

Android アプリのレビューを用いた ユーザーへの権限説明の補完

小島 智樹^{1,a)} 酒井 哲也¹

概要 :

Android 端末においてユーザーがアプリケーションをダウンロードするとき、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要求する。権限はときに悪用されるため、ユーザーはアプリケーションがなぜ権限を要求するかわからず、不安を覚えることがある。本論文ではアプリケーションのレビューなどを利用し、権限の説明を補完するために関係があるとされる単語を提示し不安の軽減を行う手法について提案する。検証の結果、提案手法ではレビュー特有の頻出単語への対応などの点で改善の余地があるという結論となった。

キーワード : Android, 情報検索, 自然言語処理, UX

1. 導入

Android は Google が開発しているスマートフォン向けの Operating System である。2017 年、アクティブな端末数は全世界で 20 億台を超える [1]。公式のアプリケーションストアである Google Play には 300 万以上のアプリケーション [2] が有料、または無料で公開されている。アプリケーションの総ダウンロード回数は 2017 年の 4 半期のみで 190 億回 [3] を超える。

ユーザーがアプリケーションをダウンロードする際に、アプリケーションはユーザーに権限を要求する。権限とは、デバイスのセンシティブなデータや機能をアプリケーションが使用することへの許可である。例として、ユーザーの位置情報の取得への権限などが挙げられる。アプリケーションは通常、権限をユーザーにとって有益な機能の提供のために使う。しかし一方で、悪用することでユーザーに害をなす事態が発生している。具体的な事例として、カメラ機能の権限を不正に利用することで盗撮を行う [4] といったものが報告されている。

このように権限は悪用されることがあるため、ユーザーは権限の要求理由がわからないときに不安を感じる。この不安を払拭するために、本来であればアプリケーションの説明文において権限を要求する理由を述べるべきである。

例えば、Yahoo!乗換案内 [5] のアプリケーションの説明文を見ると各権限について要求する詳細な理由が記述されている。しかしながら、多くのアプリケーションにおいて、説明文中には十分な説明が無く、ときにユーザーの不安を引き起こす。

この論文では、そういった説明が不足しているアプリケーションについて、権限の説明を補完することで、ユーザーの不安を和らげる手法について提案する。

先行研究では、自然言語処理や API 呼び出しを使用することにより、アプリケーションの過剰な権限の検知を行っている。それに対して、本論文では各アプリのユーザーのレビュー、権限、説明文などを用いることで、アプリケーションの権限使用への説明を追加するための単語抽出の方法について提案する。

以下に本論文の構成を述べる。

第 2 章で不正権限の検知やアプリレビューを用いたアプリケーション改良の事例などについて述べる。

第 3 章では提案手法について説明する。

第 4 章では提案手法の有用性を検証する実験の方法、及びその結果について述べる。

第 5 章では得られた結果についての考察を行う。

第 6 章では今後の展望について述べる。

2. 従来研究

2.1 過剰な権限の検知

先行研究において、様々な方法による過剰な権限の検知が

¹ 早稲田大学
早稲田大学基幹理工学研究科情報理工・情報通信専攻 〒169-8555 東京都新宿区大久保 3-4-1

^{a)} frkojima512@ruri.waseda.jp

行われている。自然言語処理を用いた例としては Pandita らが提案した WHYPER [6] というフレームワークや Qu による AutoCog[7] などがある。

Gorla らの研究 [8] ではアプリケーションの記述から分類を行い、その後 API 呼び出しを利用して異常なアプリを判別している。

その他の検知方法としては Zhang らによる API の呼び出しグラフを利用する方法 [9] や Yerima らによるベイズ分類器を用いた静的解析を利用する方法 [10] がある。

これらの研究は、不正な権限の検知という観点でユーザの安心に寄与している。一方で、本研究においては正常な権限の説明という観点でユーザの安心させることを目指す。自然言語と言う形でユーザに提示することで、従来手法に比べて納得感が増すと考えられる。

2.2 Android アプリのレビューの活用に関する先行研究

Android アプリのレビューを活用した研究には Maalej らによるレビューの種類を分類した研究 [11] や、Fu らによるレビューからアプリケーションの改善点を抽出した研究 [12] がある。

Maalej らはレビューの長さや評価のメタデータ、Bag of words 表現などを使用し、アプリのレビューをバグの報告や要望に分類した。この分類は、アプリ開発者のレビューの利用の手助けとなる。

Fu らはレビューからユーザーがアプリケーションのどこに不満を抱いているかを分析し、開発者のアプリの改善の一助とした。

従来研究はアプリケーション開発者に対してレビューを用いて改善案を提案することで間接的にユーザの体験を改善する。一方で、本研究ではレビューを用いて不安を取り除くという形で直接的にユーザの体験を改善させることを目指す。

2.3 ユーザのセキュリティ意識について

ユーザのセキュリティへの意識に関する研究には Chin ら [13] や Lin ら [14], Vidas ら [15] の研究がある。Chin らは複数の観点におけるユーザのセキュリティ意識についてコンピュータとスマートフォンの違い等に関して調査した。

Lin らはアプリケーションの権限へのユーザの感じ方をクラウドソーシングを利用して調査した。そしてその結果を用いて、ユーザーに追加の警告を出すことを提案した。

Vidas らはソースコードの API 呼び出しなどを解析することにより、開発者に最低限の権限を要求することを提案している。これによって、不要な権限をユーザに要求することを減らせ、不安を軽減することにつながる。

3. 提案手法

本章ではユーザに対してアプリケーションの権限への説

明につながる単語を抽出する方法について提案する。まず、提案手法を実行する際に用いられる幾つかの先行研究について説明し、その後に提案手法について説明する。

3.1 relevance weight

relevance weight (rw) は Robertson と Sparck Jones が考案した、確率情報検索モデルにおける単語の重みである [16]。N を総文書数、n をある単語 t の出現した文書数とし、R をあるクラスに属する文書数、r はその中で単語 t が出現する文書数とすると rw は

$$rw(t) = \log\left(\frac{(r+0.5)(N-n-R+r+0.5)}{(n-r+0.5)(R-r+0.5)}\right)$$

で計算することができるこの値は単語 t があるクラスに大してどれだけ特有のものかを表現する値である。

3.2 offer weight

offer weight (ow) は Robertson と Sparck Jones が考案した、クエリ拡張のための重みである [17]。これは、前節で述べた relevance weight を用いて

$$ow(t) = r * rw(t)$$

の形で表される。単語に対してこの値が大きいほど、その単語が特定の文書のクラスとの関係があると見なせる。

本論文では、これを利用してある権限に対する単語を抽出する。

同様の方法による権限に対する単語の抽出は Watanabe らの研究でも行われている [18]。Watanabe らの研究では抽出した単語を不正な権限の検知に利用しているが、本研究では正常な権限の説明に利用している。

3.3 word2vec

word2vec は Mikolov らが開発した単語をベクトル化するための手法である [19]。この手法の特徴としては

- ある単語に対して、意味が近い単語と遠くない単語があったとき、近い単語の方がベクトル空間上の距離が近くなる (e.g. programming という単語に対して computer と pen という単語があったときベクトル空間上で computer の方が近くなる)。
- 単語同士の加算、減算ができる (e.g. king-man=queen-woman といった計算ができる)。
- 比較的小さい次元で単語が表現できる (e.g. Bag of Words では語彙が 100 万語であったら、100 万次元のベクトルが必要だった)。

といったことが言える。この手法を実現するための方法として、CBOW と Skip-gram がある。本研究においては単語の拡張を行うために、上記の 1 番目の特徴を利用する。

今回、この単語ベクトルの作成のためにデータとして収集したアプリの説明文のデータ、及びレビューのデータを使用した。

3.3.1 説明文の品詞の計測

権限に対する説明に用いる単語の品詞を決定するために説明に使われている品詞を以下の手順で計測した。

- (1) アプリの説明文の権限に関する説明と思われる部分を抽出
 - (2) その部分を MeCab[20] を用いて形態素解析
 - (3) 品詞の数を計測
- 計測結果が以下の表である。

品詞	出現回数
名詞-一般	2997
名詞-サ変接続	2614
助詞-格助詞-一般	2335
記号-一般	2112
名詞-固有名詞-一般	1939
動詞-自立	1637
助詞-連体化	1349
助動詞	1109
記号-読点	933
記号-句点	700
名詞-固有名詞-組織	527
助詞-係助詞	524

この中から頻度が高く、なおかつ追加説明に有用であると考えられる名詞-一般、名詞-サ変接続、動詞-自立を使用する品詞として選択した。

3.3.2 レビューの使用について

提案手法ではユーザのレビューを利用している。これは、レビューにはアプリケーションの説明文に載っていない機能についてユーザが書き込むと仮定することで説明文から読み取れない機能の情報を追加情報として利用できるからである。

3.4 提案手法

提案手法は

- 権限に対応する単語を抽出
- その語を拡張
- フィルタリング
- 拡張された語をクエリとしてレビューから検索

の4ステップからなる。

3.4.1 権限に対応する単語の抽出

まず、offer weight の各値を以下のように設定し、単語 t と権限 P に関する $ow(t, P)$ を求める

N 総アプリケーション数

n 単語 t が出現したアプリケーションの説明文数

R ある権限 P を含むアプリケーション数

r ある権限 P を含み、かつ説明文中に単語 t が出現するアプリケーション数

上記の用に計算をして ow が大きい上位 100 個の単語がある権限 P に対応する単語の集合とした。説明文に対しての形態素解析は MeCab を用いて行い、単語の品詞としては名詞-一般、名詞-サ変接続、動詞-自立を選択した。これは先ほど述べたように、アプリケーションの説明文中の権限の説明文にはこれらの品詞の単語が多く使われているからである。

3.4.2 単語の拡張

次に前節で求めた単語について以下の手順で単語の拡張を行った。

- (1) word2vec の学習を説明文、及びレビューの文章を用いて行う
- (2) 前節で得た単語について類似度の高い 10 単語を単語のリストに加える
- (3) 重複した単語を除去

単語の拡張は、同じ機能を表す単語でも表記の揺れが発生し、検索が上手くいかない可能性があるため行われる。

3.4.3 フィルタリング

今回取り出したい単語はある権限に対して特有の単語である。offer weight による重み付けで、ある程度特有の単語を抽出することに成功してはいるが、出現頻度が高い単語は権限に関係なく抽出されてしまう場合がある。そこで、以下の手順で権限に関係なく出てくる単語を抽出し、一般語とする

- (1) 権限 P について ow が大きい単語上位 100 件の集合を $Words(P)$ とする
- (2) 各単語について何種類の $Words(P)$ に含まれているかをカウントする
- (3) カウント数が一定以上の単語を一般語とした。

今回全権限の 2/3 以上で特徴的と判断された単語を一般語とした。

3.4.4 拡張された語をクエリとしてレビューから検索

最後に、拡張された単語をクエリとして利用し、あるアプリケーションのある権限に対応する単語を以下の手順で求めた。

- (1) 3.4.2 で拡張した単語のリストから 3.4.3 で求めた一般語を除去し新たなリストを作成
- (2) リストの各単語のレビュー中の出現回数をカウント
- (3) 名詞-一般、名詞-サ変接続、動詞-自立における上位 3 件に対応する単語とした。

これによって各品詞につき 3 単語が得られ、計 9 単語が決定される。

4. 実験と評価

4.1 使用したデータセット

単語の重み付けには Android アプリの公式ストアである

GooglePlay から入手した 42 カテゴリ, 13487 個の権限と説明文のペアを使用した. その中から仕事効率化とツールの 2 カテゴリについて合計 1069 個のアプリケーションについてレビューを使用し, 評価実験に用いた. アプリケーションの機能の想像しやすいと考えられるため, この 2 カテゴリを選択した.

4.2 実験の方法

早稲田大学基幹理工学研究科情報理工・情報通信専攻の学生 9 人を被験者として以下の手順で実験を行った

- (1) 用意しておいたアプリケーションの中からランダムに選ばれたアプリケーションについて, 説明文を提示して読んでもらう
- (2) 権限とそれに対応するとされる単語 3 つを読んでもらう
- (3) その単語の中に権限に関係する単語があると思うかを答えてもらう
- (4) 1~3 を 10 回繰り返す

今回対象とする権限は被験者が判断しやすいマイク, 位置情報, カメラの 3 種類に絞った. 各権限について名詞-一般の単語, 名詞-サ変接続の単語, 動詞-自立の単語, の 3 問を課した. 問題の選択は完全にランダムのため重複の可能性がある. 図 1 が質問の例である. 赤文字が権限の種類であり, その下の 3 つの単語の中に一つでも関係がある単語があれば黄色いセルの中に 1 を, 一つもなければ 0 を黄色いセルに入力してもらった.

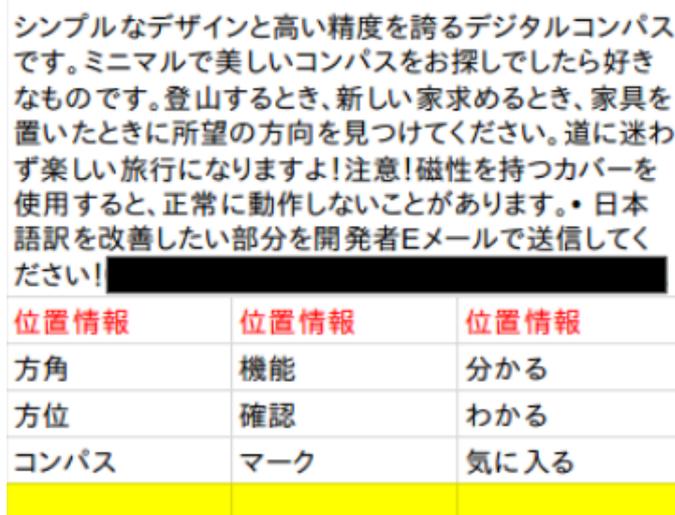


図 1 実験に用いた質問の例

4.3 実験結果

表 2 に実験結果を示す. 表の数字は (単語があると回答した数)/(全回答数) を表す. 位置情報に関するものはあると回答された割合が低く, 品詞に着目して見ると名詞-サ変接続の単語が関係あると回答された割合が高かった. 表 3

表 2 ある権限に単語が関係あると答えた数

	名詞-一般	名詞-サ変接続	動詞-自立
カメラ	18/53	19/53	10/53
位置情報	15/63	23/63	23/63
マイク	12/25	14/25	15/25

は回答の一例である. 一番上の単語が権限を表しており, それ以下の 3 つの単語が関係しているものが一つでもあるかを回答してもらった. あると思えば 1, そうでなければ 0 と回答してもらった.

表 3 回答例

カメラ	カメラ	カメラ
読み取り	機能	読み込む
手間	付属	読み取る
発	説明	読み取れる
1	0	1

5. 考察

ある権限に関係あるとされない単語が抽出されてしまったケースについて考察する.

まず, レビュー特有の「わかる」, 「思う」, 「気に入る」などの単語が抽出されたケースがある. 本来, これらの単語は一般語として除去されるべき対象である. しかし, 権限と語句の関係を計算する際にアプリの説明文を使用したためレビュー特有の単語は上手く一般語として認識されず, 除去すべき対象とならなかったと考えられる.

また, 本研究において, word2vec を用いた単語拡張を行ったが, その際には学習のデータとして説明文とレビューを利用した. そのために十分なデータ量が確保されず適切な単語拡張が行われなかったことも抽出が上手くいかなかったことの原因の一つとして考えられる.

他の失敗原因として, アプリの実際の機能と権限に関係が無いケースがある. これは, アプリ製作者が過失で, または故意に本来は不要な権限を入れているために発生する. そのため, 先行研究のアプリの過剰権限の検知などを組み合わせるなどの対策が考えられる.

実験結果について, 権限に着目して見ると位置情報に関する単語が関係あると答えられた割合が他に比べて低い. これはマイクやカメラに比べて具体的な機能が想像しづらいことに起因すると考えられる.

また, 品詞に着目して見ると名詞-サ変接続の単語が関係あるとされている割合が高い. このことから, ユーザは (名詞-サ変接続)+する の形でアプリケーションの機能を述べることが多いと考えられる.

6. 結論と今後の課題

本研究では, Android アプリケーションが権限を要求する理由への補足説明を試みた. そのために, あるアプリ

ケーションの権限に関係があるとされる単語の抽出を行った。データとしてレビュー、説明文、権限を利用し、offer weight によるランク付けで単語の抽出を行った。検証の結果、提案手法ではレビュー特有の頻出単語への対応などの点で改善の余地があるという結論となった。

以下で今後の課題について述べる。

まず、レビューの扱い方を考える。レビューは貴重な情報源になりうるが、その一方で誹謗中傷を述べただけのレビューなど情報として使えないものが多く存在する。そのため、今後は感情分析等で過度なマイナスなレビューは除外する等でレビューのフィルタリングを行うことが考えられる。

また、結果の提示方法についても改良の余地がある。今回は単語を抽出することのみに留まったが、今後は単語では無く、文を提示することでより実用的なものにしていきたい。例えば、機能の説明文を説明文やレビューから抽出しておく、もしくは作成しておき、その上でクエリとして権限やレビューを用いて検索という形で権限説明を補足するといった方法が考えられる。

本論文の手法では、レビューが少ない場合、有効に特徴語を抽出できないという欠点がある。そのため、豊富にレビューがあるアプリケーションとの類似度を用いるなどして推測を行うなど改良を行っていく必要がある。

謝辞 本研究について様々なご助言を研究室の先輩方、同輩から頂いた。ここに深謝する。

参考文献

参考文献

- [1] Google announces over 2 billion monthly active devices on android.
- [2] Number of available applications in the google play store from december 2009 to june 2018.
- [3] Record levels of app downloads & app store consumer spend in q4 2017.
- [4] Yisha Luo Weijia Jia Dong Xuan Nan Xu, Fan Zhang and Jin Teng. *Stealthy Video Capturer: A New Video-based Spyware in 3G Smartphones*. ACM, 2009.
- [5] Yahoo!乗換案内 無料の時刻表、運行情報、乗り換え検索.
- [6] Wei Yang William Enck Tao Xie Rahul Pandita, Xusheng Xiao. *WHYPER: Towards Automating Risk Assessment of Mobile Applications*. USENIX, 2013.
- [7] Xinyi Zhang Yan Chen Tiantian Zhu Zhong Chen Zhengyang Qu, Vaibhav Rastogi. *AutoCog: Measuring the Description-to-permission Fidelity in Android Applications*. ACM, 2014.
- [8] Alessandra Gorla Ilaria Tavecchia Florian Gross Andreas Zeller. Checking app behavior against app descriptions. *ICSE*, 2014.
- [9] Mohammad Nauman, Sohail Khan, and Xinwen Zhang. Apex: Extending android permission model and enforcement with user-defined runtime con-

- straints. In *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '10, pp. 328–332, New York, NY, USA, 2010. ACM.
- [10] S. Y. Yerima, S. Sezer, G. McWilliams, and I. Muttik. A new android malware detection approach using bayesian classification. pp. 121–128, March 2013.
- [11] W. Maalej and H. Nabil. Bug report, feature request, or simply praise? on automatically classifying app reviews. In *2015 IEEE 23rd International Requirements Engineering Conference (RE)*, pp. 116–125, Aug 2015.
- [12] Bin Fu, Jialiu Lin, Lei Li, Christos Faloutsos, Jason Hong, and Norman Sadeh. Why people hate your app: Making sense of user feedback in a mobile app store. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, pp. 1276–1284, New York, NY, USA, 2013. ACM.
- [13] Erika Chin, Adrienne Porter Felt, Vyas Sekar, and David Wagner. Measuring user confidence in smartphone security and privacy. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, SOUPS '12, pp. 1:1–1:16, New York, NY, USA, 2012. ACM.
- [14] Jialiu Lin, Shahriyar Amini, Jason I. Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. Expectation and purpose: Understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pp. 501–510, New York, NY, USA, 2012. ACM.
- [15] Timothy Vidas, Nicolas Christin, and Lorrie Faith Cranor. Curbing android permission creep. In *In W2SP*, 2011.
- [16] Stephen E. Robertson and Karen Sparck Jones. Document retrieval systems. chapter Relevance Weighting of Search Terms, pp. 143–160. Taylor Graham Publishing, London, UK, UK, 1988.
- [17] Stephen E Robertson and Karen Spärck Jones. Simple, proven approaches to text retrieval. Technical report, University of Cambridge, Computer Laboratory, 1994.
- [18] Takuya WATANABE, Mitsuaki AKIYAMA, Tetsuya SAKAI, Hironori WASHIZAKI, and Tatsuya MORI. Understanding the inconsistencies between text descriptions and the use of privacy-sensitive resources of mobile apps. 2015.
- [19] Greg Corrado Jeffrey Dean Tomas Mikolov, Kai Chen. Efficient estimation of word representations in vector space. 2013.
- [20] 京都大学情報学研究科—日本電信電話株式会社コミュニケーション科学基礎研究所共同研究ユニットプロジェクト. MeCab: Yet Another Part-of-Speech and Morphological Analyzer. <http://taku910.github.io/mecab/>.