Functionality-Based Mobile Application Recommendation System with Security and Privacy Awareness

Thiago Rocha[1], Eduardo Souto[1], Khalil El-Khatib[2]

[1] Institute of Computing– Federal University of Amazonas
69067-065 – Manaus – AM – Brazil

[3]Advanced Networking Technology and Security
Ontario Tech University, Oshawa, L1G 0C5, Canada

thiago.rocha@icomp.ufam.edu.br, esouto@icomp.ufam.edu.br, khalil.el-khatib@uoit.ca

Thesis available at: https://tede.ufam.edu.br/handle/tede/7793

Abstract. *In this thesis, we propose a functionality-aware system to evaluate and recommend mobile applications with security and privacy awareness. The proposed system has a security layer that evaluates an application and classifies it as being malign or benign. In this way, only applications classified as benign are considered for the functionality-aware recommendation. Also, we employ a technique, called Logical Predicate Mapping (LPM), which allows users to understand the permissions and API calls requested by the app, as well as privacy risks. This information is grouped with other metrics retrieved such as popularity, usability and privacy and shown to users. This way they can decide what to do and understand what can happen.*

1. Motivation and objectives

The number and usage of mobile devices has increased dramatically over the last decade and has changed the way users execute their daily tasks and do business and will continue to do so. The number of global smartphone users has already exceeded 2 billion and is expected to reach 3 billion by 2020 [Edwards et al. 2016]. As a result, the number of applications developed to help users perform many tasks has also grown considerably. Such applications are used daily and provide various functionalities such as phone calls, e-mail sending, GPS service, and camera to list a few.

Due to the large number of apps, recommendation systems have been used by users to find applications that suit their needs and interests. For example, Google Play recommends applications based on aspects of the app being searched, such as store category and the name of the app developer(s). However, security and privacy aspects are not satisfactorily considered by the recommendation systems, both in official stores and in recommendation systems in general [Xu et al. 2018].This is a concern since applications frequently request users sensitive or private data such as logins, passwords, location and financial information to accomplish their objectives. Therefore, these applications became potential targets for malicious developers.

There are some recent studies about recommendation systems that consider some security and privacy aspects. However, these works also have several problems and limitations. First, most of them only check app permissions configuration. Permission use is an important feature for calculating application security risks. However, they are not sufficient to guarantee that an application is safe [Akhuseyinoglu and Akhuseyinoglu 2016], [Martín et al. 2018]. Some papers and security forums have already demonstrated attacks that can happen without any Android permission usage [Kywe et al. 2016], [Paloalto 2017]. Besides that, most users cannot understand how permissions work, what they do or do not pay attention when permissions are requested [Liu et al. 2016]. Such fact creates a gap between user expectations and application behavior [Wang et al. 2015]. Moreover, there is also a problem with over permission [Xu et al. 2018], [Wang and Chen 2014] when an app requests more permissions than necessary. This situation can lead to the appearance of attacks from permissions that are not even necessary by the apps.

Another limitation is related to the way application recommendations are made. Most of the time, these recommendations are based on a set of applications that belong to the same category as the official store. This may not satisfy the needs of a user who is looking for a specific functionality. For instance, when a user is searching for an application that is similar to WhatsApp, if the recommendation system returns Facebook application, as it happens in the Play Store, it would be unsatisfactory, although both belong to the same category. In mobile app recommendation, metrics should be calculated inside a functionality context and not by category. For instance, if a user wants to change the app that is used to order food, the recommendation system should consider only safe applications that can also order food.

Privacy must also be calculated inside a functionality context to detect the leak of sensitive information. In some cases, the permissions and API calls to access sensitive information that may be considered malicious in one app could be a feature in another app [Gorla et al. 2014]. For instance, calculating the privacy score using location permissions and API calls from an app that tracks current user location may be considered if it is an app with bank functionalities but should not be considered if it is a navigation app, because in this case the request is benign and the location information is necessary for the correct execution of its service. Moreover, most users are not aware of the data collected by apps [Shklovski et al. 2014].

To overcome these problems, we propose a system to evaluate and recommend mobile applications, inside Android operational system environment, with security and privacy awareness. For that, some challenges had to be addressed. The first one is the capability to create a method that can evaluate a target application and classify it as being malign or benign before a recommendation is made. Second, the possibility to apply topic extraction techniques on applications descriptions to suggest only apps with the same functionality. Moreover, the system analyzes the information gathered from the target application and calculates metrics such as privacy, usability, popularity and checks the permissions and API calls to map all possible behaviour that could cause privacy and security risks or any behavior that is not aligned with the application description. To reach these goals this thesis has the following specific objectives:

1. Creation of a method to download apps from an application store;
2. Creation of a mechanism that extracts the features from the applications to create and train a model that classifies applications into benign or malign;

3. Development of a functionality-based recommendation engine that is able to suggest apps.
4. Development of functionality-based algorithms that can calculate usability, privacy and popularity metrics to build a ranking of applications to be suggested to the users;
5. Creation of an understandable summary with the information gathered during the evaluation in a way that it is possible for technical and non-technical users to understand.

From the objectives, this work offers the following contributions:

- A mobile application recommendation system with a machine learning security layer that evaluates apps and only suggests the ones classified as benign;
- A functionality-based app scoring system that was created to obtain functionality, privacy, usability and popularity metrics to later rank the apps that are suggested to the users. Since the scoring system is based on the purpose of the apps, all the metrics are calculated inside a functionality context. With that it is possible to only recommend applications that perform similar functionalities as the application being evaluated and also check the privacy, popularity and usability in different conditions. For instance, in relation to privacy, in some cases the permissions and API calls that may be considered malicious in one application could be a feature in another app [Gorla et al. 2014];
- Creation of a novel Logical Predicate Mapping (LPM) that aims to clarify the behavior that a given application can execute inside mobile devices so mitigation actions can be taken and problems such as overpermission can be detected and faced.

2. Main Results

Quantitative and qualitative experiments were carried out, due to space limitations not all will be shown. In quantitative experiments it is verified which classification model is more efficient in the identification of malicious applications and which parameters are best for the recommendation model. In qualitative experiments some malicious applications were chosen to be evaluated inside the prototype and in other frameworks to compare the results such as RSPSA [Jisha et al. 2018] and Google Play. RSPSA was selected because it has characteristics that are similar to the proposed system.

The metrics used for the quantitative experiments were Precision, Recall and F1-Score for classification while coherence score is used to evaluate the recommendation model. Table 1 shows the classification model evaluation results with four algorithms.

Table 1: Machine Learning Model Evaluation.

Algorithm	Precision	Recall	F1-Score
J48	96,50%	96,50%	96,50%
Logistic	95,60%	95,70%	95,70%
SMO	94,20%	94,40%	94,42%
Naïve Bayes	90,60%	84,60%	84,55%

Precision is related to the question “of all applications labelled as malware, how many applications actually were malware?” while Recall is related to “of all the applications that are really malware, how many did we label?” and F1-Score combines

Recall and Precision to reach a ratio that measures the overall quality of the model. J48 was chosen because it had the best results.

The recommendation algorithm was created using Latent Dirichlet Allocation (LDA) and evaluated using the Coherence score because it takes a topic and measures the degree of semantic similarity between the words with the highest score in the topic [Stevens et al. 2012]. These calculations help differentiate between semantically interpretable topics and the ones that are only artifacts for statistical inference.

The evaluation was made with the list of descriptions from the Play Store communication category. Figure 1 shows an average of coherence values with a different number of generated topics.

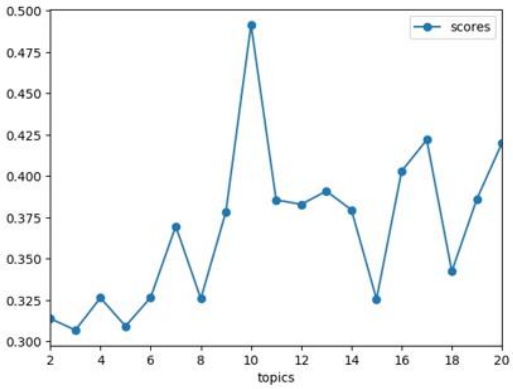

Figure 1: Coherence Score Values by Number of Topics.

Since LDA training is non-deterministic each coherence value was obtained through an average of executions and the best model was the one with 10 topics with an average coherence score value of 0.49. In order to compare the created Prototype results with RSPSA, 10 malicious applications from different categories and with different objectives were selected.

Spam Guard was the first malware application evaluated, which is from the productivity category and its goals are described as an application that automatically detects and moves spam emails from the user inbox to the spam folder. However, the application accesses users' sensitive information such as contacts and sends it through SMS messages.

The recommendation strategy used in RSPSA receives a list of applications from the same category (productivity in this case) and calculates user ratings and security scores based on permissions configuration. Then, these results are used in a clustering algorithm to perform the suggestions. Table 2 shows the top 3 applications related to Spam Guard after RSPSA evaluation and the top 3 applications suggested from the proposed prototype.

None of the applications from RSPSA results have the same goal as Spam Guard. For instance, Xodo PDF is a PDF reader and editor while Business Calendar 2 is a calendar. Meanwhile, the proposed prototype recommendation strategy discards the malicious app and suggests applications with similar functionalities. The Email Spam Filter application stands out as it is used to control and restrict which emails are added to a user inbox.

Email Spam Filter had a 2.6 user rating score and 11.0 privacy score. A low user rating score shows that the application is poorly accepted by users regarding features related to usability, such as user interface. However, the application has few revisions (193) and can improve this score over time through new user reviews and application updates. The privacy score shows that the app does not require many permissions, with seven normal

permissions out of a total of 33 and three dangerous permissions (READ_PHONE_STATE, WRITE_EXTERNAL_STORAGE and READ_EXTERNAL_STORAGE) out of 27.

Table 2: Prototype Versus RSPSA Results with Spam Guard.

Approach	*Application*	*User Score*	*Privacy Score*
RSPSA	Xodo PDF Reader & Editor	4.72	7.0
	Business Calendar 2	4.61	28.0
	Password Safe - Secure Password Manager	4.61	4.0
Prototype	Email Spam Filter	2.60	11.0
	Email - Fast & Secure mail for Gmail Outlook & more	4.60	18.0
	Microsoft Outlook	4.33	16.0

The first dangerous permission setting allows the application to retrieve information about the mobile device, such as network information and device number, while the other two permissions configuration allow the reading and writing to external storage. The privacy score value is not normalized because RSPSA has not done any normalization. Thus, the higher the privacy score is, the greater the risk of data leakage is. Figure 2 provides a screenshot from the proposed prototype with results referring to Email Spam Filter showing the application name, category, developer, and scores. Regarding the LPM, access to IMEI and writing to external storage behaviors were found and mapped, which is important as the IMEI access is an uncommon behavior in the in the Email Spam Filter similar apps cluster. This information is grouped and passed on to users so they are aware that the app can access this confidential information and may cause potential data leakage.

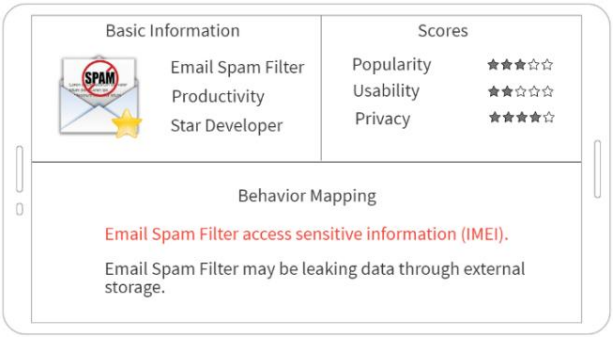

Figure 2: Email Spam Filter Results.

Table 3 shows the results with the other 9 malicious applications. Banco do Brasil application is a repackaged version of the legitimate bank application with the goals to steal information and show adds. RSPSA top application recommendation was Canadian Mortgage App, while the prototype recommended the real version of Banco do Brasil application. The description provided from the malicious application caused the real and secure version of Banco do Brasil to be suggested because of the functionality score.

Since RSPSA gets the applications from a category and calculates the overall user rating and permission scores, it always returns the same top suggested application. For instance, for the Communication category it always returns Pleymojis application, while for the Finance category it always returns Canadian Mortgage App. On the other hand, the prototype considers the apps descriptions to check its functionalities. Therefore, it returns different results depending on the app being evaluated. For instance, in Table 3, inside the

Communication category two different apps were suggested: Write Voice SMS: write SMS by voice for Voice SMS and the real Opera mini browser for Opera Mini 6.5.

Table 3: Prototype Versus RSPSA Results with Malicious Applications.

Malicious Apps	*Category*	*RSPSA*	*Prototype*
Cut The Rope	Puzzle	I love Hue	Cut The Rope (Benign)
Banco do Brasil	Finance	Canadian Mortgage App	Banco do Brasil (Benign)
Deal&Be Millionaire	Trivia	Millionaire Trivia: Who Wants To Be a Millionaire?	Millionaire 2019 - Trivia Quis
Fish Aquarium Live Lock	Personalization	New Year 2019 countdown	Fish Live Wallpaper 2018
Voice SMS	Communication	Pleymojis	Write Voice SMS
Where is My Water?	Puzzle	I Love Hue	Where's My Water? 2
ClockPlus	Tools	Post	Multi Timer StopWatch
Sberbank	Finance	Canadian Mortgage App	Sberbank Mobile Bank
Opera Mini 6.5	Communication	Pleymojis	Opera Mini - fast web browser

The last comparison is made with Google Play Store since it is the official Android store and most of the users download their applications or get suggestions from it. Table 4 shows the related apps that are returned if a user searches for Viber, a famous messaging application, inside the store and also the results with the proposed system.

For Viber, Google Play suggests IMO free video call and IMO beta free call that are applications created from the same developer, as well a Mail.ru that is an email application. Meanwhile, the proposed system suggests Telegram, Messenger and GO SMS Pro - Messenger, Free Themes, Emoji, which are all messaging apps with similar functionalities as Viber.

Table 4: Prototype Versus Google Play Results with Viber.

Approach	*Application*
Google Play	IMO free video call IMO beta free call Mail.ru - Email App
Prototype	Telegram Messenger GO SMS Pro - Messenger, Free Themes, Emoji

Besides not considering any security aspects, Google Play apparently makes its recommendations by prioritizing the app developers over functionality. An improvement in the store would be adding filters for the users to choose how they want their recommendations.

In addition, Google Play categories could be broken down into more specific categories to prevent apps with different functionalities from falling into the same category. To prove that fact, the topics distribution generated from the communication Google Play category are show in Figure 3, each bubble represents a topic and the size of it measures how prevalent the topic is relative to the data.

There are certain topics such as 2, 4 and 9 that are far away from the position that concentrates most bubbles and could be a new category in Google Play Store while the

overlapping bubbles could be analyzed and merged into a category that covers them. Figure 4 shows the world clouds with the top 10 most frequent words in topics 5,6,7 and 1,8,3 because these topics are overlapping and share worlds and from topics 4 and 9. All these topics suggest that the communication category could be broken into other smaller categories that could be more specific.

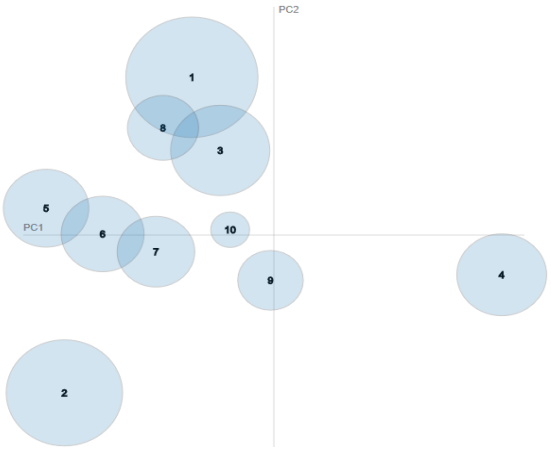

Figure 3: Topics Distance Mapping.

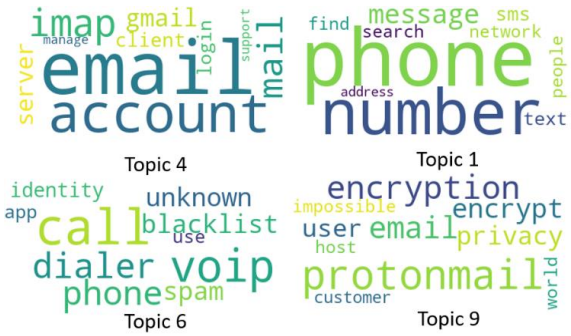

Figure 4: Top 10 Most Frequent Words in Topics 1,4,6 and 9.

2. Scientific Production

• Rocha, T., El-Khatib, K. & Souto, E. (2020), “Functionality-Based Mobile Application Recommendation System with Security and Privacy Awareness”, Computers & Security (COSE).

• Rocha, T., El-Khatib, K. & Souto, E. (2019), “Techniques to Detect Data Leakage in Mobile Applications”, International Journal of Security and Networks (IJSN), february;

• Rocha, T., & Souto, E. (2019), "Avaliação e Recomendação de Aplicativos para Dispositivos Móveis com Foco em Segurança e Privacidade.", XIX Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg 2019 - Artigos Completos / Full Papers), São Paulo (SP), september;

References

Alessandra Gorla, Ilaria Tavecchia, Florian Gross and Andreas Zeller (2014) “Checking App Behavior Against App Descriptions”, International Conference on Software Engineering, pages 1025-1035.

Elizabeth Edwards, Joanna Lumsden, Julian Rivas Gonzalo, et al. (2016) “Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps” BMJ Open, vol. 6, pages 1-9.

Haoyu Wang, Jason Hong and Yao Guo (2015) “Using Text Mining to Infer the Purpose of Permission Use in Mobile Apps”, ACM International Joint Conferen ce on Pervasive and Ubiquitous Computing, pages 1107–1118.

Ignácio Martín, José Alberto Hernández, Alfonso Muñoz and Antonio Guzmán (2018) “Android Malware Characterization Using Metadata and Machine Learning Techniques”, Security and Communication Networks, pages 1-11.

Irina Shklovski, Scott Mainwaring, Halla Skúladóttir and Hóskuldur Borgtho rsson (2014) “Leakiness and Creepiness in App Space : Perceptions of Privacy and Mobile App Use”, Conference on Human Factors in Computing Systems, pages 2347-2356.

Jiayu Wang and Qigeng Chen (2014) “ASPG : Generating Android Semantic Permissions”, International Conference on Computational Science and Engineering, pages 591-598.

Kun Xu, Weidong Zhang and Zheng Yan (2018) “A privacy-preserving mobile application recommender system based on trust evaluation” Journal of Computational Science, vol. 26, pp. 87–107.

Nuray Baltaci Akhuseyinoglu and Kamil Akhuseyinoglu (2016) “AntiWare: An Automated Android Malware Detection Tool based on Machine Learning Approach and Official Market Metadata” Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pages 1-7.

Paloalto, “Cloak and Dagger attack with no permission.”. Available in: https://unit42.paloaltonetworks.com/unit42-android-toast-overlay-attack-cloak-and-dagger-with-no-permissions/. [Accessed in November 10 2017].

Pulkit Rustgi, Carol Fung, Bahman Rashidi and Bridget McInnes (2017) “DroidVisor: An Android secure application recommendation system” IEEE Symposium on Integrated Network and Service Management (IM), pages 1071–1076.

Keith Stevens, Philip Kegelmeyer, David Andrzejewski and David Buttler (2012) “Exploring Topic Coherence over many models and many topics”, Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pages 952–961.

Rui Liu, Junbin Liang, Jiannong Cao, Kehuan Zhang, et al. (2016) “ Understa nding Mobile Users ’ Privacy Expectations : A Recommendation-based Method through Crowdsourcing”, IEEE Transactions on Services Computing, vol. 12, pages 304–318.

R. C. Jisha, Ram Krishnan and Varun Vikraman (2018) “Mobile Applications Recommendation Based on User Ratings and Permissions” International Conference on Advances in Computing, Communications and Informatics (ICACCI), pages 1000–1005, 2018.

Su Mon Kywe, Yingiju Li, Kunal Petal and Michael Grace (2016) “Attacking Android Smartphone Systems without Permissions”, Conference on Privacy, Security and Trust (PST).