

プログラマー向け QA サイトにおける質問形式の調査・検討

小瀬川博将[†] 佐藤 哲司[‡]

[†] 筑波大学 情報学群 知識情報・図書館学類 〒 305-8550 茨城県つくば市春日 1-2

[‡] 筑波大学 図書館情報メディア系 〒 305-8550 茨城県つくば市春日 1-2

E-mail: [†]{hiromasa,satoh}@ce.slis.tsukuba.ac.jp

あらまし プログラマー向け QA サイトに Stack Overflow がある。Stack Overflow では、質問や回答の質を向上するために、不明確な質問には質問の編集またはコメントを追加できる機能が提供されている。しかし、このような機能では質問解決までに多くの時間と労力が必要となる。本論文では、質問の意図を明確することを目的に未回答質問における記述の不明確さや回答が付けられていない原因を明らかにする。投稿される質問の多くは、動作の不具合や使い方に関する解決法を求めていることから、動作環境や不具合の症状、質問するまでに調べた状況などの情報が不可欠である。これらの情報が十分に記述されているかを調査し、未回答質問と認証済み質問の比較を行った。

キーワード QA サイト, 特徴付け, 質問解析, Stack Overflow

1. はじめに

現代社会において、インターネットを介したコミュニケーションや情報共有は必要不可欠であり、QA サイトを利用することは一つのツールとして知られている。QA サイトとは、電子掲示板上で質問の投稿や解答が利用者によって行われ、質問・回答が蓄積される Web サイトである。また、QA サイトを用いた情報共有は問題解決に有効な手段である。プログラマー向け QA サイトに Stack Overflow がある。Stack Overflow では、回答と質問に関してそれぞれ方針^(注1)を示している。回答に関しては「回答を得ることがこのサイトの全て」、「将来その問題を検索する人のためにもなるような回答ができればベスト」と定めている。質問に関しては「主観性の高い質問は避ける」、「質の低い質問を避ける」と定めている。このようなことから、投稿される質問・回答の数より質を重要視している。また、どの質問に対しても回答が用意されており、利用者が閲覧したとき納得できる情報の共有・蓄積していくことを目指しているといえる。Stack Overflow には良質な質問・回答を残すために内容を編集する機能が付けられている。また、Yahoo!知恵袋における最も良い回答に付けられる「ベストアンサー」に対して、Stack Overflow では「認証済み」が付けられる。認証済みとは、その回答が基準は質問者に役に立ったかどうかであることからベストアンサーとほぼ同等の機能といえる。

本論文では Stack Overflow を扱う。なぜなら、Stack Overflow には良質な質問・回答を残すために内容を編集する機能が付けられているにもかかわらず、Stack Overflow に向かない質問が未解決や未回答のまま残されていることによる。特に、主観性の高い質問や質の低い質問（Stack Overflow に向かないタイプの質問）では編集が必要である。Stack Overflow に向かないタイプの質問の問題解決には質問の編集や内容を編集するようなコメントが行われている。しかし、このような機能では、

質問解決までに多くの時間と労力が必要となる。また、Stack Overflow に向かないタイプの質問と判断された場合、その質問がクローズ（一定の基準を満たしたユーザがクローズ票を投票することで回答が受け付けられなくなる状態）される。このため、未回答質問と認証済み質問にどのような違いがあるのかを明確にすることが課題となっている。また、プログラマー向け QA サイトであれば、未解決や未回答のまま残されている状況の解決は共通の課題だといえる。

本論文は、未回答質問における記述の不明確さや回答が付けられていない原因を明らかにすることを目的とする。質問解析として未回答質問における特徴付けを行い、質問として成り立つには何が書かれてればよいのかを調査する。

具体的には、Stack Overflow において投稿される質問の多くは、動作の不具合や使い方に関する解決法を求めていることから、動作環境や不具合の症状、質問するまでに調べた状況などの情報が不可欠であると考えられる。これらの情報が十分に記述されているかを調査し、回答を得るために有効であるかを検討する。回答が付けられていない原因を明らかにすることによって、利用者間にかかる時間と労力を軽減し、QA サイトをより良い問題解決の場としていくことに貢献する。

本論文の構成を以下に示す。第2章では、研究を行う準備として関連研究について説明する。第3章では、Stack Overflow におけるデータの収集および分析手法について述べる。第4章では、分析により明らかになった結果を示す。第5章で分析で得られる特徴について考察した後、第6章でまとめと今後の課題を述べる。

2. 関連研究

本論文では、未回答質問における質問解析を行うことから、QA サイトの分析に関する先行研究と QA サイトにおける推薦に関する先行研究について説明する。

2.1 QA サイトの分析に関する先行研究

本研究の分析に関する先行研究として、片山ら [6] は、回答

(注1) : <http://ja.stackoverflow.com/tour>

者が実際に回答している質問間にどのような特徴があるかを回答者の履歴に着目し分析を行っている。分析手法として、質問の履歴と質問の文書をベクトル化し類似度計算を行っている。また、使用する質問には実際に回答している質問とそれ以外の質問に分け比較実験を行っている。その結果、実際に回答している質問のほうが回答していない質問よりも類似度が高くなる傾向から、過去に回答した質問(履歴にある質問)と回答する質問では、回答していない質問よりも語の重なりが多いことを明らかにしている。ユーザーの特徴を履歴データから抽出できる可能性を示している。

甲谷ら[4]は、教えて!Gooにおいて、三つのコミュニケーションタイプを定義し教えて!Gooのデータのログを用いて、各カテゴリの分類を行っている。また、コミュニケーションは時間とともに変化することがわかっており、QAネットワークが時間とともにどう成長するのか分析を行っている。評価実験として、各カテゴリを「知識交換」「相談」「議論」という三種類のコミュニケーションタイプに分類した上で、先行研究でわかっているYahoo Answersにおける同様の分析結果との比較を行っている。その結果、一回答あたりの平均語彙数に関してふるまいが異なることを明らかにしている。

栗山[7]は、投稿者の質問・回答(OKWaveの「恋愛・人生相談」カテゴリに投稿された質問・回答)と読売新聞における質問の特徴について比較を行っている。比較では、タイトルと本文に分け、質問中に出現する語の抽出と特徴語の抽出を行い、TF-IDFで重み付けを行っている。扱うデータは読売新聞の「人生案内」の質問とOKWaveの「恋愛・人生相談」の質問であり、それぞれのデータについて形態素解析を行っている。また、質問者がどのような悩み・質問を抱えているかを、インターネット以外の媒体を参照することで、より適切な回答が得られる可能性があるのかについて考察している。考察として、共通点と相違点をまとめている。共通点として、人間関係に関する質問、特に自分のパートナーに関する質問が多く、質問文にはパートナーや身近な人間を表す語や評価や感情を表す語がよく用いられることを明らかにしている。また相違点として、新聞の相談者は女性の割合が高く年齢層があまり偏っていないのに対し、QAサイトの質問者はほぼ男女半々で10~20代の若年層に偏っており、回答者はそれよりも少し上の20~30代が多いことがわかっている。これらの共通点と相違点から相補的に利用することで有用な回答を広く探すことが可能であることを明らかにしている。

Liら[1]は、Research Gate(RG)というQAサイトを使用し、質問の質の特徴付けを行っている。手法としては、SVM分類アルゴリズム、三予測モデルの対照分析、特長の対照分析の分類モデルを提案している。各提案手法に対して、回答品質のための最適化を行った結果、挨拶や他の感情の言葉は実際の回答品質にネガティブな影響があると示している。

2.2 QAサイトにおける推薦に関する先行研究

本研究の背景や課題に関する先行研究として、甲谷ら[5]は、これまでのQAサイトの問題点として回答者は回答したい質問を自分で見つけてくる必要があるという点を解決するために、

回答者に適切な質問を推薦する手法を提案している。QAサイトでは教えて!gooを扱っており、ユーザーをノード、回答をエッジとして、QAサイト上のコミュニケーションをグラフにモデル化したものQAネットワークに着目している。特に、3ノード部分グラフに着目している。推薦手法として、ある時点 t_0 のQAネットワークに含まれる3ノードからなる構造を全て抽出し、回答までの時間を指標に、次の回答が発生しやすいようなユーザーを発見している。提案手法の有効性の評価について、カイ二乗検定を行った結果、上位にランクインされたユーザー間には、実際に回答が発生する傾向があることを明らかにしている。また、提案手法とランダム選択した場合を比較し、回答者に適切な質問を推薦するのに有効である可能性を示している。

石川ら[3]は、QAサイトのベストアンサーを計算機が推定可能かの検証を行っている。推定実験の質問(2~4件の回答付きに限定し、全50件)を作成し、推定結果が人間の推定結果にどの程度接近かの調査を行っている。結果として、「パソコン」においては判定者らの結果を上回っており、「一般教養」「政治」はほぼ同等であり、「恋愛相談」においては判定者らの結果を下回ったという結果を示している。

見市ら[2]は、QAサイト内において質問者の質問意図と検索利用者の検索意図に差が生じるため、求める回答にも差が生じ、結果的に検索利用者の意図した回答が得られないという問題点を解消するために、検索利用者の検索意図は複数存在することに着目し、検索利用者の意図を考慮した回答検索システムを提案している。使用するデータはYahoo!知恵袋を使用している。検索利用者における複数存在する検索意図をS項目と定義しており、二つの提案について比較を行っている。S項目を用いた回答検索手法における、一つ目の提案はベクトル間類似度を用いた回答検索手法である。二つ目の提案は回答抽出による回答検索手法である。Yahoo!知恵袋の検索結果と提案1,2の比較では、一つ目の提案は全ての検索クエリにおいて検索精度の上昇を示している。二つ目の提案は三つの検索クエリにおいて検索精度の上昇を示している。このため、S項目を考慮することでより回答検索の際に検索利用者の検索意図を反映することが可能であることを示している。また、提案システムによる検索利用者の意図を反映した回答の推薦が可能であると示している。

3. 質問形式の分析手法

3.1 使用するデータ

本論文ではStack Overflowが提供するアーカイブデータの分析手法を提案する。使用するアーカイブデータは2015年8月18日に更新されたデータ^(注2)である。Stack OverflowではAppleやCodereview、Unix、Texなど様々な専門性に特化したサイトを複数運営している。本論文ではStack Overflowを構成する個々のサイトをカテゴリと称する。

アーカイブデータはカテゴリ分けされている。表1に、Stack Overflowにおける質問データについて、表2に、Stack Overflowにおける回答データについて示す。また、使用するデータ

(注2) : <https://archive.org/download/stackexchange>

ではアーカイブデータの apple カテゴリとする。アーカイブデータ以外のデータとして Stack Overflow の実際のサイトからコメント文を取得したので、表 1 の下部に追加した。追加した理由として、コメント文は質問の補足情報や意見を受け付けるために作られた機能なのでコメント文を分析することで、未回答質問においてどのような情報が不足しているのかが明らかになると考えたことによる。

表 1 質問における収集データ

ラベル (タグ名)	説明
投稿 Id(Id)	投稿順に付与される Id を表す
投稿タイプ (PostTypeId)	質問か回答かを判断するタイプを表す。タイプが 1 の場合は質問を表し、タイプが 2 の場合は回答を表す
認証済み回答 Id(AcceptedAnswerId)	認証済みとされた回答の Id が付けられる。また投稿タイプが 1 のものに付けられる
投票スコア (Score)	それぞれのユーザ自身の判断で+か-の票を入れることができ、数値が高いほど上位に、低いほど下位となる
質問文 (Body)	質問の本文を表す
最終編集者 Id(LastEditorUserId)	最後に質問を編集したユーザの Id を表す
質問タイトル (Title)	質問のタイトルを表す
タグ (Tags)	質問に関連する単語を最大五つ付けられる。また関連する単語については Stack Overflow が定めている単語となる
コメント数 (CommentCount)	コメントの数を表す
お気に入り数 (FavoriteCount)	お気に入りの数を表す
コメント文 (comment-text(実サイトから取得))	質問者とその他のユーザによる質問に対してのコメントの文

表 2 回答における収集データ

ラベル (タグ名)	説明
投稿 Id(Id)	投稿順に付与される Id を表す
投稿タイプ (PostTypeId)	質問か回答かを判断するタイプを表す。タイプが 1 の場合は質問を表し、タイプが 2 の場合は回答を表す
回答の親 Id(ParentId)	どの質問に対する回答かを表す Id が付けられる。また投稿タイプが 2 のものに付けられる

表 3 2015 年 8 月 18 日における Stack Overflow の質問件数

データ種別	データ数	%
質問総数	60,327	-
未回答質問数	11,553	19.2
認証済み質問数	23,549	39.0

本論文における質問分析について、未回答質問では、表 1, 2 に示す投稿タイプが 1 であるものかつ回答の親 Id が質問における投稿 Id に存在しないものとする。認証済み質問では、表 1 に示す投稿タイプが 1 であるものかつ認証済み回答 id が付いたものとする。また、未回答質問と認証済み質問との比較を行う。

3.2 回答件数に関する調査

Stack Overflow において未回答質問がどのくらいの件数なのか調査する。調査を行う理由としては、未回答質問の件数を調査することで、研究の有用性がより明確となることによる。apple カテゴリにおける全質問数、未回答質問数、認証済み質問数を調査することで、全体の質問数に対して研究対象となるデータ数がどのくらい占めているのかわかる。また、表 3 は、apple カテゴリの質問の件数を示す。

未回答質問が全体のどのくらい占めているのか調査した。この調査を踏まえた上で、質問内容が不明確な質問にはどのような特徴が見られるのかを調査する。

表 3 から、未回答質問数は全体の質問の 19.2% であることがわかる。また、認証済み質問数は全体の質問の 39.0% である。認証済み質問数の割合については、質問全体の約 39% を除いた 61% の質問は問題解決していないや未回答質問の可能性がある。未回答質問数の割合については、質問全体の約 19% が回答が付けられていないので、少なくとも約 19% 中に質問内容が不明確な質問も含まれている可能性がある。

図 1 に apple カテゴリにおける回答のデータ分布（回答が 0 件も含む）を示す。図 1 では回答が 1 件付与されている質問が最も多く、次いで 0 件と 2 件付与されている質問が多い。2 件以降は回答数が増えるたびに減少していく傾向にある。図 1 に示した 0 件が未回答質問である。より良い質問・回答を閲覧するサイトにしていくためには、未回答質問として残されている原因を明らかにする必要がある。

未回答質問の本文と未回答質問に付与されたコメントの分析を行う。また、コメントやプラスマイナスの評価投票やお気に

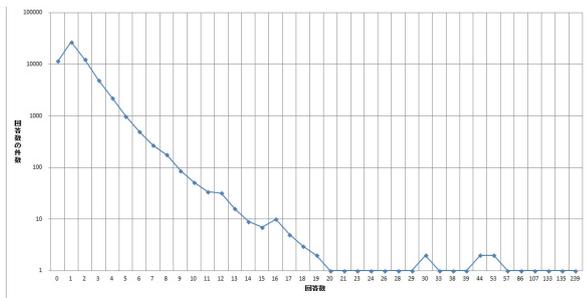


図 1 apple カテゴリにおける回答のデータ分布

入りが付与されている場合でも回答が付与されていなければ未回答質問として扱う。未回答質問は Stack Overflow のアーカイブデータから抽出を行うので、閲覧数が 0 でない質問を対象としている（アーカイブデータに存在しない投稿されたばかりの新規質問は未回答質問として扱わない）。

3.3 未回答質問の特徴抽出

未回答質問の特徴付けを行った。apple カテゴリのデータからランダムに 100 件を抽出する。ランダム抽出では、投稿日時が古いものと新しいものからそれぞれ 50 件ずつ抽出する。未回答質問において着目した点を以下に箇条書きで列挙する。

- 質問文 (Body)
- タグ (Tags)
- コメント文 (comment-text)
- 最終編集者 Id (LastEditorUserId)
- 投票スコア (Score)
- お気に入り数 (FavoriteCount)

未回答質問となる原因を明らかにするために特徴付けを行う。また、箇条書きで示した点に注目し 100 件の未回答質問に特徴付けを行う。未回答質問の特徴として、質問文に環境、症状、状況についての内容が含まれるという仮説を立てる。また、環境、症状、状況の全ての内容が少しずつ含まれている質問より環境、症状、状況のうちどれか一つだけが詳しく書かれている内容の方が未回答質問になるという仮説を立てる。環境、症状、状況について、どのようなことが述べられているとそれぞれに判断されるのかの基準を以下に箇条書きで示す。

• 環境について

開発環境や OS やソフト名が記載されていないと回答が困難であるとする。また、記載されていない場合プログラマー向け QA サイトを使用する際の周辺知識が必要であるとする。

• 症状について

エラーメッセージや問題が起こる前に何をしたかやどのようなコードを書いているのかに記載されていないと回答が困難であるとする。また、コードの変数名や使用されるデータをそのまま使うことについて、Stack Overflow ではサンプルを用意するように定めている。プログラミングの構造についての理解が必要であるとする。

• 状況について

問題解決のために何を調べたのかやどこまで問題を把握しているのかや解決してほしいポイントが明確であるかが記載され

ていないと回答が困難であるとする。また、関連する質問を調べるための方法やどんな質問が関連しているかプログラマー向け QA サイト以外で情報収集する能力が必要であるとする。

4. 未回答質問となる原因の検討

4.1 未回答質問の特徴付け

Stack Overflow が提供するアーカイブデータにおける、apple カテゴリの未回答質問 100 件に対して特徴付けを行った。具体的にどのような特徴に着目したか以下に箇条書きで示す。列挙する項目として、未回答質問の本文、その質問に付与されたコメント、最終編集者 Id の三点を挙げる。

• 質問文

文と文をつなぐときの単語が本文中に出現の有無およびその単語頻度に着目した。質問文の内容を読み、質問に付与されているタグと一致する単語はどのくらいあるのかに着目した。また、?マークや改行・空行の数や他者からの編集がされたのかについて着目した。

• コメント文

何人のユーザからコメントがされているかの件数に着目した。コメント内容に書き加える指示がされているかに着目した。また、別の方法を試しているかどうかの確認がされているかに着目した。最終投稿されているコメントに関して、そのコメントが質問者で?マークで終わっているか着目した。最終投稿されているコメントに着目した理由としては、質問に加えてコメントでも問題が発生し、その問題も解決されていない可能性があることによる。

• 最終編集者 Id

Stack Overflow では投稿した質問や回答やコメントをすべての利用者が編集を行うことができる。編集の有無に関しては、未回答質問にはどのくらい質問者以外のユーザから編集されるか着目した。

4.2 未回答質問の特徴付けを行った結果

表 4 は 100 件中に含まれた各特徴の件数を表す。表 4 に未回答質問と認証済み質問に対する特徴付けの比較を示す。また、表 4 の OS 名、Ver. 名、コード等、関連質問 URL、関連 URL、箇条書き、他者による質問編集が質問文の特徴付けであり、書き加えコメント、別の方法の確認がコメント文の特徴付けである。表 4 に示す関連質問 URL とは、Stack Overflow に投稿された質問で関連している質問の URL であり、関連 URL は Stack Overflow 以外のサイトへの URL である。

4.3 コメントによる補足情報の追加調査

追加調査で用いるアーカイブデータは 2016 年 1 月 5 日に更新されたデータ^(注3)である。追加調査理由として、コメント機能は回答に至るヒントやきっかけの情報が投稿されるためコメントの分析が重要だと考えることによる。また、apple カテゴリには削除された質問や別リンクで回答付き質問に統合された質問があり、サイト運営が未回答質問を減らすような操作が行われている。このため、時間の経過で未回答質問がどのくらい

(注3) : <https://archive.org/download/stackexchange>

表 4 特徴付けした結果の比較

特徴	未回答質問	認証済み質問
OS 名	32	62
Ver. 名	20	28
コード等	16	31
関連質問 URL	5	6
関連 URL	16	11
箇条書き	10	7
書き加えコメント	16	10
別の方法の確認	13	4
他者による質問編集	21	50

変化するか追加調査を行った。

2015 年 8 月 18 日～2016 年 1 月 5 日の変化として、3 件未回答質問に回答が付与された（未回答：78 件→75 件，回答付：2 件→5 件）。また，当初特徴付けに想定していなかった情報が追加調査によりわかった。表 5 はに示す。

表 5 特徴付けに必要なコメント文

補足情報	類型
どうやってインストールしたのか	不具合に至る過程の確認
この方法はやってみたか	方法の提案
その質問はどういうこと	質問内容の確認 または質問者と コメント者の相互理解
OS は何かなどの追記要請	質問状況の確認

5. 未回答質問の特徴についての考察

5.1 特徴付けした結果の比較についての考察

未回答質問の本文と未回答質問に付与されたコメントと未回答質問に対して付けられた評価の考察を行う。

前節で，未回答質問より認証済み質問の方が OS 名やコード，関連質問 URL が多いことがわかった。このことから未回答質問にとって OS 名やコード，関連質問 URL は，必要な情報であると考えられる。

未回答質問に付与されたコメントに関しては，回答者になりうるユーザから開発環境はどのようになっているか，OS は何か，バージョンは何かを質問文に追記させるコメントや質問解決に近づく作業は既に行っているかの確認がされていることがわかった。このことから，未回答質問には関連する情報が不足すると回答することが困難だと考えられる。

他人による質問編集では，未回答質問より認証済み質問の方が多くことがわかった。このことから，認証済み質問は必要な情報を含むため，ユーザが積極的に参加したと考えられる。

5.2 コメントによる補足情報の追加調査についての考察

補足情報の追加調査では 2015 年 8 月 18 日～2016 年 1 月 5 日の約 5 カ月間の変化として，3 件未回答質問に回答が付与されたことがわかった。このことに関して，本研究の特徴付けのみでは困難であるため，どのようなプロセスを経て変化したか

コメントや質問者の対応状況の類型化などの調査が必要だと考えられる。

6. 結論

本論文は，プログラマー向け QA サイトである Stack Overflow における未回答質問の特徴付けを行い，質問として成り立つには何が書かれていれば良いかを調査し検討した。また，apple カテゴリにおけるアーカイブデータの 2015 年 8 月 18 日～2016 年 1 月 5 日期間の追加調査を行った。

質問として成り立つには何が書かれていれば良いかを調査したことによって，未回答質問における不明確さは関連する情報が不足していることによるものだとわかった。具体的には，未回答質問には関連する情報が不足することにより，動作環境や不具合の症状，質問するまでに調べた状況の情報が不可欠であるということがわかった。また，これらの情報が回答を得るために有効であるかを検討した。

未回答質問における約 5 カ月間変化として，3 件回答が付与され，どのようなプロセスを経て変化したかについて別の視点を踏まえる必要があるとわかった。

最後に，未回答質問における記述の不明確さや回答が付かない原因が明らかになったので，今後プログラマー向け QA サイトに質問を投稿するユーザや質問を投稿したが回答が付かないユーザに対して，どのように不明確なのかの確認や具体的な原因の提示を行うことができる。そこで，今後の研究として未回答質問における内容の不明確さや回答が付与されていない原因に基づき，質問解決支援を行う。質問解決支援とは，質問内容が不明確であっても Stack Overflow からベストアンサー付き質問・回答を質問仕方事例集として提示することである。

謝辞

本研究は，JSPS 科研費 25280110 および 25540159 の助成を受けたものです。ここに記して謝意を示します。

文献

- [1] Lei Li, Daqing He, Wei Jeng, Spencer Goodwin, and Chengzhi Zhang. Answer quality characteristics and prediction on an academic Q&A site: A case study on researchgate. In *Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion*, pp. 1453–1458, 2015.
- [2] 見市高一，川越恭二. QA サイトにおける s 項目による回答検索システム. Technical Report 25, 立命館大学大学院，立命館大学情報理工学部，nov 2010.
- [3] 石川大介，栗山和子，酒井哲也，関洋平，神門典子. Q&A サイトにおけるベストアンサー推定の分析とその機械学習への応用. 情報知識学会誌，Vol. 20, No. 2, pp. 73–85, may 2010.
- [4] 甲谷優，川島晴美，藤村考. QA サイトにおける質問応答グラフの成長パターン分析. 日本データベース学会論文誌，Vol. 7, No. 3, pp. 61–66, dec 2008.
- [5] 甲谷優，川島晴美，藤村考. QA コミュニティの成長パターンに基づく回答者への質問推薦. 日本データベース学会論文誌，Vol. 8, No. 1, pp. 89–94, jun 2009.
- [6] 片山亮，川村秀憲，鈴木恵二. QA サイトにおける質問推薦へ向けた履歴データの分析 (社会システムと情報技術研究ウィーク). 電子情報通信学会技術研究報告. AI, 人工知能と知識処理, Vol.

109, No. 439, pp. 11–16, feb 2010.

- [7] 栗山和子. 新聞と Q&A サイトにおける人生相談の分析. 電子情報通信学会技術研究報告. 研究報告デジタルドキュメント (DD) , Vol. 2015, No. 8, pp. 1–8, mar 2015.