Implementation of Anti-Profanity Words in Mobile Application Platform

F Razali*, Z Jusoh1, A Salleh@Omar1 and N Azizan1
1 Faculty of Computing and Multimedia, Kolej-Universiti Poly-Tech MARA, Kuala Lumpur, Malaysia

*Corresponding e-mail: nurfarahbalqishrazali@gmail.com

Abstract. Communication technology increases day by day in our daily life due to the rapid growth of variety mobile application platforms. For instance, mobile application users are using chat rooms for sharing data, pictures, videos and documents. However, the users are exposed to receive something inappropriate such as bad words, bad photos and sensitive data in their social media platforms. The developed chat profanity filtering application can be used to detect and filter the bad words before the sender send any messages via the application. One way to detect and filter profanity words is by using the PurgoMalum web service. Therefore, the recipients are unable to view any bad words sent to the chat application. This research has shown that the developed chat profanity filtering application will minimize the potential of users abusing the chat platform in daily communication. Result from this study had successfully demonstrated that the developed chat profanity filtering application could minimize the potentials for users to abuse the chat platforms in communicating various forms of obscenity. Therefore, chat profanity filtering application was developed to minimize the abusing of social media platform in daily communication.

Keywords: mobile apps, social media, chatting, profanity, PurgoMalum

1. Introduction

Social media platform using mobile application play prominent role in our daily life due to the advancement of the mobile phone technologies and network accessibility. The popularity of social media platform is supported by the rapid growth of internet technology and social media platform. Commonly, the most popular social media platforms such as Facebook, WhatsApp, WeChat and Twitter provide facilities to communicate with other people by the sharing data such as pictures, videos and documents. In contrast, the negative effects of the social media to the people such as their perception to society and the way they view themselves and to the other people. By using the social media, the users tend to send any kind of content without considering whether the content or text is physically threatening or humiliating and not just offensive to the recipients. On top of that, the offensive contents might be affected other people feelings when the recipient received an inappropriate text or messages. Consequently, they might get anxiety and bad feeling which could influent to their behaviour and daily life.

In order to encounter the discourteous communication such as abusive words in text or media content between lecturers and students in Kolej University Poly-Tech MARA (KUPTM), a mobile social media platform was developed to curb the problem. This platform implements a chat profanity filtering application that able to detect and filter the bad words before the sender send any messages to the recipient. The application is integrated with a PurgoMalum web service that being used to detect and
filter the profane words. Moreover, the platform offers the easy and secure ways of communication between the lecturers and students by using smartphones. A secure environment is very important to ensure only authorized users able to use and access this application.

1.1. Problem Statement
Today more than 3 billion people worldwide use a smartphone [1]. They use smartphone not only for communication but also can be used to surf website, take photos, record videos and play games. There are many online applications being used for communication such as Facebook, WhatsApp, WeChat and Twitter. These applications are free and can be used by everyone. The applications offer the fastest way for people to communicate with each other. However, all these applications cannot filter any offensive contents or words from the sender to the recipient. According to Prof Timothy, the professor of psychology at Massachusetts College of Liberal Arts and his assistant Kristin, people who received the abusive words in text or verbal might have the significant impact both emotionally and physically, for instance, they will get anxiety or sleep disorder [1]. Understandably, some of the social media applications do not provide any authentication methods to authenticate the authorized users. However, this feature is very important to ensure the communication is in a secure environment and to prevent from anyone might access the platform and read the confidential message.

1.2. Objective
There are two objectives of this research, namely 1) Develop a mobile social media platform for KUPTM lecturers and students to communicate in a healthy way with filtering profanity. This application provides a feature will be used to ensure the recipient will not get any bad words from the sender. If the sender intends or wants to send offensive contents or bad words, he/she will get the warning notification because the system detects the bad words in the message. 2) Create a unique username and high secure password for authentication purposes before the user is able to use the application.

2. Literature Review
Social Media is a platform that gives individuals the opportunity to interact, using two way communication; meaning, that anyone who has online accounts can share their opinions with other social media users [2]. Hartshorn defines that Social Networking is “the act of engagement,” while social media is the tool used to communicate with mass audiences [3]. In social media, people chat with each other without necessary face-to-face interaction, hence promoting a sense of being liberal without breaking the cultural norms, resulting in an increased in the use of profanity in cyber space. Although most social media platforms have developed profane filters to censor profanity, but the cyber bullies have continuously improved on their profanity techniques to make the existing filters less and less effective. This has led to a major problem in social media, resulting in potentially serious adverse effects on young users if undetected [4].

A study conducted by [5], has found that an online community are often plagued with negative content user-generated content that is negative in tone, hurtful in intent, mean, profane, and/or insulting. However, according to [6], most research works have encountered several challenges in solving the profanity problem in social media. The three main challenges affecting profanity detection are namely; category of language spoken from region to region, types of profanity filters, and techniques used for detecting profanity. Meanwhile, according to [7], the categorization of commercialized profane filters is classified into three categories: blacklist filtering (profane words), free form whitelist filtering (non-profane words), and restricted entry whitelist filtering (text prediction of non-profane words).

Aibelardo Pardo believes that technology offers a platform for innovation, and allows its users to express their opinions about how they feel towards the information being published. He adds that, social media is also a platform that allows students to interact with one another, with their teachers and communities that share their same education. Moreover, Pardo also stated that these types of interaction
are “an essential part of how humans learn” [8]. Mobile Application Development (MDA) refers to the process of making application software for handheld devices such as mobile phones and Personal Digital Assistants (PDA). Through the usage of mobile apps, the user is provided with various features that will enable him to fulfil all his needs and much more. Apps should be interactive to the users. Apps can be downloaded from various platforms such as Google Play Store and iOS App Store [9].

3. Methodology

Systems development life cycle (SDLC) describes a process of planning, analysis, design, implementation, testing & integration and maintenance as shown in Figure 1. This model is easy to understand and has a clear view of an entire project, the tasks involved, estimated costs and project timelines. Furthermore, the advantage of using this model makes the development of the project more efficient in order to achieve the objectives and standards.

![Figure 1. System Development Life Cycle (SDLC) model.](image)

3.1. Planning

The planning phase involves the process required to define the problem statements, objectives, costs needed and the duration of the development process. The sufficient time needs to allocate included conducting extensive interviews with the target users and distribute questionnaires to them. It is better to structure all the incoming data and analyse them. Not only that, all technical limitations that may arise on the user’s side also need to be clarified and come out with a ready-to-follow specification to meet the user requirements.

3.2. Analysis

The analysis phase involves the process to analyse all data gathered from dissemination of questionnaires to the KUPTM lecturers and students. The user requirements such as the features of the apps needs to be a focus on this phase. Hence, all the requirements will be finalized before proceed to the design phase.

3.3. Design

The design phase involves the process to transform the analyse data to system design. Usually the ideas come from the sketches, including the user interface design, database design, system modelling, system components and security level. The whole system will model by using use-case diagram and flowchart
diagram. The advantages of using system modelling are assisting business analysts, project leader, the assistant project leader, the tester, or related organization and also the user that uses the system. Moreover, the model must fulfil all the user requirements and the application design should be user-friendly.

3.4. Implementation
The implementation phase involves the process to write and compile programming codes by using a software Visual Studio Code. JavaScript is a client-side programming language and used as the front-end of the apps. PurgoMalum is a simple, free, RESTful web service for filtering and removing content of profanity, obscenity and other unwanted text. PurgoMalum interface input several parameters for customization and can return results in plain text, XML and JSON [10].

This application will require a unique username and high secure password to log in. Besides, this application will work to detect the bad words when the sender tries to send it. Only registered users can log in into the application. The users need their username and password to log in into the system. An unauthorized person is not allowed to log in into the system.

3.5. Testing and Integration
The testing and integration phase involves the testing process to the whole system where there will be a system checking on the specification and also the user requirements. During this phase, bugs and glitches will be detected so that it can be corrected to ensure the system runs smoothly, able to perform all the functional tasks and meets the user requirements. The system will be tested by the specific users which are KUPTM lecturers and students. They are required to give any feedback and comment after completed the test. The testing phase includes functional and non-functional requirements. After the system passes each testing phase, the system is deployed to the user environment. However, this paper only shows the results of non-functional testing.

3.6. Maintenance
The maintenance phase involves the activities of modifying the system module after it is delivered to the target user, fixed any bugs or coding error and making changes to improve the system’s performance. In reality, it is inevitable there are some defects or issues come up, especially for a new system which leads to modifications to the system in improving the performance and efficiency. Furthermore, regular maintenance and support will be provided to the user.

4. System Design
Data and system modelling is a process of creating a data model for an information system by applying certain formal techniques. Besides, it is also a process of organizing and structuring data to be implemented in software or system development. Data modelling is a process used to illustrate and analyses the data requirements needed to support the processes within the scope of the corresponding information system in an organization. System model represents the view of the whole system in a visual diagram and clearly shows the functionality of the whole system.

4.1. Use-Case Diagram
The use-case diagram can be defined as a representation of a user's interaction with the system. For chat profanity filtering application, there are two actors involved, which are lecturer, student and system.

Based on Figure 2, it shows that lecturers and students are able to sign up and log in to the system. Furthermore, they can view messages, compose messages and send messages. The system is able to detect any type of bad words and display a warning.
4.2. Flowchart Diagram
Flowchart is a type of diagram that represents a workflow or process in sequential order. It can be used to standardize a process for effective analysis, efficiency and to ensure the procedures are being followed.

Figure 3 depicts the flowchart for chat profanity filtering application. The process flow begins with start symbol to indicate that the process is started, then it’s connected to the first task which is a lecturer and student need to sign up to become members of the chat group. Then, they can log into the chat group. After that, they can view messages and compose messages. Before messages being sent to the recipient, the system will detect any bad words. If any bad words are detected in a message, the warning message will be displayed and send button will be disabled automatically. Then, the user has to delete the bad word and again the system will check the message. If there is no bad word detected, so user will able to send the message to the recipient.

5. Results and Discussion
This section discusses the findings and the result of data gathering from dissemination of questionnaires to KUPTM lecturers and students. The questionnaires were distributed to the respondents through WhatsApp and it was made by using Google Form. The feedback is important to thoroughly understand the needs to develop a chat profanity filtering application.

5.1. Result from Questionnaire
The objective of the questionnaire is to obtain new ideas and information about the development of profanity filter application for KUPTM lecturers and students. Two types of questionnaires have been constructed, one for the KUPTM lecturers and another one for KUPTM students as shown in Table 1 and Table 2. There are 30 lecturers and 40 students were responded to the questionnaires.

Based on the survey, both lecturers and students preferred to use smartphone rather than using laptop to communicate with each other. Communication using a smartphone is much more convenient and easier. In addition, they preferred to use mobile application to communicate rather than a website. Both

Figure 2. Use-case diagram for chat profanity filtering application.
of the target users were also agreed to filter bad words on their social media platform, hence to avoid from inconvenience relationship.

Figure 3. Flowchart for chat profanity filtering application.

Table 1. Response from the lecturers.

| Questions                                                                 | Response |
|---------------------------------------------------------------------------|----------|
| 1. Do you think using a mobile social media platform is the fastest way to communicate with students? | Yes      |
|                                                                            | 100%     |
| 2. Which device do you prefer to use when communicate with students?      | Smartphone |
|                                                                            | 97%      |
| 3. Which platform do you prefer to communicate with your students?        | Mobile application |
|                                                                            | 100%     |
| 4. Have you ever received harmful/bad words when communicate with students through online social media platform? | Yes |
|                                                                            | 71%      |
| 5. Sometimes, in group chat/personal chat, students accidentally sent bad words. Do you agree chat application should be able to filter the bad words? | Yes |
|                                                                            | 100%     |
Table 2. Response from the students.

| Questions                                                                 | Response        |
|---------------------------------------------------------------------------|-----------------|
| 1. As a student, how do you prefer to communicate with your lecturer?     | Social Media    |
|                                                                            | 80% Face to face|
| 2. Do you think using a mobile social media platform is the fastest way to communicate with lecturers? | Yes 99% 1% |
| 3. Which device do you prefer to use when communicate with lecturers?     | Mobile application 100% Website 0% |
| 4. Sometimes, in group chat/personal chat, students accidentally sent bad words. Do you agree chat application should be able to filter the bad words? | Yes 48% 52% |

6. Conclusions
This research has shown that the developed chat profanity filtering application will minimize the abusing of social media platform in daily communication. Furthermore, result of this study has successfully demonstrated that the developed chat profanity filtering application could minimize the potentials for users to abuse the social media platforms in communicating various forms of obscenity.

Acknowledgments
The authors gratefully acknowledge the support provided by the supervisor, Puan Noraliza Azizan, the Senior Lecturer of Faculty of Computing and Multimedia (FCOM), Kolej Universiti Poly-Tech MARA (KUPTM). The authors would like to thank all the lecturers from the Faculty of Computing and Multimedia (FCOM) for their contribution, especially Puan Zurina Jusoh and Puan Nor Azura Salleh@Omar.

References
[1] Timothy J and Kristin 2012 The Science of Swearing
[2] Tarek A. El-Badawy and Yasmin Hashem 2015 International Journal of Business Administration
[3] Hartshorn S 2010 Social Media Today 5
[4] Abdulrehan A Mohamed, Dr George O Okeyo and Dr Michael W Kimwele 2018 International Journal of Scientific & Engineering Research 9 (10) 2229-5518
[5] Sood S O, Antin J and Churchill E 2012 Conference on Human Factors in Computing Systems ACM 978-1-4503-1015
[6] Benedictus L 2014
[7] King, M. 2013 Types of Profanity Filters for Online Safety
[8] Pardo A 2013 Int. J. Social Media and Interactive Learning Environments, 1(1), 43-58.
[9] Kishore Baktha 2017 International Journal of Computer Science and Mobile Computing 6 (9) 15-20
[10] Purgomalum.com, 2019 PurgoMalum
[11] Harish Yenala, Ashish Jhanwar, Manoj K Chinnakotla, Jay Goyal 2017 Deep Learning for Detecting Inappropriate Content in Text
[12] K R Subramanian 2017 International Journal of Scientific Progress and Research (IJSPR) 38 (109) ISSN: 2349-4689
[13] W Akram, R Kumar 2017 International Journal of Computer Sciences and Engineering 5(10) E-ISSN: 2347-2693