Survey on Mining Feedback from App User Reviews

*** INTRODUCTION ****
Our research group at the University of Zurich, in collaboration with the University of Sannio research group, is performing a study aimed at designing and develop tools to help developers of mobile app digest the users feedback and transform the user reviews of mobile app into development tasks (fixing issues or building features).

Thus, our goal is to define and evaluate a **Feedback Mechanism** that through the mining of user reviews of mobile Apps could _summarize feedback in form of development/maintenance tasks_ such as, fixing issues, building features, improve existing features etc.

We would thank you for participating to this study. Now we would be grateful if you could spend a little bit of your time in analyzing the summary we sent you and answering few questions.

*** SUMMARIES' STRUCTURE ***
Together with this survey, you received via e-mail a HTML file containing a summary of the app user reviews automatically generated by the tool we are currently developing.

Our tool tries to automatically recognize sentences in users' reviews that could help developers in designing and accomplishing specific maintenance and evolution tasks or sentences that could assist developers in understanding users' needs. Each sentence may belong to one of the following semantic categories:

- **BUG**: _sentences describing issues with the app or unexpected behaviors;_
- **REQUEST**: _sentences expressing ideas, suggestions or needs for improving or enhancing the app or its functionalities;_
- **INFO**: _sentences informing or updating users or developers about an aspect related to the app;_
- **QUESTION**: _sentences related to attempts to obtain information or help from other users or developers._

Furthermore **our approach groups sentences** extracted from reviews which are likely to **treat the same topics**. Sentences may belong to one or more of the following **topics:**

- **GUI:** sentences related to the Graphic User Interface or the look and feel of the App
- **APP:** sentences related to the entire app, such as descriptions of generic crashes, rating information, or feedback related to the entire app.
- **CONTENTS:** sentences related to the contents of the application.
- **COMPANY:** sentences containing feedback related to the company/team which developed the app.
- **DOWNLOAD:** sentences reporting feedback about app download
- **FEATURE/FUNCTIONALITY:** sentences related to specific features or functionalities of the application.
- **IMPROVEMENT:** sentences related to explicit requests of improvement.
- **MODEL:** sentences reporting feedback about specific devices or OS versions.
- **PRICING:** sentences related to application prices.
- **RESOURCES:** sentences dealing with devices' resources such as battery consumption, storage, etc.
- **SECURITY:** sentences related to security of the app or personal data privacy.
- **UPDATE/VERSION:** sentences related to specific versions of the app

The summaries are structured for reporting, in each of the above topics, the sentences that are estimated to be the most informative among all the sentences constituting the reviews. Each sentence is also tagged with a label indicating the estimated semantic category (i.e. BUG, REQUEST, INFO, QUESTION) in which it falls.

To make easier the reading of the summary, for each topic are shown firstly all the bugs, then all the requests, then all the questions and finally all the info.

Since one sentence could treat more than one topic, the same sentence may appear more times in the document under different topics. This was a precise design choice to allow multiple developers taking charge of different aspects of the same app.

IMPORTANT: ***By clicking on each of the sentences reported in the summary, a pop-up containing the data of review (i.e. app version, user who wrote the review, review date, star rating, review title and review text)*** from which the specific sentence has been extracted will be shown.

*** LET'S START THE EVALUATION***

To assess the quality of the tool outputs and collect feedback to improve them, we need that you follow the instructions reported in the beginning of the next page answering all the questions contained in the survey.

THE DEADLINE TO COMPLETE THE QUESTIONNAIRES IS February 10th, 2016

Please send all the completed survey to: panichella@ifi.uzh.ch or disorbo@unisannio.it

All of the information that you provide will be treated as confidential and will only be used for research purposes. Some personal information may be collected about you if you choose to participate in the survey. In particular, your responses to survey questions and your personal details will be collected for the simply reason of linking your responses with the summaries we generated. We will not disclose your personal information to third parties.

Thank you very much for your effort.

Dr. Sebastiano Panichella and
Andrea Di Sorbo, M. Sc. in Computer Engineering.

Step1 - EVALUATION OF THE SUMMARY

Survey participant has to write down a report that should contain the list of reviews incorrectly (only the incorrect) classified in the automatically generated summary. For example, if a sentence that talk about a BUG of the GUI is classified as a REQUEST about the GUI, the participant should copy the incorrectly classified sentence and specify which is the real category. The Report can be written as part of this document (If so, please add at the end of this document your report).

Example of report:

"The gui of the app is not well formed for my android mobile, when I click the button POWER, the button itself disappear and in few second the app crash." Classified as: REQUEST in the topic: GUI -> Real category: BUG in the topic: GUI.

->IMPORTANT 1: A user review in a summary can correctly belong to more than one category.
In this example the review is assigned to more than one category as in the following example:
GUI
"[INFO] What I'm really missing: (Auto) Rotation of the photos"
……

FEATURE/FUNCTIONALITY
"[INFO] What I'm really missing: (Auto) Rotation of the photos"

In this example the user give INFORMATION about the GUI and say that is missed the "(Auto) Rotation…"
In such situation the review is correctly assigned to both of the categories GUI and FEATURE/FUNCTIONALITY.

Clearly, in situation in which one or both the assignment are not correct report them in the report.

➔ **IMPORTANT 2:**
➔ *For participants who have to evaluate more than one summary, we need that you compile a survey for each app, indicating the specific summary to which the answers are connected*.

➔ Remember to **sign the time** that was required to validate the reviews (minutes and seconds).

START THE EVALUATION OF THE USER REVIEWS SUMMARIES..

Remember to sign the TIME required to perform the task before answer the post-questionnaire.

Insert the text here

Thank you very much for your effort.

Please, report the time required to perform the validation task and report it in minutes and seconds (e.g. "25 minutes and 45 seconds") before to go in the next step.

Type the text here

Step2 - POST QUESTIONNAIRE:

1) How do you judge the usefulness and comprehensibility of the provided summary? Choose one of the following options.

Select a value

----

2) How difficult is to analyze user feedback contained in reviews WITH the summary? Choose one of the following options.

Select a value

----

3) How difficult is to analyze user feedback contained in reviews WITHOUT the summary? Choose one of the following options.

Select a value

----

4) Proportionally, how much time you can save by analyzing feedback contained in user reviews WITH the proposed summary (if compared with the time required WITHOUT the summary)? A possible answer can be "half of the time", "25% of the time", "nothing".

Type the text here

----

5) Please rank the categories of reviews reported in the summary in order of importance (from 1 to 12 where 1 is most important to you and 12 is least important to you) from a development point of view:

☐ GUI: sentences related to the Graphic User Interface or the look and feel of the App

☐ APP: sentences related to the entire app, such as descriptions of generic crashes, rating information, or feedback related to the entire app

☐ CONTENTS: sentences related to the contents of the application

☐ COMPANY: sentences containing feedback related to the company/team which developed the app

☐ DOWNLOAD: sentences reporting feedback about app download

☐ FEATURE/FUNCTIONALITY: sentences related to specific features or functionalities of the application

☐ IMPROVEMENT: sentences related to explicit requests of improvement

☐ MODEL: sentences reporting feedback about specific devices or OS versions

☐ PRICING: sentences related to application prices

☐ RESOURCES: sentences dealing with devices' resources such as battery consumption, storage, etc.

☐ SECURITY: sentences related to security of the app or personal data privacy

☐ UPDATE/VERSION: sentences related to specific versions of the app

5.1) Please justify the ranking that you provided*:

Type the text here

.

5.2) Are missed potential useful categories in the taxonomy above?

If so write down what is missing in the taxonomy.

Type the text here

.

----

6) WHITOUT the proposed summary, evaluating the user feedback contained in reviews is prohibitively difficult and/or tedious. Choose one of the following options.

Select a value

6.1) Please justify the Answer

Type the text here   .

Content adequacy

7) Considering only the content of the summary of user feedback and not the way it is presented, do you think that the report? Choose one of the following options.

Select a value

7.1) If some important information are missing, can you specify which information are missed

----

Type the text here

.

Conciseness

8) Considering only the content of the summary and not the way it is presented, do you think that the report? Choose one of the following options.

Select a value

8.1) If the summary has some unnecessary information can you specify which information is useless?

Type the text here   .

Expressiveness

9) Considering only the way the summary of user feedback is presented and not its content, do you think that the report? Choose one of the following options.

Select a value

9.1) Please justify the Answer

Type the text here   .

----

10) Do you have any suggestions to improve the summary and make it more understandable?

Type the text here   .

----

11) General question: are the summaries useful for understanding user reviews feedback of mobile apps? Choose one of the following options.

Select a value