平成 31 年度 修士論文

Android アプリの権限要求に対する説明十分性の自動確認システムの提案

指導教員　酒井 哲也 教授

早稲田大学基幹理工学部情報理工学科

学籍番号 5118F0402

小島智樹

概要

Anrdoid 端末においてアプリケーションは端末上のセンシティブな情報にアクセスする権限をユーザーに要求する．最近では Runtime Permission システムという，実際に情報にアクセスする際にユーザに許可を求める方式が主流である。それに付随して，アプリの開発者は使用理由に関する Rationale(理論的根拠) をダイアログで提示することができる．これによって．ユーザはより安心してアプリによるアクセスを受け入れることができる．しかしながら，Liu らの調査 [11] によればこのダイアログによって提示される説明 (以後本研究ではこの説明文を Rationale Text，略して RT と呼ぶことにする) は未だに不正確なものが多いとされている。そこで、本研究ではダイアログによって提示される説明の不足について自動判定を行うシステム，Permission Rationale Checker (PRC) の構築を行う．その際に，従来研究で用いられてきた分類対象の文章だけでなく，アプリケーションの説明文 (description) を利用することで精度の向上を試みた．
その結果，description を利用した word2vec の潜在表現の学習は分類における F1 値の改善に寄与することが判明した．また，description に対して適切な前処理を施すことにより更なる F1 値の改善が見られた．

目次

第１章　導入

Permission とは，本論文では Android OS の一部機能にアプリがアクセスするための許可のことを表す．例として，位置情報の取得のための許可であったりマイクを使用することに対する許可といったものが挙げられる．これらはどれもユーザーにとってセンシティブな情報である．Bonné らの研究 [4] で行った実験では調査対象の 44% がアプリによる権限の要求に不快感を示した．そのため，ユーザーが安心してアプリを使用するためにはアプリの開発者は十分な説明を行うことが望まれる．

そこで必要とされるのが Rationale である．Rationale とは日本語で理論的根拠のことであり，本研究においては「権限をユーザーに求める際に行われる説明」という意味で使われる．開発者が Rationale を提示することによって，ユーザーはアプリケーションに権限を与えることへの安心感を高め，承認率の向上を期待することができる．

Rationale は主に 2 種類の方法で提示される．

1 つ目の方法としては、 Google Play などのアプリストア上 での説明がある. 図 1 は Google Play で公開されている Yahoo!乗換案内の説明文の一部である．ユーザーはこういった説明に

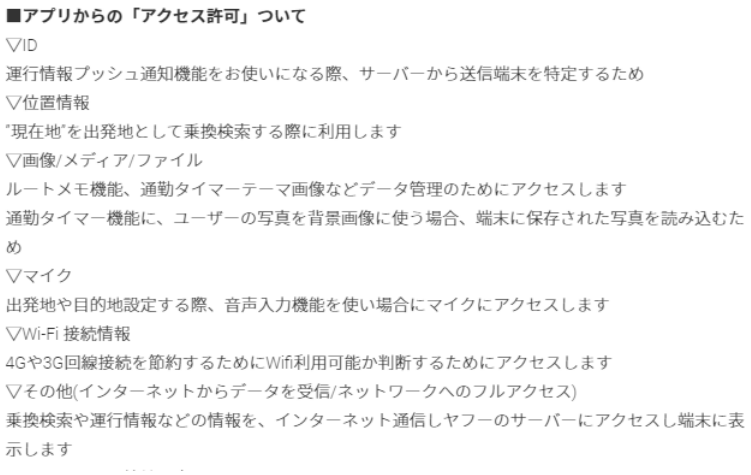
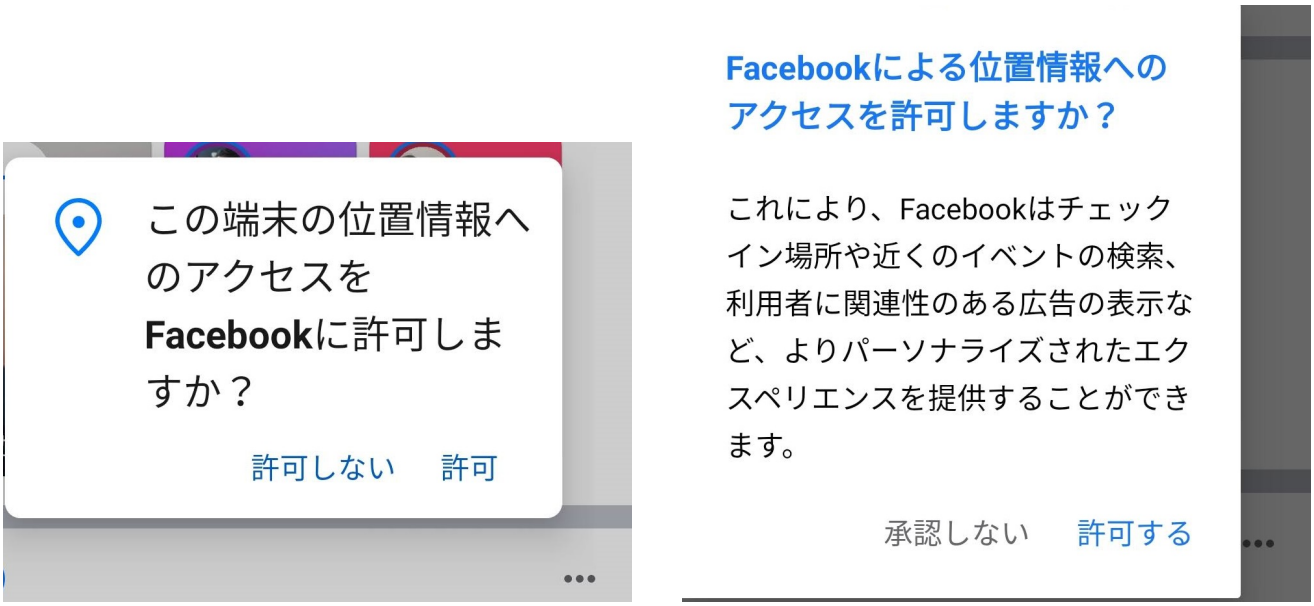

■アプリからの「アクセス許可」ついて
▽ID
運行情報プッシュ通知機能をお使いになる際、サーバーから送信端末を特定するため
▽位置情報
"現在地"を出発地として乗換検索する際に利用します
▽画像/メディア/ファイル
ルートメモ機能、通勤タイマーテーマ画像などデータ管理のためにアクセスします
通勤タイマー機能に、ユーザーの写真を背景画像に使う場合、端末に保存された写真を読み込むため
▽マイク
出発地や目的地設定する際、音声入力機能を使い場合にマイクにアクセスします
▽Wi-Fi 接続情報
4Gや3G回線接続を節約するためにWifi利用可能か判断するためにアクセスします
▽その他(インターネットからデータを受信/ネットワークへのフルアクセス)
乗換検索や運行情報などの情報を、インターネット通信しヤフーのサーバーにアクセスし端末に表示します

図 1.1　Yahoo!乗換案内の説明文

よって，より正しくアプリに権限を与えるかの決定をすることができる．

別の方法として，バージョン 6.0 以降の permission の機構である Runtime permission を拡張したものを使う場合がある．本論文では「Rationale dialog」と呼ぶことにする．



(a) 開発者が独自に設定したダイアログ　　(b) システムのダイアログ

図 1.2　Facebook 内で表示先ダイアログ

図 2, 図 3 は Google Play で公開されているアプリ「Facebook」を起動した際に表示される画面である．図 2(a) が開発者が権限を何のために要求するかを説明するために，独自に設定した Rationale dialog である．本研究ではダイアログに記されているこのような文を Rationale Text (RT) と呼ぶ.

このダイアログの表示後，図 2(b) のようなシステムのダイアログが出現し，実際にその権限を許可/不許可を選択する．このように実際に機能を使う際に権限を使用する理由を提示することによりコンテキストを理解した上での説明であるためユーザーの納得を得られ，また承認率の向上にもつながる.

しかしながら，Liu らの研究 [11] によればこの RT が適切でない場合が多々存在する．適切な文言の設定による承諾率の変化は Tan らの研究 [12] によって示されている．使用する権限の種類とその理由を明記したとき，単に使用するということを提示したときに比べ，承諾率が最大で 20% ほど向上したという結果がこの研究では示されている．このように，適切な Rationale の設定はユーザーにとっては納得感を持った権限の受け入れの助けになり，それによって開発者は自信が開発したアプリの使用に繋がるため重要である.

そのため，本研究では RT として書かれている文言が適切な説明であるかどうかの自動判定機構，Permission Rationale Checker (PRC) の提案を行っていく.

本研究における分類対象である RT は一般に短いものが多い．そのため，単に RT のみを教師データとして用いることは分類精度の低下を招くと考えられる.

そのため，今回の PRC の提案において，従来の分類対象の文のみを使用した方法に加えて，今回の研究においては分類対象のアプリケーションの説明文 (description) を学習の補助データ

として使用する．

実験の結果，description の補助的なデータとしての使用の効果が認められた．この補助的な利用は Permission Text 以外のレビューの分類などにも応用することが期待される．

以下に本論文の構成を示す．2 章において，本研究に関連した Android のセキュリティやユーザーのプライバシーへの意識，Permission システムなどに関する研究を紹介する．3 章において，PRC に関する構成について述べる．4 章において，それら機構が上手く機能しているかを実験した結果，及びその結果からの考察を示す．5 章においてまとめを記す．

第 2 章　関連研究

2.1　ユーザーのアプリケーションのセキュリティに対する意識

アプリケーションのユーザーのセキュリティに対する意識を調査した研究として Bonné らの研究 [4] や Golbeck らの研究 [6] が挙げられる．Bonné らの研究では Runtime permission が権限の承諾にどのような影響を与えるかを様々な権限に対して調査している．

また Golbeck らの研究では Facebook を例にユーザーはどのような情報へのアクセスについての理解しているのか，またそれに対してどのような懸念を抱いているのかなど，セキュリティへの意識について調査している．

2.2　Android アプリに関する自然言語の分類

アプリの Rationale に関する先行研究はまだ多くはない．そのため，ここではアプリと自然言語処理に関する論文について述べる．

Nayebi らの論文 [5] で言及されている，アプリの Review は短くて，かつしばし口語的であると言った特徴は Rationale text にも共通して見られる．そのため，これらの論文で用いられている特徴は Rationale の分類にも有用であると考えられる．

Review を利用した Android アプリの分類に関しては Gómez らの研究 [7] がある．これは Latent Dirichlet Allocation (LDA) による教師なし学習を利用し，Review からバグの発生しやすい権限のパターンについて解析する研究である．

また，別の研究として Jha らの研究 [8] では Bag Of Frame という手法が使われている．これはアプリの Review を構成する単語を抽象度の高い Frame という単位に変換することで，表現の揺れに強く過学習を抑制した特徴を作成し高い分類精度を実現している．

これらの研究にあるように，アンドロイドアプリに付随する文は表現が多様である．そのため，単語という具体的な特徴で捉えるよりも，より抽象度の高い特徴で捉える方が正確に文意を捉えることが可能になると考えられる．それを行うため，本研究では ord2Vec(w2v) を用いた．

2.3 権限の付与の提案に関する研究

ユーザの権限に関する Liu らの研究 [10] がある．これは，ユーザーの権限の許可/不許可に関するログデータを利用し，ユーザー意思を推測することで複数権限の On/Off に関する提案を行い，ユーザーの決定の手間を減らすシステムである．これによって意思決定の回数を減らし，より考えて権限の付与を行うことが可能になる．しかし，これについてはなぜその権限を On/Off にするかの理論的な説明はユーザーに行われない．また，ユーザにとってそもそもこのロギング機構をスマホに入れること自体が不快感である．そのため以下のようなことがシステムに求められる．

- ユーザーのログによらないシステム
- 使用理由に関して適切な説明を行う

2.4 Rationale の有無の自動判定について

アプリ以外に対する Rationale の有無について行われている研究として，Kurtanovic らの研究 [9] や Alkadhi らの研究 [3] がある．Kurtanovic らの研究においてはユーザによるソフトウェアの Review を対象に，どのような理由でソフトウェアが評価されているか，また改善を求められているかについてなどの Rationale の研究を行っている．その記述の有無，及び種類の判別．においては を対象に行う．Alkadhi らの研究においてはソフトウェア開発者のバグレポートについて同様の行為を行う．これにより を目指している．

これらの研究は以下の点で本研究と類似している．

- Rationale の有無の判定を行っている
- 対象の文が短い (各データは概ね 1 文)

第 3 章　提案手法

Permission Rationale Checker (PRC) は

1. テキストの前処理
2. 文のベクトル化
3. 学習と分類

の 3 ステップで構成される．

テキストの前処理について，description と RT の共通の前処理に関しては，先行研究と同様の以下のものを実施した．

- stemming
- レンマ化
- ストップワードの除去
- 文字の小文字化

それに加えて description の前処理に関しては追加で以下を実施した．

- URL の除去
- 正規表現を用いた記号の除去
- 形容詞の除去

形容詞除去は本研究独自の試みである．description に含まれる形容詞は主にアプリケーションの特徴の強調やユーザーへの宣伝のために用いられる．そのため，word2Vec の学習の際にノイズになると判断して除去を行った．

記号の除去に関しては以下の通りに行った

- 記号の後が大文字であれば，それは文を区切っている記号と判断し. に置換
- 記号の後が小文字であれば，単語を区切っている記号と判断し" " (スペース) と置換

文のベクトル化については以下の 2 つの方法で行った

1. 単語ごとの w2v の平均

2. 1-gram,2-gram を用いた Bag of words

1 つ目の，単語ごとの w2v の平均を取ることによるベクトル化ついては以下の手順で行った

1. (既存のモデルを使わない場合) w2v の学習を行う
2. RT に含まれる各単語を w2v を用いてベクトル化する
3. その平均をとる

その後，作成した RT のベクトルを学習データとして RnadomForestClassifier を学習，分類を行った．

第 4 章　評価・実験

以下で PRC の評価実験の概要について述べていく．

4.1　データセット

4.1.1　Rationale Text

RT のデータとして Liu ら [11] が論文内で使用している，公開データ [1] を利用する．このデータセットは権限毎のダイアログに対するテキストとされるものの集合である．これに対して，ユーザーが権限を承諾する際にプラスになるか，十分な説明になっているかを自動で判定する．

最初に英語以外のデータを除外する．

その後データに対してラベリングを行う．このラベリングを行う際の基準に関しては Rationale のダイアログが表示される流れは以下の通りである．

1. 機能を起動する (例:SNS アプリで投稿をしようとする)
2. Rationale ダイアログが表示される
3. permission のダイアログが表示される

そのため，正例のラベルは以下の基準を満たすものに付けられ，残りを負例とした

1. どんな機能のために権限を利用するか明示されている
2. その機能が自然言語で具体的に明記されているか
3. その機能が権限に適切か
4. 対象となる権限 (今回はマイク) のための説明であるか

正例

- to use voice search , allow android tv remote permission to record audio .
- voice based typing cannot work without audio recording and storage permissions
- without microphone and location permission the cruise finder voice search will not work .

負例

- to use the application , the microphone and storage permission is required . (1 つめの基準を満たしていない)
- uh oh ! we can t access your microphone ! (1,2 つ目の基準を満たしていない)
- triller would like to use your camera to record your videos (4 つ目の基準を満たしていない．カメラへの権限の要求の理由は説明されているが，マイクについての言及が無い)
- you must accept the microphone permission to be able to use this feature . (2 つ目の基準を満たしていない)

先述の基準に沿って分類された各データは以下の通りの個数である．

表 4.1　RT データの内訳

	カウント	平均単語数
正例	629	14.37
負例	987	12.19
正 + 負	1616	13.03

4.1.2　description データ

今回の実験では分類対象の RT のデータ以外に description のデータを用いた．これを用いることによって w2v に Android のドメイン知識を加えることが目的である．また，そのために分類対象のアプリ以外の説明文 (extra-description, extra-desc と表記する) も追加のデータとして使用した．内訳は以下のとおりである．

表 4.2　description データの内訳

	アプリ数	文の平均長
分類対象	1616	290.56
その他	52718	217.38

4.2　分類評価

評価はデータを 3:1 で学習データとテストデータに分類し，それを用いた 2 分類の Precision, Recall，F1 で行う．分類器 RandomForestClassifier を用いた．

4.2.1 ベースライン

ベースラインの一つとして，従来研究で用いられた方法である bag of words + 2-gram による RT のベクトル化を用いる．

また，今回用いたアプリ情報学習を行った w2v との比較対象として Google が公開している word2Vec のモデル [2] を用いる．これは Google ニュースデータセットの一部で学習された 300 次元，300 万単語に関するモデルである．これを用いて文章をベクトル化行う．

結果を表 2 に示す．

表 4.3　分類結果

	Precision	Recall	F1score
1 bow	**0.782**	0.477	0.592
2 w2v(Google)	0.639	0.609	0.624
3 w2v(RT)	0.595	0.609	0.602
4 w2v(RT + desc)	0.638	**0.633**	**0.635**
5 w2v(RT + desc + extra-desc)	0.622	0.578	0.599

4.2.2 形容詞の除去における分類性能の変化

アプリケーションの description において 形容詞が多い文はアプリケーションの宣伝，ユーザーへのアピールといった RT の分類に対して効果の小さいとされる文が，それは w2v の学習の際にノイズとなりうる．

そのため，description から形容詞を除去し w2v の学習を行った．それによる分類性能の変化は表 4.4 の通りであった．以下の表のモデル 4，モデル 5 は先述の表の 4，5 に対応している．

表 4.4　形容詞除去を行った分類結果

	Precision	Recall	F1score
モデル 4	0.623	0.633	0.628
モデル 4 + 形容詞除去	0.617	**0.641**	0.628
モデル 5	0.622	0.578	0.625
モデル 5 + 形容詞除去	**0.638**	0.633	**0.646**

実験の結果，提案した方法である説明文から w2v を学習させる方法，及び形容詞除去の利用は分類性能を改善に貢献していると考えられる．

第５章　まとめ

本研究では Permission Rationale Checker (PRC) の提案と作成を行った．これは Rationale Dialog に表示される文言 (Rationale Text,RT) がユーザーにとって権限を与えるための一助となりうるかを判定するものである．

本研究における分類対象である RT は概して短く，さらに先行研究のようにストアにおけるユーザーからの評価といったメタデータを直接用いることは難しい．そのため，今回はその足りないデータを補うために説明文を利用した．これを用いることで潜在表現を学習させ擬似的にデータ量を増やした．

分類機としては先行研究に習って RnadomForest を使用し，文章のベクトル化としては 2-gram や複数の方法で学習させた w2v の平均を取ったものを用いた．

その結果，RT と Description を用いて w2v を学習させたものが最高の精度を記録した．比較対象の Google ニュースデータセットを用いて学習させた w2v よりも高い精度を記録しており，このことからドメインに特化した文章による学習が有効に作用していると考えられる．

今回の実験ではマイクの権限にデータのを使用したが，これは比較的ユーザにとって理解しやすい権限でありセンシティブではあるものの，なぜ使われるか想像しやすい場合が多い．そのため，今後の実験においてはより表面的に何に用いられるかわかりにくい権限についても分析を行う予定である．

さらに，Rationale Text はユーザーの目に直接触れるものであるため，今後は多くの人からのユーザーインタビューを行い，定性的な結果を得る必要がある．

また，本研究の方法は RT 以外にもアプリのレビューやその他文が短いデータなど別の種類のテキストデータにも応用が可能であると考えられる．

謝辞

本論文の執筆にあたり、様々なご指導，ご支援をして頂いた指導教員の酒井哲也教授に深く感謝致します．また，貴重なご意見，ご提案を頂いた酒井研究室の先輩方にもお礼申し上げます．

参考文献

[1] Runtime Permission Project. https://sites.google.com/view/runtimepermissionproject/home?authuser=0.

[2] Google Code. word2vec . https://code.google.com/archive/p/word2vec/.

[3] R. Alkadhi, T. Lata, E. Guzmany, and B. Bruegge. Rationale in Development Chat Messages: An Exploratory Study. In *IEEE International Working Conference on Mining Software Repositories*, volume 0, pages 436–446. IEEE Computer Society, 2017.

[4] B. Bonné, S. T. Peddinti, I. Bilogrevic, and N. Taft. Exploring decision making with android's runtime permission dialogs using in-context surveys. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 195–210, Santa Clara, CA, July 2017. USENIX Association.

[5] N. Genc-Nayebi and A. Abran. A systematic literature review: Opinion mining studies from mobile app store user reviews. *Journal of Systems and Software*, 125:207–219, mar 2017.

[6] J. Golbeck and M. Mauriello. User Perception of Facebook App Data Access: A Comparison of Methods and Privacy Concerns. *Future Internet*, 8(4):9, mar 2016.

[7] M. Gómez, R. Rouvoy, M. Monperrus, and L. Seinturier. A Recommender System of Buggy App Checkers for App Store Moderators. In *Proceedings - 2nd ACM International Conference on Mobile Software Engineering and Systems, MOBILESoft 2015*, pages 1–11. Institute of Electrical and Electronics Engineers Inc., sep 2015.

[8] N. Jha and A. Mahmoud. Mining user requirements from application store reviews using frame semantics. In *International working conference on requirements engineering: Foundation for software quality*, pages 273–287. Springer, 2017.

[9] Z. Kurtanovic and W. Maalej. Mining User Rationale from Software Reviews. In *Proceedings - 2017 IEEE 25th International Requirements Engineering Conference, RE 2017*, pages 61–70. Institute of Electrical and Electronics Engineers Inc., 2017.

[10] B. Liu, J. Lin, and N. Sadeh. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd international conference on World wide web*, pages 201–212. ACM, 2014.

[11] X. Liu, Y. Leng, W. Yang, W. Wang, C. Zhai, and T. Xie. A large-scale empirical study

on android runtime-permission rationale messages. In C. Kelleher, G. Engels, J. Fernandes, J. Cunha, and J. Mendes, editors, *Proceedings - 2018 IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC 2018*, Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC, pages 137–146. IEEE Computer Society, 10 2018.

[12] J. Tan, K. Nguyen, M. Theodorides, H. Negrón-Arroyo, C. Thompson, S. Egelman, and D. Wagner. The effect of developer-specified explanations for permission requests on smartphone user behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’14, page 91–100, New York, NY, USA, 2014. Association for Computing Machinery.