# PULSE: An Auditory Display to Provide a Social Vibe

David McGookin, Stephen Brewster
Glasgow Interactive Systems Group
School of Computing Science
University of Glasgow
Glasgow G12 8QQ
firstname.lastname@glasgow.ac.uk
www.dcs.gla.ac.uk/~mcgookdk

## ABSTRACT

An increasing amount of social media is being tagged with the location of its creation. However, little investigation of how these tagged media can be used has been undertaken. We seek to exploit their auditory presentation in a system called PULSE. PULSE attempts to provide an understanding of the people, places and activities that are happening in the user's current locale. We outline the design of PULSE and how both message and meta-data can be implicitly and explicitly incorporated into an auditory display. We outline our plans for future evaluations to further consider how social geo-data can be aurally presented to users.

## Keywords

Auditory Display, Sonification, PULSE, Tourism, Vibe

## Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Interaction Styles*

## General Terms

Design, Experimentation, Human Factors

## 1. INTRODUCTION

In the last few years the capability of mobile handsets has increased significantly. Devices such as the iPhone, Android handsets and Windows phones provide always on connectivity to the Internet, access to thousands of applications and a rich set of technologies to allow sensing of the environment. Such devices are expected to emerge as the dominant computing platform, overtaking the personal computer. At the same time many social networking services, such as Facebook (www.facebook.com) and Twitter (www.twitter.com), which focus on User Generated Content (UGC) have emerged. Together, these mean that an increasing amount of UGC is associated with the physical location where it was created (geo-tagged). This has lead to new social networks, such as Foursquare (www.foursquare.com), that more heavily centre on location-based interactions. Whilst the amount of user generated geo-tagged content increases, there has been limited investigation into the opportunities that this creates in exploring and using the content. Searching over social networks is still primarily based on textual queries, or by filtering on user added metadata such as hashtags. Where location-based queries are supplied, they tend to be buried in the user interface with results displayed as pins on a map (see Figure 1 (A)). Other media, such as photographs, indicate the potential of different presentation techniques. Figure 1 (B) shows a screenshot of a search for nearby (to the user's current location) photographs performed by the 'Mobile Photos' flickr client. This provides a rich feeling of the area, an intrinsic understanding of the people, places and activities that exist in the environment around the user's current location. This intrinsic 'vibe' or 'feel' for a physical location is a key reason why people would wish to visit it. Guidebooks often try to describe this feeling to communicate what it is like to live in a location and understand a place from the perspective of a local. Lonely Planet [9] for instance prints discussions with locals about their lives as a means of trying to communicate the feeling of living there: *"I was born here in Savo, and I live a third of each year here in my summer cottage. I feel this is the most valuable thing about the area, that in the nature here you can really get away from all the troubles of a normal life, a busy life"* [9, p.136].

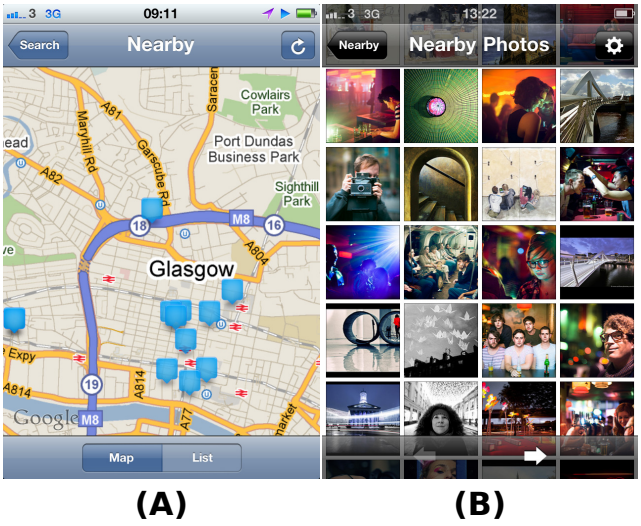


**Figure 1: How the Twitter iPhone application presents the results of a geo-search (left). How the "Mobile Fotos" iPhone app presents nearby photographs to a user (right).**


*IWS '11*, August 30, 2011 Stockholm, Sweden

However, guidebooks can be cumbersome to use during exploration [2], and because of their static nature, the information provided may not be congruent with the user's current experience of the environment. In our work, we are seeking to explore how user generated social media can be exploited to provide the understanding that a local would have. We term this the 'vibe' or social feel of an area. In the following sections we outline important related work, before discussing the design and implementation of PULSE – an auditory display to provide a contextual feeling of an area – before outlining our plans for its evaluation.

## 2. RELATED WORK

As described, there has been little work investigating how geotagged social networking data can be presented to users. However, work on the analysis of social media indicates that the content generated is likely to be useful in the context of providing a vibe. Java *et al.* [5] analysed the content of messages on Twitter and found that the majority concerned the activities and feelings of users at the time the messages were posted; a key feature of a vibe. Bollen, Pepe and Mao [1] showed that through an analysis of messages it was possible (given enough sample points) to extract the mood of users. This highlights that it is possible to extract higher levels of activity or mood from messages, such as when important local events are occurring. If we know social messages were created within a certain geographic area, it is at least possible that we can use them to provide the vibe of that area.

A second area of work concerns how we might communicate the vibe to the user. An obvious approach is to present it visually. However, there are a number of issues with this approach. Brown and Chalmers [3] found that whilst traditional visual tour guides were useful, they were not often accessed when actually visiting an area. Users tended to either pre-study the guidebook to try to absorb information, or browsed it afterwards to help fill in the blanks. Visual interfaces also require the user to make an explicit decision to look at the device, taking them away from the sights and sounds of the physical environment. As a vibe is more of an underlying feeling that augments real world information rather than replacing it, occupying the visual channel is likely to be inappropriate. Another approach is to use audio or sonification. Here the user could wear earbuds which would allow background noise to be heard, but would require no explicit interaction in order to receive the vibe. Such approaches have been evaluated in the context of tour guides [7] and have more recently been applied to message based data. Dingler and Brewster [4] have investigated how RSS feeds, Twitter messages and Facebook update activity can be communicated to a user. They used a number of thematically linked sounds to represent different events, e.g. bubbles to represent Facebook newsfeed activity, with splashes representing direct messages. The overall aim of their AudioFeeds system was to communicate overall levels of social network activity, rather than semantic content about the messages. Users would still need to activate the visual display to read the messages. Whilst not concerning social networks, a relevant approach is the Rider Spoke system by Rowland *et al.* [8]. Their study involved users cycling around town and occasionally stopping, finding a location that was important to them and answering personal, confessional questions. Riders could also opt to listen to other rider's answers. Whilst the locations users recorded messages in were logged, these did not form a part of the delivery mechanism and were used only for later analysis.

Whilst work has not been carried out to investigate if geo-coded social media can provide the vibe of an area, the work described here clearly indicates that much of the information necessary to generate it is available in the information posted on social networking sites. In addition, because of the underlying nature of the vibe, auditory presentation is a suitable means of communication.

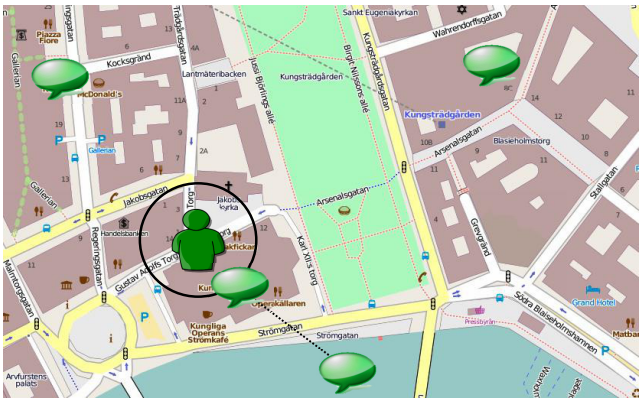


**Figure 2: The closest message is selected for presentation at a point 15m from the user along the same axis.**

## 3. PULSE

As a means of investigating how to generate and present a vibe through the use of social media, we have developed an application called PULSE to help establish what attributes of information are important and how they can best be communicated. PULSE runs on the iPhone platform, and uses the Twitter social media service to provide content. From the earlier discussion of guidebook use it is important that active interaction is not required. Users interact with PULSE passively as they walk around a physical location. Information is directly or indirectly presented through a spatialised auditory display via earbuds. These do not block out external sound, and our aim is to augment, rather than replace, the existing auditory environment in an area with a vibe. Twitter provides a geo-search API allowing developers to search for messages within a defined physical location. In PULSE we search over a relatively small area (around 100m radius) near the user. Each message obtained is processed to remove acronyms or urls. Common contractions, such as *'lol'*, are replaced with longer equivalents. At periodic intervals we select the closest message to the user's current location subject to a maximum cutoff distance of 100m. This message is then rendered using the Cereproc (www.cereproc.com) text-to-speech engine and presented geocentrically in a 3D auditory space as an audio bubble [6]. Because the messages are so short (each is a maximum of 140 characters), we move them so that they emanate from the same direction, relative to the user, but are presented at a distance of 15m. This means that users can either attend to the message or continue to walk, causing the message to naturally fade out as the distance between it and the user increases (see Figure 2). In such a way we are providing whispers of conversations around the user and one element of a vibe.

Simply presenting the content of a few messages, whilst important, is not sufficient to represent a vibe as we previously described. Such messages provide only the concrete examples of what activities occur in an area. They do not provide an overview. In addition, depending on user walking speed and the number of messages downloaded, it is very unlikely that a user would hear all of the messages available. In order to provide the full vibe we must also communicate more implicit aspects of the messages in a timely fashion. We consider the following to be key implicit aspects.

*Message Volume:* A clear metric is the number of messages that have been downloaded in a particular 100m radius area. If few

**Figure 3: During Vappu (May Day) Finnish students often wear brightly coloured overalls. The significance of these is not clear to the outside visitor. Photograph courtesy of 'Ms L' at flickr.com under creative commons attribution.**

messages have been retrieved then it is clear that there is not a lot happening in that area.

*Message Density:* Related to the volume of messages that are downloaded, the rate at which messages have been generated and their age is also relevant. An area where there are long time gaps between messages would be indicative that not a lot is happening. This is not necessarily a bad thing, and its meaning would need to be determined based on the user's other senses. For example, if the user is wandering around a sleepy seaside village, message density may be lower than if wandering around a city centre. However, the meaning of each of these would be different. In the case of the seaside village it may heighten the laid back nature of where the user is. In the city it may indicate that the user is wandering into an area where there is little to see or do, and changing direction may be more appropriate.

*Topic Diversity/Agreement:* A final important area is in the consistency and diversity of the messages in a given area. We argue both are important. Consistency can help better inform the user of large scale or common activities taking place in the current area. For example, during the traditional Vappu (May Day) celebration in Finland, groups of students hang around in towns wearing brightly coloured overalls (see Figure 3). To a visitor the reason for this may not be obvious, but the information obtained via PULSE should show a peak in discussion or comment around the Vappu celebration. I.e. a larger proportion of the messages would be identifiable as related to each other through the added hash tags. Communicating emergence of a topic is an important part of the vibe and would allow the user to integrate the unclear sight of jump suited students with a clear trend in the auditory soundscape. It is not hard to imagine similar situations such as parades, sporting or cultural events where a clear trend exists in messages that would augment a user's understanding. It is also equally important to be aware of the diversity of topics and activities that exist in an area. This would provide an impression of the richness and variety of a place. For example, a bussling city center may have lots of different activities that have a similar number of generated messages.

In all cases, this information is relatively straightforward to obtain from the messages downloaded. It is much less straightforward to consider how it should be integrated into the auditory display. In our current version of PULSE we are considering two approaches: implicit presentation and explicit presentation.

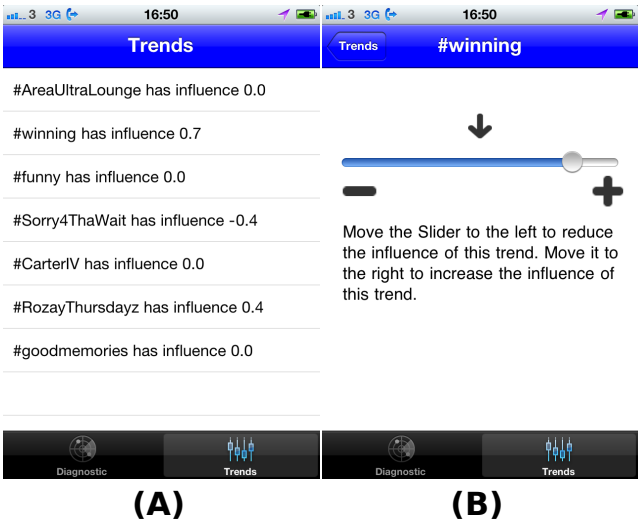


**Figure 4: Detected trends are displayed in a list (A). This allows the user to access and manipulate the influence that a trend has in the selection of messages to hear (B).**

## 3.1 Implicit Presentation

In the implicit presentation version of PULSE, rather than provide additional auditory cues to present the implicit information, we use modification of the presentation characteristics of the messages themselves to communicate the data. Here message density is reflected in the separation gap between messages. In our initial version a new message was presented every 30 seconds, plus or minus a maximum 10 second random variation. In the implicit presentation version we map the presentation rate to the mean time between the generation of messages in the current area. The shorter the time between messages, the more frequently those messages will be presented. We cap the rate so that messages are not presented any more frequently than one every 15 seconds. As the rate falls we increase the time between successive message presentations. However, it will take a few message presentations before a user is aware of this increase or decrease. This might be useful when the user is relatively static, such as enjoying a coffee at a street cafe, but it may be too slow if the user is moving through the city and the message density is changing quickly.

Whilst density is obvious to present implicitly, it is less clear how the diversity or agreement of a topic can be realised. Our initial approach, when a topic was identified, was to elevate messages from that topic so their presentation was prioritised over other messages. This is unsatisfactory for two reasons. Firstly, as with presentation of message density, the emergence of a trend will take time to be apparent to the user. As the user moves around trends may come and go, something the user is unlikely to be aware of. Secondly, the user may not consider the topic messages to be of greater importance to be elevated. For example, a user may not be interested in a nearby football game and wish to exlude those messages. To counter this, we have taken the approach of changing the speaker voice when a message from a topic is presented. In this way the user is quickly made aware of one or more emergent topics and, via a slider on the visual display, can decide whether to elevate it (see Figure 4). Future messages on the same topic would then be prioritised over the closest message, even if they were slightly further away.

Implicit notification via the modification of message presentation

has the advantage that we are not obtrusively modifying the existing soundscape with new auditory cues. We cannot be sure what environment the user is in when using PULSE, so picking auditory cues that fit with the environment is difficult. What would be appropriate in a rural setting may be different to an urban city. However, as shown above, implicit cues also have disadvantages. They are likely slower to be understood by users which may limit their effectiveness. Implicit cues also provide a courser grained feedback. The use of a different speaker for emergent trends provides only binary feedback: there is a trend or not. It cannot provide the amount or significance of that trend as a proportion of all messages in the area.

## 3.2 Explicit Presentation

Our second approach is to employ explicit presentation of the additional information. This allows the volume, density and topic agreement to be more rapidly presented to the user, overcoming many of the issues of the implicit notifications discussed, and allowing a finer grained vibe to be presented. However, as mentioned, the sounds need to work within the existing, natural, auditory environment. In our initial version we have based these explicit cues around a water-based soundscape. Water exists in both urban and rural settings, so is a good metaphor to draw upon. It also allows us to present a structured soundscape [4] to complement the spoken messages.

In this version of PULSE, the volume of messages is represented by the sound of dropping stones into water. The number of splashes heard is proportional to the number of messages downloaded. Message density is linked to the auditory volume of flowing water. As the message density increases, the volume also increases, providing the impression that the water has increased it's flow rate. Conversely, a decrease in message density will reduce the audio volume. We limit the audio level so that it does not become too prominent, and we are currently carrying out tests to determine what the mapping function from message density to volume should be. Finally, emergence of topic is presented via a bubbling sound, analogous to a common topic bubbling up. As with message density, auditory volume is increased based on the number of messages in the current highest trending topic. This provides a finer grained indication of the popularity of a particular topic than is available with implicit notifications. As with implicit notifications the user can manipulate a visual slider to request greater weight be given to messages from determined topics.

As the assumption behind PULSE is that the system might be active for long periods, we do not want to constantly provide auditory cues to the user as this may become annoying. Therefore in designing the cues we have tried to ensure that sounds will not be constantly presented for long periods. We are currently calibrating our mappings to determine appropriate levels to minimise annoyance.

# 4. DISCUSSION

Our future plans involve a two part evaluation of PULSE. Firstly, we want to further investigate the implicit and explicit feedback techniques to present the summarised data. It seems that both feedback types have advantages and disadvantages, and these need to be investigated in more detail. We need to determine both the ability of users to derive a feel for the metrics we are trying to communicate with each technique as well as user preference for each. Once the feedback techniques are established, we intend to carry out longer term studies investigating how the use of PULSE affects visitors' understanding of an area and the people that live there. There are several open issues that we still need to consider. For example, if the amount of social activity in an area is not reflected by the amount of online social activity, will this distort a user's understanding or impression of an area? The results of both of these should significantly enhance our understanding of how to communicate a vibe.

Our current implementation of PULSE allows us to leverage social networking sites, and the increasing amount of geo-coded content they generate, to provide a potentially useful and informative aid in understanding the locations and places a person may visit. The nature of the passive interaction used in PULSE, and the need for it to evolve over time, greatly suits the auditory modality for presentation. Users do not need to remember to explicitly check their mobile devices, but rather simply wear headphones. Our work so far has only considered short, text based message services, but we have shown a relatively sophisticated overview of activity can be generated from these. In conclusion, we believe that employing sound to integrate the increasing amounts of geo-content back into the real world is a rich area of investigation that deserves much greater study than has been so far employed.

# 5. ACKNOWLEDGMENTS

This work is supported by EU FP7 Project No.224675 "Haptimap".

# 6. REFERENCES

[1] Bollen, J., Pepe, A., and Mao, H. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)* (Barcelona, Spain, 2011), AAAI.

[2] Brown, B. Working the problems of tourism. *Annals of Tourism Research 34*, 2 (2007), 364–383.

[3] Brown, B., and Chalmers, M. Tourism and mobile technology. In *Eighth European Conference on CSCW* (Helsinki, Finland, 2003), vol. 1, Kluwer Academic, pp. 335–354.

[4] Dingler, T., Brewster, S., and Butz, A. Audiofeeds - a mobile auditory application for monitoring online activities. In *ACM Multimedia 2010* (Florence, Italy, 2010), ACM Press, pp. 1067–1070.

[5] Java, A., Song, X., Finn, T., and Tseng, B. Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* (San Jose, California, USA, 2007), vol. 1, ACM Press, pp. 56–65.

[6] McGookin, D., Brewster, S., and Priego, P. Audio bubbles: Employing non-speech audio to support tourist wayfinding. In *HAID 2009* (Dresden, Germany, 2009), Springer, pp. 41–50.

[7] Reid, J., Geelhoed, E., Hull, R., Cater, C., and Clayton, B. Parallel worlds: immersion in location-based experiences. In *CHI 2005* (Portland, USA, 2005), ACM, pp. 1733–1736.

[8] Rowland, D., Flintham, M., Oppermann, L., Marshall, J., Chamberlain, A., Koleva, B., Benford, S., and Perez, C. Ubikequitous computing: Designing interactive experience for cyclists. In *MobileHCI 2009* (Bonn, Germany, 2009), ACM Press, pp. 1–11.

[9] Symington, A., and Dunford, G. *Finland*, 6th ed. Lonely Planet. Lonely Planet Publications Pty Ltd, London, UK, 2009.