Master Thesis

Twitter User Classification Based on Specificity of their Information Dissemination Target

Supervisor Professor Keishi Tajima

Department of Social Informatics
Graduate School of Informatics
Kyoto University

Hikaru TAKEMURA

February 7, 2014

Twitter User Classification Based on Specificity of their Information Dissemination Target

Hikaru TAKEMURA

Abstract

With the widespread of social media, such as blogs and social network services (SNS), today we have become able to publish information on the Web more easily than previously. Especially recently, microblogging services, which are a kind of blogs coupled with some characteristics of SNS, has been growing explosively. As of 2013 Octover, Twitter, which is one of the most popular microblogs, has over 218 million active users in the world, and as of June, more than 500 million messages are posted on it per day.

Twitter has many characteristics of conventional social medias, so it is used for many purposes. Some publish information to the wide public, some publish information on some specific topics, and some communicate with their friends. Because of this characteristic, Twitter has attracted great attention as a new type of social media.

As explained above, Twitter is used for various purposes. As a result, target specificity, the wideness of target scope of information publishing, varies greatly among users. So, in this study, we propose a method to classify Twitter users based on target specificity of their information publishing. In the method, we focus on the followers of the user, and classify him whether his followers are consistent in some noticeable characters or not. If his followers are consistent in some noticeable characters, it suggests that his target specificity is high, i.e., he publishes information in which particular users are interested. On the other hand, if they are not consistent in any noticeable characters, there is a high possibility that the user is followed by a wide variety of users and it suggests that his target specificity is low, i.e., he publishes information in which the public is interested.

In addition, in this study, we focus on Twitter users whose target specificity is high, and we propose the method to determine why their target specificity is high. In a large number of Twitter users, their target specificity is high and the causes vary from user to user. For example, users publishing technical in-

formation about programming is supposed to publish information to unspecific users, but their target specificity is considered high because the topic of their publishing information is specified to certain users, i.e., users who are interested in programming. And also, users communicating with their friends or announcing to club members are supposed to publish information to the users specified extensionally. So their target specificity is supposed to be high, regardless of contents of their publishing information. In the method, first, we roughly classify causes of target specificity into two categories: (1) because they publish information specified for certain topics, and (2) because they publish information to the users specified extensionally. Then we construct classifiers which determine whether users only belong to the category (1), only belong to the category (2), or belong to both (1) and (2), based on various features which correlate with each category.

On the Web, it is hard to know what kinds of users each Web page targets to. In Twitter, however, we can guess what kinds of users each user targets to by examining the followers of the user. By using this information, we can determine whether a given user has high target specificity or not.

There are many existing studies on the classification of Twitter users. These studies, however, do not concern about target scope of information publishing. But Twitter is used for various purposes, so the wideness of target scope of their information publishing varies greatly among users. Following this observation, we propose a new classification scheme of Twitter users that focuses on target specificity of their information publishing.

The classification scheme of this study is supposed to apply to Twitter search. In current Twitter search, messages in search results have various target scope of information publishing. So it frequency happens that messages of certain target scope we need are buried in many other messages. At that time, by using the classification of this study, we can search messages based on what kinds of users they target to. So we can prevent messages we need from being buried in many other messages and find messages we need easily.

We implemented our methods with using real Twitter data, and our experimental results confirmed that our proposed method is effective.

情報発信の対象限定性に基づく Twitter ユーザの分類

竹村 光

内容梗概

ブログや SNS というソーシャルメディアの普及に伴い，誰もが簡単に Web 上で情報を発信できるようになった．特に近年では，マイクロブログと呼ばれる，SNS の性質を併せ持ったブログサービスが，爆発的な成長を遂げている．最も普及しているマイクロブログの 1 つである Twitter では，2012 年 12 月現在，ユーザ数が 2 億人を超えており，同年 6 月現在，1 日に 4 億以上もの記事が投稿されていると言われている．

Twitter は，従来の多くのソーシャルメディアの性質を兼ね合わせており，その利用目的が多岐に渡っている．社会のニュースのように，広く一般のユーザが興味を示すような情報を発信するユーザもいれば，あるトピックに特化した情報を発信するユーザや，友人とのコミュニケーションを行うユーザもいたりと，多様な利用目的のユーザが混在している．このような性質のため，Twitter は現在，新たな情報発信メディアとして大きな注目を集めている．

上記のように，Twitter は多様な目的で利用されるため，ユーザによって情報発信の対象範囲が様々である．本研究では，この点に着目し，Twitter ユーザを，広く一般のユーザが興味を示す情報を発信するのか，それとも一部のユーザのみが興味を示す情報を発信するのかという観点から分類する手法を提案する．提案手法では，ユーザのフォロワーに着目し，フォロワー内に何か一貫した傾向あるかどうかに基づいて分類を行う．もしあるユーザのフォロワー全体が何か一貫した傾向を持っていれば，そのユーザは，ある特定のユーザが興味を示す記事を投稿していると考えられる．逆に，フォロワー内の各ユーザの傾向がばらばらであれば，そのユーザは多様なユーザからフォローされている可能性が高く，広く一般のユーザが興味を示す記事を投稿していると考えられる．

さらに本研究では，上記の分類の結果，対象が一部のユーザに限定されていると判定されたユーザに関して，どのような要因でその対象範囲が限定されているのかを判定する手法を提案する．情報発信の対象範囲が限定されているユーザは数多く存在するが，その要因は，ユーザによって様々である．例えば，プログラミングに関する技術的な情報を発信するユーザは，情報自体は不特定のユーザに向けて発信しているものの，その内容の専門性から対象範囲が限定さ

れていると考えられる．また，友人とのコミュニケーションを行うユーザや，ある特定のクラブメンバーに向けてアナウンスを行うユーザなどは，その内容にかかわらず，クローズドなコミュニティを対象に記事を投稿しているため，対象範囲が限定されていると考えられる．提案手法では，まず，対象範囲が限定される要因を，(1) あるトピックに特化した情報を発信している，(2) ユーザを外延的に特定して情報を発信している の大きく 2 つに分類する．そして，対象範囲が限定されている要因が (1) のみであるのか，(2) のみであるのか，それとも (1) と (2) の両方の要因を併せ持つのかを判定する識別器を構成し，それらの要因を特徴付ける様々な特徴量を学習させることにより，判定を行う．

Web では，各ページがどのようなユーザを対象としているのかは分かりにくい．しかし，Twitter では，そのフォロー関係を利用することで，あるユーザの記事がどのようなユーザに読まれているのかが分かり，この情報を活用することで，このように情報発信の対象範囲の広さに基づいて分類することができる．

Twitter ユーザの分類に関する研究は盛んに行われている．通常，これらの研究では，情報発信の対象範囲については考慮していない．しかし，Twitter は多様な目的で利用されるため，ユーザによって情報発信の対象範囲が様々である．そこでわれわれは，情報発信の対象範囲の広さに着目することで，既存の研究とは異なる観点から分類を行った．

本研究で実現する対象範囲によるユーザ分類は，Twitter 検索に応用できると考えられる．現在の Twitter 検索は，様々な対象範囲の記事が検索結果に混ざってしまっているため，自分の求めている対象範囲の記事が他の多くの記事に埋もれてしまうといった事態が頻繁に発生する．そこで，本研究で実現するユーザ分類を利用することで，どのようなユーザを対象としている記事がほしいのかを基に検索を行うことが可能になり，上記のような事態の防止につながる．

提案手法を実装することで評価実験を行い，精度を測定した結果，提案手法が有効であることを確認した．また，その結果を基に考察を行った．

Twitter User Classification Based on Specificity of their Information Dissemination Target

Contents

Chapter 1 Introduction

With widespread use of social media, such as blogging services or social network services (SNS), today we can publish information on the Web more easily than before. Recently, microblogging services have especially been growing rapidly.

Microblogs are a new type of services which have both characteristics of blogs and SNS. In microblogs, users can post short messages more easily and quickly than in conventional blogs or SNS. Microblogs are not necessarily regarded as media for publishing useful information to the public, and many microblog users post messages more casually than in conventional blogs or SNS. Because of these characteristics of microblogs, a large number of messages are posted every day, and the messages contain various types of contents, from personal notes or life logs to useful information or discussion on specific topics.

Among many microblogs, Twitter[1)] is especially growing explosively. As of 2013 Octover, Twitter has over 218 million monthly active users in the world, and more than 500 million messages are posted on it per day[1]. In Twitter, users can post short messages with at most 140 characters, which are called tweets. The most distinctive feature of Twitter is its mechanism of *"follow"*. In Twitter, if a user follows other users, all tweets by these followee users are retrieved in real time, and are shown in a list sorted in the reverse chronological order, as shown in Figure 1. This list is called the *"timeline"* of the follower users. The mechanism of follow is more casual than user-linking functions in ordinary SNS; it does not basically require the permission by the followee, and does not necessarily imply reciprocal relationship. Another important function in Twitter is the *"reply"* function, by which a user can post a message as a reply to another user. By using this function, users can use Twitter for conversation, as in instant messaging services.

As Twitter has characteristics similar to various conventional social media, so it is used for many purposes. Some users, e.g., Twitter accounts owned by major news companies, publish general information on various topics to the wide public, while some users publish information specific to certain topics. There are

1) `http://twitter.com/`

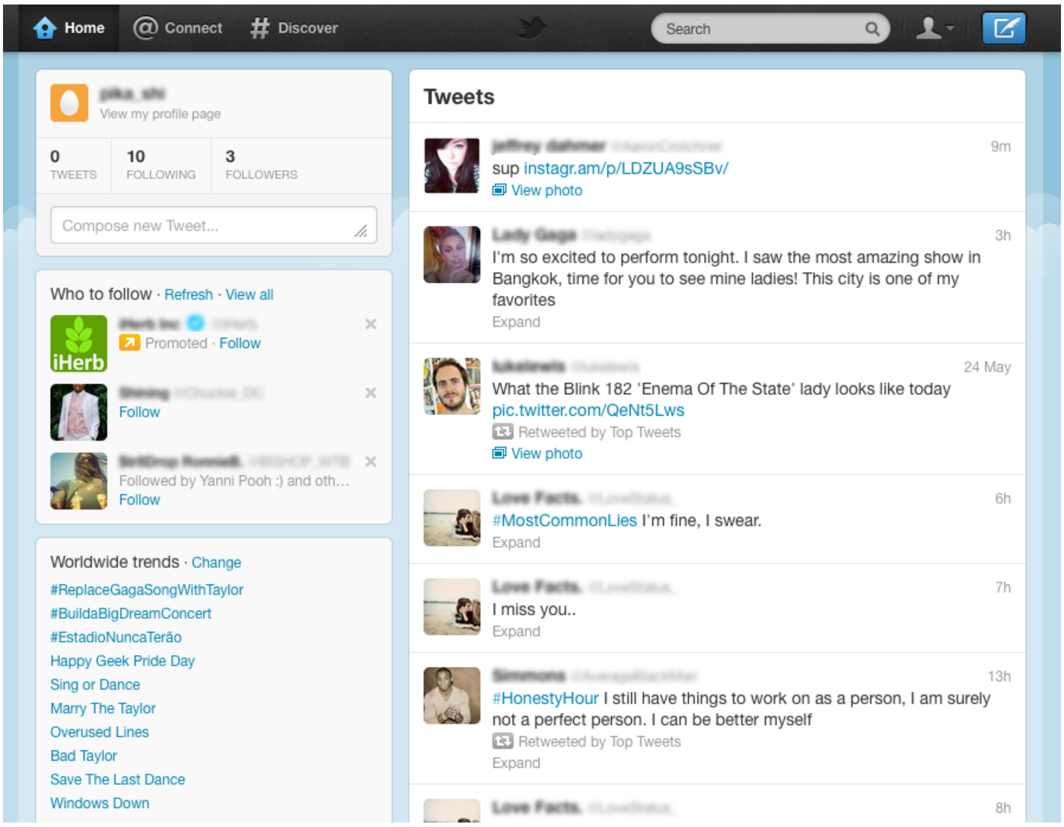

Figure 1: An example of a user's timeline in Twitter

also users that use Twitter for the communication with their friends or others. Because of this characteristic of Twitter, it has attracted great attention as a new type of social media.

As explained above, Twitter is used for various purposes. As a result, the breadth of target scope of information publishing varies greatly from users to users. In this study, we propose a method to classify Twitter users based on how broad the target scope of their information publishing is, i.e., whether they publish information to the wide public or publish information to specific groups of users. In the former case, we say *"target specificity is low"*, and in the latter case, we say *"target specificity is high"*.

In our classification method, we use the information on the consistency among the followers of each user. If most of the followers of a user have some noticeable characteristics in common, we determine that the user publishes information only to users that have some specific interests or characteristics. On the other hand, if the followers of the user have no common noticeable characteristics, it means that the user is followed by a wide variety of users, and we determine that the user publishes information to the wide public.

In this study, we also focus on the causes of such target specificity. We propose a method of determining why target scope of a given user, which has been classified as a user with high target specifity, is restricted to some specific type of users. A large portion of Twitter users have high target specificity, and the causes of their target specificity vary from users to users. For example, some users use Twitter for the communication with the friends, or for announcing some information to members of a certain organization. Messages by such a user may include various topics and may not be restricted to some specific topic, but they intend to publish information only to specific users. Therefore, their target specificity is high. On the other hand, a user who publishes technical information about computer programming does not have specific users in his mind, but their messages include only specific topics, i.e., programming, and they publish information only to users who are interested in programming. Such a user also have high target specificity.

We roughly classify the causes of target specificity into two types: (1) topic specificity, i.e., target specificity of a user is high because the user publishes information on some specific topics, and (2) user specificity, i.e, target specificity of a user is high because the user publishes information to a specific group of users. A "specific group of users" means a set of users which can be defined extensionally, such as friends of some users or members of some organization. A set of users which can be defined only intensionally, e.g., a set of users that are interested in programming, is not regarded as a "specific group" here.

We classify users based on the causes of their high target specificity by constructing a classifier which determines whether a user belongs only to category (1), only to category (2), or belong to both category (1) and (2). Our classifier uses various features of users which correlate with each category. Figure 2 shows main components of our method and information flow among them.

On the Web, it is hard to know what kinds of users each Web page targets to. In Twitter, however, we can guess what kinds of users each user targets

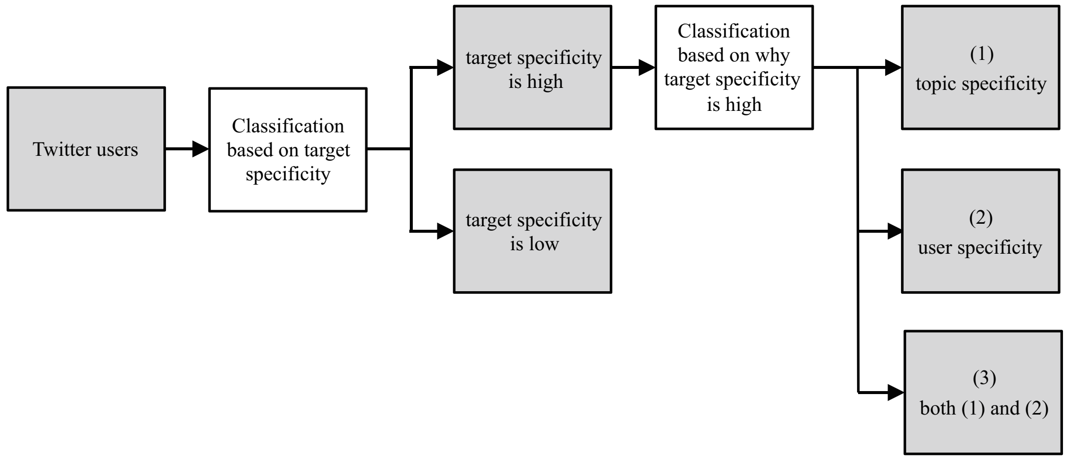

Figure 2: Main components of our method and information flow among them

to by examining the followers of the user. By using this information, we can determine whether a given user has high target specificity or not.

Twitter user classification by our method can be utilized in Twitter search systems. In current Twitter search systems, we input query keywords and receive messages including the keywords. In these systems, however, a search result includes messages with various target scope. As a result, it frequently happens that messages relevant to the querying user are buried in many messages whose target scope do not include the querying user. For example, when a user submits a query "MacBook Air" to the Twitter search system, what kinds of messages he needs depends on the situation. The user may need public news about MacBook Air, may need technical information about MacBook Air, or may need ordinary users' reviews of MacBook Air in order to decide to or not to buy a MacBook Air. In the current Twitter search systems, however, these different types of messages are mixed in a search result. In such a case, we can use our Twitter user classification method in order to classify messages in the search result, so that messages that the user wants are not buried in many other types of messages.

The contribution of this paper is summarized as follows.

- We propose a new classification scheme of Twitter users which is based on target specificity their information publishing.
- We show a method of classifying Twitter users in the scheme above.

- We also show a method of determining the causes of high target specificity for users that are classified as users with high target specificity.

The rest of this paper is organized as follows. Next chapter explains some related work and clarify the relationship between our study and the related work. We then define target specificity of information publishing and formulate our problem in Chapter 3. In Chapter 3, we also discuss why target scope of some user is restricted to some specific type of users. In Chapter 4, we explain the method of classifying Twitter users based on their target specificity of information publishing. In addition, we explain the method of determining the causes of target specificity of the users classified as high target specificity users in Chapter 5. Then in Chapter 6, we show the results obtained from experiments we conducted to evaluate our methods. Chapter 7 concludes the paper.

Chapter 2 Related Work

With explosive widespread use of microblogs, studies about them have recently become frequently performed. There are various studies of microblogs, e.g., studies of the classification microblogging messages from various point of view[2], studies of ranking microblogging messages by the content relevance and so on[9], or studies of focusing on real-time nature of migroblogs[25, 21].

In this study, we focus on the fact that Twitter is used for various purposes. We attempt to classify Twitter users based on the breadth of target scope of their information publishing, and apply this classification scheme on Twitter search and so on. We explain the following previous studies: studies about the purpose of use of Twitter in 2.1, studies of classifying Twitter users and measuring their influence in 2.2, studies useful for finding messages related to only certain Twitter users in 2.3, and studies about Twitter search in 2.4. We make the position of this study clear by introducing these studies.

2.1 Purpose of Use of Twitter

There has been many studies about the purpose of use of Twitter. Java et al.[14] analyzed the topological and geographical structure of Twitter's social network and attempted to understand the user intentions and community structure in microblogging services. As a result, they found that the main types of user intentions are daily chatter, conversation, sharing information and reporting news. Kwak et al.[16] reported that Twitter is used both as a social network service and as a media for disseminating or gathering information, and in its follower-following topology analysis they have found a non-power-law follower distribution, a short effective diameter, and low reciprocity, which all mark a deviation from known characteristics of human social networks. There are many more studies about the purpose of use of Twitter[28, 30].

Ehrlich et al.[10] conducted a content analysis and examined the use of public microblogs (Twitter) for public and private use by comparing internal microblogs (in the workspace). As a result, there were significant differences in content. The internal microblogs were generally used to solicit technical assistance or as part of a conservation, and the public microblogs were used for status updates and to share general information.

In recent years, Twitter, one of public microblogs, is often used for not only publishing information to the public but also having a relationship to only a certain community. It is able to be said that this study focuses on the fact that we use Twitter for various purposes.

2.2 Classification of Twitter Users and Measuring their Influence

There are many studies focusing on Twitter users, e.g., studies which classify them from various point of view, and studies which measure their influence.

Studies focusing on the classification of Twitter users are performed frequently and they have a wide variety of classification schemes, e.g., the classification based on their attributes such as political orientation or ethnicity by leveraging observable information such as the user behavior, network structure, and linguistic content of their posting messages[24], the classification into spam users or not by extracting observable features from the collected candidate spam profiles, e.g., number of friends, text on the profile, age, and so on[17], and the classification into human users, bots, and cyborgs using entropy measures, machine learning, and so on[8]. Bots refer to automated programs posting on Twitter, and cyborgs refer to either bot-assisted humans or human-assisted bots, i.e., interweave characteristics of both humans and bots.

The classification scheme proposed by Yan et al.[29] deeply relates to ours. They proposed methods to classify Twitter users into open accounts and closed accounts. An open account is the account with a purpose for advertising or spreading information such as a shop, a singer, a news agency, and so on. On the other hand, a closed account is the account with a purpose for making friends or communication such as a user publishing messages about daily log, feeling show, and so on. This classification scheme is close to ours, but does not coincide with ours because open accounts do not often publish information to the wide public. For example, a user publishing technical information about programming to unspecific users is an open account though he is a target user.

In microblogs like Twitter or other social network services like Facebook, users correspond to nodes in social network graphs. As well as the classification of users, i.e., nodes in the graphs, the classification of edges, i.e., relationship between a user and his followers, is related to our study. Leskovec et al.[18, 19] classified edges in SNS into positive edges such as friendship, and negative edges such as antagonism. Kunegis et al.[15] also use positive edges and negative edges in Slashdot, a message board service, in order to rank the users. Cheng et al.[7] and Hopcroft et al.[12] studied the problem of predicting reciprocity between two given Twitter users.

There are also studies focusing on measuring influence of Twitter users. Jianshu et al.[27] focused on the problem of identifying influential users of microblogs. Cha et al.[6] analyzed the influence of them by employing three measures that capture different perspectives: indegree, retweets, and mentions. Then they measured the dynamics of influence across topic and time. If target specificity of Twitter users defined in this study is high, there is a high probability that they have a big influence on Twitter, but how low target specificity of a user is does not necessarily coincide with how big his influence is.

2.3 Find Messages Related to Certain Twitter Users

There are also studies useful for finding of microblogging messages related to only a part of Twitter users.

Sakaki et al.[11] proposed a method of monitoring messages in Twitter and detecting occurrences of a specific kind of event in the real world, such as earthquakes or typhoons. They produced a probabilistic spatiotemporal model for the target event that can find the center and the trajectory of the event location. Ikawa et al.[13] attempted to discover the location where a message was generated by using its textual content. They learned associations between a location and relevant keywords from past messages, and guessed where a new message came from. It is able to be said that these studies are useful for finding messages in Twitter related to certain geographical areas.

Sriram et al.[5] proposed approach which effectively classifies the message to a predefined set of generic classes such as News, Events, Opinions, Deals, and

Private Messages. They proposed to use a small set of domain-specific features extracted from the user's profile and text. Nishida et al.[23] proposed a method which uses data compression for classifying an unseen tweet as being related to an interesting topic or not. It is able to be said that these studies are useful for finding messages to some specific topics in Twitter.

As mentioned above, there are many studies useful for finding microblogs related to only a part of Twitter users in various points of view. But these points of view exist in great number, so it is not an effective approach to find these messages from each point of view. Thus in this study, we attempted to measure Twitter users' target specificity of information publishing in an integrated way. In addition, we roughly classified various causes of target specificity into two categories.

2.4 Twitter Search

The characteristic of search on microblogs is different from that of Web search[3] in that search on microblogs can get information in real time[4] and not only information published by the mass media but also much casual information published by individuals[14]. Thus, a purpose of use of search on microblogs often becomes a subject of study.

Teevan et al.[26] observed that people use Twitter search to find temporally relevant information, e.g., breaking news, real-time content, and popular trends, and information related to people, e.g., content directed at the searcher, information about people of interest, and general sentiment and opinion. Furthermore, they compared Twitter search with Web search and found that search results on Twitter included more social chatter and social events, and those on the Web included more basic fact and navigation content. Massoudi et al.[20] proposed a retrieval model for searching messages on microblogs for a given topic of interest and a dynamic query expansion model for messages retrieval. And Nagamoti et al.[22] described several strategies for ranking messages of microblogs in a real-time search engine.

As mentioned above, there are many studies about search on microblogs, and contents of them are greatly various. In this study, we focus on the purpose of use of microblogs, and attempt to apply it to Twitter Search. It is able to be said that this study also focuses on search on microblogs just like studies explained above, but there has not been studied based on Twitter user's target specificity of information publishing so far.

Chapter 3 Target Specificity of Twitter Users

In this chapter, we discuss target specificity of Twitter users, the measure of to what extent target scope of their information publishing is specific, and define it. Then we also discuss why target scope of their information publishing is specific.

3.1 Definition of Target Specificity

In this study, we consider target specificity of Twitter users, as the measure of to what degree target scope of their information publishing is specific. More formally, we define *target specificity* of a Twitter user as to what extent the user set supposed to be included in target scope of his information publishing is inclining to a part of all Twitter users, i.e., to what extent this user set deviates from the user set randomly sampled from all Twitter users. In this paper, we express target specificity of the Twitter user u as $\mathit{TargetSpecificity}(u)$. This formula takes a range of $[0, 1]$.

For example, a user mainly publishing technical information about programming is supposed to publish information to programmers. So users who are interested in this information are inclining toward a part of Twitter users. Thus, it is supposed that target specificity of the user is high.

On the other hand, a user publishing information about world news publishes information useful for the wide public. So the public is supposed to be interested in this information, and the deviation between users who are interested in this information and users randomly sampled from all Twitter users may be very small. Thus, it is supposed that target specificity of the user is low.

As mentioned above, target specificity of Twitter users is defined as to what extent the user set supposed to be included in target scope of his information publishing deviates from the user set randomly sampled from all Twitter users. Thus, the fact that target specificity of a user is high does not necessarily coincide with the fact that there is high similarity between users supposed to be included in target scope of his information publishing each other. For example, users supposed to be included in target scope of information publishing of a user publishing information about earthquake in a certain area are probably consistent in the area they live in, and so his target specificity is supposed to be high. But their other characteristics, e.g., age, sex, interests, communities they belong to, and so on, vary from user to user. Thus it is not be able to be said that they have high similarity each other. In other words, we consider that even if there are various types of users in target scope of his information publishing, target specificity of the user is high as long as the majority of users in the target scope are consistent in at least one attribute.

In this paper, we determine a threshold δ. If target specificity of a user is higher than δ, we call him a *target user*, and if lower, we call him a *non target user*. More formally, we define them as follows:

$$\begin{cases} u \text{ is a } \textit{target user}, & \text{if } \textit{TargetSpecificity}(u) > \delta \\ u \text{ is a } \textit{non target user}, & \text{otherwise.} \end{cases}$$

3.2 Causes of High Target Specificity

In this subchapter, we discuss what causes target specificity of a user, i.e., why target scope of his information publishing is specifiic. As a result of our analysis, this is roughly classified the causes into two types: (1) because he publishes information on some specific topics, and (2) because he publishes information to a specific group of users. We discuss the causes of the target specificity in the follow.

(1) Topic Specificity

The first cause of target specificity of a user is because he publishes information to some specific topics, regardless of whether he publishes information to a specific group of users or not. In this case, we say "topic specificity". For example, a user who publishes technical information about computer programming does not have specific users in his mind, but their messages include only specific topics, i.e., programming, and they publish information only to users who are interested in programming. Furthermore, a user mainly publishing information about a certain conference is supposed to publish information to the users who attend the conference or are interested in it. Thus, it is considered that target scope of his information publishing is specific.

The way to specify topics of information is roughly classified into two cases. In the first case, a user specifies topics based on demographic data, e.g., age, settled areas, sex, occupation, career, and so on. It is able to be said that a user publishing weather information in a certain area specifies topics based on demographic data. In the second case, a user specifies topics based on psychographic data e.g., taste, hobby, values, and so on. It is able to be said that a user publishing information about cooking specifies topics based on psychographic data.

(2) User Specificity

The second cause of target specificity of a user is because he publishes information to a specific group of users, regardless of whether he publishse information on some specific topics or not. In this case, we say "user specificity". A "specific group of users" means a set of users which can be defined extensionally, such as friends of some users or members of some organization. A set of users which can be defined only intensionally, e.g., a set of users that are interested in programming, is not regarded as a "specific group" here. For example, a user communicating with his friends publishes various contents of information, but it is considered that target scope of his information publishing is specific because he specifies a group of users, i.e., he publishes information to the closed group of users, such as his friends. Furthermore, a user mainly getting in touch with members of a certain club publishes information to the closed groups of users, i.e., the club members. Thus, it is considered that target scope of his information publishing is specific.

Sometimes, both (1) and (2) cause target specificity of a user simultaneously. For example, a user publishing information to the members of the artist's fan club publishes information specified not only users of his information publishing, i.e., the members of the artist's fan club, but also the topic of information, i.e., the latest news about the artist and so on. Furthermore, it is also true in case of a user notifying students in a certain university of the news toward them because he publishes information on some specific topics, i.e., the news toward them, to a specific group of users, i.e., students of the university. In addition, some users use Twitter for the both purpose of publishing information of a certain topic and communicating with their friends. It is able to be said that such users are also an example of the case that both (1) and (2) cause target specificity of a user simultaneously.

Chapter 4 Classifying Users Based on Target Specificity

In this chapter, we explain the method of classifying Twitter users based on target specificity of their information publishing defined in Chapter 3.

4.1 Assumptions and Outline of the Method

In this study, we assume that a follower set of a user is the user set randomly sampled from users included in target scope of his information publishing, and focus on the follower set of a user we intend to classify.

A user publishing information to the wide public, e.g., a user publishing information about world news, is supposed to be followed by various types of users. On the other hand, followers of a user publishing information specified in certain users are supposed to be consistent in a certain noticeable character. For example, a user publishing technical information about programming is supposed to be mainly followed by programmers, and a user communicating with his friends is supposed to be mainly followed by his friends.

Based on the above, we classify a user whether his followers are consistent in a certain noticeable character and difficult to suppose to be randomly sampled from all Twitter users, or his followers are not consistent in any noticeable characters. Figure 3 (a) shows the case of followers being consistent in the noticeable character A. In such a case, his followers are supposed to incline toward a part of all Twitter users, thus we consider that the more consistent his followers are in a certain noticeable character, the higher his target specificity is.

In addition to this parameter: whether followers of a user are consistent in a certain noticeable character or not, we consider whether his follower set are covered with consistency subsets covering it intermediately widely. Here, a *consistency subset* denotes a subset which have consistency in a certain noticeable character. Figure 3 (b) shows the case that half of followers are consistent in the noticeable character A and the others are consistent in the character noticeable B. It is not able to be said that they are consistent in one noticeable

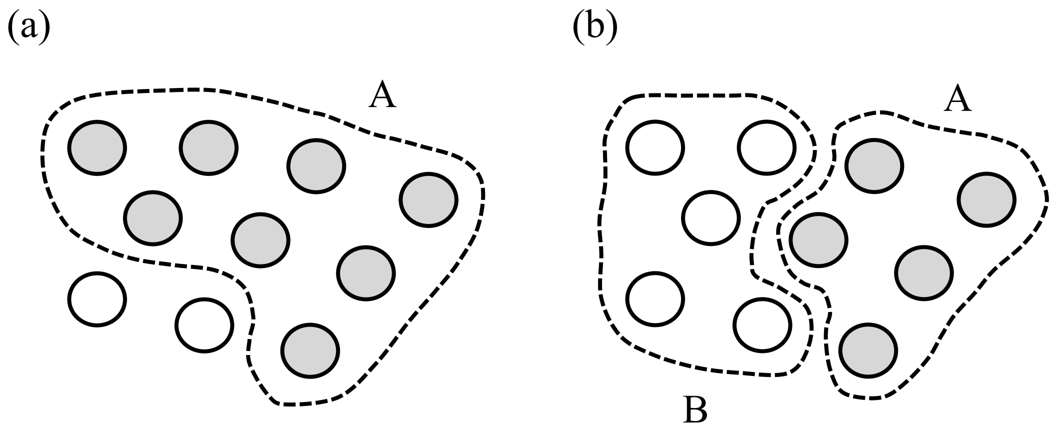

Figure 3: Two examples of high target specificity

character, but his follower set are covered with two consistency subsets covering it intermediately widely. In such a case, we consider that his target specificity is high.

These two parameters are summarized as follows:

(a) The more consistent followers of a user are in a certain noticeable character, the higher his target specificity is, and

(b) The more covered his follower set is with consistency subsets covering it intermediately widely, the higher his target specificity is.

Target specificity of a user is quite high when his followers are consistent in one noticeable character, and it becomes lower as they do not become consistent in any noticeable characters.

4.2 Classification Algorithm

In this subchapter, we explain the algorithm of computing a score of target specificity of a user u based on the outline mentioned in the above subchapter.

First, we collect all consistency subsets included in F_u, the follower set of u. Then, in regard to each subset $S_{F_u c}$, which is consistent in the character c, we compute $SubsetScore(S_{F_u c})$ which denotes to what extent users in $S_{F_u c}$ are consistent in c. We will propose two models computing $SubsetScore(S_{F_u c})$ in 4.3.

Second, in the descending order of $SubsetScore(S_{F_u c})$, we give this score to

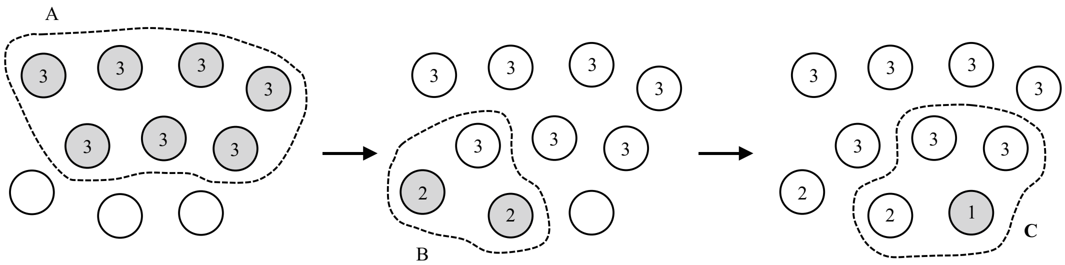

Figure 4: A case of the algorithm flow

each follower f in $S_{F_u c}$ as $\mathit{UserScore}_{att}(f)$, where att means a attribute measuring consistency, and we will explain two attributes in 4.4. Here, we do not give a score to f if he already has a score. Then, we repeat this step over all $\mathit{SubsetScore}(S_{F_u c})$. In regard to users who are not given a score after the above repeat, we set 0 to them. In other words, in regard to each f of u, we give $\mathit{UserScore}_{att}(f)$ to the largest $\mathit{SubsetScore}(S_{F_u c})$ of $S_{F_u c}$ in which f is included, as follows:

$$\mathit{UserScore}_{att}(f) = \max_{S_{F_u c} \in F_u} \{\mathit{SubsetScore}(S_{F_u c}) |\ f \in S_{F_u c}\}.$$

Then, we take the average of all $\mathit{UserScore}_{att}(f)$ for $\mathit{SpecificityScore}_{att}(u)$, a score of target specificity of u using the attribute att, as follows:

$$\mathit{SpecificityScore}_{att}(u) = \frac{1}{|F_u|} \sum_{f \in F_u} \mathit{UserScore}_{att}(f).$$

For example, suppose the case shown in Figure 4. We assume that a follower set of a user u have three consistency subsets: A, B, and C, and $\mathit{SubsetScore}(A)$, $\mathit{SubsetScore}(B)$, and $\mathit{SubsetScore}(C)$ are 3, 2, and 1 respectively. In the descending order of $\mathit{SubsetScore}(S_{F_u c})$, i.e., in order of A, B, and C, we give scores to each follower f as $\mathit{UserScore}_{att}(f)$ as shown in Figure 4. Finally, we compute the average of $\mathit{UserScore}_{att}(f)$ and take 2.6 for $\mathit{SpecificityScore}_{att}(u)$.

This is how we compute a score of target specificity of u using the attribute att. Below is a summary of the algorithm:

1. we collect all consistency subsets included in the follower set,

2. we score each consistency subset based on to what extent users in it are consistent in a certain noticeable character,

3. in the descending order of the above scores, we repeatedly give the score to each user in the subset, and

4. we take the average of scores for a score of target specificity.

4.3 Scoring Models of Consistency Subsets

In this subchapter, we explain a couple of models: the probabilistic model and the subtracting model which compute $SubsetScore(S_{F_u c})$ mentioned in 4.2.

4.3.1 Probablistic Model

In this model, in regard to the consistency subset $S_{F_u c}$ being consistent in the character c in the follower set F_u, we consider that how low the probability that the user set of the same size as F_u randomly sampled from all Twitter users includes the subset being consistent in c and whose size is $|S_{F_u c}|$ and over. The lowness of the probability means that S_{F_u} is inclining to a part of all Twitter users, so it is able to be said that $S_{F_u c}$ has consistency in a noticeable character. Thus, the lower the probability is, the higher score we give. On the other hand, if the probability is not so low, the deviation between $S_{F_u c}$ and the user set randomly sampled from all Twitter users may be small, and it is difficult to say that $S_{F_u c}$ has consistency in a noticeable character. Thus, we give a low score in this case.

In addition to this parameter: how low the probability mentioned above is, we consider the covering rate of $S_{F_u c}$ to F_u. The higher the covering rate of $S_{F_u c}$ to F_u is, the higher score we give. Based on these two parameters, we compute $SubsetScore(S_{F_u c})$.

These two parameters are summarized as follows.

- The lower the probability that the user set of the same size as F_u randomly sampled from all Twitter users includes the subset being consistent in the same character as the character of $S_{F_u c}$ and whose size is $|S_{F_u c}|$ and over, the higher score we give.
- The higher the covering rate of $S_{F_u c}$ to F_u is, the higher score we give.

Then, we define this model more formally. First, we compute $P(S_{F_u c})$, the probability that the user set of the size of $|F_u|$ randomly sampled from all

Twitter users includes the subset being consistent in c and whose size is $|S_{F_u c}|$ and over, by the formula below:

$$P(S_{F_u c}) = \int_{|S_{F_u c}|}^{n} \binom{n}{x} p^x (1-p)^{n-x} dx, \qquad n = |F_u|, \qquad p = \frac{|S_{Ac}|}{|A|}$$

where A is the set of all Twitter users, and S_{Ac} is the subset which is consistent in the character c included in A. The parameter x in the above formula is based on a binomial distribution.

In addition, by adding the covering rate of $S_{F_u c}$ to F_u to the above parameter, we compute $\mathit{SubsetScore}(S_{F_u c})$. However, there is some possibility that $P(S_{F_u c})$ is excessively low because $S_{F_u c}$ has a remarkable character, e.g., the character only a few users of all Twitter users have, in spite of the size of $S_{F_u c}$ is very small. In this model, we expect that the consistency subset covers the follower set at least intermediately widely. So in such a case, we do not give any scores to $S_{F_u c}$. More formally, we determine a threshold γ which cuts down the above case.

Then, we compute $\mathit{SubsetScore}(S_{F_u c})$ by the formula below:

$$\mathit{SubsetScore}(S_{F_u c}) = \begin{cases} \dfrac{|S_{F_u c}|}{|F_u|} \log(1 - P(S_{F_u c})) & \text{if } \dfrac{|S_{F_u c}|}{|F_u|} > \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

This is how we compute $\mathit{SubsetScore}(S_{F_u c})$ by using the probabilistic technique.

4.3.2 Subtracting Model

In this model, in regard to the consistency subset $S_{F_u c}$, we consider a covering rate of $S_{F_u c}$ to F_u in comparison to a covering rate of the subset which is consistent in the character c to the set of all Twitter users. We call the former a *local rate* and the latter a *global rate*. The fact that a local rate is high and a global rate is low means $S_{F_u c}$ has consistency in a noticeable character and is inclining toward a part of all Twitter users. If a local rate is low or a global rate is high, it is difficult to say that $S_{F_u c}$ has consistency in a noticeable character. Based on the above, we compute $\mathit{SubsetScore}(S_{F_u c})$ by the formula below:

$$SubsetScore(S_{F_u c}) = \max\{\frac{|S_{F_u c}|}{|F_u|} - \frac{|S_A c|}{|A|}, 0\}.$$

In summary, we give a high score in the case of having the two features simultaneously as follows:

- a covering rate of $S_{F_u c}$ to F_u is high, and
- a covering rate of the subset being consistent in the same character as the character of $S_{F_u c}$ to the set of all Twitter users is low.

This is how we compute $SubsetScore(S_{F_u c})$ by using the subtracting technique.

4.4 Attributes for Extracting Consistency Subsets

In this subchapter, we explain a couple of attributes measuring consistency: common terms in profiles and location information, and common followees. By using these attributes, we extract consistency subsets from the follower set of a user. We explain these two attributes in the follow.

4.4.1 Common Terms in Profiles and Location Information

As the first attribute measuring consistency, we consider common terms included in profiles and local information of followers.

There is a high possibility that users belonging to the same community or having the same interest have the same term in their profiles or location information in common. Thus, we extract such terms for measuring consistency. Here, we extract only noun phrases, which characterize their profiles or location information more strongly than other phrases.

Based on the above, we define the method of extracting consistency subsets more formally. We compute the consistency subset $S_{F_u t}$ which is consistent in the term t in F_u by the formula below:

$$S_{F_u t} = \sum_{f \in F_u} \{f |\ t \in Demography(f)\}$$

where $Demography(f)$ is the profile and location information of f. This is how we extract consistency subsets by using common terms in profiles and location information.

4.4.2 Common Followees

As the second attribute measuring consistency, we consider common followees of followers.

Users being consistent in a certain noticeable character often have the common tendency of the follow. Users belonging to the same community are dense on the social graph, so there is a high possibility that they follow common users in the community. In addition, followers of a user publishing technical information about programming is supposed to follow another user publishing useful information about programming in common. Thus, we focus on the tendency of the follow and extract such followees for measuring consistency.

Based on the above, we define the method of extracting consistency subsets more formally. We compute the consistency subset $S_{F_u e}$ which is consistent in the followee e in F_u by the formula below:

$$S_{F_u e} = \sum_{f \in F_u} \{f |\ e \in E_f\}$$

where E_f is the followee set of f. This is how we extract consistency subsets by using common followees.

4.5 Final Classification Based on Target Specificity

In this subchapter, we explain the method of computing target specificity of a user by using scores computed up to this point. In addition, we also explain the method of classifying users based on target specificity.

The $\mathit{SpecificityScore}_{att}(u)$ computed in 4.2 is higher in the cases that

- the more consistent followers of a user are in a noticeable character, or
- the more covered his follower set are with consistency subsets covering it intermediately widely.

Then, we compute a score of target specificity of u. We first compute $\mathit{SpecificityScore}_{term}(u)$, which is using common terms in profiles and location information as an attribute measuring consistency mentioned in 4.4.1, and $\mathit{SpecificityScore}_{followee}(u)$, which is using common followees mentioned in 4.4.2.

Next, we propose a couple of approaches computing target specificity of u by using the above two scores.

(1) Average and maximum of two scores

In the first approach, we take the average and the maximum scores of two scores. We first normalize them by computing the deviation values of them.

Then, we take the average score of two scores for target specificity of u by the formula below:

$$\mathit{TargetSpecificity}(u) = \text{avg}\{\mathit{SpecificityScore}_{\mathit{term}}(u), \mathit{SpecificityScore}_{\mathit{followee}}(u)\}$$

In addition to the average score, we also take the larger one of two scores for target specificity of u, because the larger one is supposed to characterize target specificity more strongly than the other. More formally, we compute $\mathit{TargetSpecificity}(u)$ by the formula below:

$$\mathit{TargetSpecificity}(u) = \max\{\mathit{SpecificityScore}_{\mathit{term}}(u), \mathit{SpecificityScore}_{\mathit{followee}}(u)\}$$

This is how we compute a score of target specificity of u using the approach taking the average and the maximum of two scores.

(2) Binary classifier with the features of two scores

In the second approach, we construct a binary classifier with the features of two scores. If target specificity of u is high, we set $\mathit{TargetSpecificity}(u)$ to 1, and otherwise we set it to 0. We adopt SVM and the decision tree as a classifier.

Then we classify users based on target specificity. We determine a threshold δ which can classify target users and non target users accurately the most, and we classify them by δ. More formally, we classify them as follows:

$$\begin{cases} u \text{ is a } \mathit{target\ user}, & \text{if } \mathit{TargetSpecificity}(u) > \delta \\ u \text{ is a } \mathit{non\ target\ user}, & \text{otherwise.} \end{cases}$$

This is how we classify users into target users and non target users based on target specificity.

Chapter 5 Classifying Users of High Target Specificity

In this chapter, in regard to target users classified by the method mentioned in the above chapter, we explain the method of determining why their target specificity is high based on the result of our analysis mentioned in 3.2.

5.1 Architecture of the Classifiers

Here, we focus on the two causes of high target specificity mentioned in 3.2 as follows:

(1) because they publish information on some specific topics, and

(2) because they publish information to a specific gropu of users,

and we determine whether a user we intend to classify correlates with each cause mentioned above.

We first determine various features of the user which potentially correlate with each cause. Then, based on these features, we construct the classification method which classifies users into three categories: (1) their target specificity is high because they publish information on some specific topics, (2) because they publish information to a specific group of users, and (3) in the cause of both (1) and (2). By classifying users with the above method, we determine why their target specificity is high.

We adopted SVM and the decision tree as a classifier. Next, we propose a couple of approaches to classify users into three categories.

5.1.1 3-class Classifier

In the first approach, we construct a single 3-class classifier using one-against-one method. Each result class corresponds to each category: (1), (2), and (3). Figure 5 (a) shows the architecture of a 3-class classifier.

5.1.2 2 Binary Classifiers

In the second approach, we construct 2 binary classifiers, each of which determines (i) whether users publish information on some specific topics, and (ii) whether users publish information to a specific group of users, respectively. When the results of two classifiers are *yes* and *no* respectively, we classify them

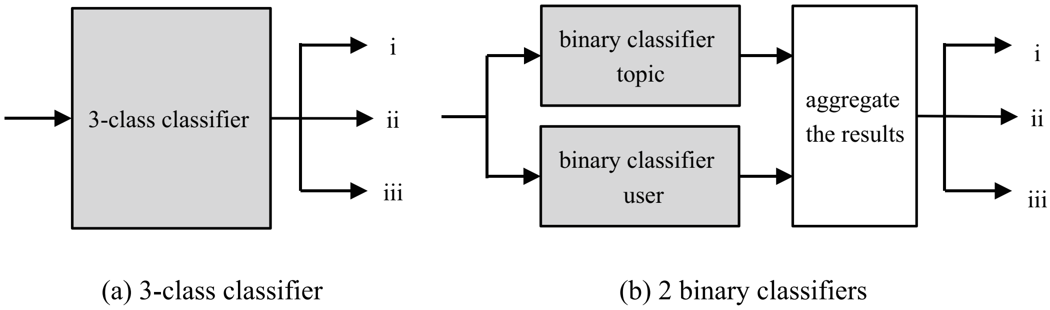

Figure 5: Architecture of a couple of classifiers: (a) 3-class classifier and (b) 2 binary classifiers

into the category (1): their target specificity is high because they publish information on some specific topics. When the results are *no* and *yes*, we classify them into the category (2): because they publish information to a specific group of users, and when the results are *yes* and *yes*, we classify them into the category (3): in the cause of both (1) and (2). Figure 5 (b) shows the architecture of 2 binary classifiers.

5.2 Features Used for the Classification

In this subchapter, we explain what features of users we used for the classification. All the feature values shown below are normalized to values between 0 and 1.

(i) numbers of followees and followers, and their ratio

If the user publishes information to unspecific users, there is a high possibility that a number of his followers is quite large or a number of his followees is quite small. So in such a case, a ratio of a number of his followees to a number of his followers is supposed to be very small. Furthermore, if the user publishes information to the closed users, i.e., his friends, his club members, and so on, there is a high possibility that numbers of his followers and followees are very close because the user is supposed to have a reciprocal connection with them. Thus, numbers of followees and followers, and their ratio are expected to be useful for determining why their target specificity is high.

We take a logarithm of numbers of followers and followees because the difference of large numbers of followers and followees are not as important as the difference of small numbers of followers and followees.

(ii) mutual follow ratio

There is a high possibility that the user publishing information to the closed users has a large mutual follow ratio, i.e., a number of users with whom one follows one another is large, because the user is supposed to have a reciprocal connection with them. Therefore, a mutual follow ratio is expected to be useful for determining why their target specificity is high.

(iii) frequency of replies by "@"

There is a high possibility that the user publishing information to the closed users has a high frequency of replies by "@", i.e., the user replies to his followers frequently. In regard to a mutual follow ratio mentioned in the feature (ii), there are some users publishing information to unspecified users in spite of a large mutual follow ratio, who are called socializers. But in regard to a frequency of replies by "@", there is a high possibility that the user publishes information to the closed users. This is because a high frequency of replies by "@" demonstrates that the user is supposed to have a reciprocal connection with them. Therefore, a frequency of replies by "@" is expected to be useful for determining why their target specificity is high.

(iv) partialness of topics in messages

In regard to a user publishing information on some specific topics, topics in his messages are often partial. Thus, we use the partialness of topics in his messages as a feature for the classification.

Then, we explain how to compute a partialness of topics in messages of u. We first collect up to the latest 200 messages from each user we intend to classify. Second, we extract noun phrases from them and we use their phrases as a corpus. We extract only noun phrases because they characterize contents of the messages more strongly than other phrases. Then, in regard to each user u, we determine the topic of each message by using Latent Dirichlet Allocation (LDA), which is a generative probabilistic model for collections of discrete data such as text corpora, by using the above corpus. Finally, we compute the partialness of topics in messages $partialness(u)$ as follows:

$$partialness(u) = -\sum_{t \in T_u} p_t \log p_t, \qquad p_t = \frac{|\{m|\ m \in M_u,\ topic(m) = t\}|}{|M_u|}$$

where M_u is a message set of u, T_u is the topic set we use for computing this feature, and $topic(m)$ is the topic of a message m. The partialness of topics $partialness(u)$ is the entropy of M_u on the topic. This is how we compute a partialness of topics in messages.

Chapter 6 Experiments and Discussion

In this chapter, we conduct experiments to evaluate our methods proposed in Chapter 4 and Chapter 5, and present the results obtained from them. In addition, we discuss our methods based on the results.

6.1 Data Set

We collected the data set from the real Twitter data by using Twitter API.

We first randomly selected 1,000 Twitter users whose timezone is Japan. At this time, we omitted users followed from nobody or posting no tweet in order to select only active users. Then, we divided them in two sets equally, i.e., each of which include 500 users.

Second, we had 6 experienced Twitter users as participants, all of whom are male graduate students in engineering, from 23 to 25 years old. We assigned each set to 3 participants, and we asked each participant to determine one of the following categories each user in the assigned set is supposed to be in:

(i) the user publishes information to the wide public,

(ii) the user publishes information to some specific topics,

(iii) the user publishes information to a specific group of users, and

(iv) the user publishes information (ii) on some specific topics (iii) to a specific group of users.

These categories correspond to the category "target specificity is high", (1), (2), and (3) in Figure 2 respectively.

Then, we selected users whose category at least 2 out of 3 participants coincide with, and as a result, we were able to collect 93, 320, 375, and 30 users in the category (i), (ii), (iii), and (iv) respectively. We randomly selected 90 users from the category (i), and 30 users from each category of (ii), (iii), and (iv). We collected these 180 users in total, and we used them as the data set. Table 1 shows the breakdown of the data set: the average and the standard deviation of numbers of followers, followees, and tweets in each category.

Then, for each user, we collected at most 1,000 followers, their profiles and local information, and at most 1,000 followees of them. When collecting this

Table 1: Breakdown of data set: average and standard deviation of numbers of followers, followees, and tweets in each category shown in the format of *Average (SD)*

category	**follower**	**followee**	**tweet**
(i)	475,679 (535,894)	11,274 (37,906)	9,763 (14,607)
(ii)	58,142 (171,784)	3,353 (7,218)	9,992 (23,572)
(iii)	573 (1,389)	598 (1,545)	8,829 (29,505)
(iv)	82,942 (262,161)	1,568 (3,594)	5,677 (6,600)

data, we removed terms and followees whose local rate is 0.01 or below to reduce calculation costs. In addition, we also removed stop words from profiles and local information. Especially in this experiments, we included Twitter original words, e.g., tweet, follow, and so on, in the stop word list. We used this data in order to evaluate our methods.

6.2 Experimental Settings and Libraries

First, we conducted the experiments evaluating the method of classifying users based on target specificity of their information publishing mentioned in Chapter 4. We used 90 users in the category (i) as target users, and 90 users in the categories (ii), (iii), and (iv) as non target users. We first computed $\mathit{SpecificityScore}_{\mathit{term}}(u)$ and $\mathit{SpecificityScore}_{\mathit{followee}}(u)$ for each user u using a couple of models, i.e., the probablistic model and the subtracting model mentioned in 4.3, and computed the accuracy of the classification of target users and non target users using each attribute separately. In regard to the probablistic model, we take 0.01, 0.03, 0.05 for γ, the threshold which cuts down the case that a local rate is very small, and compared them. When computing $\mathit{SpecificityScore}_{\mathit{followee}}(u)$, we assumed that the rate of active users in Twitter is 0.01.

Next, we computed $\mathit{TargetSpecificity}(u)$ based on the above scores using a couple of baselines and four types of proposed methods. A couple of baselines are as follows:

follower: numbers of followers arranged in the descending order, and
SVM: a binary SVM with the features of the max rate of common terms in profiles and location information and common followees in followers.
For types of proposed methods are the average and the maximum of the above two scores and a binary classifier with the features of two scores adopting SVM and the decision tree mentioned in 4.5.

Then, we determined a threshold δ which can classify target users and non target users accurately the most, and evaluated the classification results with δ.

Second, we conducted the experiments evaluating the method of determining why target specificity of the users is high in regard to target users mentioned in Chapter 5. We extracted 30 users from each category of (ii), (iii), and (iv), and used 90 users in total. We first extracted features mentioned in 5.2 from the user, which are normalized to a value between 0 and 1. Then, based on these features, we constructed two types of classification approaches: a 3-class classifier and 2 binary classifiers mentioned in 5.1, which classify users into three categories: (ii), (iii), and (iv), and evaluated the classification results using 3-fold cross validation. We used two learning algorithms: SVM and the decision tree as a classifier, and compared them. For SVM, we used LIBSVM[1)], which is a popular SVM library, with the Gaussian kernel, which using one-against-one method for multiclass classification. For the decision tree, we used scikit-learn [2)], which is a Python module for machine learning.

We used twpro search API[3)] in order to get a number of users having a certain term in their profiles. We also used MeCab[4)] for morphological analysis of Japanese sentences in profiles, local information, and tweets of users. Furthermore, we used gensim[5)] for using Latent Dirichlet Allocation (LDA).

---

1) http://www.csie.ntu.edu.tw/~cjlin/libsvm/
2) http://scikit-learn.org/stable/modules/tree.html
3) http://twpro.jp/doc/api/search
4) http://mecab.sourceforge.net/
5) http://radimrehurek.com/gensim/

Table 2: Accuracy of the classification of target users and non target users using a couple of attributes measuring consistency separately

attributes	**Probablistic Model**			**Subtracting Model**
	$\gamma = 0.01$	$\gamma = 0.03$	$\gamma = 0.05$	
common terms	**0.861**	0.850	0,839	0.850
common followees	0.828	0.828	0.817	**0.833**

6.3 Results and Evaluation of Classifying Users Based on Target Specificity

In this subchapter, we show the results of the experiments classifying users based on target specificity of their information publishing.

We first show the accuracy of the classification of target users and non target users using a couple of attributes measuring consistency, i.e., common terms and common followers, separately. We also used a couple of models, i.e., the probablistic model and the subtracting model, which compute $SubsetScore(S_{F_u c})$. The 3rd and 4th rows of Table 2 show the accuracy of the classification using common terms and common followees as attributes, respectively. For each attribute, a bold number shows when the accuracy becomes the highest.

In each case of using common terms and common followees as attributes, we achieved the highest accuracy when we used the probablistic model with $\gamma = 0.01$ and the subtracting model respectively. In regard to the probablistic model, we can see that the smaller γ is, the higher the accuracy is. This suggests that rare attributes characterize target users though the covering rate of $S_{F_u c}$ to F_u is not high, and can contribute the classification of target users and non target users. But for both attributes, the difference between the accuracy of the probablistic model and that of the subtracting model is not so big. So it is supposed that these two models are about the same effect.

In addition, the accuracy using common terms as attributes is higher than that using common followees regardless of models. This suggests that common terms in profiles and location information are more useful for extracting consistency subsets than common followees. This is supposed to be because not only

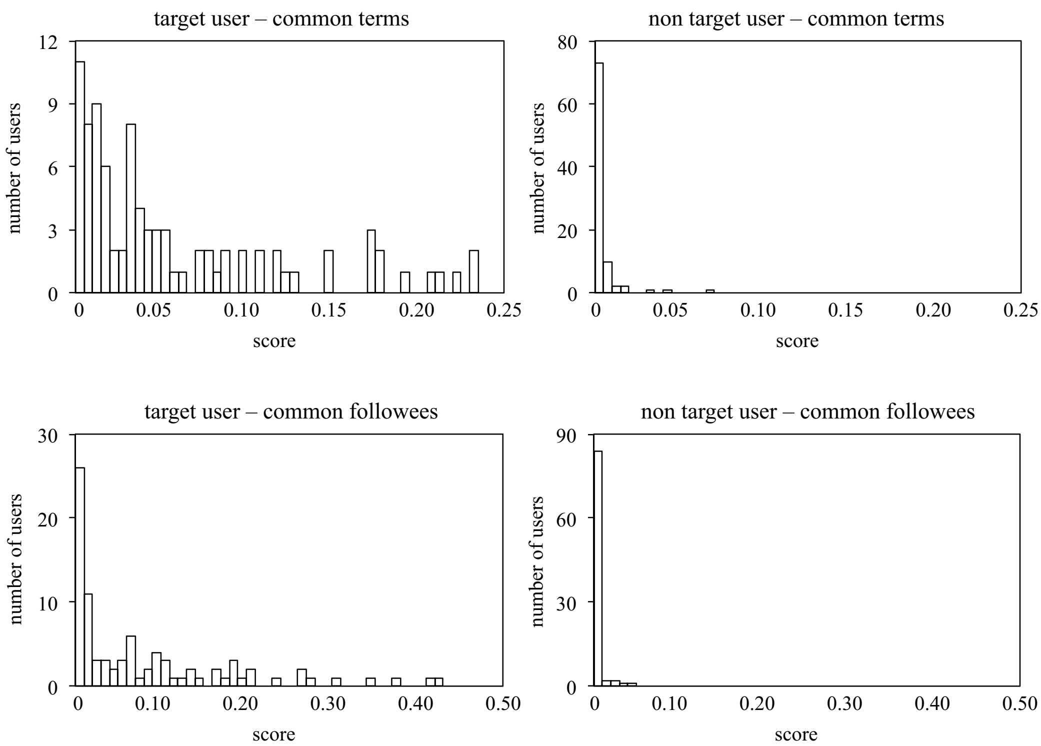

Figure 6: The histogram of specificity scores of each attribute on target users and non target users

target users but also non target users sometimes have common followees in their followers, though there is a little possibility that non target users have common terms in their followers. For example, a non target user publishing information about world news sometimes have common followees publishing world news in his followers. But the global rate of such followers is usually large, so there is no necessity to notice about them in most cases.

Now for each attribute, we use the model which is the highest accuracy, i.e., the probablistic model with $\gamma = 0.01$ for common terms and the subtracting model for common followees. Figure 6 shows the histogram of specificity scores of each attribute on target users and non target users. In the case of using common terms as attributes, specificity scores of most non target users are very small, with the average 0.004. On the other hand, those of most target users are large, and even the average, 0.065, is as same as the maximum score for non target users, i.e., 0.067. A similar trend is apparent in the case of using

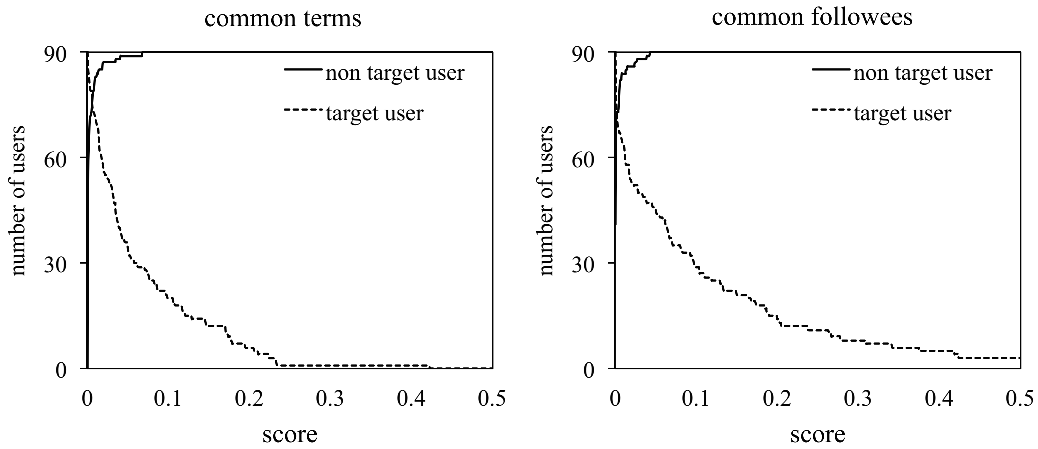

Figure 7: The cumulative histogram of specificity scores of each attribute on target users and non target users

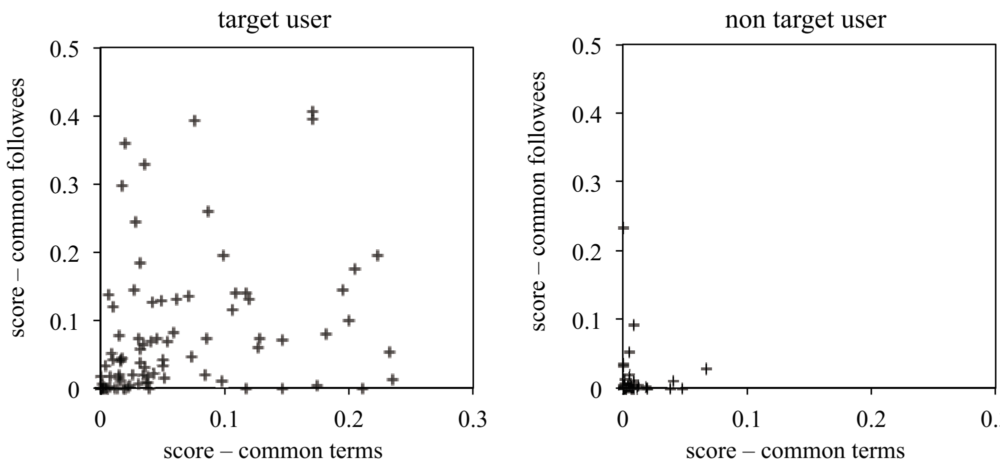

Figure 8: A scatter diagram based on specificity scores of a couple of attributes, common terms and common followees, on target users and non target users

common followees as attributes: even the median for target users, i.e., 0.050, is larger than the maximum score for non target users, i.e., 0.043.

Figure 7 shows the cumulative histogram of specificity scores of each attribute on target users and non target users. More and more target users cumulate as scores become smaller, and more and more non target users cumulate as scores become larger. Therefore, the score on which the sum of numbers of target users and non target users is the largest is the score which can classify

Table 3: Accuracy of the final classification of target users and non target users

Baseline		**Proposed Method**			
follower	SVM	max	avg	SVM	decision tree
0.878	0.828	0.856	0.872	**0.944**	0.906

target users and non target users accurately the most, and we set this score to the threshold δ. In the cases of using common terms and common followees as attributes, we take δ for 0.009 and 0.012 respectively.

Next, we show the relationship between two attributes, i.e., common terms and common followees. Figure 8 shows a scatter diagram based on specificity scores of these two attributes on target users and non target users. On target users, points of scores of these attributes are widely distributed and have weak positive correlation, with the correlation coefficient 0.507. That is, there is a high possibility that the higher one score is, the higher the other score is. Of course, as shown in Figure 8, there are some outliers, i.e., one score is high and the other score is low. On the other hand, points of scores of these attributes concentrate on the point $(0, 0)$ on no target users, and have no correlation, with the correlation coefficient 0.232.

Then, we show the accuracy of the final classification of target users and non target users. The 1st and 2nd columns of Table 3 show the accuracy of a couple of baselines: a follower method and a SVM method mentioned in 6.2, respectively. The accuracy of the follower method achieved 0.878, which suggests that target specificity is mainly related to a number of followers. This accuracy is quite high for a baseline, but notice that our final goal is classifying outliers: target users who have many followers or non target users who have a few followers, with high accuracy. While the accuracy of the SVM method was 0.828.

The following columns show the accuracy of four types of proposed methods: the maximum and the average of scores of each attribute, and a binary classifier with the features of these scores adopting SVM and the decision tree, respectively. The accuracy of the maximum score was 0.856, which was lower than the accuracy using only common terms as attributes as shown in Table 2, i.e., 0.861. This suggests that some non target users may have high scores in either attribute, and they are able to become noises for the classification. The accuracy of the average score was 0.872, which was higher than the accuracy using common terms or common followees separately as attributes, but lower than a follower method, i.e., 0.878. The accuracy of a binary classifier adopting SVM and the decision tree achieved 0.944 and 0.906 respectively. The former was the highest accuracy of four types of proposed methods, and 0.66 point higher than the follower method. This suggests that a binary classifier adopting SVM is also able to classify users who is difficult to classify by only a number of followers, i.e., target users who have many followers or non target users who have a few followers with a certain degree of high accuracy.

Finally, we show the details of the results of a part of users. Upper and lower half of Table 4 show details of the results of target users and non target users respectively: their screen names, noticeable terms and followees extracted for measuring consistency, their local rates, global rates, and scores of consistency subsets, and specificity scores of each attribute.

The first rows of upper half show the details of the results of the target user @MCstaff_Fukuoka, the account publishing information about the concert hall in Fukuoka. In regard to common terms in profiles and location information, our method extracted terms related to the account, e.g., Fukuoka, music, Hakata, and so on. In regard to common followees, our method also extracted followers related to the account as with common terms, e.g., @fukuoka_yokane: the account publishing information about the CD shop in Fukuoka, @f_sunpalace: the other account publishing information about the same concert hall, and so on. These terms and followees have high local rates in spite of low global rates, so the scores of consistency subset become high. As a result, the specificity score of each attribute on this account is high. The other target users, @Jars0830 and @pa_ko065, are just alike.

The first rows of lower half show the details of the results of the non target user @tenkijp, the account publishing weather information in Japan. In

Table 4: Details of the results of a part of users

target user	**term & followee**	*local rate*	*global rate*	*SubsetScore*	*SpecificityScore*
@MCstaff_Fukuoka	Fukuoka	0.563	0.047	0.390	0.235
	music	0.114	0.096	0.076	
	Hakata	0.095	0.003	0.076	
	@fukuoka_yokane	0.069	0.023	0.046	0.010
	@f_sunpalace	0.063	0.023	0.040	
	@mbc_o2_eiji	0.059	0.028	0.031	
@Jars0830	Arashi	0.439	0.039	0.304	0.233
	participation	0.190	0.026	0.131	
	line	0.130	0.045	0.090	
	@Yamnos5	0.249	0.004	0.246	0.067
	@ars_762	0.109	0.072	0.037	
	@nino_xoxo_	0.070	0.033	0.037	
@pa_ko065	piano	0.528	0.009	0.366	0.223
	music	0.250	0.096	0.173	
	Sapporo	0.139	0.022	0.096	
	@mofu_co	0.500	0.001	0.499	0.277
	@miko3535	0.500	0.003	0.497	
	@shimagaraneko	0.472	0.002	0.470	
non target user	**term & followee**	*local rate*	*global rate*	*SubsetScore*	*SpecificityScore*
@tenkijp	Tokyo	0.033	0.196	0	2.51e-5
	hobby	0.022	0.095	9.28e-16	
	music	0.021	0.096	1.18e-15	
	@tenkijp_jishin	0.421	10.4	0	0
	@Kantei_Saigai	0.358	14.8	0	
	@bouei_saigai	0.277	6.31	0	
@masason	fan	0.059	0.676	2.62e-17	1.20e-12
	Tokyo	0.027	0.196	0	
	music	0.024	0.096	1.34e-15	
	@shigeruishiba	0.050	0.939	0	0
	@WSJJapan	0.066	3.56	0	
	@HeizoTakenaka	0.05	3.62	0	
@Kantei_Saigai	Tokyo	0.034	0.196	0	3.52e-9
	hobby	0.027	0.095	1.14e-15	
	movie	0.019	0.043	1.85e-7	
	@MofaJapan_ITPR	0.056	0.597	0	0.001
	@CAO_BOUSAI	0.203	0.834	0	
	@MofaJapan_jp	0.062	0.950	0	

Table 5: Accuracy of the classification of target users

SVM		**decision tree**	
3-class	2 binary	3-class	2 binary
0.678	**0.689**	0.556	0.533

Table 6: Accuracy of SVMs without each feature

Removed Feature	**3-class**	**2 binary**	**topic**	**user**
with all	0.678	0.689	0.856	0.833
(i)	**0.644**	**0.656**	**0.811**	0.844
(ii)	0.722	0.711	0.867	0.844
(iii)	0.678	0.678	0.844	0.833
(iv)	0.655	0.678	0.844	**0.822**

regard to common terms in profiles and location information, our method extracted terms which is not rare, e.g., Tokyo, hobby, music, and so on. In regard to common followees, our method also extracted followees who is not rare, i.e., users who have a large number of followers, as with common terms, e.g., @tenkijp_jishin: the account publishing earthquake information, @Kantei_Saigai: the account publishing disaster information, and so on. These terms and followees have high local rates to a certain degrees, but their global rates are much higher than them, so the scores of consistency subsets become very small. As a result, the specificity score of each attribute on this account is small. The other non target users, @masason and @Kantei_Saigai, are just alike.

6.4 Results and Evaluation of Classifying Users of High Target Specificity

In this subchapter, we show the results of the experiments classifying target users. Table 5 shows the accuracy of four types of classification methods: a combination of two approaches and two classifiers, i.e., a 3-class SVM, 2 binary SVMs, a 3-class decision tree, and 2 binary decision trees respectively, classify-

Table 7: A part of topics extracted by LDA

id	words
1	news, program, broadcast, morning, night, tonight
2	update, blog, picture, smart phone, weather
3	worst, typhoon, Sea of Japan, electricity
4	earthquake, observation, focus, concern, teacher

ing into three categories with all features mentioned in Chapter 5. The accuracy of the classification methods using SVMs as classifiers are more than 10 points higher than that using decision trees. When using SVMs as classifiers, the accuracy of 2 binary classifiers are a little higher than that of a 3-class classifier, and it is the opposite when using decision trees. But for both methods using SVMs and decision trees, the difference between the accuracy of 2 binary classifiers and that of a 3-class classifier is not so big, which suggests that the two causes of the high target specificity mentioned in 3.2 are highly independent of each other.

Next, we show the details of the results of the classification methods using SVMs as classifiers. The 2nd row of Table 6 shows the accuracy of each classification method with all features, and the following rows show the accuracy when we remove each feature from the data. Each feature of (i), (ii), (iii), and (iv) correspond to those in Chapter 5 respectively. For each method, a bold number shows when the accuracy becomes the lowest.

The 2nd and 3rd columns show the accuracy of a 3-class SVM and 2 binary SVMs respectively. For each method, the accuracy became the lowest when we removed the feature (i), i.e., numbers of followees and followers, and their ratio. It is able to be said that numbers of followees and followers, and their ratio are mainly useful for determining the cause of high target specificity. The accuracy also became lower to some extent when we removed the feature (iv), i.e., the partialness of topics in messages. This suggests that partialness of topics in messages is related to high target specificity. Table 7 shows a part of topics out of 20 extracted by LDA. We were able to extract topics about news programs, disasters, and so on. On the other hand, the accuracy became higher when we removed the feature (ii), i.e., mutual follow ratio. This suggests that mutual follow ratio and whether the user publishes information to the closed users are not necessarily correlating.

The 4th and 5th columns show the accuracy of each binary SVM used for 2 binary SVMs: a binary SVM determining whether users publish information on some specific topics or not, and that determining whether users publish information to the users to a specific group of users or not, respectively. The accuracy of the former became the lowest when we removed the feature (i), i.e., numbers of followees and followers, and their ratio, and that of the latter became the lowest when we removed the feature (iv), i.e., the partialness of topics in messages, which is the noticeable matter. This suggests that partialness of topics is more useful for determining whether users publish information to a specific group of users or not than whether users publish information to some specific topics or not.

Chapter 7 Conclusion

In this study, we focused on the fact that the breadth of target scope of information publishing in Twitter varies greatly among users, and proposed the method to classify Twitter users from the point of view of how broad target scope of their information publishing is, i.e., whether they publish information to the wide public or publish information to specific groups of users.

First, we defined target specificity of a user, as the measure of to what extent target scope of his information publishing is specific. Second, based on this definition, we proposed the algorithm computing a score of target specificity. In this algorithm, we focused on two parameters: (a) whether followers of a user are consistent in a certain noticeable character or not, and (b) whether the follower set of a user is covered with consistency subsets covering it intermediately widely or not. For computing the score, we proposed a couple of models computing scores of consistency subsets: the probabilistic model and the subtracting model, and a couple of attributes for extracting consistency subsets from the follower set of a user: common terms in profiles and location information and common followees. Then, we finally proposed four types of methods of classifying users into target users and non target user based on the above scores. We conducted experiments to compare the performance of the methods, and the results suggested that the method using a binary SVM classified users with the highest accuracy. The results also suggested that common terms in profiles and location information were more useful for extracting consistency subsets than common followees.

In addition, in this study, in regard to users classified into "target specificity is high" by the above method, we proposed the the method of determining why their target specificity is high. We first analysed the causes of high target specificity, and we classified users into three categories based on them: (1) topic specificity, i.e., their target specificity is high because they publish information to some specific topics, (2) user specificity, i.e., their target specificity is high because they publish information to a specific group of users, and (3) in the causes of both (1) and (2). Then, we constructed a couple of approaches to classify them into three categories: a 3-class classifier and 2 binary classifiers, based on various features of the user which potentially correlate with each cause. Our experimental results suggested that a 2 binary SVM classified users with the highest accuracy. The results also suggested that the two causes (1) and (2) were highly independent of each other.

We believe that our proposed methods are applicable to some applications e.g., Twitter search systems, and they will enrich users' experience on Twitter.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Professor Keishi Tajima for his valuable advises and helpful discussions. I deeply appreciate Proffessor Yasushi Sakurai and Associate Professor Yasuhito Asano for spending precious time for valuable discussions. I would also like to thank the other lab members for their supports for our research.

References

[1] Protalinski, E.: Twitter sees 218m monthly active users, 163.5m monthly mobile users, 100m daily users, and 500m tweets per day, http://thenextweb.com/twitter/2013/10/03/twitter-says-it-sees-215-million-monthly-active-users-100-million-daily-users-and-500-million-tweets-per-day/ (2013).

[2] Irani, D., Webb, S., Pu, C. and Li, K.: Study of trend-stuffing on twitter through text classification, *Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS)* (2010).

[3] Broder, A.: A taxonomy of web search, *ACM Sigir forum*, Vol. 36, No. 2, ACM, pp. 3–10 (2002).

[4] Busch, M., Gade, K., Larson, B., Lok, P., Luckenbill, S. and Lin, J.: Earlybird: Real-time search at twitter, *Data Engineering (ICDE), 2012 IEEE 28th International Conference on*, IEEE, pp. 1360–1369 (2012).

[5] Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H. and Demirbas, M.: Short text classification in twitter to improve information filtering, *Proceeding of the 33rd international ACM SIGIR conference on research and development in information retrieval*, ACM, pp. 841–842 (2010).

[6] Cha, M., Haddadi, H., Benevenuto, F. and Gummadi, K.: Measuring user influence in twitter: The million follower fallacy, *4th international aaai conference on weblogs and social media (icwsm)*, Vol. 14, No. 1, p. 8 (2010).

[7] Cheng, J., Romero, D. M., Meeder, B. and Kleinberg, J.: Predicting reciprocity in social networks, *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)*, IEEE, pp. 49–56 (2011).

[8] Chu, Z., Gianvecchio, S., Wang, H. and Jajodia, S.: Who is tweeting on twitter: human, bot, or cyborg?, *Proceedings of the 26th Annual Computer Security Applications Conference*, ACM, pp. 21–30 (2010).

[9] Duan, Y., Jiang, L., Qin, T., Zhou, M. and Shum, H.: An empirical study on learning to rank of tweets, *Proceedings of the 23rd International Conference on Computational Linguistics*, Association for Computational Lin-

guistics, pp. 295–303 (2010).

[10] Ehrlich, K. and Shami, N.: Microblogging inside and outside the workplace, *Proc. ICWSM*, Vol. 10 (2010).

[11] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake shakes Twitter users: real-time event detection by social sensors, *Proceedings of the 19th international conference on World wide web*, ACM, pp. 851–860 (2010).

[12] Hopcroft, J., Lou, T. and Tang, J.: Who will follow you back?: reciprocal relationship prediction, *Proceedings of the 20th ACM international conference on Information and knowledge management*, ACM, pp. 1137–1146 (2011).

[13] Ikawa, Y., Enoki, M. and Tatsubori, M.: Location inference using microblog messages, *Proceedings of the 21st international conference companion on World Wide Web*, ACM, pp. 687–690 (2012).

[14] Java, A., Song, X., Finin, T. and Tseng, B.: Why we twitter: understanding microblogging usage and communities, *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, ACM, pp. 56–65 (2007).

[15] Kunegis, J., Lommatzsch, A. and Bauckhage, C.: The slashdot zoo: mining a social network with negative edges, *Proceedings of the 18th international conference on World wide web*, ACM, pp. 741–750 (2009).

[16] Kwak, H., Lee, C., Park, H. and Moon, S.: What is Twitter, a social network or a news media?, *Proceedings of the 19th international conference on World wide web*, ACM, pp. 591–600 (2010).

[17] Lee, K., Caverlee, J. and Webb, S.: The social honeypot project: protecting online communities from spammers, *Proceedings of the 19th international conference on World wide web*, ACM, pp. 1139–1140 (2010).

[18] Leskovec, J., Huttenlocher, D. and Kleinberg, J.: Predicting positive and negative links in online social networks, *Proceedings of the 19th international conference on World wide web*, ACM, pp. 641–650 (2010).

[19] Leskovec, J., Huttenlocher, D. and Kleinberg, J.: Signed networks in social media, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 1361–1370 (2010).

[20] Massoudi, K., Tsagkias, M., de Rijke, M. and Weerkamp, W.: Incorporating query expansion and quality indicators in searching microblog posts, *Advances in Information Retrieval*, Springer, pp. 362–367 (2011).

[21] Mathioudakis, M. and Koudas, N.: Twittermonitor: trend detection over the twitter stream, *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, ACM, pp. 1155–1158 (2010).

[22] Nagmoti, R., Teredesai, A., De Cock, M. et al.: Ranking approaches for microblog search, IEEE Computer Society (2010).

[23] Nishida, K., Banno, R., Fujimura, K. and Hoshide, T.: Tweet classification by data compression, *Proceedings of the 2011 international workshop on DETecting and Exploiting Cultural diversiTy on the social web*, ACM, pp. 29–34 (2011).

[24] Pennacchiotti, M. and Popescu, A.: A machine learning approach to twitter user classification, *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)* (2011).

[25] Takemura, H. and Tajima, K.: Tweet Classification Based on Their Lifetime Duration, *Proceedings of the 21st International Conference on Information and Knowledge Management (CIKM)*, ACM, pp. 2367–2370 (2012).

[26] Teevan, J., Ramage, D. and Morris, M.: # TwitterSearch: a comparison of microblog search and web search, *Proceedings of the fourth ACM international conference on Web search and data mining*, ACM, pp. 35–44 (2011).

[27] Weng, J., Lim, E.-P., Jiang, J. and He, Q.: Twitterrank: finding topic-sensitive influential twitterers, *Proceedings of the third ACM international conference on Web search and data mining*, ACM, pp. 261–270 (2010).

[28] Wu, S., Hofman, J. M., Mason, W. A. and Watts, D. J.: Who says what to whom on twitter, *Proceedings of the 20th international conference on World wide web*, ACM, pp. 705–714 (2011).

[29] Yan, L., Ma, Q. and Yoshikawa, M.: Classifying Twitter Users Based on User Profile and Followers Distribution, *Database and Expert Systems Applications*, Springer, pp. 396–403 (2013).

[30] Zhao, D. and Rosson, M. B.: How and why people Twitter: the role that micro-blogging plays in informal communication at work, *Proceedings of the ACM 2009 international conference on Supporting group work*, ACM, pp. 243–252 (2009).