Analyzing tweets to enable sustainable, multi-modal and personalized urban mobility: approaches and results from the Italian project TAM-TAM

A. Candelieri[1] & F. Archetti[1,2]
[1]*Consorzio Milano Ricerche, Italy*
[2]*Department of Computer Science, Systems and Communication, University of Milano-Bicocca, Italy*

Abstract

This paper presents the approaches proposed in the Italian project TAM-TAM to support a smarter, personalized and sustainable urban mobility, by taking into the loop the users of the transportation services, in particular citizens, tourists and commuters in Milan. A computational module has been defined and developed in order to collect and analyze relevant tweets posted by users as well transport operator. The main goals are two: identifying events (e.g. accidents, sudden traffic jams, service interruptions, etc.) and evaluating the overall sentiment about the service as well as mobility options. Detected events are used by other computational modules of TAM-TAM in order to support a more effective travel planning; on the other hand, the perceived quality of service is provided both to users, enabling more personalized choices, and to transport company, supporting them in the management of mobility supply.

Keywords: smart urban mobility, sentiment analysis, crowd sourcing.

1 Introduction

The wide diffusion of smart mobile devices together with the large use of social networking applications are rapidly changing the life style of the modern humankind. User-generated contents and crowd-sourcing represent huge sources of usually unstructured information which can be mined to enabling services and products innovation while improving quality of life for citizens.

The Italian project TAM-TAM, co-funded by the Italian Ministry of Education, University and Research together with Regione Lombardia, is aimed at designing developing an ICT-based platform for supporting a smarter, personalized and sustainable urban mobility.

Recent research work [1] points out the relevance of "*narrative aware design framework*" in the design and implementation of smart urban environments. TAM-TAM aims at further specify the vision proposed in [1], related to the combination of information from sensors data and *users' storytelling*, by exploiting the potential benefits provided by the huge amount of unstructured information generated through crowd-sourcing and related to transportation services.

Overall, TAM-TAM aims at providing innovative services through the design and development of:

- technological modules able to integrate data and information coming from different sources (e.g. time-tables, on-line positioning data, traffic estimation, etc.);
- applications to support and integrated and intermodal transport services, by taking into account urban and peri-urban mobility options;
- computational applications for cross-media contents analysis, based on sentiment and opinion mining techniques [2];
- applications for travel planning, able to provide users with information on costs, time, environmental impact and perceived quality of service;
- decision support functionalities to identify and address criticalities in the proposed transport service supply, enabling more effective and efficient plans according to variations in mobility users preferences.

TAM-TAM will benefit from the combination of data and information from official sources with the contents generated by the users, even on the move. Citizens, commuters and tourists may use the socially aware and collective intelligence based functionalities to make more personalized and informed mobility decisions. The project is going to be validated on a relevant use case, the urban and peri-urban area of Milan, in Italy.

The contribution of this paper is focused on the computational module that has been proposed by the authors in order to collect tweets both from users of the urban mobility services and official sources, such as the Twitter account of the public transportation company in Milan, *Azienda Trasporti Milano* (ATM), and then analyze their content through machine learning techniques.

The benefits provided by the automatic analysis of tweets have been already investigated and proved in other domains, such as the automatic detection of anomalies related to power outage events during hurricane Irene on August 27, 2011 [3].

The computational module developed in TAM-TAM has two main goals: the former consists in the automatic identification of events (e.g. accidents, sudden traffic jams, etc.), as posted by the users, the latter consists in selecting tweets containing some opinion about transport service (e.g. delays, inefficiencies, perceived security, dirt, etc.) in order to evaluate the overall sentiment about it.

2 TAM-TAM: general architecture

In the following Figure 1 the overall architecture of the TAM-TAM platform is reported. This paper is focused on the system component devoted to the retrieval and analysis of the contents (i.e. tweets) generated by users and related to urban transport service. This component is highlighted in gray and consists of the following sub-modules:

- A *crawler* which may be configured to retrieve, regularly, the most recent tweets according to specified criteria;
- An *analyzer* devoted to perform a preliminary analysis of the retrieved tweets and assess if the correspondent content is associated to the communication of some events (e.g. delay, interruption, traffic jam) or opinion about service;
- An *event analyzer* aimed at storing the relevant information on the event (e.g. mobility option, line, location, etc.) into the systems central repository;
- A *sentiment analyzer* to evaluate the opinion/sentiment reported by the users, in particular with respect to each specific mobility option. The information is then stored into a dedicated database to enable evaluation over time;
- A *query executor* aimed at performing specific query on the dedicated database in order to provide information about the perceived quality of service of the different mobility options over time (e.g. last week, last month, last three months, etc.).

Finally, the other components of the TAM-TAM platform are devoted to provide all the functionalities of data retrieval and management, optimal travel planning, users login and profiling, integrated ticketing and decision support for mobility service managers.
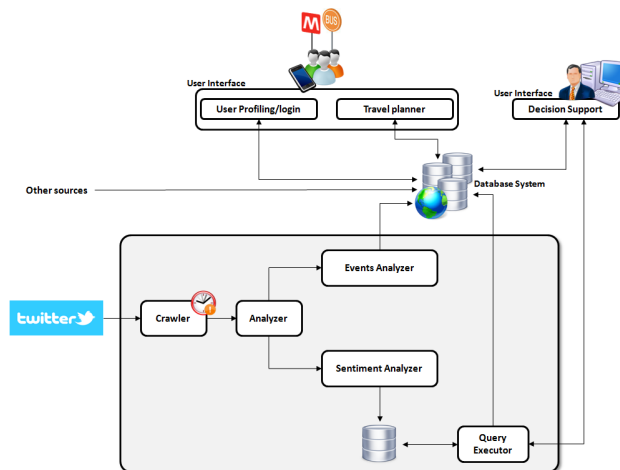

Figure 1: Overall architecture of the TAM-TAM system.

3 Materials and methods

Design and development of the analytical module of TAM-TAM has initially required the identification of searching criteria to retrieve relevant tweets and, consequently, collect them into a database. Results presented in this paper are related to tweets in the period 12 June 2013 to 25 February 2014, posted from/to the account of the public transport company in Milan (ATM, Azienda Trasporti Milano) as well as generated by communities of commuters taking trains every day to travel from/to Milan.

In more detail, searching criteria are the following:

- all the tweets *from* "@atm_informa" (Twitter account of the public urban transport company in Milan)
- all the tweets *to* "@atm_informa"
- all the tweets *from* "@CoordPendPV" (Twitter account posting news and notifications for the community of commuters travelling by train between Milan and Pavia)
- all the tweets *from* "@PendolariVCPV" (Twitter account posting news and notifications for the community of commuters travelling by train between Vercelli and Pavia)
- all the tweets containing the *hashtag* "#treniPV" which is used by the commuters travelling by train from and to Pavia (most of them works in Milan)

Retrieved tweets have been labeled by a human supervisor according to four possible alternatives: event, positive, negative and neutral. As tweets are short messages usually unstructured and informally written, techniques like parsing, pattern matching, complex grammars are usually ineffective. In [4] the content is represented by the frequency of each word or n-gram word pair; in [5, 6] words are counted as Boolean; in [7] words are weighted by their correspondent Inverse Document Frequency (IDF) score, that is the logarithm of the number of documents in the collection (tweets) divided by the number of documents containing a specific word. An alternative score is the Term Frequency-Inverse Document Frequency (TF-IDF) that is the IDF score multiplied by the frequency of a specific word divided by the number of words in the document [8].

In a more recent study [9] an extension of the TF-IDF approach is proposed, consisting in weighting words by the difference of their TF-IDF scores (delta TF-IDF) with respect to the associated positive or negative sentiment. The Support Vector Machine (SVM) classification learning technique [10, 11] has been used to identify a reliable model able to detect the polarity of a document with respect to the computed delta-TF-IDF.

This approach proved to be more accurate with respect to the other ones, and is the core of the application presented in this paper. In particular, a multi-level schema has been applied:

- delta TF-IDF has been computed with respect to tweets related to events or not and a correspondent SVM classifier has been trained.

- for all the tweets previously classified as "not event", delta TF-IDF has been computed with respect to neutral and not neutral contents; a correspondent SVM classifier has been trained.
- for all the tweets previously classified as "not neutral", delta TF-IDF has been computed with respect to positive and negative contents; a correspondent SVM classifier has been trained.

Taking into account the overall architecture depicted in previous Figure 1, the *analyzer* uses the first SVM classifier to detect tweets containing information on events and send them to *event analyzer*. Then the *analyzer* adopts the second SVM classifier to detect the tweets containing some sentiment/opinion and send them to the *sentiment analyzer*.

4   Results

Preliminary results obtained in the project TAM-TAM are presented in this section. As a first result, some considerations about volume and rate of tweets are reported with respect to the different searching criteria.

In the following Figure 2, daily tweets rates *from* and *to* "@atm_informa" are depicted. It is easy to note "anomalous" peaks within the time period considered; moreover, the peaks in the two time series are overlapping. These peaks are associated to specific events: strikes and interruption of 1 of the 4 underground lines (due to technical difficulties). The interruption is easy to recognize: it occurred in correspondence of the third peak. Users post to "@atm_informa" in order to ask for information and news during the strike and "@atm_informa" reply to provide users with such information; on the contrary, when the interruption occurred, most of the users post gripes to "@atm_infoma" who – logically – reply only to request of information.
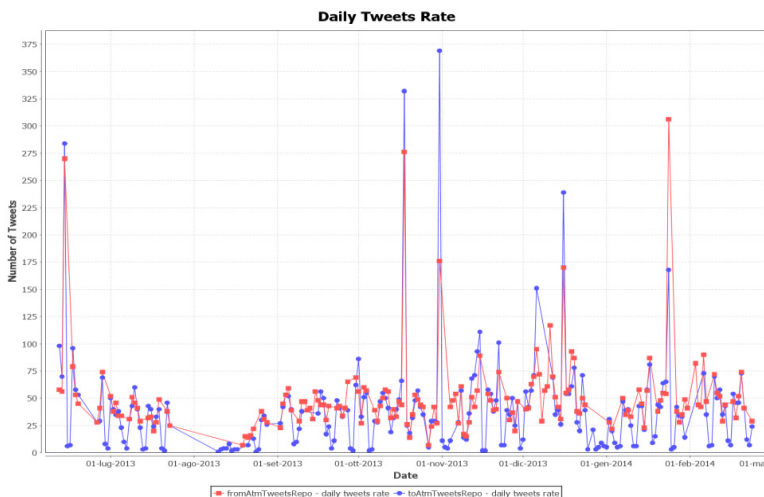

Figure 2:   Daily tweets rate *from* and *to* "@atm_informa".

In the following Figure 3 the daily tweets rates related to the searching criteria *from* "CoordPendPV" (red), *from* "PendolariVCPV" (blue) and containing the *hashtag* "#treniPV" are depicted.

Compared to the previous Figure 2, the volume of tweets posted by commuters in the peri-urban area of Milan is lower than the one in the urban area.

Although in this case there is not a reply by the transport company to the tweets posted by commuters, peaks in the time series are still the main indicator of a possible sentiment reported by the users of the transport service.

In particular, each commuter is used to post tweets in order to share his negative experience and opinion with other commuters, more than positive sentiment. Even in this case, peaks are associated to strikes impacting the train service and interruptions.
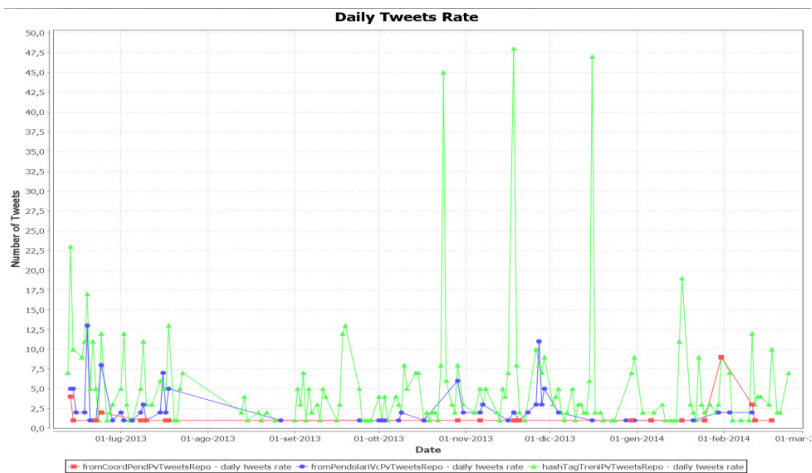

Figure 3:   Daily tweets rate *from* "CoordPendPV" (red), *from* "PendolariVCPV" (blue) and containing the *hashtag* "#treniPV".

Finally, performances related to the accuracy in classifying tweets are reported in the following Table 1. The technique named 10 fold cross validation has been used in order to estimate the capability of the system in classifying, correctly, future tweets.

As positive tweets are really few, classification appears more difficult showing a lower, even if relevant, level of accuracy.

Table 1:   Classification performances.

Classes	Accuracy on 10 fold cross validation
Event/not Event	87.7%
Neutral/not Neutral	85.3%
Positive/Negative	82.3%

5 Conclusions

This paper presented the activities performed to develop an analytical model able to collect and analyze tweets about urban and peri-urban transport service in Milan, as generated by citizens, tourists and commuters.

The results obtained prove the benefits to keep in the loop crowd-sourcing in order to enable a smarter and personalized urban mobility management. Every user is able to make more informed mobility decisions, taking into account also the quality of service as perceived and reported by other users. At the same time, this information may be used by the transport companies in order to monitoring the sentiment/opinion about their mobility supply and, possibly, plan effective and efficient modifications to the available options.

This analytical functionality is only one of those provided by the platform TAM-TAM, which is still an ongoing real-world experience on how social and collective intelligence may change transportation services in modern cities.

References

[1]   Srivastava, L., Vakali, A., Towards a narrative aware design framework for smart urban environment, F. Álvarez et al. (Eds.): FIA 2012, LNCS 7281, 166–177, 2012.

[2]   Pang, B., Lee, L., Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval, 2(1, 2): 1–135, 2008.

[3]   Thom, D., Bosch, H., Koch, S., Worner, M., Ertl, T., Spatiotemporal Anomaly Detection through Visual Analysis of Geolocated, IEEE Pacific Visualization Symposium 2012.

[4]   Joachims, T., Text Categorization with Support Vector Machines: Learning with Many Relevant Features, Springer, 1997.

[5]   Pang, B., Lee, L., Vaithyanathan, S., Thumbs up? Sentiment classification using machine learning techniques, In Proceedings of EMNLP, 2002.

[6]   Whitelaw, C., Garg, N., Argamon, S., Using appraisal groups for sentiment analysis, In Proceedings of the 14th ACM International Conference on Information and Knowledge Management, 625–631, 2005.

[7]   Kim, S., Pantel, P., Chklovski, T., Pennacchiotti, M., Automatically assessing review helpfulness, In Proceedings of EMNLP, 423–430, 2006.

[8]   Yates, B., Modern Information Retrieval, Addison-Wesley Longman Publishing Co., 1999.

[9]   Martineau, J., Finin, T., Delta TFIDF: An Improved Feature Space for Sentiment Analyis, In Proceedings of the Third International ICWSM Conference, 258–261, 2009.

[10] Scholkopf, B., Smola, A. J., Learning with kernels. Support Vector Machines, regularization, optimization and beyond. Massachussetts Institute of Technology, USA, 2002.

[11] Vapnik, V., Statistical Learning Theory. New York, Wiley, 1998.