Text Mining of Personal Communication

Understanding the Technical and Privacy Related Challenges

Håkan Jonsson
Corporate Technology Office
Sony Ericsson
Lund, Sweden
hakan1.jonsson@sonyericsson.com

Pierre Nugues, Christofer Bach, Johan Gunnarsson
Computer Science
Lund University
Lund, Sweden
pierre.nugues@cs.lth.se, buffyin@gmail.com,
johan.gunnarsson@gmail.com

Abstract— **This paper reports on the work on a new service using text mining on SMS data: SMSTrends. The service extracts trends in the form of keywords from SMS messages sent and received by ad hoc location-based communities of users. Trends are then presented to the user using a phone widget, which is regularly updated to show the latest trends. This allows the user to see what the user community is texting about, and makes her aware of what is going on in this community.**

Privacy considerations of the service are governed by user expectations and regulations. Brenner and Wang [1] discussed mining of personal communication in operator bit pipes. We expand on this by looking deeper into privacy and regulatory aspects through the specific example of SMSTrends. Especially, the use of adaptive location granularity selection is introduced.

Keywords-text mining; messaging; location; context awareness; collective awareness; privacy

I. INTRODUCTION

Personal communication such as SMS is considered highly private. This combined with privacy and data protection regulations makes it very hard to develop services and applications or do research which require a priori access to large amounts of SMS messages. Examples of such services are text prediction engines and marketing analytics on SMS.

A. Background

The work on the SMSTrends service was started as a research project to extract named entities from SMS messages (SMS). When we discovered the problems of finding or collecting a relevant corpus of SMs to carry out the project, the corpus collection became a topic in itself: Under what conditions are users ready to give others access to their SMS?

As SMS messages are private data exchanged between two parties, a classical approach to corpus collection – automatic gathering from machine-readable documents or transcriptions from printed sources – is not applicable. A first naïve request to our colleagues to hand us their SMs for the sake of science miserably failed. We started the SMSTrends application in an attempt to offer them a benefit to sharing their SMS data. After a small group of users had tried it (about a third of the people asked), few wanted to continue using it unless it was made possible to mark messages as secret, to make sure they were not used by the service. After this feature was introduced, a small group continued to use the service. However, the user group is yet too small to make any conclusions regarding the end user value of the service compared to the cost of the user information, and further studies with larger groups are needed.

B. The Service

The service extracts trends in the form of keywords from SMS messages sent and received by the users of the application. Trends are then presented to a user using a phone widget, which is regularly updated to show the latest trends. This allows her to see what a user community is texting about, and makes her aware of what is going on in this community.

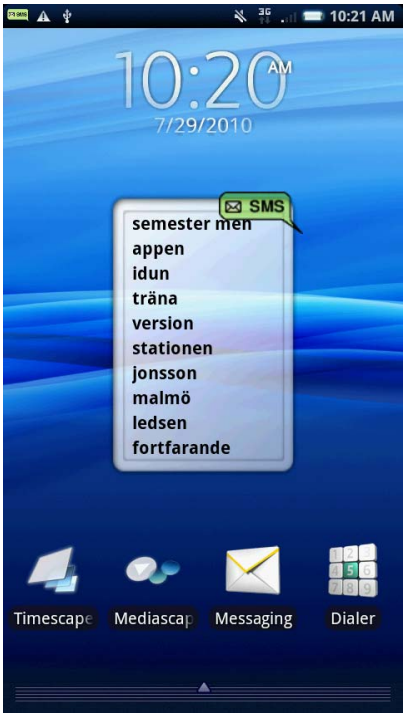

Figure 1. SMSTrends widget screenshot

978-1-4244-7445-5/10/$26.00 ©2010 IEEE

By collecting location information and filtering the trends retrieved by the widget by user location, the list of trending words shown to a user becomes context dependent. When the user is in Stockholm, she sees the currently trending words for Stockholm. When in Berlin, she sees the Berlin trends.

C. User Value

The end user value of SMSTrends consists of getting a contextual up-to-date view of what is going on in an area, as seen through the messages people in the area sent. From interviews with users, this is used in one of two ways. It can either be seen as an ever-changing poem or message about the area. People want to believe that the trends are actually a message and will try to see a pattern in it. They will try to interpret it as a SMS.

The other way is to understand it as up-to-date news indications, very similar to trending topics on Twitter. Users look at Twitter trends because it is entertaining, and it can give immediate notifications about important news. For example, it is often possible to see reports about earth quakes in Twitter trends 15 minutes before they appear on regular news sites.

II. Trend Extraction

We evaluated three methods for extracting trends from messages. Due to the difficulty in obtaining SMS data, the algorithms were developed and evaluated using Twitter, which is sufficiently similar for this type of text mining. For more details on the evaluation of trend extraction methods, see [2].

a) We computed the weight of a word using frequencies of documents containing the word and carried out a regression analysis to identify changes over time.

b) We also used a simpler method based on the mean frequency of documents containing the word over time.

c) The classical vector space approach to term extraction using TF-IDF proved to not work optimally for this service as it identifies static words and not changes over time.

Trends are computed for each location at each granularity level. The top ten trends for the users location are shown in the widget (see section Figure 1).

III. Privacy

Privacy is an important aspect for the user to build trust in a service, especially when data are considered to be highly personal and private.

A. Managing User Expectations

The vast majority of users does not read EULAs, terms and conditions or privacy policies [3]. They just click trough them when signing up for a new service. If the user discovers that information she considers private or sensitive has been disclosed to parties she did not expect, she will be surprised and annoyed. Annoyed users can be dangerous to business in more ways than through legal actions. Detractor influencers [4] can cause widespread defection from a service and generate negative publicity. Thus, the terms and conditions of the service only protect the service provider from a legal perspective, and not from a business or user perspective.

Several social network providers have learned this the hard way [5]. The main lesson to learn from these experiences is: "Don't surprise the user". Managing privacy is about managing expectations of the user, and not about making sure the Terms and Conditions is correct. Neither does it help to provide fine grained controls of privacy control settings if the user is not aware of them or if default privacy control settings are surprising.

Another valuable lesson is that users are willing to share more of sensitive data if they know they can control the access to it, even if they never exert these controls [6].

B. Regulations

Most regulations in the area of privacy and personal information only cover Personally Identifiable Information (PII). The text content of a SMS message only falls under regulations if it contains information that is personally identifiable. SMS may contain PII, and using automatic processing it is in general not possible to know if it does or not.

Within the EU privacy of data and electronic communication is regulated by The Privacy and Electronic Communications Directive [7] and the EU Data Protection Directive 95/46/EC [8]. Each member country implements its own laws and regulations following these directives. The legislations of each member country, e.g. the Swedish Personal Data Act, are sometimes more restrictive that the EU directives.

The EU regulations and directives in general dictate that data collection and processing must be fair, specific and explicit to the user:

(28) Whereas any processing of personal data must be lawful and fair to the individuals concerned; whereas, in particular, the data must be adequate, relevant and not excessive in relation to the purposes for which they are processed; whereas such purposes must be explicit and legitimate and must be determined at the time of collection of the data; whereas the purposes of processing further to collection shall not be incompatible with the purposes as they were originally specified; [8]

Swedish Personal Data Act [9] (compliant with EU Data Protection Directive):

Personal data may […] in principle, only be processed if the registered person gives his or her consent. However, there are several exceptions to this rule, for example, if it is necessary […] in order that a contract with the registered person may be performed.

Specifically for operators, The EU Privacy and Electronic Communications Regulations state that:

Traffic data relating to a subscriber or user may be processed and stored by a provider of a public electronic communications service if such processing and storage are for the purpose of marketing electronic communications services,

or for the provision of value added services to that subscriber or user;

Also, the regulations dictate that the user must be allowed access to data about her, and to be allowed to modify inaccurate data. Also, the user must be allowed to know the logic of any processing done. This has consequences for the design and implementation of services. They must all be implemented in such a way that individual data can be extracted, modified and deleted, and that any processing results based on them must be reprocessed. Also, the algorithms used in the processing must be well documented.

(41) Whereas any person must be able to exercise the right of access to data relating to him which are being processed, in order to verify in particular the accuracy of the data and the lawfulness of the processing; whereas, for the same reasons, every data subject must also have the right to know the logic involved in the automatic processing of data concerning him, at least in the case of the automated decisions referred to in Article 15 (1); whereas this right must not adversely affect trade secrets or intellectual property and in particular the copyright protecting the software; whereas these considerations must not, however, result in the data subject being refused all information;

However, there are exceptions to this rule in case it involves "disproportionate efforts":

(40) Whereas, however, it is not necessary to impose this obligation of the data subject already has the information; whereas, moreover, there will be no such obligation if the recording or disclosure are expressly provided for by law or if the provision of information to the data subject proves impossible or would involve disproportionate efforts, which could be the case where processing is for historical, statistical or scientific purposes; whereas, in this regard, the number of data subjects, the age of the data, and any compensatory measures adopted may be taken into consideration;

Regarding text messages the following is stated in [9]:

Processing of personal data in unstructured material, for example running text, may take place as long as this processing does not entail a violation of the registered person's personal integrity.

In conclusion, regulations allow operators to provide value added services such as SMSTrends if

- the user gives her consent
- the user can access and modify the data
- the user can access documentation about the processing logic
- the data is used only for the agreed purpose

There are exceptions though, that allows processing without additional consent for scientific or statistical purposes:

(29) Whereas the further processing of personal data for historical, statistical or scientific purposes is not generally to be considered incompatible with the purposes for which the data have previously been collected provided that Member States furnish suitable safeguards; whereas these safeguards must in particular rule out the use of the data in support of measures or decisions regarding any particular individual;[8]

SMSTrends could potentially be considered a statistical service, which immediately publishes the statistics, and thus in theory would not need the user's consent if the messages had been collected for other purposes. However, complying with regulations is secondary to managing user expectations, which in most cases would require consent anyhow.

Also, the statistics and research exceptions in the EU directive are an area where member country legislations may be more restrictive. For example, the Swedish Personal Data Act states that collecting personal data for research purposes must be approved by a Scientific Ethics Council to be considered an exception.

IV. Privacy and Features in SMSTrends

SMSTrends was designed to exploit personal communication while protecting the privacy of the users, in order to investigate the privacy issues and regulations encountered. This section covers how SMSTrends was designed to fulfill regulations and user expectations.

A. Terms and Conditions

The terms and conditions for SMSTrends clearly state that data will be collected to provide the service as well as for research purposes. The user has to agree to the terms and conditions to use the service.

B. Filters

By introducing a filter which requires a minimum number of occurrences in unique messages we make sure that any word that is extracted as a trend is not personally identifiable, since the trending words always comes from two independent sources. Thus, words in a single message form a single user can never end up as a trending topic even if there are no other users.

Additionally the following filters are applied:

- Number filter e.g. phone numbers, PIN codes.
- Credit card number filter
- Social security number filter

C. Adaptive Location Granularity Selection

Using location as a context tag to calculate trends creates sparsity problems. Depending on the sending location, we observed large variations in the number of available messages. Calculating trends on few messages results in trend words revealing much of the message content for those messages. To alleviate this, we introduced an adaptive location granularity selection. Coarse-grained location coordinates are identified using Android APIs and a lookup towards SERCPOS, a Sony Ericsson proprietary location database. A lookup is then made towards the GeoNames web service, from which the place name and place hierarchy is retrieved.

Starting at the bottom of the tree, e.g. city, two thresholds on the number of messages sent of different time intervals are

evaluated. If the number of messages is above the thresholds, that location level is used. If not, the next level is checked, and so on until all levels have been checked. Figure 2 shows an aggregated geotree over a subset of messages. Bold locations are eligible for selection, while the others are not, since too few messages are available in those locations.

- **Earth** (2012)
  - **Europe** (1363, 68%)
    - **Sweden** (1363, 100%)
      - Västra Götaland (692, 51%)
        - Sparsör (135, 20%)
        - Skalle (15, 2%)
        - Horred (540, 78%)
      - **Skåne** (574, 42%)
        - Östra Odarslöv (240, 42%)
        - Lunds Kommun (158, 28%)
          - Lund (158, 100%)
        - Värpinge (14, 2%)
        - Håstad (12, 2%)
        - Vallkära (16, 3%)
        - Staffanstorps Kommun (107, 19%)
          - Hjärup (106, 99%)
      - Halland (70, 5%)
        - Ledsgård (70, 100%)
      - Jämtland (22, 2%)
        - Åre (15, 68%)

Figure 2. Aggregated geotree (number of messages, percentage)

D. Message Exclusion

A feature that was explicitly requested by users was the ability to exclude specific messages from being uploaded to the server and used when calculating trends. This was introduced in such a way that users can at any time view the current messages queue, and unmark any messages that is about to be uploaded. This has to be done manually for each SMS and there is no reminder functionality. Messages are uploaded every 15 minutes.

E. Logic Documentation and Data Access

No specific features were implemented to comply with regulations related to modifying or deleting already collected data. Any user requests to delete or modify data would have to be handled by manual database operations.

Since trend extraction is done over limited time intervals there is no need to recalculate trends if a message changed or deleted on user request after the time interval has expired. It is very unlikely that a user would like to change any data related to SMSTrends, especially within the time interval used for trend extraction.

Regarding documentation to comply with the regulations regarding information about "the logic involved in the automatic processing of data", documentation has been produced, including this article.

For a commercially deployed large scale service, a different approach would have to be taken to make it cost efficient. Features for user access to data, including modification and deletion, would have to be built in the system from the start.

Regarding documentation of logic, the regulations are not very clear regarding to which extent the logic has to be detailed. The logic used by the service provider in the data mining logic may be a business secret, and publishing information about it according to the regulations may be in conflict with the business interests of the service provider. Thus, this factor must be taken into account as a cost or risk when designing the service.

F. False Privacy Breach

A particular risk for SMSTrends is that the user sees a word she recently sent in a SMS, which then appears in trending topic list, and thinks that it appeared in the trending topics list only on basis of her communication, thus perceiving it as a breach of privacy. Even if she knows that this is not the case, it may make her uncomfortable. We call this a false privacy breach. If a false privacy breach occurs, it means that the management of user expectations of the service has failed.

Modifying the thresholds used for location granularity selection may reduce the risk of false privacy breaches, but it is not clear that it will actually solve the problem. A user may think that she is somewhat unique in using certain words or expressions, or participating in a specific activity, while this may not be the case.

We did not find any other way to address this problem other than educating the user on how the service works, i.e. making the users aware of the logic documentation dictated by the EU regulations.

V. BUSINESS MODELS

Under what conditions are users ready to give others access to their SMs? The question can also be formulated economically and more generally: What is the cost of buying personal information from a user? From our study, the privacy value of SMS is high, and would thus incur a high price for someone buying that data. The utility value of the SMSTrends application is very low. We need further studies on a larger user base to determine if the value of SMSTrends is enough for the user to allow collection of SMS in the longer term.

In any case, it seems unlikely that users would be willing to pay for a service like SMSTrends. This does not mean it has no value. Similar to several internet companies, collected data is used for improvement of existing services and to create new services, which is usually also stated clearly in terms and conditions and privacy policies. The most common business model applied in these cases is advertisement based.

SMSTrends could utilize an advertisement business model. Using the whole body texts of messages collected from users to serve relevant ads is not likely to be approved of by the user, since it may result in ads being served on obviously private information. Marketing based on PII is not regulated in EU directives, but is often strictly regulated by the member states' laws, e.g. Swedish Personal Data Act [9]. However, by using the trending topics instead, which do not contain PII, this

makes it both allowable and relevant to the user, since she can associate the ads with the trending topics. For each update of the trends list, it is possible to know to what degree each user contributed to each trending topic, and thus how interesting ads related to that topic may be for a specific user. We have not used advertising in SMSTrends.

The value of using collected SMS to improve existing services, create new services or do research is very hard to estimate, even though it is real. For our research project into named entity extraction, it has been very valuable since we did not have access to SMS text in any other way.

VI. FUTURE WORK

Further study is needed on larger user groups, to verify that the measures taken to protect user privacy actually meets the user expectations.

REFERENCES

[1] M. Brenner and D. Wang, "Mining the bit pipes: Discovering and leveraging users' behavior", 13th International Conference on Intelligence in Next Generation Networks, 2009. ICIN 2009.

[2] C. Bach and J. Gunnarsson, "Extraction of trends is SMS text", Master's Thesis Report ISSN: 1650-2884 / LU-CS-EX: 2010-16, Lund University, April 2010.

[3] Y. Bakos, F. Marlotta-Wurgler and D. Trossen, "Does Anyone Read the Fine Print? Testing a Law and Economics Approach to Standard Form Contracts", NYU Law and Economics Research Paper No. 09-40, 2009.

[4] F. Reichheld, "The Ultimate Question: Driving Good Profits and True Growth", Harvard Business School Press, 2006.

[5] B. Slattery, "Google Hits Restart on Buzz Privacy Settings", PC World, April 6th, 2010.

[6] B. Schneier, "Google And Facebook's Privacy Illusion", Forbes.com, April 6th, 2010.

[7] Statutory Instrument 2003 No. 2426, The Privacy and Electronic Communications (EC Directive) Regulations 2003.

[8] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal L 281 , 23/11/1995 P. 0031 - 0050.

[9] Swedish Personal Data Act (SFS 1998:204)