Title: Sensing Emotions with Heterogeneous Mobile Phone Data

Dr Weisi Guo and Dr Maria Liakata

Objectives

There is a lack of research into how to correlate low level data from mobile devices (signal strength, GPS location) with higher level data (e.g. application data from social media) to derive useful patterns. Our research proposes to take a series of high level data output over time and combine it with the corresponding low level data to infer an understanding of user emotional and behavioural profiles and how these correlate to the users' environment.

Background

Recent work has used twitter to capture the personality of an individual [1] and the general mood of a population [2], with applications such as stock-market prediction [3]. However, we are not aware of work on the emotional fluctuation of users based on their tweets over time. Low-level data can reveal hidden trends about the user environment, which is generally discarded by higher level applications. Variations in the instantaneous signal strength can reveal the dynamism and structural surroundings of that location. Existing work has shown that instantaneous variations in signal variations can more accurately yield the location of a mobile user than traditional GPS methods [4].

Methodology

The low and high level data sets can be coupled to create a heterogeneous fingerprint on a user's emotion and behaviour. This mapping process can yield a richer understanding into not only how people feel over time, but if there is any correlation between emotions and the physical environment. To achieve our objectives we envisage the following steps:

1) Acquire data from an urban laboratory comprised of a number of smartphone volunteers. This will generate a rich set of data on behavioural patterns and their corresponding physical environment, which is inferred from the heterogeneous data set. [Before the start of the project-involvement of MSc student optional]
2) Classify the application layer data (i.e., tweets) in terms of emotions [2 weeks]
3) Correlate the variation of signal strength with the physical environment using statistical modelling techniques. The statistical properties will then be mapped to the application layer data. [3 weeks]
4) Build a discrete-time series model of human emotions and behaviours. [3 weeks]
5) Extract patterns of how the discrete-time emotions data relates to the physical environment data. [4 weeks]

This work will require skills in analysing time series data, statistical modelling and data mining. Classification of tweets will build on existing work on the classification of emotions in tweets [Current MSc student at Warwick supervised by Dr Liakata].

Applications and Future Plans

Being able to correlate a users' emotional state over time and in relation to the user's environment will increase our understanding of both emotional cycles, as well as how the environment plays a role in affecting these. We see this work as a pilot study for a larger project which would apply these methods on a large scale and would have important applications. Our plan is to seek council as well as industrial funding to extend the implications of this work.

References

[1] Golbeck J, Robles C, Edmondson M, Turner K, *'Predicting Personality from Twitter'*, 2011 IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, pp. 149-156, 2011.

[2] Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM (2011) Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. PLoS ONE 6(12): e26752. doi:10.1371/journal.pone.0026752

[3] Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter mood predicts the stock market." Journal of Computational Science 2, no. 1 (2011): 1-8.

[4] Kaemarungsi K, Krishnamurthy P. *Modelling of Indoor Positioning Systems Based on Location Fingerprintin*g, Computer Communications Conference (INFOCOM), IEEE. Vol. 2, pp.1012-1022, 2004