UNIVERSITI PUTRA MALAYSIA

AUTOMATED DATA PROCESS IN PARTICIPATORY SENSING USING QR-CODE AND EAN-13 BARCODE

MOHAMAD FAKHRUL SYAFIQ BIN CHE YA

FSKTM 2018 30

AUTOMATED DATA PROCESS IN PARTICIPATORY SENSING USING QR-CODE AND EAN-13 BARCODE

MOHAMAD FAKHRUL SYAFIQ BIN CHE YA

MASTER OF INFORMATION SECURITY

UNIVERSITI PUTRA MALAYSIA

2018

MOHAMAD FAKHRUL SYAFIQ MASTER OF INFORMATION SECURITY 2018

AUTOMATED DATA PROCESS IN PARTICIPATORY SENSING USING QR-CODE AND EAN-13 BARCODE

By

MOHAMAD FAKHRUL SYAFIQ BIN CHE YA

Thesis submitted to the School of Graduate Studies,

Universiti Putra Malaysia,

in Fulfilment of the Requirements for the Master of Information Security

JANUARY 2018

DEDICATIONS

To my parents, project supervisor, lecturers, friends and internet. Thank you for making this possible.

ABSTRACT

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Information Security

AUTOMATED DATA PROCESS IN PARTICIPATORY SENSING USING QR-CODE AND EAN-13 BARCODE

By

MOHAMAD FAKHRUL SYAFIQ BIN CHE YA

January 2018

Supervisor: Dr. Sharifah Binti Md. Yasin

Faculty: Faculty of Computer Science and Information Technology

Abstract:

Advancement of digital technology nowadays has led to the creation of various type of mobile devices such as smartphone, tablet, phablet, computer and many more. Internet also is one of an important element to either connecting people or spreading of an information. This contributes to the creation large amount of data or information such as big data. Big data is a phrase for huge data sets having large, more variety and complicated element with the challenges of storing, analyzing and visualizing for further actions and obtaining the results. However, maintaining data integrity for specific item or information is always being a challenge. In this paper, Quick Response Code (QR code) and EAN-13 barcode was used to enhancing the previous work. The QR code was used as a mechanism to activating the function for mobile application and determining the location, while EAN-13 barcode was used as a product identification. Both mechanism was used to maintain data integrity between the prices

corresponding to the product. Thus, correct and updated crowdsourced data are stored in the database are based on real-time data and location that was submitted by the user or known as crowdsourcer or crowdworker for this work. The enhanced algorithm was evaluated using a developed prototype which is an Android mobile application of a crowdsourcing data submission based on product price and information, WE+Price, in which, the algorithm was embedded. The results showed that the algorithm was able to preserving data integrity with 99.13% and up to 100% accuracy.

ABSTRAK

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Sarjana Keselamatan Maklumat

PROSES DATA AUTOMASI DALAM PENGINDERAAN PENYERTAAN MENGGUNAKAN KOD-QR DAN KOD BAR EAN-13

Oleh

MOHAMAD FAKHRUL SYAFIQ BIN CHE YA

Januari 2018

Penyelia: Dr. Sharifah Binti Md. Yasin

Fakulti: Fakulti Sains Komputer dan Teknologi Maklumat

Abstrak:

Kemajuan teknologi digital masa kini telah menyumbang kepada penghasilan pelbagai jenis peranti mudah alih seperti telefon pintar, tablet, phablet, komputer dan lain-lain. Internet juga merupakan salah satu elemen penting dalam menghubungkan manusia atau menyebarkan maklumat. Ini menyumbang kepada penghasilan data atau maklumat yang banyak seperti data raya. Data raya adalah frasa untuk set data dalam jumlah besar, kepelbagaian variasi dan elemen rumit dengan cabaran seperti menyimpan, menganalisa dan visualisasi untuk tindakan selanjutnya dan memperoleh keputusan. Walau bagaimanapun, mengekalkan integriti data untuk perkara atau maklumat tertentu sentiasa menyumbang kepada masalah. Dalam kertas ini, Kod Tindak Balas Pantas (Kod-QR) dan kod bar EAN-13 telah digunakan untuk meningkatkan hasil kerja sebelumnya. Kod-QR digunakan sebagai mekanisma untuk mengaktifkan fungsi dalam aplikasi mudah alih dan menentukan lokasi, sementara kod bar EAN-13 digunakan sebagai pengenalan produk. Kedua-dua mekanisme digunakan untuk mengekalkan integriti data antara harga yang tepat untuk produk yang betul. Oleh itu, data bersumber khalayak yang tepat dan dikemaskini yang disimpan di dalam

pangkalan data adalah berdasarkan data dan lokasi secara masa nyata yang dihantar oleh pengguna dikenali sebagai penyumbang khalayak atau pekerja khalayak dalam kerja ini. Algoritma yang dipertingkatkan telah dinilai menggunakan prototaip yang dibangunkan iaitu aplikasi mudah alih Android untuk penghantaran data secara sumber khalayak berdasarkan harga dan maklumat produk iaitu WE+Price, di mana algoritma itu tertanam. Keputusan menunjukkan bahawa algoritma berupaya untuk mengekalkan integriti data dengan kadar 99.13% sehingga 100%.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents for their countless supports in my journey to complete my years as a Master student.

Notably, I would like to express my appreciation to my supervisor, Dr. Sharifah Binti Md. Yasin who have been guiding and assisting me in my journey. Thank you for all the knowledge, experience, advice, teachings and acquaintance that have been enlightened me during the research period. Indeed, your encouragement and expertise have shaped and trained me in these research works.

I also would like to extend my thankful to participants who have involved as prototype tester during the data collection phase for this project.

Not to forget, special thanks to all lecturers and students of Universiti Putra Malaysia and my friends who involved directly and indirectly in this work. It has been a wonderful moment to learn and share knowledge with many people during the research period. With their help, opinions and expertise, all of these have been made possible.

Lastly, thank you to all internet resources especially from Google, Stack Overflow and Simplified Coding for all of useful information especially during programming work for prototype development phase of this project. With their help, programming become more interesting and easy.

APPROVAL

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Information Security. The members of the Supervisory Committee were as follows:

DR. SHARIFAH BINTI MD. YASIN

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Supervisor)

Date: 26 January 2018

DECLARATION

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials asstated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012.

Signature: _____ Date: _____

Name and Matric No.: **MOHAMAD FAKHRUL SYAFIQ BIN CHE YA**

GS47423

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Big data is a phrase for huge data sets having large, more variety and complicated element with the challenges of storing, analyzing and visualizing for further actions and obtaining the (Sagiroglu & Sinanc, 2013). Throughout this paper, we deal with price information as a data. An automated data process is a way to manage uncountable data in a system with very least human supervision. Retrieval of data from crowdsourcing activity from various source will contribute to data collection as in (Syafiq et al., 2016) anyone can participate to supply the data into the system. This problem lead to the inconsistency of the accuracy. According to (Howe, 2006), crowdsourcing is the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call. Crowdsourcing activities are expanding to mobile platforms with the expectation of getting larger crowd. This process led to the existence of mobile crowdsourcing. Mobile crowdsourcing is a form of crowdsourcing where tasks are being advertised and submitted through mobile crowdsourcing application (MCA) installed in mobile devices (Väätäjä, Vainio, & Sirkkunen, 2012). Mobile crowdsourcing from crowd's contribution has two types which are participatory sensing and opportunistic sensing. In participatory sensing involvement, the crowd manually compute or generate the data as the input. The data is then submitted using the applications through manual key-in information. Different with opportunistic sensing, the data is being generated automatically by the sensors that

attach to the mobile devices (Chatzimilioudis, Konstantinidis, Laoudias, & Zeinalipour-Yazti, 2012) as in such as global positioning system (GPS), gyroscope and accelerometer. To improve the accuracy of similarity of data between the original source and data collected from crowdsourcing activity, we enhance it by applying a security mechanism which are Quick Response code (QR code) and EAN-13 barcode. The use of QR code will grant an access for the crowd to enter the system while EAN-13 barcode will automatically determines the specific items before an information can be submitted. Both of the codes need to be scanning using the mobile applications in order for the crowd to submit the data.

To improve the accuracy of similarity of data between the original source and data collected from crowdsourcing activity, it been enhanced by applying a security mechanism which are Quick Response code (QR code) and EAN-13 barcode. The use of QR code will grant an access for the crowd to enter the system while EAN-13 barcode will automatically determines the specific items before an information can be submitted. Both of QR code and EAN-13 product barcode need to be scanning using the mobile applications in order for the crowd to submit the data. However, only a crowd with mobile device with built-in camera will able to scan the QR code and then perform next actions. This is to avoid from inaccurate, wrong and tampered data since the data may come from all types of society with different background (Syafiq et al., 2016). After pass the first step, the crowd will scan the barcode of the item and the application will automatically determine type of item that being scanned. Lastly, the information of scanned items will be displayed and the crowd will enter and submit the data to the system to be processed based on the specific criteria as explained on the literature review section. The application of these techniques will improves the

similarity of accuracy of the information after being processed automatically in the system.

1.2 PROBLEM STATEMENT

In a today's world, people are depending on getting information through the internet since it fast and reliable. To be more exactly, the usage of mobile devices together with several application, the process of information delivery and retrieval become more faster. Some of the examples are useful application that deals with data collection are Facebook, TripAdvisor, Waze, and etc. The process of updating an information is normally done by the individual or when the information being shared widely, the process of spreading it done by the crowd. The source of information is known as crowdsource or crowdsourcing. Some of the challenges for mobile crowdsourcing application for data collection are data erroneous for the existing work as in (Syafiq et al., 2016) where users can potentially enter wrong data by mistake such as human error. In [2], the data involved is a price information of product. The major human error that can be explained from previous work is when the users submitting a price information for wrong product. As example, when the user submit a price information of product A to product B. This cause product B has received a wrong price information by mistake. This might affect the data accuracy between the actual price and the price recorded by the system from user's submission. As well it may tampered the data integrity since the price information must be correct corresponding to the actual product because this information will be checked and compared by the crowd as the purpose to obtain an updated information. The suggested algorithm also does not have the mechanism to deal with this kind of erroneous data.

Another issue from previous work are the application can be accessible by a user at any place with an internet connection. This cause a data submission can take into place, thus the data validity cannot be guaranteed since the originating source were doubted due to the data was not in a real-time state considering that the data must be submitted based on where and when the data was displayed. Literally, to ensure the data was arrived in real-time state, the data which is price information must be submitted from where it was displayed which is supermarket location. The price information usually been displayed on the price tag at the product rack in the supermarket. This reflect to the previous research study that the data which is price of an item may change daily.

Moreover, every supermarket has their own operation hours, as example from 10 a.m till 10 p.m daily. Instead a price information should be arrived from where it supposed to be which is the supermarket location, it also must arrived during the supermarket operation hours to ensure the correctness of the data. So, there is necessity to ensure the data validity is at real-time state.

For this work, we will implementing a mechanisms that could improve the data recordation process application based on the previous research.

1.3 RESEARCH OBJECTIVE

The aim of this research is to improve the data collection process based on the previous work by developing an Android based mobile application. In order to achieve this goal, the following are the objectives of this research:

i. To maintain the data integrity between the product details and its data which is price based on EAN-13 barcode identification.

ii. To improve the way how the user can access the application by using QR-code activation.

1.4 RESEARCH SCOPE

This research will focus on improving the technique of automated data collection, process and store in a database by using EAN-13 product barcode identification, QR code as method for application activation and participatory sensing as crowdsourcing method for collection of data. To test those proposed techniques, an Android mobile application will be developed. However, since the data collection involving price information of a product, during testing process, a numbers of random application tester and also known as crowdworker or crowdsourcer will be hired to collect the price information using the mobile application. Since this work is not mainly aimed to improve the data accuracy like previous work, we will only covered one supermarket with radius of 5 kilometer from the research location which is Faculty of Computer Science and Information Technology (FSKTM), Universiti Putra Malaysia (UPM) Serdang, Selangor.

1.5 RESEARCH SIGNIFICANCE

We are living in a digital world that requires us to depend on a technology to gain an information around us. In another way of explanation, we can sit back and relax while using a smartphone to do anything such as online shopping, reading a news, book a flight ticket, gathering some information and so on. The high dependencies of the human beings toward the technologies has influence us to come out with an application that is able to record, process and store the price information from various

supermarket. All the process will be done automatically by the application itself. We are also aimed to produce a smart consumer community such as the user is able to compare the best price of specific item from a various supermarket. For the current work, we are focusing on automated data process part before we can go further into other elements such as comparing the price of specific item with others supermarket. Therefore, people lives becoming more intelligent as technology progresses.

1.6 RESEARCH LIMITATION

Due to time constraint, the data collection process period will be conducted only for one month and taking the research objecives into consideration, it will only involving one supermarket and limited to 10 selected items. Furthermore, the data collection which is price information of product will only referred to the displayed price tags on the product shelf.

1.7 THESIS STRUCTURE

The summary of thesis structure was shown below:

Chapter 1 – Briefly describes about introduction, problem statement, objectives, research scope, research significance, research methodology and research limitation in conducting a research work on the Automated Data Process in Participatory Sensing using QR Code and EAN-13 Barcode.

Chapter 2 – This chapter focused on extensive literature review from relevant publications to understand more about the perspective of crowdsourcing, QR code mechanism, barcode types, data collection and management processing, and

comparison of existing mobile application that having QR code and barcode as part of their features that available at digital market such as Google PlayStore and Apple AppStore.

Chapter 3 – This section covers a full phase of methodology that will be using throughout this research such as survey, construction of algorithm, prototype development, prototype testing for data collection and analysis of result.

Chapter 4 – Further details on the proposed features of planned algorithms that will be construct and deploy into the mobile application. It also covers the technical requirements and specifications that will be needed in order to develop the application such as tools, platform and cloud computing storage.

Chapter 5 – The details analysis of result after obtaining real-time data from the crowdworker by using the developed application. The process including the final result were elaborated in this chapter.

Chapter 6 - As the final chapter for the thesis, the summary of research works will be elaborated here.

REFERENCES

8 Crowdsourcing apps (besides OpenSignal) we love. (2015). Retrieved from https://opensignal.com/blog/2015/07/09/8-crowdsourcing-apps-besides-opensignals-love/

About oBike. (n.d.). Retrieved from https://www.o.bike/my/about/

Anagnostopoulos, I., Zeadally, S., & Exposito, E. (2016). Handling big data: research challenges and future directions. *The Journal of Supercomputing*, *72*(4), 1494–1516. https://doi.org/10.1007/s11227-016-1677-z

Android Studio. (2017a). Run Apps on the Android Emulator. Retrieved from https://developer.android.com/studio/run/emulator.html

Android Studio. (2017b). SDK Platform Release Notes. Retrieved from https://developer.android.com/studio/releases/platforms.html

Bitnami. (2017). LAMP Overview. Retrieved from https://bitnami.com/stack/lamp

Chatzimilioudis, G., Konstantinidis, A., Laoudias, C., & Zeinalipour-Yazti, D. (2012). Crowdsourcing with smartphones. *IEEE Internet Computing*, *16*(5), 36–44. https://doi.org/10.1109/MIC.2012.70

Church, K., & de Oliveira, R. (2013). What's up with whatsapp?: comparing mobile instant messaging behaviors with traditional SMS. *15th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'13)*, 352–361. https://doi.org/10.1145/2493190.2493225

EAN Code. (2017). Retrieved from https://www.scandit.com/products/barcode-scanner/symbologies/ean-code/

Farkas, K., Nagy, A., Tomas, T., Szabó, R., Nagy, Z., & Tom, T. (2014). Participatory sensing based real-time public transport information service. *2014 IEEE International Conference on Pervasive Computing and Communications Demonstrations*, 141–144. https://doi.org/10.1109/PerComW.2014.6815181

Garcia-Molina, H., Joglekar, M., Marcus, A., Parameswaran, A., & Verroios, V. (2016). Challenges in Data Crowdsourcing. *IEEE Transactions on Knowledge and Data Engineering*, *28*(4), 901–911. https://doi.org/10.1109/TKDE.2016.2518669

GSC Mobile App. (2017). Retrieved from http://www.gsc.com.my/html/default.aspx?ID=60&PID=69

Guo, B., Chen, C., Zhang, D., Yu, Z., & Chin, A. (2016). Mobile crowd sensing and computing: When participatory sensing meets participatory social media. *IEEE Communications Magazine*, *54*(2), 131–137. https://doi.org/10.1109/MCOM.2016.7402272

Hodson, H. (2013, February). Crowdsourcing grows up as online workers unite. *New Scientist*. Retrieved from https://www.newscientist.com/article/mg21729036.200-crowdsourcing-grows-up-as-online-workers-unite/

Howe, J. (2006). The Rise of Crowdsourcing. Retrieved from https://www.wired.com/2006/06/crowds/

International Article Number. (2017). Retrieved from https://en.wikipedia.org/wiki/International_Article_Number

Kanhere, S. S. (2011). Participatory Sensing: CrowdSourcing Data from Mobile Phones in Urban Spaces. *12th IEEE International Conference on Mobile Data Management (MDM)*, 3–6. https://doi.org/10.1109/MDM.2011.16

Khan, M. A. A. H., Hossain, H. M. S., & Roy, N. (2015). SensePresence: Infrastructure-Less Occupancy Detection for Opportunistic Sensing Applications. *Proceedings - IEEE International Conference on Mobile Data Management*, *2*, 56–61. https://doi.org/10.1109/MDM.2015.41

Maestre, D. P. (2012). QRP : An improved secure authentication method using QR codes, 1–11.

Qibla Connect - Find Qibla Direction. (2017). Retrieved from http://qiblaconnect.com/

Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. https://doi.org/10.1109/CTS.2013.6567202

Scholar, I. (2016). Developing Mobile Application To Navigate Blind People Using Sensors, 80–84.

Shamal, S., Monika, K., & Neha, N. (2014). Secure Authentication for Online Banking Using QR Code, *4*(3), 778–781.

Soon, T. J. (2008). QR Code. *Synthesis Journal*, 59–78.

Surowiecki, J. (2004). The Wisdom of Crowds. *How Collective Wisdom Shapes Business Economies Societies and Nations New York Doubleday*, 296. https://doi.org/10.3174/ajnr.A3417

Syafiq, F., Aris, H., Ismail, H., & Yusof, S. (2016). Automated Update of Crowdsourced Data in Participatory Sensing: An Application to Crowdsourced Price Information. *PERTANIKA Journal of Science and Technology*, *25*, 1–14.

The Star Online. (2017). Maybank launches cashless mobile payment option using QR codes. Retrieved from https://www.thestar.com.my/business/business-news/2017/12/15/maybank-launches-cashless-mobile-payment-option-using-qr-codes/

Torres, S., Lalanne, F., Del Canto, G., Morales, F., Bustos-Jimenez, J., & Reyes, P. (2016). BeCity: Sensing and sensibility on urban cycling for smarter cities. *Proceedings - International Conference of the Chilean Computer Science*

Society, SCCC, 2016–Febru. https://doi.org/10.1109/SCCC.2015.7416587

Väätäjä, H., Vainio, T., & Sirkkunen, E. (2012). Location-based crowdsourcing of hyperlocal news. *Proceedings of the 17th ACM International Conference on Supporting Group Work - GROUP '12*, 85. https://doi.org/10.1145/2389176.2389189

Wang, G., Liu, F., & Yan, W. Q. (2016). 2D Barcodes for visual cryptography. *Multimedia Tools and Applications*, *75*(2), 1223–1241. https://doi.org/10.1007/s11042-014-2365-8

Waze official website. (2017). Retrieved from https://www.waze.com/

WhatsApp Web. (2015). Retrieved from https://blog.whatsapp.com/614/WhatsApp-Web

Witt, L., Green, R., Cross, D., McFadden, B., & Zhang, K. (2003). Automatic deletion in data storage management. misc, Google Patents. Retrieved from https://www.google.com/patents/US20030220949

Xiao, Z., Lim, H. B., & Ponnambalam, L. (2017). Participatory Sensing for Smart Cities: A Case Study on Transport Trip Quality Measurement. *IEEE Transactions on Industrial Informatics*, *138632*(c), 1–1. https://doi.org/10.1109/TII.2017.2678522

ZXing ("Zebra Crossing") barcode scanning library for Java, Android. (2017). Retrieved from https://github.com/zxing/zxing

ZXing Team. (2015). Barcode Scanner. Retrieved from https://play.google.com/store/apps/details?id=com.google.zxing.client.android&hl=es_419