An Integration System of Communication App on Image Recognition

Wei Wu\textsuperscript{1}, Chen-Li Huang\textsuperscript{2}, Shiu-Tung Chen\textsuperscript{3}, Hung-Min Sun\textsuperscript{2,4}

\textsuperscript{1} Fujian Provincial Key Laboratory of Network Security and Cryptology, School of Mathematics and Computer Science, Fujian Normal University, China
\textsuperscript{2} Department of Computer Science, National Tsing Hua University, Taiwan
\textsuperscript{3} Industrial Technology Research Institute, Taiwan
\textsuperscript{4} Research Center for Information Technology Innovation, Academia Sinica, Taiwan

weiwu@fjnu.edu.cn, yoie782112@livemail.tw, hsuantung@itri.org.tw, hmsun@cs.nthu.edu.tw

Abstract

Recently, mobile phones have become ubiquitous, therefore, there is a huge market for a variety of applications. Among all, communication applications have received a special attraction as they facilitate the communication with family and friends. In order to utilize such a service, users are expected to have the same communication application on their mobile phones. To remedy this shortcoming, we integrate communication applications to facilitate the communication. In this paper, we aim to solve this problem to let users enjoy communication with each other without being worried about the application brand. Moreover, our solution results in less space occupation as users do not demand to install many similar applications on one mobile phone.

Keywords: Shareable keyword search, Multi-user keyword search, Encrypted data

1 Introduction

1.1 Background

The number of mobile users is increasing dramatically where Figure 1, depicts the annual growth rate from 2005 to 2014. Experts expect the mobile phone population reaches a peak by 2018 as in Figure 2. As it is shown in Figure 3, users are engaged in different communication applications which could become inconvenient as users need to install many applications in order to communicate with their friends in different countries. For example, nowadays users are installing and maintaining many applications including LINE, WeChat, FB’s Messenger, Skype, Google Talk etc. in order to keep in touch with each other. However, this strategy not only wastes some of their phone resources but also leaves the user in distress of losing some important messages.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{The global mobile user growth curve from 2005 to 2014 by ITU}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.png}
\caption{Prediction of global mobile user growth curve from 2013 to 2018, 2014 by eMarketer}
\end{figure}

1.2 Features of Communication Application

The statistics in Figure 4 shows that roughly one-third of daily used applications on mobile phone is for communication purpose. This is a non-negligible portion, therefore, any assist in this domain will influence millions of users worldwide. Over years, many attempts have been done in this area to attract the users. We briefly introduce some of these important features.
Diversity communications content. Transmitting text messages and emotions are the most basic means of communication. Several existing applications also offer some other services such as voice, video, photographs, maps information, contact information, etc. Group chat. Also known as “chat room”, a predefined number of contacts form a group are equipped to send messages at the same interface. Support of computer. Besides the mobile version (Android, iOS), a majority of users utilize computers in their offices. Therefore, it would be counted as an advantage if a desktop or a web-based version is prepared to enable the communication on any platform. The amount of resources. Mobile phones are still believed to be resource-intensive. Hence, their resources such as computing, memory, and storage are relatively limited compared to computers. Therefore, there always exists a restriction on services that application can deliver on mobile phones.

1.3 Contribution

In this paper, we aim to facilitate the means of communication via mobile phones. Nowadays, users have to install many communication applications in order to keep in touch with their friends worldwide. We present a novel system that integrates the variety of communication applications into a stream of messages sent to users’ mobile phone through the web application. Therefore, users can enjoy interacting with others via a single interface without being interrupted or switching among applications.

1.4 Organization

We will introduce related products and applications in Section 2. Section 3 and Section 4 describe the structure of the whole system and the use of related technologies, respectively. We will make the relevant experiments and comparison in Section 5. Section 6 summaries the whole paper and prospects for the future.

2 Related Works

2.1 Product Category

The main purpose of this paper is to enable users from different communication applications to interact with each other via a single interface. Generally, nowadays we have three different categories in which the applications can be implemented: Communication App. Communication App is a basic messaging system that can convey the messages, establish contacts. Web App. Web App is mainly a browser-based application on a mobile device. It is also known as mobile web App, which is basically the same App but does not need to be installed in the mobile phones and consumes the minimum resources. We discuss more details in Section 3. Integration App. Integration App collects and displays all the notifications from different communication applications into a single interface. In fact, users are still required to install all the applications, but in a sense, they receive and manage their communications via a single interface.

2.2 Product Introduction

In some existing products, they integrate some fixed applications, but users are not allowed to add a new application. Therefore, they put another barrier for the users to include more applications. These products mainly cooperate with some fixed number of vendors. We introduce some actual products as following:

- IM+ [1]:
  - Feature:
    - Integration of multiple accounts, and support for multi-functional background, Push notifications and log in more than one place at the same time and built in browser.
  - Contained Applications: MSN/Facebook/Yahoo! Messenger/Skype/Twitter/GoogleTalk/AIM/iChat/ICQ/MySpace/Jabber
Support OS: iOS/Android

- Feature:
  IMO belongs to the Web App, uses SSL encryption to increase security and the share a small capacity.
  - Contained Applications:
    MSN/Facebook/GoogleTalk/Yahoo!Messenger/AIM/ICQ/Jabber/VKontakte/Steam
  - Support OS: iOS/Android

- SnowBall [3]:
  - Feature:
    Snowball itself is not a communications application, instead, its main function is to integrate communication application on the phone. Hence, user needs to install the required communication application, and overall capacity.
    - Contained Applications:
      Facebook Messenger/WhatsApp/Snapchat/GoogleTalk/Hangouts/Twitter/LINE/WeChat/SMS/Slack
    - Support OS: Android

- Ebuddy [4]:
  - Feature:
    Ebuddy is suitable for a temporary use and designed to work with restricted user network environment. EBuddy let the user break through the firewall restrictions and also can be used without downloading the software.
    - Contained Applications:
      MSN/Yahoo!Messenger/ICQ/AIM/GoogleTalk/Facebook/MySpace
    - Support OS: iOS/Android

As shown in Table 1, none of the integrating apps can include all the communications applications. Moreover, they do not let the users add their favorite applications. In addition, Table 2 indicates that some of them are suitable for both iOS and Android platforms.

<table>
<thead>
<tr>
<th>Table 1. Table of integration APP’s contained communication App</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB’Msg</td>
</tr>
<tr>
<td>IM+</td>
</tr>
<tr>
<td>IMO</td>
</tr>
<tr>
<td>SnowBall</td>
</tr>
<tr>
<td>Ebuddy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Table of integration APP’s support OS</th>
</tr>
</thead>
<tbody>
<tr>
<td>iOS</td>
</tr>
<tr>
<td>IM+</td>
</tr>
<tr>
<td>IMO</td>
</tr>
<tr>
<td>SnowBall</td>
</tr>
<tr>
<td>Ebuddy</td>
</tr>
</tbody>
</table>

Therefore, in this paper, we proposed a novel system that not only accommodates all communication applications but also integrates new communication applications. Meanwhile, our solution enjoys less resource consumption.

3 Design

3.1 Design Goal

We propose a new integrated system for two main purposes, namely scalability, and reducing capacity. We will discuss scalability in more detail in Section 5. We now briefly introduce the intention of these two purposes:

- **Scalability.** Users are allowed to include their favorite communication applications without restriction of application’s vendors under this system.
- **Reducing capacity.** No need to be installed on the mobile phone. Hence, it consumes fewer resources, and the effectiveness of the phone can be increased.

3.2 The Structure of the System

Our system consists of 4 parts, namely, reception, backstage, cloud storage, and interface. The reception is directly operated by the user. The backstage simplifies the local operations and integrates actions and all messages from diverse communication applications. The cloud storage is used to store all the information. The interface is used to link the reception and the backstage. The whole communication among these four parts is through a secure channel, which is to ensure the confidentiality and privacy of user’s communication.

Reception. The reception provides user with operations such as: authentication, adjust the settings, add contacts, convey messages, we describe these operations as following (Figure 5):

![Figure 5. Reception](image)
(2) Adjust setting: Users according to their preferences can set the communication interval, font size, color, background, etc.

(3) Add contacts: Users are allowed to add contacts. In doing so, users only need to add the new contact into a communication application, and the backstage updates the contact list automatically upon arrival of the next message.

(4) Convey message: Choose the contact from the contact list and type the message content, or reply directly by tapping a contact from the package, as shown in Figure 6.

![Figure 6. Message package](image)

**Backstage.** The backstage integrates communications applications, in addition to information collection. The first step is to set up the environment, then the communication applications that users desire are added to the group. Moreover, captor helps to capture the message from the screenshot. Finally, the integrator collects all messages at the interval and transmits them to the cloud storage. We introduce the main operations are as following:

1. Set up environment: First, we require developing the simulation environment for the smartphone on the private host. There are some existing tools such as Andy [5], Bluestacks [6], Droid4X [7], DuOS [8], Genymotion [9] and Windroye [10]. In order to enable the user to include the desired communication application, mobile phone simulation environment is essential. Besides, we can also use a virtual environment to add multiple accounts.

2. Add communication App: Once the environment is available, users need to down-load and install their desired applications. Meanwhile, they need to sign in to each application in order to receive and transmit messages to their friends.

3. Capture the image: When the desired communications applications installed, if there is a message delivering, then the captor will capture the information on the screen, which contains information source, contacts and message content (Figure 7).

4. Image processing: The image from the captor cannot use directly, so integrator needs to do some image processing as following (Figure 8):

   - Optical Character Recognition:
     OCR (Optical Character Recognition) is used to process the contact’s name and the message component. Therefore, we leverage [17] and [18] as an image recognition library to retrieve the message sender information.

   - Chinese Optical Character Recognition:
     Similarly, in order to recognize Chinese messages, we use [17] and [18] as our preferred image recognition libraries.

   - Image Recognition:
     We apply Image Recognition to the source image, like communication application’s image.

5. Collect and arrange: After an interval of time, we organize the information and sort the message by time stamp. Finally, we collect all the information into a package.

6. Send and receive: The backstage uploads the contacts list and the messages to the cloud storage. Furthermore, it collects the message content, the contact name and the communication application’s name to help the user to receive and send the right message via the appropriate application.

7. Renew: When a user adds a new contact, the contact’s information is transmitted from the phone to the backstage. Later, the backstage renews the contact

![Figure 7. Format of the message](image)

![Figure 8. The captor and the integrator](image)
list and sends the information to the cloud storage to maintain the updated contact list.

**Cloud storage.** The cloud storage uses the Data Storage Cloud technology to store the historical message and contact list. Therefore users can update the whole information, and consequently, reduce the amount of data which the users have to store. Whenever a user requires accessing the information, the user needs to go through an authenticating step. As one can see in Figure 9.

![Figure 9. The content of the cloud storage](image)

Figure 9. The content of the cloud storage

Cloud storage uses the Data Storage Cloud technology to store the historical message and contact list. Therefore users can update the whole information, and consequently, reduce the amount of data which the users have to store. Whenever a user requires accessing the information, the user needs to go through an authenticating step. As one can see in Figure 9.

Figure 9. The content of the cloud storage

**Interface.** We use the Web App technology [15] to design the system interface. Here are the reasons we choose such app to develop the interface.

1) Excellent cross-platform: Since mobile phones have different operating systems and also we are not sure about the market in the future, we would rather use the cross-platform technology. Moreover, sometimes due to the limitations in developing phase, there are some difficulties in implementing our system on various platforms.

2) Development of low threshold: The potential of using the existing standard technology in future is very limited. Many experts believe that web-based application has more chance to be seen in future.

3) Simple development environment: After developing a web application, we only need to install a browser to reach the minimum requirements. Generally, web applications do not need to go through an intensive test environment before delivering to the end users.

4) Without update: In a case of any changes on any functions of a web application, these changes only need to be applied to the backstage side and end users are not aware of this layer. Therefore, the changes can be applied with minimum cost and users do not need to apply any changes.

5) A brief definition of the Web App: Use Embedded Browser to open web pages by applications, and using Web technology implementations for mobile device applications.

Most Web developers choose the Web App, because the programming language is easier than others. Take HTML5 as an example, we can see from Figure 11 that almost every major browsers support HTML5, no matter in windows or mac.

![Figure 11. The browsers support HTML5 or not by HTML5 test](image)

Figure 11. The browsers support HTML5 or not by HTML5 test

From Figure 11, it is clear that cross-platform capability is extremely powerful and supported by almost all the major browsers.

We divided the basic architecture of the Web App into the following four parts: Application, Web View Component, JavaScript Bridge and OS Application (Figure 12):

![Figure 12. The structure of the Web App](image)

Figure 12. The structure of the Web App
(1) Application: Application on the mobile phone’s OS is the main basis of the Web App, which can directly operate on the OS application.

(2) Web view component: Web View Component is a web page, which is developed by using web language so that the user can operate and browse on the page by opening a browser.

(3) JavaScript bridge: JavaScript Bridge is a link between the web and the phone. Hence, a web page can also use OS application which is a special technology for the Web App, where we can take the JavaScript bridge in [16] as an example.

(4) OS application: OS Application is one of the basic phone applications such as device vibration, gyroscopes, accelerators and phonebook.

In this system, we apply the Web App technology to build the front-end and the back-end. The front-end is a page and a user can operate at the back-end, which is a connection to the cloud storage as shown in Figure 13.

4 Implementation

4.1 Basic Model

Figure 14 shows four parts as described in Section 3: Reception, Backstage, Cloud Storage and interface. The Reception is operated by the user and the Backstage is responsible for operations of the phone simulation environment. Cloud Storage acts as a database to store a massive amount of information. Web App is the interface connecting the front-end and back-end. In this section, we describe the connection between four parts by the operations.

4.2 Operation

In this system, we have four main operations such as setup, add contacts, notice and convey the messages. We introduce each operation as the following:

**Setup.** The purpose of Setup is to set the background environment. The steps are shown in Figure 15 and described as following:

- Steps:
  1. Establish mobile simulation environment, and set up the relevant operating system (Android / iOS).
  2. Download the desired communication Apps.
  3. Users log into communication Apps, and stay connected smoothly.
  4. Back up the contact list and store them into the cloud storage.
  5. Users send the request to the interface to get the contact list after authentication.
  6. Send the contact list to the mobile devices and finally, users can start the communication.

**Add Contacts.** The purpose of Add Contact is to add a new contact to the list and refresh the contact list on the cloud storage. The steps are shown in Figure 16:
**Figure 16.** The operation of Add contacts

- **Steps:**
  1. Send the request and the corresponding contact to the backstage via log into the web through the browser.
  2. Add the new contact to the contact list.
  3. Refresh the data on the cloud storage.
  4. Send the new contact list to the mobile devices.

**Notice.** The purpose of the notice is to aware the user who has received a message. Steps are shown in Figure 17, and then described as following:

**Figure 17.** The operation of notice

- **Steps:**
  1. Receive the message from a contact with the simulation environment.
  2. Captor captures the image from the screen on the backstage.
  3. Integrator recognizes the message from the image and indicates the data type.
  4. Store the message on the cloud storage.
  5. Send a notification to the user.

**Convey Message.** The purpose of Convey Message is to enable users to send the message to the contact. The steps are shown in Figure 18 and described as following:

**Figure 18.** The operation of convey message

- **Steps:**
  1. Send the request and the message to the backstage and log into the web by a browser.
  2. Convey the message to the contact.
  3. Store the message on the cloud storage.

5. **Analysis**

In this section, we analysis our system for three major goals: scalability, efficiency, security.

5.1 **Scalability**

Scalability mainly refers to the number of the communication applications that can be supported by an integration system. For comparison, we utilize the integrators discussed in Section 3. As one can see from Table 3, the scalability of the proposed system is better than any other integration software.

**Table 3.** Table of integration APP’s contained communication APP

<table>
<thead>
<tr>
<th></th>
<th>FB*Msg</th>
<th>LINE</th>
<th>Yahoo’s Msg</th>
<th>WeChat</th>
<th>GTalk</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM+</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IMO</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SnowBall</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ebuddy</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Proposed</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

5.2 **Efficiency**

For the system discussed in this article, there is no
need to download any additional mobile applications. Users only need to connect to the interface of the system by a browser and then start to communicate. Therefore, we transfer the burden from the shoulders of mobile users to the cloud center.

5.3 Security

In this system, users store and transmit private information that needs to be protected, so we must provide the corresponding safety precautions. **Dangerous.** Under this system, we use the Web App interface. Moreover, we use a browser to connect to the web pages. A user in a public wireless network may face several critical situations as described below:

1. **Man In The Middle (MITM):** MITM is an indirect invasion attack which is controlled by the intruder between two communicating computers connected to the network. This computer is called the “middle people” (Figure 19).

![Figure 19. Man in the middle (MITM)](image)

2. **WebView JavaScript Vulnerability:** It is used in JavaScript and executes Java surface operations. Once a vulnerability is exploited, it may threaten users’ safety. The vulnerability can be located in any part of the system, such as an SD card reader, etc.

3. **Remote Code Execution (CVE-2012-6636) [19]:** RCE risk exists in all around the add Javascript Interface, which intends to create a “bridge” for the inter-loaded pages and native programs. It uses a predefined interface that allows a web page to call a specified public function and obtain results.

![Figure 20. Example of remote code execution](image)

4. **Same-Origin Policy bypass (CVE-2014-6041) [20]:** In order to prevent safety problems caused by the page while loading external resources, the browser needs to implement some policies to restrict the interaction of homologous code and resources among different domains (Figure 21).

![Figure 21. Example of same-origin policy](image)

**Precaution.** Due to vulnerabilities in the programming language as well as man-made malicious attacks, in order to protect the privacy of individuals, and reduce the risk of being attacked, we equip our system with several measures as following:

1. **Wireless Encryption:** Under this system, we choose Wi-Fi Protected Access 2 (WPA2) [11] and Advanced Encryption Standard (AES) encryption mode (Figure 22) [12].

![Figure 22. Speed of wireless encryptions, by Techbang](image)

- Wi-Fi Protected Access (WPA2): WPA2 has been replaced with WPA. WPA2 requires testing and certification by the Wi-Fi Alliance, implements the mandatory elements of IEEE 802.11i. In particular, it introduces CCMP, a new AES-based encryption mode with strong security. Certification began in September 2004. From March 13, 2006, WPA2 certification is mandatory for all new devices to bear the Wi-Fi trademark.
Advanced Encryption Standard (AES): AES is a specification for encrypting electronic data. It has been adopted by the U.S. government and is now used worldwide. The algorithm described by AES is a symmetric-key algorithm, meaning the same key is used for both encrypting and decrypting the data.

(2) Update: We operate this system in the mobile phone OS and use a web browser. Hence, the vulnerabilities mentioned above cannot be avoided. Therefore, OS and browser need to be updated continuously in order to achieve the desired security level. For instance, in the Android operating system, we can see that the newest update can prevent some vulnerabilities mentioned in Figure 23.

<table>
<thead>
<tr>
<th>Android</th>
<th>CVE2012-6636</th>
<th>CVE2014-3939</th>
<th>CVE2014-6001</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.X</td>
<td>danger</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3.X</td>
<td>danger</td>
<td>danger</td>
<td>–</td>
</tr>
<tr>
<td>4.0.X</td>
<td>danger</td>
<td>danger</td>
<td>danger</td>
</tr>
<tr>
<td>4.1.X</td>
<td>danger</td>
<td>danger</td>
<td>danger</td>
</tr>
<tr>
<td>4.2.X</td>
<td>safe</td>
<td>safe</td>
<td>danger</td>
</tr>
<tr>
<td>4.3.X</td>
<td>safe</td>
<td>safe</td>
<td>danger</td>
</tr>
<tr>
<td>4.4.X</td>
<td>safe</td>
<td>safe</td>
<td>safe</td>
</tr>
</tbody>
</table>

Figure 23. Versions of Android system prevent vulnerabilities or not

6 Conclusion and Future Work

6.1 Conclusion

In this work, we use image recognition technology to crawl down all messages. In order to achieve the scalability of the system, we utilize web-based application. Moreover, to reduce the storage costs, we leverage cloud storage technology. This provides a lighter experience for the mobile users with the same functionality even if they install several applications on their mobile phones.

In our design, the operations are hidden from users and executed in the backstage. Therefore, users can enjoy higher efficiency and freely include several communication applications.

6.2 Future Work

We hope to be able to use mobile computing technology [13] to move the computation to the background and achieve full automation. By doing so, one can reduce the local costs by having the same functionality.

Acknowledgment

This study was partially supported by the Ministry of Science and Technology, Taiwan, under the Grants MOST 104-2221-E-007-071-MY3 and MOST 106-3114-E-007-006.

References

Biographies

Wei Wu received her Ph.D. degree from the School of Computer Science and Software Engineering, University of Wollongong, Australia, in 2011. Her research focus is on public key cryptography and its applications. She has published more than 40 papers in refereed international journals and conferences.

Chen-Li Huang received his M.E. degree in Department of Computer Science from National Tsing Hua University in 2015. He works as a software engineer in ASUSTeK Computer Inc. since 2015. His research interests include graphic recognition, information security and communication software.

Shiuan-Tung Chen received his Ph.D. degree in Department of Computer Science from National Tsing Hua University in 2016. He works as a researcher in Information and Communications Research Laboratories at Industrial Technology Research Institute since 2016. His research interests include forward error correction, reliable data transmission and applied cryptography.

Hung-Min Sun received his Ph.D. degree in computer science and information engineering from National Chiao-Tung University in 1995. Currently, he is working as the Director of Institute of Information Systems and Applications, and a full professor with the Department of Computer Science, National Tsing Hua University.