

Situation-based reply candidate presentation method for mobile phone

Masafumi Matsuhara*

Faculty of Software and Information Science,
Iwate Prefectural University,
Iwate, Japan
Email: masafumi@iwate-pu.ac.jp
*Corresponding author

Norifumi Watanabe

Seiko Epson Corporation,
Nagano, Japan
Email: norifumi007@gmail.com

Goutam Chakraborty

Faculty of Software and Information Science,
Iwate Prefectural University,
Iwate, Japan
and
Sendai Foundation of Applied Information Sciences,
Miyagi, Japan
Email: goutam@iwate-pu.ac.jp

Hiroshi Mabuchi

Faculty of Software and Information Science,
Iwate Prefectural University,
Iwate, Japan
Email: mabu@iwate-pu.ac.jp

Abstract: Recently, many people use mobile devices such as a smart phone or a smart watch. Necessity and importance for creating messages on a mobile device are increasing since users on social networking service are rapidly increasing. Users need to input text for messages with small keys because the screen for the operation is small. In particular, it may be difficult to quickly respond when a user creates a reply on a mobile device. We have proposed method to automatically present multiple reply candidates. The presented candidates correspond to different situations depending on the topic of the received message. The reply process is completed just by selecting the appropriate one from the candidate replies. Rapid response is possible since the user neither needs to create nor revise a reply message on mobile phones, facilitating smooth communication. The effectiveness of our proposed method and its evaluation is done through experiments in Japanese.

Keywords: mobile phone; smart watch; social networking service; SNS; quick response; reply candidates presentation; tweet collection; tweet clustering; sentence division; sentence scoring; different situations; natural language processing.

Reference to this paper should be made as follows: Matsuhara, M., Watanabe, N., Chakraborty, G. and Mabuchi, H. (2020) ‘Situation-based reply candidate presentation method for mobile phone’, *Int. J. Information Technology, Communications and Convergence*, Vol. 3, No. 4, pp.319–336.

Biographical notes: Masafumi Matsuhara received his BE and ME degrees from Hokkai-Gakuen University, Hokkaido, Japan in 1996 and 2000, respectively. He received PhD in Engineering from Hokkaido University, Hokkaido, Japan in 2003. He was an Assistant Professor of Iwate Prefectural University in 2003–2014. He is currently an Associate Professor of the Faculty of Software and Information Science, Iwate Prefectural University. His research interests include artificial intelligence and natural language processing.

Norifumi Watanabe received his BE and ME degrees from Iwate Prefectural University, Iwate, Japan in 2018 and 2020, respectively. He is currently in Seiko Epson Corporation. His research interests include natural language processing.

Goutam Chakraborty received his PhD in 1993 from Research Institute of Electrical Communication, Tohoku University, Japan. Presently, he is a distinguished Professor of Iwate Prefectural University, Japan, and the chief executive researcher of Sendai Foundation of Applied Information Sciences, Japan. His research interests are heuristic and soft computing algorithms, their applications to solve pattern recognition, prediction, scheduling and optimization problems including applications in wired, wireless, and sensor networks. He works in time-series modelling, anomaly detection and prediction, especially for bio-signals. He also works in data mining problems. He is a senior member of IEEE, and a senior life-member of ACM.

Hiroshi Mabuchi received his ME and DE degrees from Hokkaido University, Sapporo, Japan, in 1992 and 1995, respectively. In 1995, he joined the faculty of Engineering of Tohwa University, where he was a Lecturer in 1995–1998. In 1998, he joined the faculty of Software and Information Science of Iwate Prefectural University, where he was a Lecturer in 1998–2002, and an Associate Professor in 2002–2018. From 2018, he is a Professor of Faculty of Software and Information Science. His research interests include problem solving, natural language processing and drone.

This paper is a revised and expanded version of a paper entitled ‘Method to present situation aware suitable reply candidate on mobile terminal’ presented at The 10th IEEE International Conference on Awareness Science and Technology (iCAST 2019), Morioka, Japan, 23–25 October 2019.

1 Introduction

In recent years, many people use mobile phones and smart watches. Due to availability of the internet everywhere, the opportunity and use of communication on social networking service (SNS) is increasing. Users for SNS such as Twitter (<https://twitter.com/>) and LINE (<https://line.me/ja/>) are increasing rapidly.

Since the operation screen is small on a mobile device, the key size for character input is small. It takes time to input a message and typos are frequent. With a small mobile device such as a smart watch, character input is even more difficult because compared to a smart phone. Since the operation screen is small, only a limited number of characters are displayed at a time. In particular, when creating a reply after receiving a message via SNS, it may be difficult to respond quickly. Communication via messaging is difficult.

To support smooth communication, usually a smart watch is equipped with input method like voice input, pointing input, fixed phrase selection and so on. The fixed phrase selection method does not burden the user because it can create a reply message just by selecting a sentence from a few fixed phrases. However, it is difficult to make a reply message that includes necessary information, such as emotions, since the presented sentences are simple such as ‘thank you’ and ‘OK’, prepared in advance.

There is Smart Reply in Gmail (<https://www.google.com/gmail/>). Smart Reply uses deep learning to make three probable replies and presents them to the user. Although it is possible to present simple replies using this application, the replies are too simple to be useful most of time and are without consideration of scenarios and situations for users.

In order to present meaningful responses, we have proposed a method to create multiple candidates for replies depending on different situations of the user (Watanabe et al., 2019). In this work, we propose a method to present multiple non-redundant reply candidates to the user in different situations according to the topic of the received message. By using the proposed method, it is possible to select an appropriate reply without performing character input.

The system collects tweets in order to create reply candidates. The collected tweets are divided into sentences, to make reply candidates simple and suitable. Since tweets contain sentences that indicate various situations, it is possible to group appropriate sentences as reply candidates by clustering and scoring them according to their suitability.

K-means, which is a commonly clustering algorithm, is used in our proposed method. *K*-means is chosen because it is efficient. The system based on our proposed method divides tweets into clusters that indicate specific situations. It is possible to present users with suitable replies in different situations by selecting messages from different clusters. Each reply is scored using word coverage. The word coverage is a value that indicates the degree of importance of words contained in a sentence. By scoring with word coverage, it is possible to present the user with a simple reply candidate that contains many important words that matches with the situation and does not contain unnecessary words.

The rest of the paper is as follows. Section 2 discusses related works. Section 3 explains our proposed algorithm. Section 4 is about experiment results and discussion. Finally, the paper is concluded in Section 5.

2 Related work

2.1 Rule-based response

The rule-based response is a method of outputting a response sentence by interacting, based on some rules by Weizenbaum (1966) and Wallace (2008). The method has to prepare a response rule for the text input by a user. For example, when the input is ‘How are you?’, a response sentence is generated by setting a rule in advance to respond as ‘good’. The rule-based response method requires large cost because the rules have to be set manually. It is necessary to set a large number of rules in order to present a user with reply candidates that match the situations for various received sentences. This poses a limitation to present appropriate reply candidates using the rule-based response method.

2.2 Statistics-based response

The statistics-based response is a method of presenting an appropriate response by searching a database proposed in Banchs and Li (2012) and Lasguido et al. (2014). The method has to prepare a database of pairs of input sentences and response sentences. Among statistic-based response, this particular method is called example-based. Because the method uses a database, it is not necessary to create rules manually. However, it is necessary to collect dialogue data in which input and response are paired. It is not easy to collect enough dialogue data for creating input-response pairs for every situation.

2.3 Input method on small device

In order to input sentences, character input methods have been proposed without real keyboards. Boustila et al. (2019) proposed a character input method with the head-mounted displays in virtual environments. The method uses the touch screen keyboard in the virtual environment to input characters. Jiang and Weng (2020) also proposed a character input method with the head-mounted displays in virtual environments. The method uses HiPad which is leveraging a circular touchpad with a circular virtual keyboard. The users are able to input text by one-hand with the HiPad in virtual reality. These methods need other devices in addition to mobile phones although they are able to be used on small devices.

Many character input methods specialised for small mobile devices have been proposed. A method was proposed by Oney et al. (2013) to reduce tap errors, by expanding the rough position on the tapped keyboard and then, once the intended area of the keyboard is expanded, it enables tapping the intended key. A method was proposed by Corsten et al. (2017) to use the finger pressure on the back of the device. Users are able to apply various pressure levels with each fingers on the back of the device to enter different characters. A method was proposed by Fujita (2019) to provide a user interface equally usable for all. The proposed method is a text entry system designed for operation with user’s dominant hand on mobile phones. Hong et al. (2015) proposed to input characters by switching the keyboard divided into multiple areas by swipe operation.

Although each method is devised to input characters easily, it takes time to get used to the operation, and use the input method.

2.4 Tweet clustering

Tweets are often written in a spoken language on various topics. Since the amount of information in a tweet is huge, it is necessary for users to select and obtain the required information. To solve this problem, researches on tweet clustering have been proposed. Jiang et al. (2011) proposed to cluster tweets into positive, negative or neutral groups. Go et al. (2009) proposed to get features from emoticons used in tweets and using the features to classify a tweet as positive or negative.

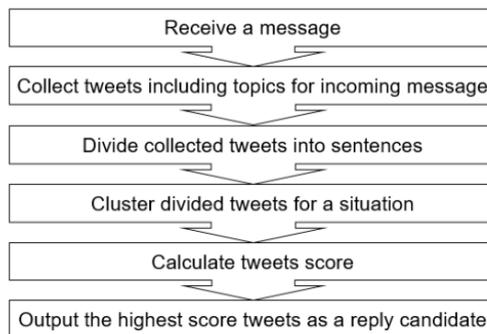
Our proposed method classifies tweets into classes that indicate specific situations. Tweets are not dialogue data. We use unpaired tweets as data. It is easy to collect unpaired data than to collect dialogue data.

3 Proposed method

3.1 Flow of proposed method

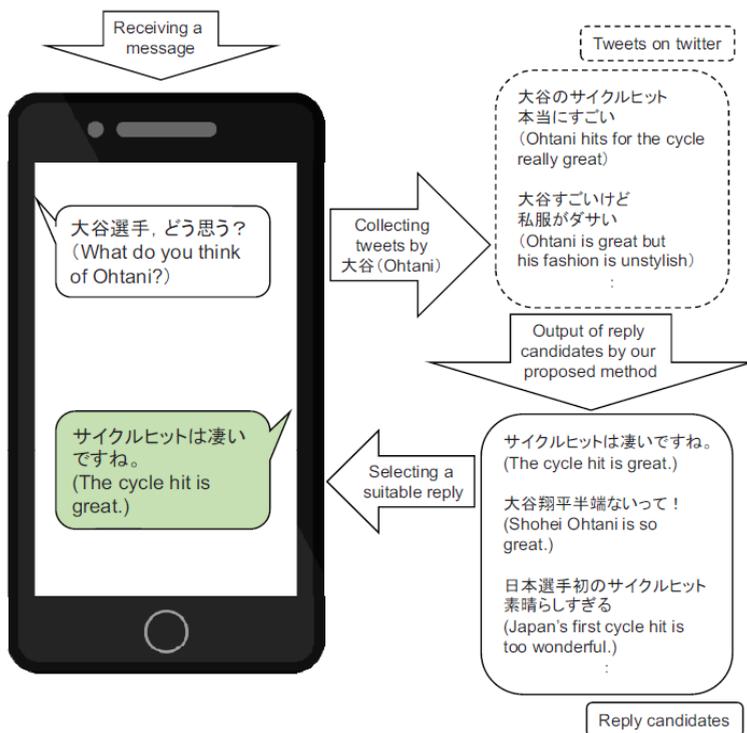
We proposed a method to present multiple non-redundant reply candidates to a user in different situations according to the topic of the received sentence. The content of the received sentence is related to questions about behaviour and emotion. For example, when the received message is ‘How was the movie?’, the user creates different reply sentences such as ‘good’ or ‘boring’. In order to be used by various users in different scenarios and situations, our proposed method presents multiple reply candidates for every situation by clustering and scoring them.

Figure 1 Processing flow



The flow of the proposed method is shown in Figure 1. The system collects tweets which include topics of the received message. Next, the collected tweets are divided into sentences. Clustering is performed on sentences, obtained by dividing the tweet, to create classes according to the situation. Finally, the system calculates the score using the word coverage for each sentence. The score indicates effectiveness of a reply. The system presents sentences, with the highest score in each class, as the reply candidates.

The overall system of the proposed method is shown in Figure 2. The user is able to respond quickly by selecting a suitable reply from presented reply candidates.

Figure 2 Overall system (see online version for colours)

3.2 Collect tweets

In order to present reply candidates according to the contents and topics of the received sentence, the proposed system collects tweets with nouns contained in the received sentence. However, if all the corresponding tweets are collected, inappropriate tweets as reply candidates will also be collected. Five filtering rules proposed by Inaba et al. (2014) are used to eliminate inappropriate tweets as reply candidates.

- 1 Retweets are excluded because they have a lot of advertising texts. They are unsuitable for reply candidates.
- 2 Tweets including URL are excluded because they are usually for image-only posts. Also news-article titles are unsuitable as reply candidates.
- 3 Reply tweets are excluded because it is likely to be a specific for other user. Reply tweets alone may not contain a comprehensible meaning, without its proper context.
- 4 Tweets including hashtag are excluded because they have a lot of advertising texts and mostly they are inappropriate for reply candidates.
- 5 Twitter with limited terms, such as 'フォロー (follow)', 'フォロワー (follower)', 'リプ (reply)' and 'direct message (DM)' are excluded because they cannot be used as general reply candidates.

Filtering is performed by using the above rules. Tweets that did not fit with the above rules are collected as reply candidates.

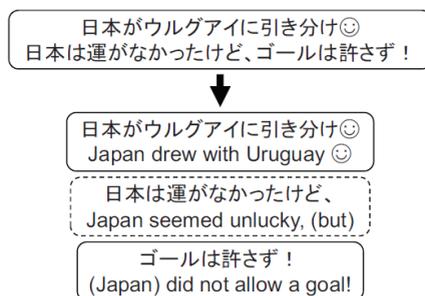
3.3 Divide tweet

A tweet with many words is likely to be not focused to a single issue. Furthermore, since the whole reply sentence cannot be displayed in a small operation screen, it is unsuitable for a small mobile device. Tweets are collected using nouns that appear in the received sentence. If a tweet is used as a reply candidate, redundant reply candidates that include the same noun will always be presented. In order not to present a reply candidate containing many characters and a redundant reply candidate, the system based on our proposed method divides the tweets into sentence units with punctuation marks and symbols. At that time, three filtering rules as set by Inaba et al. (2014) are provided for removing sentences of wrong grammar (problems specific to Japanese), from reply candidates.

- 1 Sentences beginning with the following parts-of-speech 助詞 (particle), 助動詞 (auxiliary verb) and 接続詞 (conjunction) are excluded because the sentence may contain limited or typographical errors.
- 2 Sentences ending with the following parts-of-speech 格助詞 (case particle), 係助詞 (associate particle), 接続助詞 (conjunctive particle) and 並列助詞 (parallel particle), are excluded because it is difficult to indicate the situation.
- 3 Sentences containing 終助詞 (final particle) other than at the end are excluded because they may contain typographical errors and punctuation missing.

An example of filtering using the above rules is shown in Figure 3. The tweet shown in Figure 3 is divided into three sentences by punctuation marks and symbols. At this time, two sentences enclosed by solid lines may be considered as reply candidates. However, the sentence enclosed by dotted lines is excluded because the part-of-speech at the end of the sentence is 接続助詞 (conjunction particle). We removed unsuitable sentences as reply candidates by using the above rules.

Figure 3 Example of dividing tweet

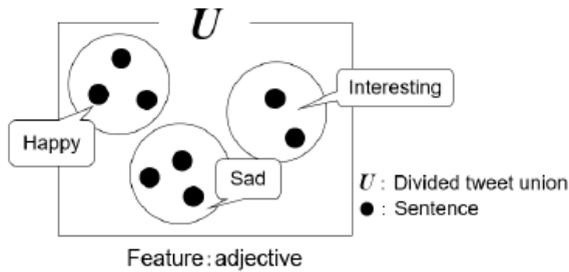


3.4 Clustering

The proposed method uses *K*-means clustering. The features for clustering are term frequency-inverse document frequency (tf-idf) values of words in the sentences. The value is high for words for which the term frequency is high and the document frequency is low, the criterion on which tf-idf is computed. The system uses only words whose parts-of-speech are verbs, adjectives for clustering. If the received sentence is a question about action, clustering with verbs as the feature is effective. If the received sentence is a question about emotion, clustering with adjectives as the feature is effective.

An example of clustering is shown in Figure 4. In adjective-based clustering, sentences are classified to indicate different emotions. In Figure 4, sentences are classified as ‘happy’, ‘interesting’ and ‘sad’.

Figure 4 Example of clustering



A Japanese word often has different spellings even though their meanings and pronunciations are same. For example, the character ‘玉子’ is different from ‘卵’, but their meanings and pronunciations are same, both mean ‘egg’ and the pronunciations are ‘tamago’. Therefore, words are translated into *Katakana* characters in our proposed method. *Katakana* is a Japanese phonetic character. For example, the word ‘玉子’ and ‘卵’ are translated into the same *Katakana* word ‘タマゴ’. The system is then able to regard as the same word.

3.5 Score calculation

The best reply candidate from each cluster is presented to the user, i.e., the number of reply candidates is the same as the number of clusters. In order to evaluate reply candidates, the word coverage for a candidate is calculated. The word coverage WC_s for candidate sentence s belonging to class c is calculated by equation (1) as follows:

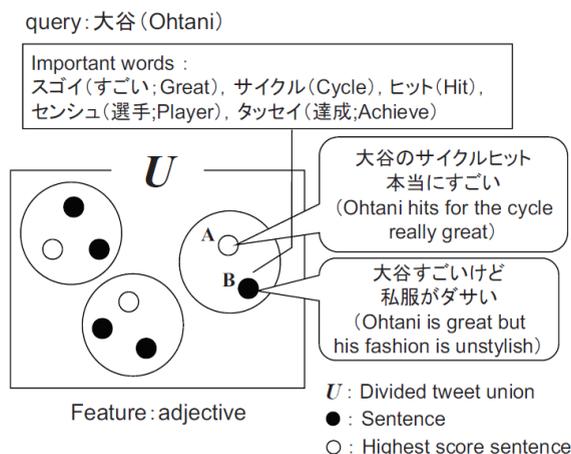
$$WC_s = \frac{|W_s \cap I_c|}{|W_s \cup I_c|} \tag{1}$$

where W_s is the set of verbs, adjectives and nouns contained in sentence s . I_c is a set of important words in class c to which sentence s belongs. Important words are those with high frequency of occurrence in the class. The top N words whose parts-of-speech are verbs, adjectives and nouns are designated as important words in the class. However, nouns used in search queries are not designated as important words. The score for a

sentence that contains many important words and few other words is high. The system is able to present suitable reply candidates with users' needs, by using the word coverage.

An example of calculating score from clustering result is shown in Figure 5. In Figure 5, it is assumed that the user received a message asking questions such as '大谷どう思う? (What do you think of Ohtani?)'. 大谷 means Shohei Ohtani. He is a young Japanese baseball player now participating in US major league. Sentences A and B belong to the same class because both contain the word 'スゴイ (すごい; great)'. Therefore, sentence A and B are considered to be in the same cluster. 'スゴイ (すごい; great)', 'サイクル (cycle)', 'ヒット (hit)', 'センシュ (選手; player)' and 'タッセイ (達成; achieve)' are selected as the set I_c (important words) of the class to which sentence A and sentence B belong when the value of N is equal to 5. In this case, we calculate the scores of sentence A and sentence B as follows.

Figure 5 Example of calculating score from clustering result



In sentence A, three important words 'スゴイ (すごい; great)', 'サイクル (cycle)', 'ヒット (hit)' appear and no other important words appear which form the set W_s . Here, $|W_s \cap I_c| = 3$ and $|W_s \cup I_c| = 5$. The score of sentence A is $\frac{3}{5} = 0.6$. In sentence B, one of the important word 'スゴイ (すごい; great)' appears and two words 'シフク (私服; fashion)', 'ダサイ (ダサイ; unstylish)' appear in the other words which form the set W_s . Here, $|W_s \cap I_c| = 1$ and $|W_s \cup I_c| = 7$. The score of sentence B is $\frac{1}{7} = 0.14$. Since the score of sentence A is higher than that of sentence B, the reply candidate from this cluster is sentence A. In this way, a reply candidate in each cluster is determined.

3.6 Output of reply candidate

The sentence with the highest score in each class is presented to the user as a reply candidate. Because our proposed method uses clustering, each reply candidate indicates a different situation. Like the fixed phrase selection method of smart watch, the user finally selects the suitable reply displayed on the screen. Since multiple reply candidates are

presented, the user finds the intended reply with high probability and communicate smoothly.

4 Evaluation experiment

4.1 Experimental conditions

In order to confirm the effectiveness of our proposed method, we conduct experiments using real tweets and presenting reply candidates in Japanese. Ten subjects of aged from 22 to 25 years evaluate the reply candidates presented by the proposed system. The subjects evaluate whether the reply candidates are suitable or not. We calculate the accuracy for the suitable reply candidates, from subjects' evaluations.

Twelve search queries are used for collecting tweets in this experiment. Ten queries among them are '地震 (earthquake)', 'なでしこ (*Nadeshiko*)', 'オリンピック チケット (Olympic ticket)', '八村塁 (*HachimuraRui*)', 'コパアメリカ (Copa America)', '山ちゃん (*Yamachan*)', '大谷 (*Ohtani*)', 'ラグビー (rugby)', '台風 (typhoon)' and 'ブラックフライデー (Black Friday)'. They are referring to the trend words of Google (<https://trends.google.co.jp/trends/trendingsearches/daily?geo=JP>) from June 14th to November 22nd, 2019. We use the trend words as search queries because the trend words are best samples of what many users talk about on SNS. The rest two queries are '授業 (lecture)' and '年末 (year-end)'. Because we consider that the two words are commonly used on everyday life, we use the words as search queries. Thus, 12 search queries are used in this experiment.

The number of collected tweets for each query is 1,000. The number of clusters for *K*-means is set to 4. We use Python 3.5.1 for the development language and MeCab (<https://taku910.github.io/mecab/>) for morphological analysis.

The top *N* words whose parts-of-speech are verbs, adjectives and nouns are designated as important words in the class. The value of *N* is equal to 3 in this experiment.

4.2 Results and discussion

The clustering result by *K*-means is shown in Table 1. After division, the number of collected sentences for '地震 (earthquake)' is 1,978, that for 'なでしこ (*Nadeshiko*)' is 2,422, and so on. In addition, when we calculate the average number of appearances of features, we can find that no single classes are formed. The average number of features in their classes is no more than 1.0. The fact that the average number of appearances of features does not exceed 1.0 means that the sentences which do not contain any features are classified into one class. Such a class does not indicate a specific situation, and of no interest to our algorithm.

A few sentences in each cluster for the search query '地震 (earthquake)' are shown in Table 2. In Table 2, sentences of class A do not indicate a specific situation because the sentences have no adjective. We can observe that the average appearance of features for the class A is 0.15 and is not more than 1.0 in Table 1. The reply candidate is not presented for the class whose average appearance frequency of feature does not exceed 1.0. In Table 1, the average number of appearances of features of the class A for the search query 'なでしこ (*Nadeshiko*)' is 0.14, that of class A for 'オリンピック チケット

(Olympic ticket)’ is 0.79 and so on. Therefore, the system presents 3 (not 4, the number of clusters) candidates per query to the user.

Important words for each search query are shown in Table 3. ‘地震 (earthquake)’ was the trend word of Google because a large earthquake occurred in Yamagata and Niigata, Japan, on June 18th, 2019. The result for the search query ‘地震 (earthquake)’ has ‘怖い (scary)’, ‘大きい (big)’ and ‘ない(not)’ in classes B, C and D respectively. Since they are adjectives, it seems that each cluster represents a specific situation. In addition, verbs ‘クル (来る; come)’ and ‘ネル (寝る; sleep)’ are selected as important words. There are many sentences including ‘クル (来る; come)’ and ‘ネル (寝る; sleep)’ because the earthquake occurred at night.

Table 1 Clustering result

<i>Query</i>	<i>Class</i>	<i># sentences</i>	<i>Average appearance</i>
地震 (Earthquake)	A	1,763	0.15
	B	114	1.13
	C	31	1.23
	D	70	1.26
なでしこ (<i>Nadeshiko</i>)	A	2,314	0.14
	B	23	1.13
	C	50	1.22
	D	35	1.17
オリンピック チケット (Olympic ticket)	A	1,669	0.79
	B	68	1.28
	C	190	1.30
	D	89	1.76
八村塁 (<i>HachimuraRui</i>)	A	1,700	0.13
	B	198	1.05
	C	33	1.24
	D	29	1.07
コパアメリカ (Copa America)	A	1,911	0.14
	B	20	1.10
	C	91	1.21
	D	35	1.31
山ちゃん (<i>Yamachan</i>)	A	2,079	0.23
	B	92	1.18
	C	67	1.52
	D	96	1.14
大谷 (<i>Ohtani</i>)	A	1,664	0.15
	B	197	1.05
	C	45	1.13
	D	18	1.11

Table 1 Clustering result (continued)

<i>Query</i>	<i>Class</i>	<i># sentences</i>	<i>Average appearance</i>
ラグビー (Rugby)	A	2,083	0.18
	B	85	1.14
	C	50	1.12
	D	64	1.11
台風 (Typhoon)	A	2,278	0.17
	B	106	1.08
	C	36	1.19
	D	17	1.24
ブラックフライデー (Black Friday)	A	1,525	0.13
	B	83	1.23
	C	85	1.19
	D	51	1.51
授業 (Lecture)	A	1,856	0.92
	B	24	1.75
	C	43	2.02
	D	81	1.52
年末 (Year-end)	A	1,946	0.84
	B	66	1.68
	C	61	1.69
	D	77	1.74

Table 2 Sentences in class of ‘地震 (earthquake)’

<i>Class</i>	<i>Sentence</i>
A	<p>ありがとうだけど英語...</p> <p>(Thank you, but English...)</p> <p>ちびちゃんをよく寝てます!</p> <p>(A child is sleeping well!)</p> <p>それ思い出してしまう。</p> <p>(I remember it.)</p> <p>:</p>
B	<p>地震のニュースが怖いのでねる。</p> <p>(I sleep because I'm afraid of the earthquake news.)</p> <p>この時間帯に地震が起こると寝るのも怖いですね。</p> <p>(It's scary to sleep if an earthquake occurs during this time.)</p> <p>怖いなあ...</p> <p>(I'm scared...)</p> <p>:</p>

Table 2 Sentences in class of ‘地震 (earthquake)’ (continued)

<i>Class</i>	<i>Sentence</i>
C	なかなか大きかったね... (It was quite big...) 新潟の方地震大きい (Niigata prefecture has big earthquake.) かなり大きかったみたいで心配です。 (I was worried that it was pretty big.) :
D	もう空中に住むしかないな (I have no choice but to live in the sky.) 恐怖しかない。 (I am scared.) 備えるしかない (I have to prepare.) :

Table 3 Important words

<i>Query</i>	
<i>Class</i>	<i>Important words</i>
<i>地震 (Earthquake)</i>	
B	コワイ (怖い; scary), クル (来る; come), ネル (寝る; sleep)
C	オオキイ (大きい; big), クル (来る; come), ヒガイ (被害; damage)
D	ナイ (ない; not), オモウ (思う; think), コワイ (怖い; scary)
<i>なでしこ (Nadeshiko)</i>	
B	キビシイ (厳しい; severe), ケツショウ (決勝; final), トーナメント (tournament)
C	ナイ (ない; not), ケツタイ (決定; decision), スキ (好き; love)
D	イイ (良い; good), カントク (監督; coach), センシュ (選手; player)
<i>オリンピック チケット (Olympic ticket)</i>	
B	ハズレル (外れる; lose), ゼンブ (全部; all), モウシコム (申し込む; apply)
C	アタル (当たる; win), ヒト (人; people), ケツショウ (決勝; final)
D	モウシコム (申し込む; apply), アタル (当たる; win), ナイ (ない; not)
<i>八村塁 (Hachimura Rui)</i>	
B	スゴイ (凄い; great), シメイ (指名; nomination), ドラフト (draft)
C	ナイ (ない; not), ドラフト (draft), ニッポンジン (日本人; Japanese)
D	イイ (良い; good), カッコ (格好; figure), センシュ (選手; player)
<i>コパアメリカ (Copa America)</i>	
B	ツヨイ (強い; strong), チリ (Chile), レンパ (連覇; consecutive victory)
C	ナイ (ない; not), オモウ (思う; think), ミル (見る; look)
D	ヨイ (良い; good), オモウ (思う; think), イウ (言う; say)

Table 3 Important words (continued)

<i>Query</i>	
<i>Class</i>	<i>Important words</i>
<i>山ちゃん (Yamachan)</i>	
B	ヨイ (良い; good), カイケン (会見; interview), ケッコン (結婚; marriage)
C	スゴイ (凄い; great), イ (井; well), ケッコン (結婚; marriage)
D	イイ (良い; good), カイケン (会見; interview), ヒト (人; people)
<i>大谷 (Ohtani)</i>	
B	スゴイ (凄い; great), サイクル (サイクル; cycle), ヒット (hit)
C	ナイ (ない; not), イウ (言う; say), ハンパ (半端; fragment)
D	スバラシイ (素晴らしい; wonderful), サイクル (cycle), センシュ (選手; player)
<i>ラグビー (Rugby)</i>	
B	スゴイ (凄い; great), ミル (見る; look), キノウ (昨日; yesterday)
C	オモシロイ (面白い; interesting), ミル (見る; look), キノウ (昨日; yesterday)
D	ナイ (ない; not), ミル (見る; look), ヒト (人; people)
<i>台風 (Typhoon)</i>	
B	ナイ (ない; not), ヒガイ (被害; damage), アル (ある; exist)
C	スゴイ (凄い; great), ヒガイ (被害; damage), コンカイ (今回; this time)
D	ヤバイ (やばい; dangerous), エイキョウ (影響; influence), ヨン (四; four)
<i>ブラックフライデー (Black Friday)</i>	
B	ナイ (ない; not), カウ (買う; buy), セール (sale)
C	ヤスイ (安い; cheap), カウ (買う; buy), カエル (帰る; return)
D	ホシイ (欲しい; want), ナイ (ない; not), カウ (買う; buy)
<i>授業 (Lecture)</i>	
B	ワカル (分かる; understand), イウ (言う; say), イイ (良い; good)
C	デル (出る; go out), オモウ (思う; think), ヒト (人; people)
D	オワル (終わる; end), ハヤイ (早い; early), キョウ (今日; today)
<i>年末 (Year-end)</i>	
B	アル (ある; exist), マエ (前; front), イウ (言う; say)
C	カウ (買う; buy), ジャンボ (jumbo), イク (行く; go)
D	イク (行く; go), コミケ (comic market), ヨテイ (予定; schedule)

The result for the search query ‘なでしこ (Nadeshiko)’ has words ‘キビシイ (厳しい; severe)’, ‘ナイ (ない; not)’ and ‘イイ (良い; good)’ in classes B, C and D respectively. ‘なでしこ (Nadeshiko)’ is a nickname for the Japan women soccer team. Since the words are adjectives, each cluster shows a specific feeling. In addition, ‘ケツシヨウ (決勝; final)’, ‘ケツテイ (決定; decision)’ and ‘センシュ (選手; player)’ are selected as important words. There are many sentences about soccer games because the World Cup was held, clustering the period.

Similarly, the result for the search query ‘オリンピック チケット (Olympic ticket)’ has words ‘ハズレル (外れる; lose)’, ‘アタル (当たる; win)’ and ‘モウシコム (申し込む;

apply)’ in classes B, C and D respectively. Since they are verbs, it seems that each cluster shows a specific action. There are many sentences including them because tickets for the Olympic Games in Tokyo 2020 were on sale.

Table 4 Reply candidates, scores and accuracies (‘地震’～‘山ちゃん’)

<i>Query, feature</i>			
<i>Class</i>	<i>Reply candidates</i>	<i>Score</i>	<i>Accuracy [%]</i>
<i>地震 (Earthquake), adjective</i>			
B	地震怖すぎた寝れない (Earthquake too scared to sleep.)	0.67	100
C	津波の被害が大きくなりませんように。 (Do not increase the damage of Tsunami.)	0.50	90
D	地震大したことなくてよかった (It was nice not to have a big earthquake.)	0.33	100
<i>なでしこ (Nadeshiko), adjective</i>			
B	このままだと決勝トーナメントは厳しそうだ... (The final tournament seems to be tough if left unchecked.)	1.0	100
C	なでしこ決定力無すぎ... (Nadeshiko is too decisive.)	0.67	100
D	いい監督だ。 (He is a good coach.)	0.67	70
<i>オリンピック チケット (Olympic ticket), verb</i>			
B	オリンピックチケット全部外れた (I lost all Olympic tickets.)	0.67	100
C	どうやらサッカー決勝が当たったらしい w (Apparently the soccer final was won.)	0.50	100
D	僕は当たるわけがないと端から申し込んでいない。 (I didn’t apply because I don’t think I could win.)	0.75	90
<i>八村塁 (HachimuraRui), adjective</i>			
B	八村塁ドラフト 1 巡目指名まじかよ凄すぎる... (The first round nomination of Hachimura Rui on the draft is so great.)	1.0	90
C	もはや日本人のパワーではない (His power is no longer Japanese.)	0.50	100
D	カッコいいな～! (That’s cool!)	0.67	100
<i>コパアメリカ (Copa America), adjective</i>			
B	コパアメリカ 2 連覇中のチリは流石に強かった! (Chile, who has won Copa America for the second consecutive year, was really strong!)	0.75	70
C	やっぱり日本はコパアメリカをなめてると思われても仕方がない (After all, it seems that Japan is licking Copa America.)	0.50	70
D	良い試合だったと思いますよ。 (I think it was a good match.)	0.50	100
<i>山ちゃん (Yamachan), adjective</i>			
B	とても良い結婚会見だったなあ... (It was a very good marriage interview.)	1.0	100
C	山ちゃんと蒼井優の結婚ほんとにすごいな～ (The marriage between Yamachan and Yu Aoi is really amazing.)	0.75	100
D	山ちゃんの会見いいね! (Yamachan’s interview is good.)	0.67	100

Table 5 Reply candidates, scores and accuracies (‘大谷’～‘年末’)

<i>Query, feature</i>			
<i>Class</i>	<i>Reply candidates</i>	<i>Score</i>	<i>Accuracy [%]</i>
<i>大谷 (Ohtani), adjective</i>			
B	サイクルヒットは凄いですね。(The cycle hit is great.)	1.0	100
C	大谷翔平半端ないって!(Shohei Ohtani is so great.)	0.67	100
D	日本選手初のサイクルヒット素晴らしすぎる (Japan’s first cycle hit is too wonderful.)	0.75	100
<i>ラグビー (Rugby), adjective</i>			
B	昨日見たラグビー凄かったなあ (The rugby I saw yesterday was great.)	1.0	100
C	昨日のラグビー見始めたけどめっちゃめっちゃ面白い (I started watching rugby yesterday but it was really interesting.)	0.75	90
D	俺もラグビー観るしかないのか (I have no choice but to watch rugby.)	0.67	100
<i>台風 (Typhoon), adjective</i>			
B	こちらは台風の被害はなかった。(There was no typhoon damage here.)	0.67	100
C	今回これだけの被害で済んでるのは凄いことだ。(It is great that this damage has been done this time.)	0.75	100
D	台風の影響で田んぼが全体的に死んでてやばい (Rice field is dead due to typhoon.)	0.33	100
<i>ブラックフライデー (Black Friday), adjective</i>			
B	Amazon のブラックフライデー特に買うもの無いな (Amazon’s Black Friday has nothing to buy.)	0.67	100
C	確かに安いからついつい買ってしまう... (I’m sure they’re cheap so I just buy them.)	0.50	90
D	本当に欲しい物を買わないとね! (I have to buy what I really want.)	0.67	100
<i>授業 (Lecture), verb</i>			
B	どこに座ったらいいか分からんから帰ってきた... (I came back because I don’t know where to sit.)	0.40	100
C	授業すら出てないのに単位出ると思ってるの? (Do you think you will get credit even if you don’t have classes?)	0.50	50
D	いつもより早く授業終わった (Classes ended earlier than usual.)	0.67	100
<i>年末 (Year-end), verb</i>			
B	そういえば年末賞与なんてワードありましたね (Yes, there was a word the year-end bonus.)	0.40	70
C	年末ジャンボ買おうかな (I will buy Year-end jumbo lottery.)	0.67	80
D	まあ年末はコミケに行つて特別な日にするけど。(Well, at the end of the year we go to Comic Market for a special day.)	0.50	90
Total		0.66	93.1

Table 4 and Table 5 show the reply candidates presented to the user, their scores and accuracies evaluated by subjects. In all search queries, the system outputs reply candidates that indicate different situations, emotions or actions. All reply candidates include different important words and are non-redundant. The number of important words is three and fixed in this experiment, i.e., $N = 3$. We need to confirm that the number is appropriate because it is decided empirically.

The accuracy of the reply candidate in class C for the search query ‘授業 (lecture)’ is 50%, the minimum accuracy in Table 4 and Table 5. The evaluation is low since the reply candidate is an interrogative sentence. We need to work further to improve this, e.g., the system does not collect the tweets which include an interrogative sentence or decreases their scores and so on.

The content of the reply candidates for the search query ‘大谷 (*Ohtani*)’ is biased. The meaning of the words ‘凄い (great)’ and ‘素晴らしい (wonderful)’ in the sentences is similar. The system is able to present unbiased candidates if the clustering is done using other features. Therefore, we will do experiments with different features in future.

The total accuracy rate is 93.1% from all results in Table 4 and Table 5. Thus, it is proved that our proposed method is effective. We consider that smooth and rapid response could be facilitated using our proposed method.

5 Conclusions

In this paper, we have proposed a method to present multiple non-redundant reply candidates to a user in different situations and time, appropriate to the topic of the received message. The system collects tweets including topics of the received sentence when receiving a message. Next, the collected tweets are divided into sentences. Clustering is performed on the tweets on divided sentences, to create clusters with different characteristic features. Then, the system calculates the score indicating the effectiveness of a reply candidate using the word coverage for each sentence. Finally, the system presents the user sentences with the highest score in each cluster. The number of reply candidates is approximately the number of clusters. The reply process is completed as the user selects one from the reply candidates.

In order to make simple and suitable reply candidates, the collected tweets are divided into sentences because a tweet contains sentences that may be redundant. It is possible to choose only appropriate sentences as reply candidates by the division of sentences.

The features for clustering are verbs, adjectives. A cluster represents a specific situation, emotion or action. Reply candidates are outputted from each class by evaluating the word coverage. The word coverage for sentences that contain many important words and no other words, is high. It is possible to present simple reply candidates indicating specific situation by eliminating redundancy.

Our proposed method is proved to be effective by the evaluation experiment. However, not complete messages but only keyword/s from the messages is used in the experiment. We need to consider how to decide and extract proper keyword/s as the topic of the message, evaluating our whole system to assess the efficacy of the topic extraction concept. In this experiment, the number of important words is three and fixed. We are going to continue experiment by dynamically adjusting the number of important words depending on the received message. Moreover, we consider applying re-clustering for presenting unbiased reply candidates.

Acknowledgements

This work was supported by JSPS KAKENHI Grant number 18K11358.

References

- Banchs, R.E. and Li, H. (2012) 'Iris: a chat-oriented dialogue system based on the vector space model', *Proc.ACL2012Association for Computational Linguistics*, pp.37–42.
- Boustila, S., Guegan, T., Takashima, K. and Kitamura, Y. (2019) 'Text typing in VR using smartphones touchscreen and HMD', *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Osaka, Japan, pp.860–861.
- Corsten, C., Daehlmann, B., Voelker, S. and Borchers, J. (2017) 'BackXPress: using back-of-device finger pressure to augment touchscreen input on smartphones', *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, May, pp.4654–4666.
- Fujita, K. (2019) 'Text entry system for smartphone operation with user's dominant hand', *2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI)*, Toyama, Japan, pp.980–984.
- Go, A., Bhayani, R. and Huang, L. (2009) *Twitter Sentiment Classification using Distant Supervision*, CS224N Project Report, Stanford.
- Hong, J., Heo, S., Isokoski, P. and Lee, G. (2015) 'A simple split soft keyboard for wristwatch-sized touch screens', *The 33rd Annual ACM Conference on Human Factors in Computing Systems(CHI2015)*, pp.1233–1236.
- Inaba, M., Kamizono, S. and Takahashi, K. (2014) 'Candidate utterance acquisition method for non-task-oriented dialogue systems from Twitter', *The Japanese Society for Artificial Intelligence*, in Japanese, Vol. 29, No. 1, pp.21–31.
- Jiang, H. and Weng, D. (2020) 'HiPad: text entry for head-mounted displays using circular touchpad', *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Atlanta, GA, USA, pp.692–703.
- Jiang, L., Yu, M., Zhou, M., Liu, X. and Zhao, T. (2011) 'Target-dependent Twitter sentiment classification', *Proc. ACL2011*, pp.151–160.
- Lasguido, N., Sakti, S., Neubig, G., Tomoki, T. and Nakamura, S. (2014) 'Utilizing human-to-human conversation examples for a multi domain chat-oriented dialog system', *IEICE Trans. Inf. & Syst.*, June, Vol. E97-D, No. 6, pp.1497–1505.
- Oney, S., Harrison, C., Ogan, A. and Wiese, J. (2013) 'ZoomBoard: a diminutive QWERTY soft keyboard using iterative zooming for ultra-small devices', *The SIGCHI Conference on Human Factors in Computing Systems(CHI2013)*, pp.2799–2802.
- Wallace, R.S. (2008) 'The anatomy of A.L.I.C.E.', *Parsing the Turing Test*, pp.181–210, Springer Netherlands.
- Watanabe, N., Matsuhara, M., Chakraborty, G. and Mabuchi, H. (2019) 'Method to present situation aware suitable reply candidate on mobile terminal', *The 10th IEEE International Conference on Awareness Science and Technology (iCAST 2019)*, Morioka, Japan, 23–25 October, pp.395–400.
- Weizenbaum, J. (1966) 'ELIZA – a computer program for the study of natural language communication between man and machine', *Commun. ACM*, Vol. 9, No. 1, pp.36–45.