ISBN 978-93-84422-51-6
Proceedings of 2016 International Conference on Urban Planning, Transport and Construction Engineering (ICUPTCE'16)
Pattaya, Jan. 2-3, 2016 pp. 7-13

Understanding Commuter Sentiments from Tweets

Hoong Chor CHIN[1], Zhaoxia WANG[2]

[1]Dept of Civil & Environmental Engineering, National University of Singapore

[2]Department of Social & Cognitive Computing, Institute of High Performance Computing

***Abstract*:** *Typically, transport operators conduct surveys to evaluate their service performance. With the rise in social media and the abundance of information exchanged through twitter, it is now possible to obtain commuter feedback more rapidly and with a larger sample size. However, deciphering messages transmitted through tweets remains a serious challenge.*

This paper seeks to understand the sentiments of commuters from their tweet messages, in the aftermath of train service breakdowns in the underground MRT system in Singapore. The work examines tweets messages arising from 5 incidents of service breakdowns for 7 weeks from August 2013 to Oct 2013 and classifies them according to Shaver generic emotions, following which Emotion Scores are obtained. The sentiments are analyzed and the changes in sentiment scores observed prior to, during and after the service breakdown period and examined along with the extent and impact of the delays.

Keywords: *MRT breakdown, tweets, commuter sentiments*

1. Introduction

Typically, transport operators conduct surveys on a regular basis to evaluate their system performance and to assess the attitudes or perception of the commuters towards their service [1]. These may be in the form of self-completion questionnaires or interview surveys, which can be costly and generally difficult to conduct because of the low return or response rates. Increasingly, operators have turned to using online surveys to supplement or even replace the traditional form of paper surveys [2]. However, while online surveys can be broadcast more widely and can be more cost effective to administer, there is little guarantee that the targeted group is well represented among the respondents. Such surveys may also yield more superficial findings because, without the interviewer, the issues raised may not be adequately probed further.

Another form of online feedback which may become useful is social media talks or discussions, e.g. in facebook, twitter and blogs. Carefully followed, these social chats can highlight issues that can be more pertinent and relevant to the transport service. The unsolicited nature of the talk also means that the online users are more likely to have a stake in the transport operations and their feedback may be more genuine. Moreover, the issues raised are more current so that any deficiencies in the service can be addressed more promptly. Furthermore, discussions on the online forums may facilitate raising of multiple perspectives from different users thus allowing issues to be better understood.

While social talks can be a rich source of information to be examined, and transport operators cannot ignore the discussions raised in such platforms, there remain numerous pitfalls in dealing with these feedback [3], [4]. Firstly, there is a massive amount of information constantly being circulated on the online platform so that digesting them can be a herculean task. While various intelligent machine learning algorithms can be developed

to collect, classify and decipher these messages, it is still very challenging to achieve high quality analyses. Secondly, that these messages come with different formats, lengths and characteristics complicate the data analytics. Limiting this to a particular platform, e.g. tweets from twitter, simplifies the problem somewhat. In the case of twitter, the restriction of tweet length means the messages can be rather crisp and the use of keywords helps to isolate the domain of discussion.

Another concern in the use of social media information, e.g. tweets, is that of representation. Certainly, unlike the typical surveys, the sourced information cannot be considered a random sample. For example, a sizeable proportion of the transport commuters may not even be familiar with twitting and among those who may read the tweets, they may not participate in discussions. Offhand, there is little knowledge about the group of the tweet users. Hence, before any attempt is made to dip into the big data of tweets and to evaluate commuter sentiments, it may be useful, first, to develop a better understanding of the behaviour of tweet users.

This paper is an attempt to understand tweets made by Singapore commuters expressing their sentiments towards the breakdown of train services in the Mass Rapid Transit (MRT) in Singapore.

2. Methodology

2.1 Data Collection

The first step in the study is to collect the necessary tweet data for the study period from 24 August 2013 to 6 October 2013. Two sources of tweets were traced. The first is the official twitter page of the train operator (SMRT) with hash tag, @SMRT Singapore. These tweets represented announcements by SMRT on the status of the train system. The second category are the public tweets on SMRT. To narrow the search for tweet messages from the public twitter accounts, the keywords “MRT” or “train” were used to identify the potential tweets. Since the study is only concerned with the MRT performance and in particular commuter sentiments on the breakdown in train services, the public tweets were further filtered to eliminate irrelevant tweets.

During the study period, there were 5 instances of breakdown in train services resulting in delays to commuters, as shown in Table I. These failures occurred on different days of the week and under a variety of traffic conditions; 3 instances during the evening peak and 2 during off-peak periods. The estimated delays range from 20 to 35 minutes.

TABLE I: Occasions of train service breakdown studied

DATE	DAY	TIME OF INCIDENT	ANNOUNCED DELAY DURATION (MIN)	AFFECTED LINE	AFFECTED STATIONS
24-08-13	**Sat**	**17:37**	**20**	**East-West**	**Tampines – Pasir Ris**
24-09-13	**Tue**	**9:02**	**20**	**North-South**	**Newton – Marina Bay**
29-09-13	**Sun**	**6:40**	**35**	**North-South**	**Woodlands – Ang Mo Kio**
03-10-13	**Thu**	**17:05**	**30**	**North-South**	**Yishun – Woodlands**
07-10-13	**Mon**	**7:49**	**20**	**East-West**	**Jurong East – Queenstown**

The official announcements given by the SMRT were divided into several types according to the nature of the announcements. The first type is an announcement of a service disruption with the estimated expected delay duration. There may be follow-up tweets to update on the delay duration, whether or not there is a change in estimate. Another type is an announcement on the reason for the delay or advisory information about available alternative transport. In some cases of shorter delays, there may be no such announcements. Finally, there will always be an announcement when the normal service has resumed.

2.2 Data filtering

On the 5 occasions of service breakdown, the public tweets were tracked just prior to the first official announcement of the delay until the end of the day of the incident. In all, a total of 5,142 tweets were gathered

but not all these tweets were related to the service performance. For example, there were numerous tweets associated with what the twitters were doing, such as "I am eating near Clementi MRT". Since the purpose of the study is on service performance, particularly in relations to the service disruption, only tweets which contain relevant sentiments were extracted.

2.3 Sentiment Analysis

After filtering, only 874 tweets from the 5 occasions were used. These tweets were then assessed based on the sentiments expressed; matched to the list of 135 emotion words by Shaver [5]. Most of the sentiments in the tweets are associated with emotions under the basic categories of *Joy, Surprise, Sadness and Anger*. From the emotion attached to the tweet message, each tweet was assigned the Shaver Emotion Score (SES). From all the tweets, SES was found to range from the most positive of 1.90 for cheerfulness to the most negative of -1.23 for bitterness. The variations in twitter's emotions prior to and into the service disruption period were then traced. The results of these will be discussed in the following section.

3. Results and Discussions

To observe the changes in sentiments, the positive and negative sentiments are traced separately. Figures 1 to 5 show the Cumulative Shaver Emotion Score (CSES) on a time chart in relation to the time of official announcements from SMRT. By plotting CSES with time, the graphs can reveal the rate of increase in emotion scores as well as the density of the tweet response.

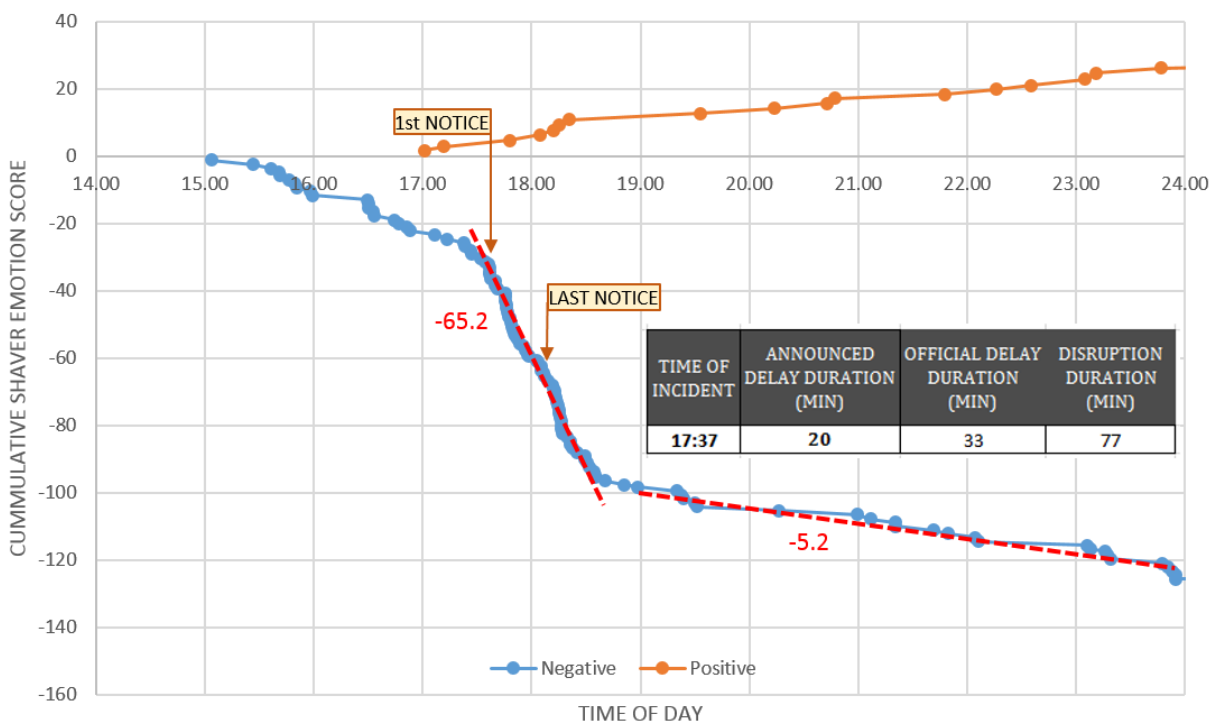
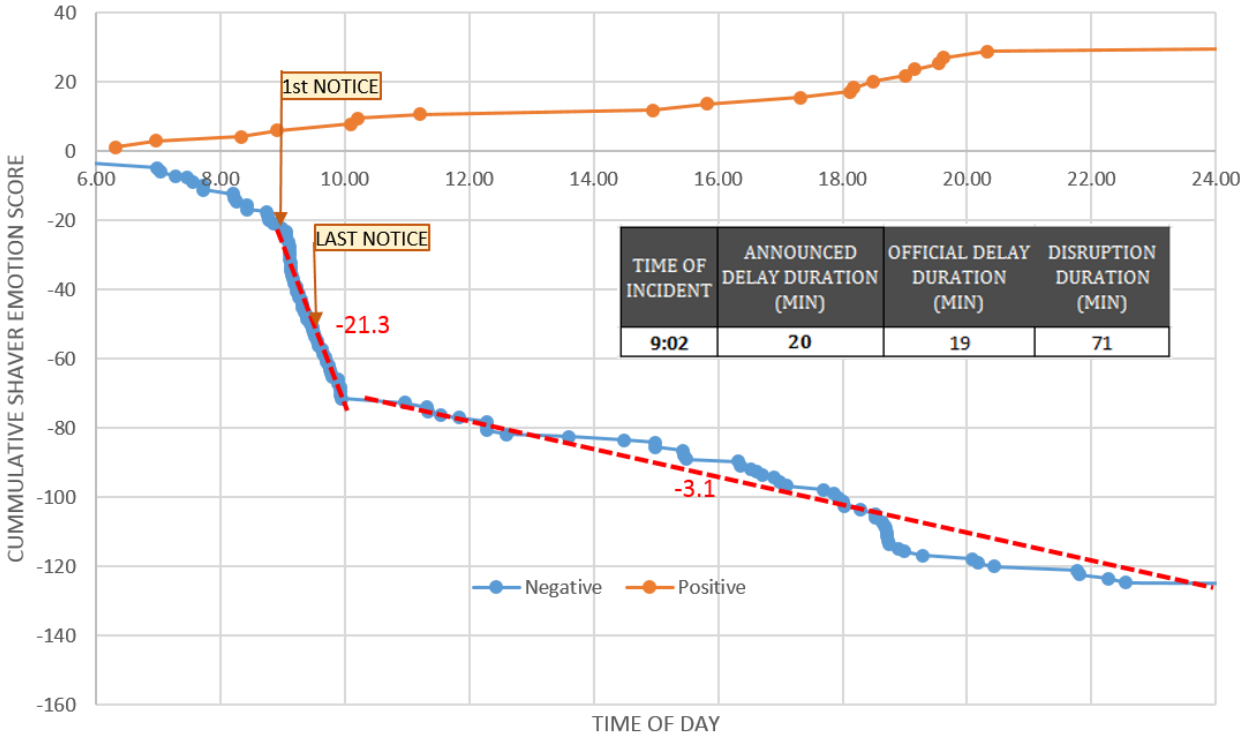


Fig. 1: Sentiment Analysis of MRT breakdown on 24 August 2013 (Saturday)

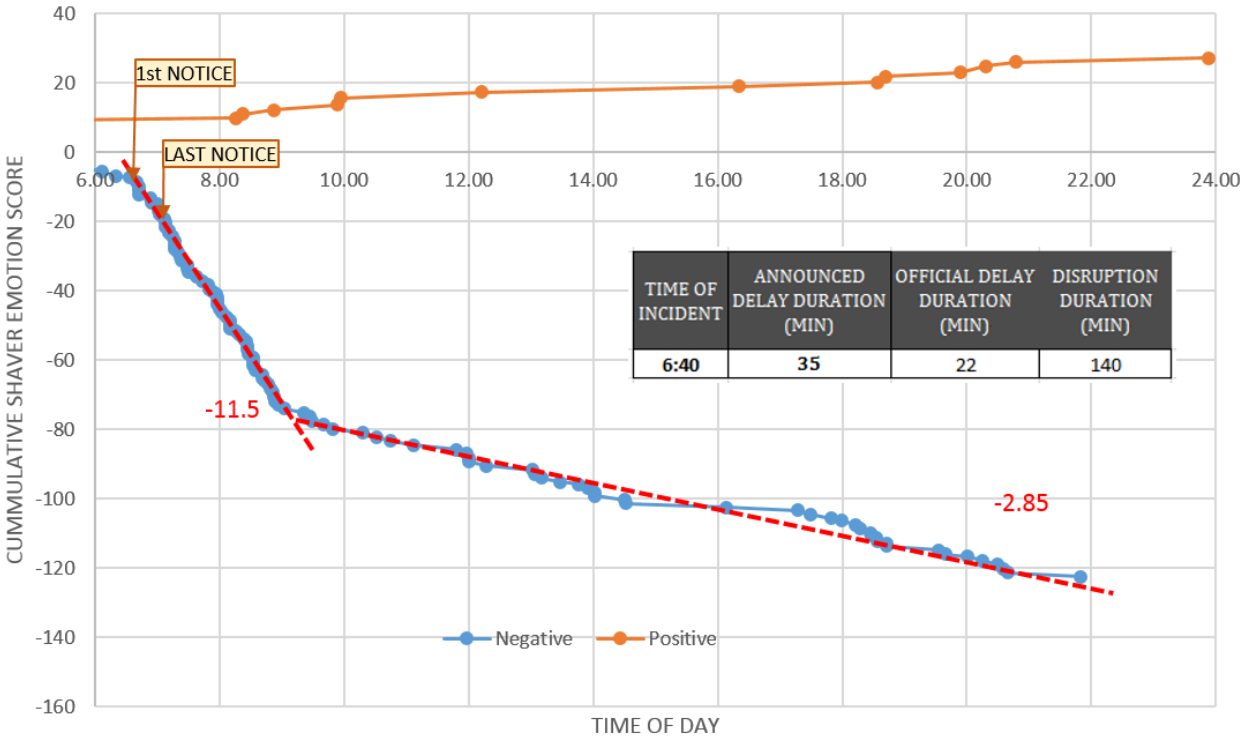
In each of the graphs from Figures 1 to 5, there are distinct periods during which the rate of growth in negative CSES is quite linear as indicated by the linear line along with the slope (measured in CSES/h). The value of growth in negative CSES during and after the disruption are also computed and shown in the graphs. Notice that for each day, the rate prior to the service disruption is almost similar to that of the pre-disruption period. The impact period of prolonged dissatisfaction appears to depend on the severity of the delay and the time of the day as well as the day of the week, all of which are related to the number of affected commuters as

well as the impact on station or train congestion. In general, this disruption or impact period is longer than the announced duration of the service breakdown. These values are tabulated in Table II for comparison. Table II also includes the computation of the mean SES for the tweets made during the disruption period.



TIME OF INCIDENT	ANNOUNCED DELAY DURATION (MIN)	OFFICIAL DELAY DURATION (MIN)	DISRUPTION DURATION (MIN)
9:02	20	19	71

Fig. 2: Sentiment Analysis of MRT breakdown on 24 September 2013 (Tuesday)



TIME OF INCIDENT	ANNOUNCED DELAY DURATION (MIN)	OFFICIAL DELAY DURATION (MIN)	DISRUPTION DURATION (MIN)
6:40	35	22	140

Fig. 3: Sentiment Analysis of MRT breakdown on 29 September 2013 (Sunday)

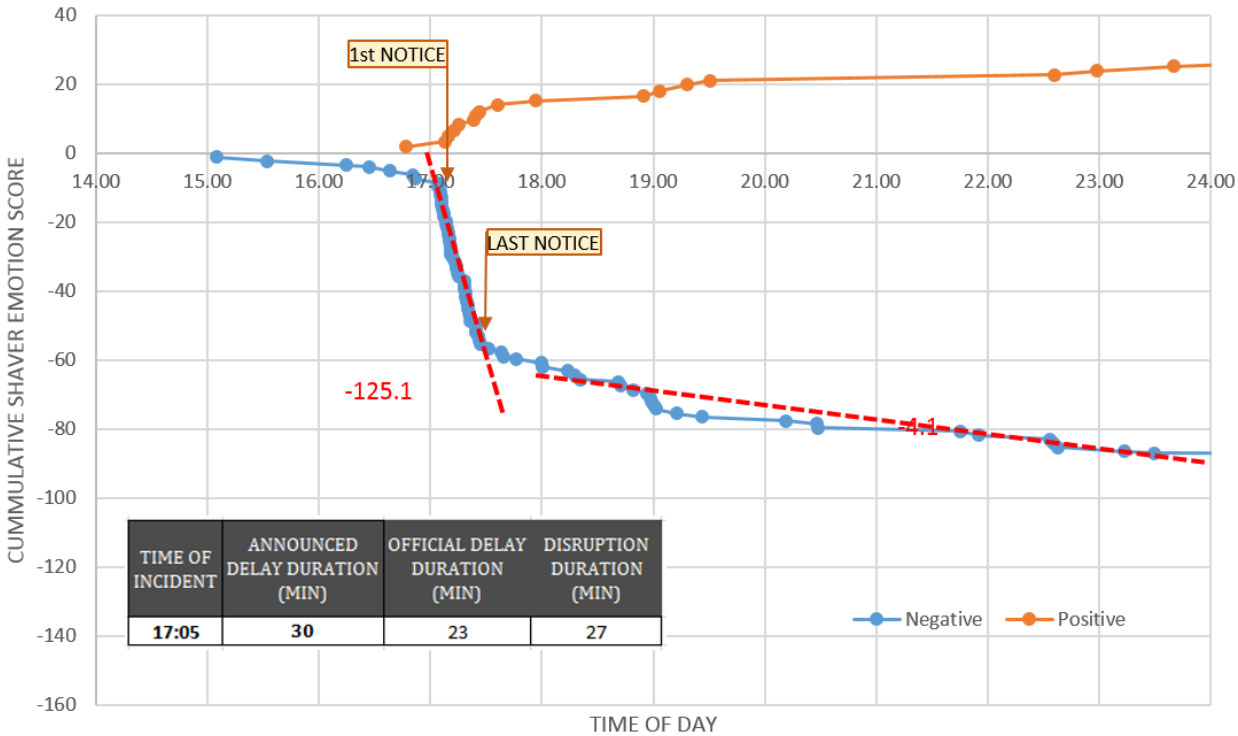


Fig. 4: Sentiment Analysis of MRT breakdown on 3 October 2013 (Thursday)

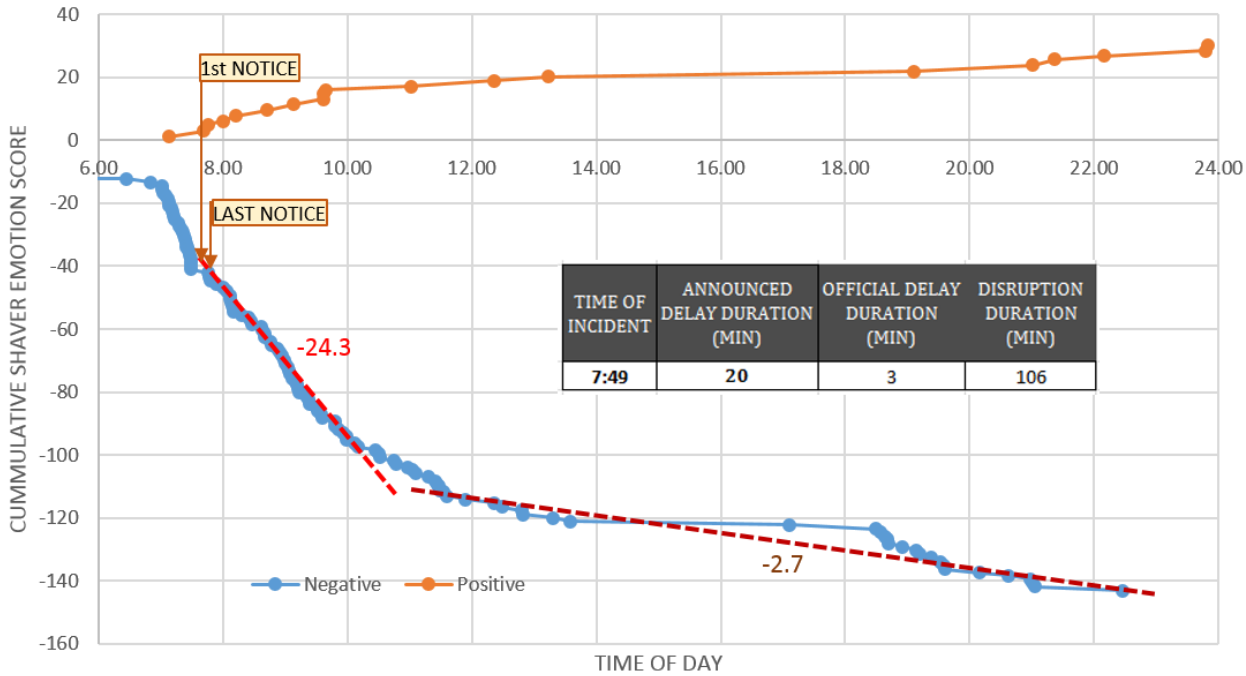


Fig. 5: Sentiment Analysis of MRT breakdown on 7 October 2013 (Monday)

Examining Figures 1 to 5 and the summarized results in Table II, we can see that the slopes of CSES are quite consistent during normal operating conditions, varying from -2.7 to -5.2 CSES/h, with the evening periods having slightly more negative values. This may represent the base level of negative sentiments towards the MRT system, which may be relatively stable in the short term during this 7-week period of investigation.

TABLE II: Summarized results of Sentiment Analysis

DATE	DAY	TIME OF INCIDENT	NUMBER OF AFFECTED STATIONS	NO. TWEETS	NO. NEG TWEETS	NO. POS TWEETS	ACTUAL DELAY (MIN)	GROWTH RATE (CSES/HR)	NORMAL RATE (CSES/HR)	IMPACT DURATION (MIN)	AVE SES
24-08-13	**Sat**	**17:37**	**2**	943	148	20	33	-46.5	-5.2	77	-1.08
24-09-13	**Tue**	**9:02**	**4**	1509	209	52	19	-21.3	-3.1	71	-1.1
29-09-13	**Sun**	**6:40**	**7**	660	113	18	22	-11.5	-2.9	140	-1.1
03-10-13	**Thu**	**17:05**	**4**	1076	131	35	23	-125.1	-4.1	27	-1.11
07-10-13	**Mon**	**7:49**	**6**	954	129	19	3	-24.3	-2.7	166	-1.12

The graphs also show that CSES changes rapidly from the base level on the onset of a service breakdown. On the 5 occasions, the rate of change varied from a low value of -11.5 CSES/h on Sunday morning of 29 Sep 2013 to a high of -125.1 CSES/h on the Thursday evening of 3 Oct 2013. Naturally, the impact is more severe on weekdays than weekends, when there are generally more commuters who are more concerned with missed appointments or late for work. With a higher number of affected commuters, there will also be more twitters who will voice out their frustrations. Moreover, the crowdedness effect during the peak conditions will further generate negative emotions among the commuters. It is interesting to note that commuters together seem to be more negative during the evening periods than the morning periods. This may be because commuters are more irritable and easily agitated by delays and congestion in the latter part of the day and especially during the evening after-work peak.

Comparing the breakdown on the Tuesday morning on 24 Sep 2013 (Figure 2) and the Monday morning of 7 Oct 2013 (Figure 5), it is interesting to observe that the CSES slopes during the service disruption are almost similar but the durations of the impact are rather different. On 24 Sep 2013, the negative impact on the commuters continued for another 50 minutes beyond the declared end of the disruption while on 7 Oct 2013, this was for a further 100 minutes. This may have to do with the actual ground situation since the affected trains and stations could have continued to be crowded until about 10:00am. Notice also that the negative sentiments already started building up at about 7:10am on 7 Oct 2013, about half hour before the first official announcement. The announcement of resumption of service which came shortly afterwards did not help to alleviate the negative sentiments resulting from the delayed notification.

The results for the Thursday evening of 3 Oct 2013 (Figure 4) is also worth noting. The negative sentiments appeared to rise quickly at -125.1 CSES/h and over a short duration. The anticipated delay of 30 minutes was higher than the 20 minute delays for the peak conditions on the other days. Naturally the notice of such a long delay tends to adversely upset many who were already in the system waiting for the trains. This may imply that there exists a threshold tolerance level somewhere between 20 to 30 minutes. Interestingly, on 3 Oct 2013, the duration of the impact on commuters was rather short; almost coinciding with the actual service breakdown duration. This is in stark contrast on other occasions when the impact period was felt long beyond the resumption of service. This can be explained that since commuters are less tolerant to the delay of 30 minutes, they may be more willing to seek alternative travels actively thereby reducing the demand on the system. Because of these reneging commuters, the crowded conditions generated by the earlier queues are more likely dissipate faster. Whereas for delays of 20 minutes, most commuters may choose to continue with the travel plans and not seek alternatives, which may not be very much better.

The average emotion score of the commuters when they were experiencing the service breakdown is also presented in Table II. Interestingly, there is little variation in the mean value. This implies that commuters generally express the same dissatisfaction of frustration, anger and disappointment with few displaying extreme outrage.

Another observation is that the positive CSES seemed to climb a little during the period of disruption. This may be due to the positive response from some potential commuters or bystanders who felt happy that they were not part of the affected party. These come later in the disruption period and are generally small in number.

4. Conclusion

It can be seen that it is possible to analyse the tweets to gain a good and quick feedback of commuter sentiments in the event of a train service breakdown. While the study is limited to 5 occasions of breakdown, it gives sufficient understanding of the sentiment behaviour of commuters. The approach can be applied over a longer study period to allow a better investigation into the correlation between commuter sentiments and conditions of delay and operator response. There is also intention to extend this by automating the process of analysis using machine learning [6] to manage the large dataset.

5. References

[1]. Wagner, J. "Measuring performance of public engagement in transportation planning: three best principles". *Transportation Research Record*. No. 2397, pp. 38-44, Transportation Research Board, 2013.

[2]. Chow, W. "Evaluating online surveys for public transit agencies using a prompted recall approach". M.S. thesis, Department of Civil and Environmental Engineering Massachusetts Institute of Technology, 2014.

[3]. Fellesson, M., & Friman, M. "Perceived satisfaction with public transport service in nine European cities". *Journal of the Transportation Research Forum*, Vol. 47, No. 3, pp. 93-103, 2008.

[4]. Wang, Z., Joo, V., Tong, C., Xin, X., & Chin, H. C. "Anomaly Detection through Enhanced Sentiment Analysis on Social Media Data". *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, pp. 917-922. IEEE, 2014.
http://dx.doi.org/10.1109/cloudcom.2014.69

[5]. Shaver, P., Schwartz, J., Kirson, D., & O'connor, C. "Emotion knowledge: further exploration of a prototype approach". *Journal of personality and social psychology*, Vol. 52, No. 6, pp.1061-1086, 1987.
http://dx.doi.org/10.1037/0022-3514.52.6.1061

[6]. Wang, Z., Tong, V. J. C., & Chin, H. C. "Enhancing machine-learning methods for sentiment classification of Web data". *Information Retrieval Technology*, Vol. 8870, pp. 394-405. Springer International Publishing, 2014.