

# 日本のスマートフォンアプリケーションにおける評価の低いユーザーレビューでの苦情内容の分析

2014SE001 安部寛生 2014SE015 波多野雅信 2014SE053 小林祐汰

指導教員：横森励士

## 1 はじめに

近年スマートフォンアプリケーションは増加傾向にあり、ユーザーが投稿するレビューは関係者にとって、ユーザーからの重要なフィードバックとみなすことができる。Khalid ら [1] は北米の無料アプリケーションを対象に、低評価ユーザーレビューを調査し、レビューにおいてどのような苦情が多く報告されるか、低評価をつけられやすい苦情は何かを調査した。本研究では、日本のアプリケーションを対象に同様の調査を行い、[1]の結果と比較することで、世界的にユーザーが共通して考えていることと、日本のユーザー固有の特徴を取得する。日本向けのアプリケーション開発を行う上で開発者が特に考えなければいけないことを提言できると考える。また、ゲームアプリケーションにおいては様々なビジネスモデルが展開されている。ビジネスモデル毎に低評価ユーザーレビューを調査することでどのような苦情が報告されることが多いのかや、星評価1をつけられやすい苦情は何かについて、調査を行うことで各ビジネスモデルにおける特徴を示すことができると考える。それぞれの特徴を知ることによって開発するアプリケーションに応じた対策を提案することができる。

## 2 背景技術

### 2.1 アプリケーションストアにおけるユーザーレビュー

スマートフォンアプリケーションを配布する仕組みとして、アプリケーションストアを通じた配布方法が一般的に利用されている。利用者はアプリケーションストアを通じて必要とするアプリケーションをインストールして使用する。利用者は、使用したアプリケーションについての評価をアプリケーションストアに投稿することができる。多くのレビューは星評価と具体的なコメントで構成され、開発者は利用者がアプリケーションに対して評価している点と不満に感じている点の両方の情報を得ることができる。これらの情報は、開発者へのフィードバックや他のユーザーへのアドバイスとなっている。

### 2.2 先行研究

アプリケーションの苦情レビューを精査することで、ユーザーがどのような要素に対して、不満をもちやすいかを調査した Khalid らによる研究がある [1]。[1]では、北米で提供されている無料 iOS アプリケーションを対象に、低評価レビューでどのようなコメントが多くなされていたか、どのような種類のコメントが低評価につながりやすいかを調査した。低評価レビューのコメント内容に基づい

て、表1で示す12種類の苦情を表現するタグが付けられ、各苦情タイプについて低評価レビューの中でどれだけ出現しやすいか(苦情頻度)、ユーザーから嫌悪され、星評価1がつきやすい項目はどれか(苦情影響力)を星評価2に対する星評価1の比率として求めている。表2はその調査結果で、最も多く発生する苦情が「機能エラー」、「機能要求」、「強制終了」であり、最も悪影響を与える苦情が「プライバシーと倫理」、「隠されたコスト」であることが報告されており、これらの情報を Khalid らは改善を目的としたリソースの配分時に役立つ情報であると結論づけている。

表1 Khalid らが特定した苦情の種類 [1]

苦情タイプ	苦情の詳細
強制終了	アプリケーションが強制終了する
互換性	アプリケーションが特定のデバイスやOSのバージョンに問題がある
機能削除	1つあるいは多くの特定の機能がアプリケーションを台無しにしている
機能要求	アプリケーションがより良い評価を得るために、機能を追加する必要があると感じている
機能エラー	アプリケーションの特定の問題に言及し、不満を感じている
隠されたコスト	アプリケーションの全てを経験するために隠されたコストが必要
インターフェース設計	デザイン、制御、映像について不満がある
ネットワーク問題	アプリケーションがネットワークに問題があるか、応答速度が遅い
プライバシーと倫理	アプリケーションがプライバシーを侵す、または反倫理的である
アプリケーションが応答しない	アプリケーションの入力の応答が遅い、または全体的に遅い
魅力のない内容	特定のコンテンツが魅力的ではない
重いリソース	アプリケーションがバッテリーまたは容量を消費しすぎる
特定できない	ただ単にアプリケーションが悪いと言っている

表2 Khalid らの調査結果 [1]

苦情タイプ	苦情頻度		苦情影響力	
	順位	中央値 (%)	順位	星1:星2
機能エラー	1	26.68	7	2.1 (-1)
機能要求	2	15.13	12	1.28
強制終了	3	10.51	4	2.85
ネットワーク問題	4	7.39	6	2.25
インターフェース設計	5	3.44	10	1.5
機能削除	6	2.73	3	4.23
隠されたコスト	7	1.54	2	5.63
互換性	8	1.39	5	2.44
プライバシーと倫理	9	1.19	1	8.56
アプリケーションが応答しない	10	0.73	11	1.4
魅力のない内容	11	0.29	9	1.5
重いリソース	12	0.28	8	2
特定できない	-	13.28	-	3.8

## 3 スマートフォンアプリケーションにおける低評価ユーザーレビューでの苦情内容の調査

### 3.1 研究の動機

先行研究 [1] では、北米のアプリケーションを対象として、それを利用するユーザーがどのような観点で低評価のレビューを行っているかを調査している。日本のアプリケーションに対して、条件をそろえた上で同じ調査を行うことで、共通点から世界的にユーザーが共通して考えていることと、相違点から日本のユーザー固有の特徴が得られるのではないかと考える。日本のユーザー固有の特性は、日本向けにアプリケーションを提供する上で、特別に考えなくてはならないことを提言できると考えた。

さらに、スマートフォン上で動作するゲームアプリケーションでは、様々なビジネスモデルの上でアプリケーションが提供されている。例えば、単純な買い切りのモデルだけでなく、分割して購入できるようにするモデルが存在する。他にも、無料でインストールできるが、課金要素を導入するモデルや広告収入を目的としたモデルなどが存在する。各ビジネスモデルをベースとしたアプリケーションの苦情の傾向を調査することで、苦情の面から各ビジネスモデルの特徴を捉えることができると考えた。

### 3.2 調査手順

本研究は、海外と日本の苦情の傾向の違い及び、ゲームアプリケーションにおけるビジネスモデルの苦情の傾向の違いについて調査する。各調査手順は実験毎で説明する。

1. 日本の App store において公表されているアプリケーションのユーザーレビューを一定期間取得する。ユーザーレビューは、「iTunes Store Web Service Search API」を通じて取得する。
2. 各ユーザーレビューのアップデート日, id, レビュータイトル, 星評価, コメントを抽出する。
3. Sample Size Calculator[2] を用いて、信頼水準 95 %, 信頼区間 5 % で各アプリケーションから抽出する苦情の件数を決定する。この件数に基づき、星評価が 1, 2 であるような低評価レビューを無作為に抽出する。
4. それぞれの低評価レビュー毎に、表 1 に基づいてどの種類の苦情が存在するかをタグ付けする。
5. アプリケーション毎にタグ付けした結果を項目毎に出現割合(苦情頻度)や、低評価レビューにおける星評価 1 のレビュー比率(苦情影響力)を求める。苦情頻度とは、各苦情数を取得したレビューの総数で割り、低評価レビューの中でどれだけ出現しやすいかを示した値をアプリケーション毎に求め、それらの中央値で評価したものとする。また、苦情影響力はユーザーから嫌悪され、星評価 1 がつきやすい度合いを示したもので、星評価 1 の苦情数を星評価 2 の苦情数で割った値とする。

## 4 実験 1(海外と日本の苦情の傾向の違い)

### 4.1 調査内容

低評価レビューが 750 件以上ある日本の無料アプリケーションを 20 個選択した。本研究では、[1] の結果と比較を行うために [1] で選択されたアプリケーションと似た条件でアプリケーションを選択した。表 3 に示すように、アプリケーション数、ジャンル数、サンプリング範囲や低評価、高評価のアプリケーション数などがほぼ一致するようにアプリケーションを選択した。海外の結果と本実験の結果を比較し、アプリケーションに対する苦情の傾向で同じ部分、異なる部分を調査する。具体的には、順位が上がったタイプ(2 以上上昇)、下がったタイプ(2 以上下降)、変

わらなかったタイプ(+1 から-1) 毎に分類し、考察する。

表 3 先行研究と本研究でのアプリケーションの条件

	アプリケーション数	ジャンル数	サンプリング範囲	低評価(星 3.5 未満)アプリケーション数	高評価(星 3.5 以上)アプリケーション数
先行研究	20 個	15 ジャンル	264~383	10 個	10 個
本研究	20 個	13 ジャンル	255~373	10 個	10 個

### 4.2 調査結果

本研究では、20 個のアプリケーションを対象に、2017 年 6 月 21 日から 12 月 01 日までの間に取得した低評価レビューの調査を行った。

苦情頻度 20 個のアプリケーションに対し、表 4 は苦情頻度の順位の比較をグループ化した表で、苦情タイプ毎に 20 個のアプリケーションにおける出現割合の順位、中央値、先行研究 [1] での順位と中央値も並べて示す。

表 4 調査結果(苦情頻度)

順位比較のグループ	苦情タイプ	本研究		先行研究 [1]	
		順位	中央値 (%)	順位	中央値
2 以上上昇	アプリケーションが応答しない	2	11.45	10	0.73
	魅力のない内容	3	7.75	11	0.29
± 1 以内	インターフェース設計	4	7.19	5	3.44
	プライバシーと倫理	8	1.26	9	1.19
	重いリソース	11	0.59	12	0.28
	機能エラー	1	31.71	1	26.68
2 以上下降	機能削除	6	5.26	6	2.73
	強制終了	5	5.77	3	10.51
	機能要求	7	4.74	2	15.13
	ネットワーク問題	9	0.95	4	7.39
	互換性	10	0.79	8	1.39
	隠されたコスト	12	0	7	1.54
-	特定できない	-	5.6	-	13.28

日本のユーザーによって、最も多く報告された上位 3 つの苦情は、「機能エラー」、「アプリケーションが応答しない」、「魅力のない内容」であった。[1] と比較して、順位が 2 以上上昇したグループの苦情は「アプリケーションが応答しない」、「魅力のない内容」であり、共に [1] では低い順位であったが、本研究では上位の順位であった。反対に順位が 2 以上下降したグループは、「強制終了」、「機能要求」、「ネットワーク問題」、「互換性」、「隠されたコスト」であり、「機能要求」、「ネットワーク問題」、「隠されたコスト」が最も大きく順位が異なった。また、「機能エラー」については先行研究と同じく、最も多い苦情であった。

苦情影響力 表 5 は苦情影響力の順位の比較をした表で、苦情タイプ毎に 20 個のアプリケーションにおける星評価 1 と星評価 2 の比率を求め、順位、比率、先行研究 [1] での順位と比率を並べて示す。

表 5 調査結果(苦情影響力)

順位比較のグループ	苦情タイプ	本研究		先行研究 [1]	
		順位	星 1 : 星 2	順位	星 1 : 星 2
2 以上上昇	アプリケーションが応答しない	2	6.79 (:1)	11	1.4(:1)
	機能エラー	5	4.23	7	2.1
	重いリソース	6	3.7	8	2
	インターフェース設計	7	3.67	10	1.5
± 1 以内	強制終了	3	5.94	4	2.85
	プライバシーと倫理	1	9.58	1	8.56
	機能要求	12	1.89	12	1.28
	機能削除	4	4.38	3	4.23
2 以上下降	魅力のない内容	10	2.88	9	1.5
	ネットワーク問題	8	3.48	6	2.25
	隠されたコスト	9	3.21	2	5.63
	互換性	11	2.7	5	2.44
-	特定できない	-	4.15	-	3.8

日本のユーザーから嫌悪され、星評価1が付きやすい項目の上位3つが「プライバシーと倫理」、「アプリケーションが応答しない」、「強制終了」である。[1]と比較して順位が2以上上昇したグループの中でも特に順位に差があった苦情は「アプリケーションが応答しない」であり、[1]では11位だが、本研究では2位と大きく異なった。反対に順位が2以上下降したグループの中で特に順位が異なった苦情は「隠されたコスト」であり、[1]では2位だが、本研究では9位である。また、「プライバシーと倫理」の苦情は、いずれの場合も最も星評価1が付きやすい苦情であり、「機能要求」は最も星評価1が付きにくい苦情であった。

## 5 実験2(ゲームアプリケーションにおけるビジネスモデルの苦情の傾向の違い)

### 5.1 調査内容

ゲームアプリケーションにおいては、インストールに料金が必要となるかどうか、課金要素の有無などによっていくつかのビジネスモデルが存在する。それらのビジネスモデルを完全無料、無料+追加課金、買い切り、有料+追加課金の4つに分ける。各ビジネスモデル毎から10個、合計40個のゲームアプリケーションの2017年7月01日から12月01日までの間に取得した低評価レビューの調査を行った。表6に各ビジネスモデルの調査対象としたアプリケーションの取得期間中の総レビュー数、低評価レビュー数、サンプルサイズの合計と各ビジネスモデルでのダウンロード数を示す。各ビジネスモデルの苦情頻度と苦情影響力を求めることで各ビジネスモデルで実際に、ユーザーが広告、課金についてどのように捉えているか、モデルによってそれら以外の部分にも影響が存在するか調査する。具体的には、実験1と同様に、ビジネスモデル毎に苦情を分類し、苦情タイプ毎に、苦情頻度、苦情影響力を求める。ゲームアプリケーション固有の苦情が出現したので、表7に示す7つの苦情タイプを新たに設定した。さらに、広告と購入については、その内容を細分化している。4つのビジネスモデル毎に結果を示し、特徴的な点を紹介する。

表6 調査対象のアプリケーションの情報

	完全無料	無料+追加課金	買い切り	有料+追加課金
総レビュー数	9,040	44,948	5,177	5,030
低評価レビュー数	2,815	29,789	1,909	2,712
サンプルサイズ	1,567	3,104	1,260	1,538
ダウンロード数	500,000~ 1,000,000	7,500,000~ 30,000,000	100,000~ 500,000	100,000~ 500,000

### 5.2 調査結果

表8に各ビジネスモデルでの上位3位の結果を示す。表8においては上から苦情頻度、苦情影響力を示す。表9に広告に関する苦情、表10にアプリケーション内での購入に関する苦情の苦情頻度・苦情影響力の上位3位の苦情を示す。広告は完全無料のみ、アプリケーション内での購入に関する苦情は無料+追加課金、有料+追加課金のみ示す。

広告の苦情は、完全無料のビジネスモデルで多くみられたが、最も多く報告された苦情ではなかった。インストー

表7 追加苦情タイプと小項目の詳細

苦情タイプ	苦情の詳細
広告	広告に関する苦情
広告が多い	広告の表示が多いことへの苦情
広告の表示位置	広告の表示位置が悪い、または邪魔なことへの苦情
広告の削除	広告を削除して欲しいという苦情
広告の不具合	広告が消えない、消えにくいという苦情
広告の内容	広告の内容に対して、不満に感じることへの苦情
不明	広告のことに言及しているが特定できない苦情
アプリケーション内での購入に関する苦情	ガチャなどの課金要素を含むことに対しての苦情
確率の上昇	ガチャなどの課金要素を含むコンテンツに対して確率を上げてほしいという苦情
金額に釣り合わない	課金要素を含むことに対して期待した結果が得られなかったことに対しての苦情
確率表記	ランダムに排出されるコンテンツに対して、確率が表記されていないことへの苦情
無料石などの配布	運営側のユーザーへの配布が少ないもしくははないことへの苦情
課金しないでできない	課金をしないと機能の拡張ができない、運営が課金をさせようとしていると感じる
購入に関する苦情	アプリケーションの購入に対して、金額に釣り合わないと感じることへの苦情
運営への要望	運営に対しての、機能要求とは別の要求
ユーザーの悪質行為	他のユーザーの悪質行為への苦情
他のゲームアプリケーションに似ている	ユーザーが他のゲームアプリケーションに似ていると感じている苦情
ユーザー人口の過剰化	他のユーザーが少ないと感じ、一部機能が使えない、または使いにくいことへの苦情

ルが有料なビジネスモデルでは、購入に関する苦情が多く報告されていた。追加課金が可能なビジネスモデルにおいて、無料+追加課金ではアプリケーション内での購入に関する苦情が最も多く報告された苦情であったが、有料+追加課金では初期の購入に関する苦情の方が報告されやすかった。広告の苦情の中でも広告の表示位置に関する苦情が最も多く報告され、嫌悪されている苦情は広告の内容に関する苦情であった。また、アプリケーション内での購入に関する苦情では、金額に釣り合わない、課金しないときできないという苦情が多く報告され、金額が釣り合わないことが嫌悪されやすい苦情の1つであった。

## 6 考察

### 6.1 海外と日本の苦情の傾向の違い

4章で得られた結果からは、傾向として以下のような特徴があるのではないかと考えた。

- 日本のユーザーは意見を言うことに消極的  
日本のユーザーの特徴として、アプリケーションへの要求が消極的であると考えられる。ユーザーの望む結果が得られなかった時に「魅力のない内容」として報告し、アプリケーションから離れてしまう傾向があると考えた。ユーザーの意見を聞く機会を増やすことでアプリケーションの成功により繋がると考えられる。
- スマートフォンの普及による改善  
本研究は、[1]の3年後の実験であるため、環境が整備され、ユーザーがビジネスモデルに対して適応していると考えられる。「ネットワーク問題」では、ネットワーク環境がより整備されたことでネットワークに接続できないということが低下し、問題の改善につながっていると考えられる。また、「隠されたコスト」では、ユーザーが課金に対しての抵抗が薄れてきたことで苦情として報告する頻度が低下したと考えられる。
- リリース前に十分なテストが必要  
「アプリケーションが応答しない」という苦情の中で多く報告されていたコメントがアプリケーションが開かないということであり、アプリケーションが正常に

表 8 ビジネスモデルにおける上位 3 位の苦情

苦情頻度	順位	完全無料		無料+追加課金		買い切り		有料+追加課金	
		苦情タイプ	中央値 (%)	苦情タイプ	中央値	苦情タイプ	中央値	苦情タイプ	中央値
苦情頻度	1	魅力のない内容	18.78	アプリケーション内での購入に関する苦情	27.77	機能エラー	21.52	購入に関する苦情	20.87
	2	機能エラー	7.53	魅力のない内容	15.91	購入に関する苦情	19.69	強制終了	19.37
	3	広告	6.52	機能エラー	8.33	インターフェース設計	13.72	アプリケーションが応答しない	18.57
苦情影響力	順位	苦情タイプ	星 1:星 2	苦情タイプ	星 1:星 2	苦情タイプ	星 1:星 2	苦情タイプ	星 1:星 2
	1	他のゲームアプリケーションに似ている	15.5:(1)	プライバシーと倫理	38.5	購入に関する苦情	7.18	購入に関する苦情	19
	2	ユーザー人口の過疎化	6.6	ユーザーの悪質行為	30.67	互換性	6.78	互換性	6.84
3	互換性	6	アプリケーション内での購入に関する苦情	9.03	強制終了	4.18	機能エラー	6.37	

表 9 広告に関する苦情

順位	完全無料			
	苦情頻度	苦情タイプ	苦情影響力	星 1:星 2
1	広告の表示位置	24.62	広告の内容	8.67:(1)
2	広告の削除	14.76	広告が多い	7.08
3	広告が多い	6.52	広告の削除	5

表 10 アプリケーション内での購入に関する苦情

順位	無料+追加課金		有料+追加課金	
	苦情頻度	苦情影響力	苦情頻度	苦情影響力
1	金額に釣り合わない	49.63	金額に釣り合わない	15
2	課金しないといけない	18.39	課金しないといけない	4.94
3	確率の上昇	11.81	課金しないといけない	4.94

利用できないことに不満を感じやすい傾向があると考えられる。また、「強制終了」の問題に対しても不満を感じやすいことから、リリース前に十分なテストを行うことで、「アプリケーションが応答しない」、「強制終了」の問題の不満解消につながると考えられる。

## 6.2 ゲームアプリケーションにおけるビジネスモデルの苦情の傾向の違い

各ビジネスモデルのゲームアプリケーションを調査することで以下の特徴を特定できた。

- 完全無料は単純なコンテンツが多い  
完全無料では、広告が収入源であり、開発費にける費用が他のビジネスモデルよりも少ないと考えられる。よって、機能の拡張が少なく単純なゲームアプリケーションになってしまうため、「魅力のない内容」でつまらないというコメントが多く報告されていると考える。また、広告の表示位置について、広告のせいでアイテムが取れないなどのコメントも多くあった。広告の表示される位置が悪いためにゲームプレイに影響を与えている。広告の表示位置についての苦情を減らすために、広告を表示する位置に注意すべきである。
- 無料+追加課金はユーザーの目に留まりやすい  
無料+追加課金のアプリケーションは他のビジネスモデルよりも 10 倍以上、多くのユーザーにダウンロードされていることから、ユーザーの目に留まりやすいと言える。また、無料+追加課金のビジネスモデルの主な収入がユーザーの課金であることから、「アプリケーション内での購入に関する苦情」の報告も多くなっていると考えられるが、アプリケーション内での購入に関する苦情はプライバシーと倫理やユーザーの悪質行為よりもユーザーから嫌悪されていないことから受け入れられていると考えられる。特に、プライバシーと倫理では運営側がユーザーレビューを消去することや、不自然なコメントが投稿されているといったコメントが多く、これらのことについてより重点を置く必要があると考えられる。
- 買い切りでは、インターフェース設計に注意すべき

買い切りのアプリケーションの特徴として、ゲーム機からの移植アプリケーションが多く、本研究で見られたインターフェース設計での苦情として操作しにくいというコメントが多くあった。スマートフォン上での操作性を考えず、ゲーム機の操作をそのまま利用することで操作しにくくなっていると考えられる。

- ユーザーが払った金額に見合わないと批判されやすい  
無料+追加課金、有料+追加課金の「アプリケーション内での購入に関する苦情」での「金額に釣り合わない」や買い切り、有料+追加課金の「購入に関する苦情」の報告が多いこと、ユーザーに嫌われていることからユーザーが支払った金額に見合わないと批判されやすいと考えられる。特に、初期に払うコストについての不満が多かったため、適切なコストに設定する必要があると考えられる。

## 7 まとめと今後の課題

本研究では、日本のアプリケーションのユーザーレビューを調査し、日本のユーザーがアプリケーションに対して報告しやすい苦情、低評価をつけられやすい苦情を調査した。結果からは、[1]の結果と同じ部分、異なる部分が存在し、それらは日本ユーザーの特徴と考えることができる。また、ゲームアプリケーションでのビジネスモデル毎の調査による結果から、各ビジネスモデルの特徴を把握することができた。これらの苦情の傾向は、開発者がアプリケーションを開発、運用する時に参考になる情報だと考える。

今後の課題として、他ジャンルのアプリケーションについての特徴や、苦情と更新の関係性を調査することでより開発者がアプリケーションを開発、運用する際に参考になる情報を提供したい。

### 参考文献

- [1] Hammad Khalid, Emad Shihab, Meiyappan Nagappan, Ahmed E. Hassan : "What Do Mobile App Users Complain About? ", In IEEE Software, Vol.32, No.3, pp.70-77, 2015.
- [2] Sample size calculator - creative research systems : <http://www.surveysystem.com/sscalc.html>