line, but the purpose is to direct customers towards a line of chatter or posted material is most appropriate to their needs. In other words, it is a more developed version of model 1 above. Examples of organisations using model 2 above were Direct Ferries and Frankfurt Airport, with examples of tailored invitations as follows. KLM was the only site to advise users not to share personal information.

*'Visit Frankfurt Airport on the popular social network platforms Facebook, Twitter and just recently also on YouTube. Connect with other guests and passengers and stay informed about what is happening at Frankfurt Airport' 'Contact us. Please share your personal details via private messages only. Ask KLM a question 27/7 via Facebook> Ask KLM a question via Twitter'.*

In general the style of communication from the organisation was far more informal than was used on the main website. Facebook and Twitter generally had a 'person' who was responding ie someone with a first name and responses that appeared to have been written by a human rather than a computerised/automated response. In this respect these media represent a far more accessible and welcoming interface with the organisation than some customer relations websites which, at best, may offer a menu of FAQ or what appear to be computer generated responses to standard customer enquiries. The main functions of social media use across the organisations sampled were as follows (not in any particular order of priority or frequency of function):

- 1) **Information/updates on main website** – this was a common type of message on Facebook from the stakeholders and generally informing travellers that a problem or updating process was happening with the main website. In general this kind of message was one-directional ie the main webpage would not advise of changes to Facebook, the public did not appear to offer responses to this type of message:

*Hi Facebook! Just to let you know that we're continuing to update information on our service disruptions page of the website, so you can find out about school closures, bin collection and gritting plans on the site. Thanks, Rob*

- 2) **Advising the public on Travel Disruption** - this function is very much part of the core business for many of the organisations surveyed, particularly those concerned with scheduling, timetables and ticket sales. The type of information posted by the stakeholders included 'best knowledge' on the current or expected weather conditions and signposting to other sites that would hold the 'official' news on the disruptions.

*Hi Facebook World! I'm going to be heading off for the weekend shortly, but just to confirm that although the snow seems to have eased off for now - it's likely to be back with a vengeance! We have updated the service disruption page on our website listed below, and you can follow us on Twitter for updates on gritting (@LCCPressOffice). Keep wrapped up this weekend. Thanks, Rob*

- 3) **Handling travel queries and complaints** - following the stream of interaction between the organisation and members of the public showed the different approaches used to deal with either straight forward and or general questions, personal queries, or complaints. Questions which didn't refer to a particular travel booking and which didn't involve a complaint were generally responded to on-line (example above). Where the public

asked about specific bookings, the organisation generally advised that they would be contacted off-line using the individual's personal Facebook account (in order to preserve security/privacy). For complaints about travel, the organisations generally offered a public apology and offered to speak with the individuals off-line by phone or in email correspondence. This 'code of practice' seemed to work well in general, although there were some issues aired around the use of social media (see 4 below). Some organisations appeared to ignore negative comments from the public which may have avoided lengthy public correspondence but which then left the comments hanging 'in the air' on the social media site in a fairly unsatisfactory way.

***I can't find how much KLM charge per extra kilo on a flight from Dublin to Hong Kong? The limit is 23k, what happens if your bag is 25kg? It seems to cost €200 extra for an additional bag?? Can you confirm how much in Euro it is?***

***That sounds like quite an interesting and nice route, Mairead. In regards to your question, we can inform you that we don't charge extra weight as per the kilo concept but as per piece. If you decide to take one additional luggage, then it will cost EUR100 for that route, for the 2nd piece (on top of your free allowance) and EUR200 for each piece in addition of those two. Let us know in case you have further doubts.***

- 4) **Responding to queries around use of social media** - inevitably some members of the public commented on the way in which the organisation was using social media. This included positive messages concerning the speed of responses, but also comments concerning non-response or unwanted presence on the users own social media site. The most constructive responses from organisations were to offer the route to addressing this through their own databases etc. It is the equivalent to telephone customers asking for their number to be removed from the databases of direct marketing organisations. This type of interaction may be best dealt with through published policies on how the organisation uses social media and how customers can choose to end the social media relationship.

***You were absent from my FB page for one wonderful week now you're back again! Please go away and stay away nothing I can do this end makes any difference***

*Like explained before, our posts will only show on the people who likes us, pages. If you don't wish to see them, you need to unlike us (same place as you liked us)..... Kind regards, Camilla*

***Why do you delete the posts from people airing their opinions ?? You should be taking action, not ignoring them. (no response from the organisation/CR team)***

- 5) **Seasonal goodwill messages** - the final category of dialogue and one which was seen on many of the sites was that of the seasonal message of goodwill. This was consistently informal in nature and seemed to be aimed at promoting the concept of timeliness, community and friendly service.

***Happy Easter to all our followers! We hope you have a great day, enjoy your Easter eggs! The Frankfurt Airport Team wishes all fans very wonderful and happy holidays***

Some general observations from this short study were as follows. The size of the organisation overall didn't solely determine the degree or type of social media use. The largest organisation didn't use social media as a strategic stance. Four main levels of use of media were observed, ranging from highly focused and directed use, to a low level social media presence with general material posted. It is conjectured that a combination of the length of time the organisation has used social media, along with the nature of their core business may determine the extent to which it is used and how. The two governmental organisations for example demonstrated different use to the airports. The longer the social media sites have been established, the more likely the use becomes focused. A pattern of interaction was seen with some organisations that used social media in a highly interactive fashion, which may set an example for other organisations seeking to either the social media arena or increase their level of activity.

### 3. Evaluating passenger perceptions through Social Media: an Exploratory Study

Performance measurement is fundamental for assessing the performance of a transport service, for example in relation to community expectations, for assessing the costs of the service, in case of relevant externalities (Nocera et al., 2012; Nocera and Cavallaro, 2012; Nocera and Tonin, 2013), and as a monitoring tool for improving the service (Transportation Research Board, 1994). In the case of service quality, this can be measured by comparing performance against particular thresholds (Nocera, 2010; Eboli and Mazzulla, 2012a), or by evaluating passengers' perceptions using Customer Satisfaction Surveys (CSS), which help to establish customer priorities (Eboli and Mazzulla, 2011, 2012b). There is, however, convincing evidence that an individual's perceptions of characteristics might differ greatly from the actual values of those characteristics (e.g. McFadden, 2001; Lyons, 2001; Zhao and Harata, 2001; Avineri and Prashker, 2003, Bonsall et al., 2004). This is especially influential on the mode choice of those segments of medium-high elasticity (Libardo and Nocera, 2008). Individuals can tend to overestimate positive and negative features of chosen alternatives compared with the non-chosen alternatives. Furthermore, perceptions of certain characteristics appear to differ widely across the population of travellers (e.g. Bonsall et al., 2004). As perceptions of reality are forceful drivers of choice and information may affect these perceptions rather than reality itself, this makes the formation and adaptation of perceptions a non-negligible issue to be dealt with in the analysis of choice behaviour, especially under conditions of uncertainty (Delavande, 2003; Manski, 2004), or when DRTS are at-stake (Nocera and Tsakarestos, 2004).

Social media is an arena for users to freely express their opinions and needs, and therefore has the potential to enhance existing information. Sterne (2010) defined seven broad categories of social media: Forums and Messages Boards, Review and Opinion Sites, Social Networks, Blogging, Micro Blogging, Bookmarking and Media Sharing. The first five are used substantially for the exchange of transport related information. Dedicated forums, such as Topix<sup>§</sup>, Facebook groups such as **\*\*LONDON-ESSEX-KENT\*\*** and social networks such as WAZE<sup>††</sup> are just a few examples of this fast growing trend.

An exploratory study was conducted to investigate what information on travellers' needs and perceptions of their travel can be harvested from social media, and what lessons can be learned as a result. In order to focus the search, the study was specifically designed to harvest information related to trips to and from Liverpool Football Club games. Twitter was chosen to be the source of information as a large portion of its content is public, and tools for extracting recently sent Tweets are available<sup>‡‡</sup>. Due to the volume of the stream of Twitter messages, an automatic process was used to identify relevant messages, i.e., messages relevant to the event of interest and transport related. Two main processing steps were applied: 1) Identification of event-related terms (including the location of the match and terms describing the nature of the event), were used to filter the harvested twitter messages between specific dates 2) Supervised machine learning techniques were applied to automatically recognize transport-related messages, based on the words in the messages. Specifically, a classifier was trained to identify transport-related messages given examples of relevant and irrelevant tweets. Another classifier was also trained to distinguish between messages originating from authorities (of lower interest for this part of the research), and those written by individuals expressing personal opinions.

The study included messages for three football matches (between the Liverpool and Chelsea, Swansea and West Bromwich teams). A pool of possibly relevant messages was filtered from Twitter using keywords associated with these events, and then processed by the learned classifiers. A set of top-scoring 1,174 (non-duplicate) example messages that were automatically found to be individually-authored and transport-related by the classifiers has been fully annotated by two domain experts. Based on these manual annotations, we evaluated the effectiveness of the automatic processing. In addition, careful inspection of the categorization results of the classifiers revealed interesting findings regarding the weak spots of the automatic filtering process. The first step

<sup>§</sup> <http://www.topix.com/forum/business/transportation>

<sup>\*\*</sup> <http://www.facebook.com/groups/369684789781652/?fref=ts>

<sup>††</sup> <http://world.waze.com/?redirect=1>

<sup>‡‡</sup> We used the Twitter Firehose streaming interface, which provides 1% randomly sampled messages of overall traffic

of the processing pipeline (aimed at identifying messages that were sent by individuals rather than by authorities), was highly effective. Specifically, less than 5% of the considered corpus (50 messages overall) was wrongly categorized as messages written by individuals. Strong indicators found by the classifier of individually-posted messages include use of personal pronouns such as 'I', 'we', as well as informal language, such as 'lol' (laughing out loud), and so forth. The following is an example of a message mistakenly identified as having been written by an individual:

***Fa cup Liverpool v Chelsea charter train direct from WAVERTREE TECHNOLOGY PARK DIRECT TO WEMBLEY. <http://t.co/Ih5TsRgb>***

The only indication that this message was posted by an authority is the URL at the end of the message. If the URL was omitted, the manual annotation would also fail in associating it with an authority. It is one among many examples demonstrating the challenge of fully automatic text mining. The classification accuracy in further classifying individually-posted messages directly related to travel in the considered corpus was 0.86, implying the classification of messages as transport-related was erroneous in 14% of the cases. Various reasons for misclassification were found, and the following example demonstrates one:

***@toshack7 I have a train ticket but no match ticket. I will be watching the game in a Glasgow pub probs.***

The automatic process identified this message as relevant as it includes the terms 'train and 'ticket', which refer to a transport mode. However, the words in the message that are directly associated with transport did not determine the relevance of the message. The following example is one of many that can verify this statement:

***@lpoolcouncil Bus stops and pavements cleaned and being rid of dog s\*\*\*. Why not look after the Anfield area like this all year round? '***

The message clearly addresses an aspect of quality-of-service, yet out of the 24 words in the message, only the words “bus stop” indicate its relevancy. Having manually filtered messages originating from authorities and those that did not address travel explicitly, messages addressing non-surface transport modes were also identified and eliminated from the corpus (11 messages). The next and most challenging aspect concerned the location referred to in the message. Locations names were used in the automatic filtering process as part of the event-related terms. However, specific place names may not be posted and are complex overall. Liverpool is the name of a UK city, a street, a station and even the surname of a person. Geographic meta-information about a user may help in resolving this, but it is only available when users choose to reveal these details and many don't. Recent work has addressed this problem by applying algorithms that analyze the content of the text in conjunction with location names databases (Amitay et al., 2004; Paradesi, 2011). However at this early stage of the research these methodologies have not yet been applied in our classifying process. Messages referring to non-UK locations were firstly eliminated (39 messages). The following example demonstrates one of the reasons for these occurrences where clearly “Chelsea” in conjunction with “traffic” is the reason for including this message in the corpus:

***Guy tries to get away from cops screws up traffic MT @MassDOT Traffic Alert: Chelsea- Carter Street off ramp Rt 1SB Accident ramp blocked***

It was also clear in a manual inspection that some messages were associated with London. In order to eliminate them from the corpus, three criteria were applied. Messages were eliminated that: 1) Contained the phrase “Liverpool Street Station” (26 messages), 2) Contained the phrase “London Bridge” (26 messages) and 3) Explicitly addressed tube and tram (54 messages). Overall, 259 messages out of 1174 were manually disqualified and omitted from the set of messages, leaving a corpus of 915 messages (78% of the original set).

### *3.1. Manual annotation of the content of messages*

The 915 remaining messages were then further annotated according to the message purpose. Three purposes were defined, each relevant to authorities in promoting transport policy goals. These purposes are: 1) To express a need for a transport service or describe a journey, 2) To report a transport related incident (which may be a planned or unplanned disruption to the ‘normal’ transport service) and 3) Expressing an opinion regarding a transport service. Messages classified according to the first purpose contain information about journeys that travelers take

or about to take and some included a need for information regarding transport modes such as availability of trains and time schedules of buses. Two examples for such messages are:

***En route from London to Liverpool to go and see Liverpool vs West Brom #anfield here we come!***

***@Everton just saw about 30 fans waiting for taxis in lime street station all looking to get to goodison or a pub ahead of the match! #COYB***

Users' reports on an incident were annotated as such only if the message indicated that the impact of the incident was still effective. Those associated with the road network typically included a report on congestion generally or a closed road due to an accident. Messages associated with public transport typically reported delays in bus or train service, cancellation of service or crowdedness at stations.

***Packed at Lime Street whats going on in Liverpool today? (@ Liverpool Lime Street Railway Station (LIV) w/ 6 others) <http://t.co/blSkdpFR>***

***@adamkay1991 FYL- signalling problem in the Hartford area services between Liverpool Lime Street & Crewe may be delayed by30 minutes***

Messages expressing travelers opinions on transport services are probably the most variant in nature. Travelers comment on a range of issues including the availability of transport services, density at stations and in public transport vehicles, travel time, cost and more.

***also liverpool centrals northen line train station smells of bad eggs***

***Packed at Lime Street whats going on in Liverpool today? (@ Liverpool Lime Street Railway Station (LIV) w/ 6 others) <http://t.co/blSkdpFR>***

For the entire group of 915 messages, Figure 1 provides the distribution of the message purpose. It is apparent that the largest proportion of the messages harvested in the study were concerned with expressing an opinion on transport services. This finding is not surprising. The desire of people to share their opinions and experiences is well reflected well in popular review and opinion sites (such as Tripadvisor) as well as in the large number of talkbacks provided for almost every article on news sites. The proportion of messages containing information about journeys taken by the users is also substantial. It is an expression of new social phenomena that "if I didn't report it, it didn't happen". The smallest proportion of messages concerned the reporting an incident. Taking into account the emerging field of referring to people as human sensors, and the high hopes for using it in various fields of life (Sakaki et al., 2010; Siqi Zhao and Zhong, 2011), it is somewhat disappointing that only 16% of the messages found belong to this category. Clearly one case study is not enough to conclude whether this proportion is typical. Once the entire set of messages was annotated according to the message purpose, each message was annotated again according to the transport mode it referred to. 53% referred to train, 20% to bus, 17% to private vehicles and 10% to all other modes of transport. In order to further investigate the content of the messages, those dealing with opinions about PT (train and bus) were annotated according to the aspect of quality-of-service they addressed. Figure 2 shows the distribution between the different aspects of quality-of-service. Density (crowdedness), availability and travel time were found to be the most prominent issues. Density is often associated with the behavior of the fans, as reflected in the next example:

***Loud football idiots on the train. Not too threatening just LOUD. Where has respect for other travellers gone? Maybe it never existed.***

The first of following two messages addresses the availability of public transport and the second its reliability:

***if liverpool fas cant get the train back from wembley they should provide a s\*\*\* load of coaches time to give it back to the fans '***

***great i couldve been at the match now but I was afraid Id miss the bus. And now the bus is late -***

From Figure 2 it is clear that apart from density, availability and travel time, many other aspects of quality-of-service are addressed by the travellers. Some of the messages do not include specific information about the



train or bus line used, while others are more specific. Although information pointing to the route itself would have been valuable in order to take specific improvement actions, this message still has value in the overall context of the set of messages. The fact that 20% of the messages dealt with density while only 7% dealt with cost can provide guidelines regarding efficient allocation of resources for improving public transport services.

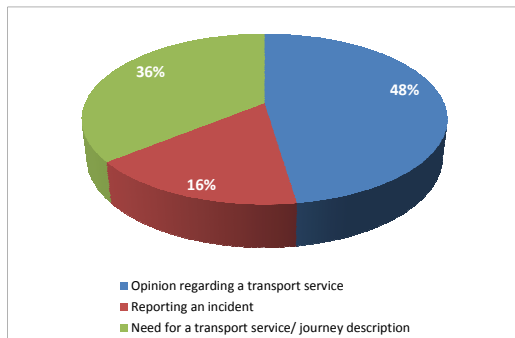


Figure 1: Distribution of the message type

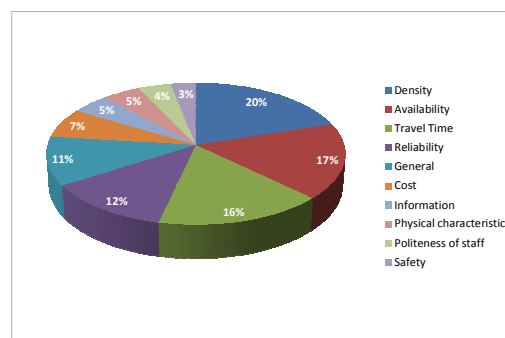


Figure 2: Distribution of quality-of-service messages

#### 4. Conclusions

Social networks now touch on every aspect of our lives and are used extensively to disseminate information and opinions. They provide unprecedented opportunities for policy makers to learn firsthand about public opinions and the transport sector is already engaging with this new communication medium. Service providers disseminate information and encourage the public to express opinions on specific topics of interest to them from a “top-down” communication channel (see table 1). This communication channel is, however, somewhat limited by the “top-down” structure – the organizations/policy makers take the initiative and bring issues for discussion. It is not known whether the public responses are collected and analyzed further – for example to understand broader trends or by population subgroups. Section 3 sheds light on what seems to be a yet unexplored information gold mine. Mining social network data may enable stakeholders to better understand public views and needs on a range of issues and then form future policies and strategies accordingly. However, as shown here, this “bottom-up” channel is fundamentally different to the “top-down” one. As it is not initiated by an organization and there is a need to apply text mining techniques to retrieve relevant information from the continuous information stream available on these networks. The process is complicated, involving a definition of information requirements followed by a process of searching for examples, training and testing classifiers and finally analyzing the results. The exploratory study here demonstrates the potential of this approach and the ability to harvest valuable transport related information automatically. Such information, when collected, aggregated and analyzed could help policy makers concerned with transport challenges such as temporal and geo-located traffic bottle-necks, parking needs etc. However harvesting relevant information is not straight forward. It requires integrated, multidisciplinary research that combines domain experts with information technology experts. Whilst having fundamental differences, it also seems that the two approaches are inter-related and can act in synergy. Specific issues can be initiated for public discussion whilst the real power of social media may be harnessed to identify needs and problems by authorities. Current knowledge concerning the most efficient means to realise both the top-down and the bottom-up approaches is far from sufficient. Further research is required to investigate the main aspects involved in integrating social media into on-going transport planning, management and operational activities, while addressing the socio-technical factors that play a role in this challenging and promising field.

#### Acknowledgement

This paper draws on work also conducted by Tsvika Kuflik and Itay Shoor (Haifa University), Ebithal Sheety

(Technion) and Frances Hodgson (University of Leeds). Without their contribution the paper would not have been possible. This research was partially funded through the Worldwide Universities Network (WUN) scheme.

## References

- Amitay E., Har'El N., Sivan R., Soffer A., (2004), Web-a-Where: "Geotagging Web Content", *SIGIR '04 Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 273-280
- Avineri, E. and Prashker, J. N. (2003) Sensitivity to uncertainty: the need for a paradigm shift. Paper presented at the 82nd Meeting of the Transportation Research Board, Washington, DC, USA.
- Bonsall, P. (2004) Traveller behaviour: decision-making in an unpredictable world, *Journal of Intelligent Transportation Systems*, 8, p. 45–60
- Bregman S., (2012), "Uses of Social Media in Public Transportation", TCRP SYNTHESIS 99
- Delavande, A. (2003) Pill, patch or shot? Subjective expectations and birth control choice. Working Paper, Northwestern University, Evanston, IL, USA.
- Eboli L., and Mazzulla G. (2011), A methodology for evaluating transit service quality based on subjective and objective measures from the passenger's point of view. *Transport Policy*, 18-1: 172-181
- Eboli L., and Mazzulla G. (2012a), Performance indicators for an objective measure of public transport service quality, *European Transport, International Journal of Transport Economics, Engineering and Law*, 51: 1-21
- Eboli L., and Mazzulla G. (2012b), An SP experiment for investigating on travellers' preferences about the use of car and bus. The influence of the service quality factors, *Italian Journal of Regional Science*, 11(3): 75-92
- European Commission (2001) *A Sustainable Europe for a Better World: A European Union Strategy for Sustainable Development*. Brussels, Belgium
- Libardo A., and Nocera S. (2008): Transportation Elasticity for the Analysis of Italian Transportation Demand on a Regional Scale. *Traffic Engineering and Control* 49-5: 187-192
- Kiely (2013) <http://businessetc.thejournal.ie/facebook-social-media-ryanair-robin-kiely-783104-Feb2013/> accessed 1/4/1.
- Lyons, G. D. (2001) Towards integrated traveller information, *Transport Reviews*, 21, pp. 217–235.
- Manski, C. F. (2004) Measuring expectations, *Econometrica*, 72, pp. 1329–1376.
- McFadden, D. (2001) Disaggregate behavioural travel demand's RUM side: a 30 year retrospective, in: D. Hensher (Ed.) *Travel Behaviour Research: The Leading Edge*. Pergamon, Amsterdam, The Netherlands
- Nocera S. (2010), Un approccio operativo per la valutazione della qualità nei servizi di trasporto pubblico/An Operational Approach for Quality Evaluation in Public Transport Services. *Ingegneria Ferroviaria* 65-4: 363-383
- Nocera S. (2011), The key role of quality assessment in public transport policy. *Traffic Engineering and Control* 52-9: 394-398
- Nocera S., and Cavallaro F. (2011): Policy Effectiveness for Containing CO<sub>2</sub> Emissions in Transportation. *Procedia - Social and Behavioral Science* 20: 703-713
- Nocera S., and Cavallaro F. (2012): Economical Evaluation of Future Carbon Impacts on the Italian Highways. *Procedia - Social and Behavioral Science* 54: 1360 – 1369
- Nocera S., and Cavallaro F. (2013): A Methodological Framework for the Economic Evaluation of CO<sub>2</sub> Emissions from Transport. *J. Adv. Transp.* doi: 10.002/atr.1249
- Nocera S., Maino F. and Cavallaro F. (2012): A heuristic Method for evaluating CO<sub>2</sub> Efficiency in Transport Planning. *European Transport Research Review* 4: 91-106
- Nocera S., and Tsakarestos A. (2004): Demand responsive transport systems for rural areas in Germany - an overview of the projects MOBINET and MOB<sup>2</sup>. *Traffic Engineering and Control*, 45-10: 378-383
- Nocera S. e Tonin S. (2013): A Joint Probability Density Function for reducing the Uncertainty of Marginal Social Cost of Carbon Evaluation in Transport Planning. *Procedia - Social and Behavioral Science* (accepted for publication)
- Paradesi S., (2011), "Geotagging Tweets Using Their Content", *Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference*
- Sakaki T., Okazaki M., Matsuo Y., (2010), "Earthquake shakes Twitter users: real-time event detection by social sensors", *Proc. ACM WWW*
- Siqi Zhao S., Zhong L., (2011), "Human as Real-Time Sensors of Social and Physical Events: A Case Study of Twitter and Sports Games", Technical Report TR0620-2011, Rice University and Motorola Labs, June 2011
- Sterne J., 2010, Social media metrics: How to measure and optimize your marketing investment, books.google.com
- Transportation Research Board [TRB] (1994). *The Role of Performance-Based Measures in Allotting Funding for Transit Operations*. TCRP Synthesis 6. National Academy Press, Washington, DC
- Transportation Research Board [TRB] (2004) *User information systems; developments and issues for the 21st century*, in: TRB Millennium Papers (Washington, DC: National Academy of Sciences).
- Zhao, S. and Harata, N. (2001) Travel information, perceived travel time, and route diversion behaviour. Paper presented at the 9th World Conference on Transport Research, Seoul, South Korea