

ANALYSING THAI SOCIAL MEDIA CONTENT TO IMPROVE CUSTOMER SATISFACTION

Jitrlada ROJRATANAVIJIT
Metropolitan Electricity Authority – THAILAND
Jitrlada@mea.or.th

Chingchai EIAMSITHIPAN
Metropolitan Electricity Authority – THAILAND
Chingchai@mea.or.th

ABSTRACT

The emergence of social media in Thailand has given millions of users a platform to express and share their opinions about products and services, and so social media platforms are considered to be a rich source of information for companies to understand their customers. This offers companies a fast and effective way to monitor public opinions on their brands, products, services, etc. The Metropolitan Electricity Authority (MEA) is concerned about this situation and developed a voice of the customer management system (VOCMS) that uses and analyses data related to customers, payments, electrical usage, power outage events, complaints and customer calls. In addition, it also retrieves feedback from customers via various social media channels. However, sentiment analysis performed on Thai social media has challenges brought about by language-related issues, such as the differences in writing systems between Thai and English, short-length messages, slang words, and word usage variation. This paper focuses on social media content classification and on solving data sparsity issues. We use lexicon-based techniques to classify them into positive, negative, or neutral sentiments. The procedure of analysing Thai social media content is subdivided into three modules: (1) data retrieval, (2) data pre-processing and (3) data classification. The results from this initiative project can help the MEA to improve its customer services and satisfaction by enabling quick responses to customer complaints by obtaining more details to help solve problems. Moreover, the information from analysing social media information could give customers knowledge and could be helpful for both long-term and short-term planning.

INTRODUCTION

Customer Satisfaction

Customer satisfaction is one of the essential parts of business achievement. It can help companies to evaluate their ability in meeting their customers' needs and expectations effectively and analyse the performance of an offering to customers in order to identify areas for improvement as well as which areas customers consider to be very important to them [1]. According to Goodman [2], "customer satisfaction is complicated to define because normally it is extenuated by expectation. Customers will be satisfied if the product or service above demands or

expectation." Customers who are satisfied with a company's offering may tell others about it (positive word-of-mouth), just as a dissatisfied customer may also badmouth the company to others. He suggested that dissatisfied customers tell ten people on average about a company compared to five people by satisfied customers and also stated that it costs five times more to attract a new customer than to maintain a current one. Therefore, it is necessary for service businesses to satisfy customers on a consistent basis.

Social Media in Thailand

The emergence of social media in Thailand has supplied millions of users with social media platforms to express and share their opinions about products and services (among other subjects), and so it is considered as a rich source of information for companies to understand and communicate with their customers. In 2018, the Electronic Transactions Development Agency of Thailand (ETDA) stated that Thais spend more than 9 hours each day using the Internet for pursuits such as social networking, chatting and messenger applications and VoIP [3]. Suchit Leesanguansuk reported that Thoth Zocial OBVOC, which monitors social media trends in Thailand, researched and summarised social media data indicating that Thailand ranks in the World's top 10 for social media usage, with video becoming more popular and entertainment the topic most talked about on social media [4]. In 2017, there were 49 million Facebook users in Thailand, which comprises a 4% growth compared to the previous year. He mentioned that among Facebook users globally, Thailand ranks 8th in the world. More specifically, he also stated that the number of Bangkok Facebook users is 24 million per day and they prefer to use the app more often between 6 pm and 11 pm.

The large amount of feedback on social media coming directly from customers has become a new source from which to mine what is referred to as customer insight. Zinner and Zhou mentioned that "Social media feedback is usually more emotional – that it is strongly slanted toward an opinion." In comparison, traditional feedback from customers via other channels typically contains less emotion and sentiment [5]. Thus, social media feedback could help companies to consider how customers really feel about an issue. However, the data from social media are unpredictable and the amount is growing exponentially, and so machine learning and text mining are being applied to the vast amount of social media data to

detect and discover new knowledge and interesting patterns. This information could be helpful for companies to understand what customers really want and what their competitors are doing [6].

Customer Analytics Project

In 2018, to improve customer satisfaction, the Metropolitan Electricity Authority (MEA) developed a voice of the customer management system (VOCMS) that uses and analyses data related to customers, payments, electrical usage, power outage events, complaints and customer calls. In addition, it also retrieves feedback from customers via various social media channels such as Facebook, Twitter and the Pantip web board. The system was developed based on the Hadoop platform and consists of four main components: (1) data collection, (2) data pre-processing, (3) data analytics and (4) visualisation [7]. However, VOCMS still has a number of limitations when extracting hidden knowledge from social media content, and so our focus is on the unstructured data from social media content.

METHODOLOGY

The main focus of this study is on Thai social media content classification and on solving data sparsity issues. The main idea is to pre-process the raw data and operate variant transformations to deal with the slang, transliterated words, abbreviations and other noise. Additionally, there are no spaces between words in Thai, so they must be segmented before being fed to the classifier. There are significant differences between written Thai and English. English has 26 letters: 5 vowels and 21 consonants whereas Thai has 59 letters: 44 consonants (Thai: พยัญชนะ; phayanchana), 15 vowels (Thai: สระ; sara) and four tone diacritics (Thai: วรรณยุกต์ or วรรณยุกต์; wannayuk or wannayut) [8]. In English, a space is used between words to separate them and there is punctuation such as full stops (.) to indicate the end of a sentence. In Thai, there are no spaces between words; spaces in Thai content demonstrate the end of a clause or sentence. Therefore, existing text mining and sentiment analysis techniques cannot be directly applied to the Thai language.

The Procedure of System

The process is subdivided into three modules: (1) data retrieval, (2) data pre-processing and (3) data classification. Social media content is obtained from the VOCMS using JDBC drivers for Impala. The data pre-processing is used to extract the social media content for text pre-processing and for Thai word segmentation. For the classification module, the main objective is to identify the polarity of the content. The proposed method is shown in **Figure 1**.

Data Retrieval Step

The VOCMS retrieves social media data from Twitter, Facebook, Google+, the Pantip web board and news feeds, which are updated every three hours. A retrieval program was developed to retrieve social media content from the VOCMS using JDBC drivers for Impala. Next, the social media content is sent to the data pre-processing step.

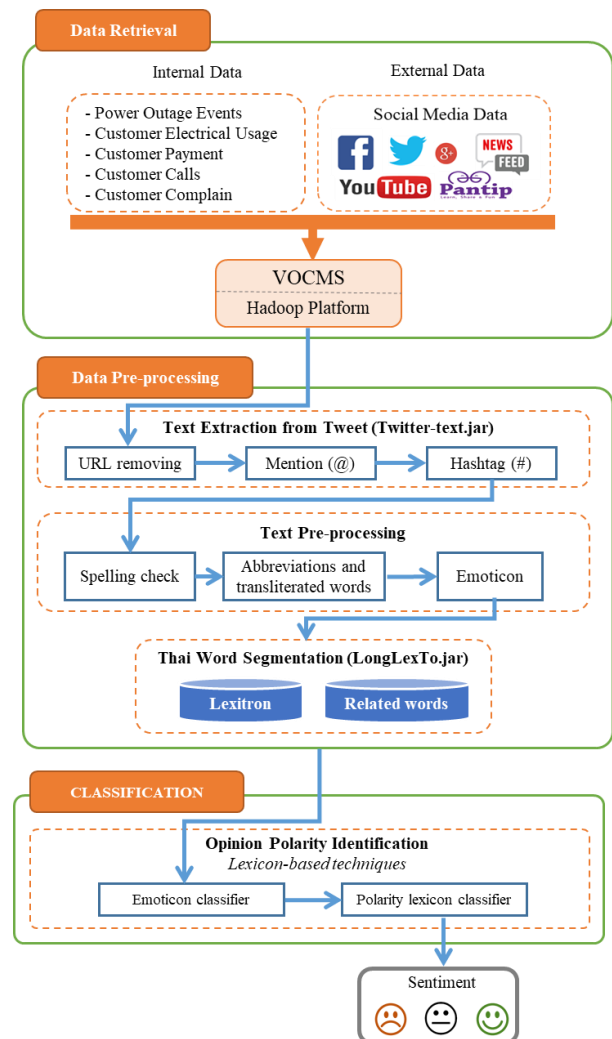


Figure 1 Process of Identifying Sentiment Thai social media content from VOCMS

Data Pre-processing Step

Social media content usually includes text possibly with special symbols such as user mentions (@), hashtags (#), URLs (HTTP links), and so on. Additionally, because of the difference in the Thai and English writing systems (e.g., there are no spaces between words in the Thai language), existing text pre-processing techniques cannot be directly applied to the Thai sentiment classification system. In order to pre-process social media content, we

used the procedure of Acquiring Sentiment for Twitter System (ASTS) [9] comprising three sub-steps as follows:

1) Text extraction is the first sub-step of the data pre-processing step. The Twitter-text java library [10] was applied to extract the identities of the social media content from text messages: URLs, user mentions (@) and hashtags (#). We removed all of the URLs. For user mentions, @ is used to indicate a user account; we removed the @ sign from these user mentions except for those that matched keywords. For the hashtags, we eliminated the # sign and kept all the hashtag text.

2) For text pre-processing, we defined four types of word, namely abbreviations, transliterated words, slang words and misspelt words [9]. The texts passed from the previous step were automatically checked against the words, then they were replaced by expansions or the original words. Emoticons are domain and language independent and have become an important token for social media content since they can be used to express the feelings of the writer in the form of icons. They were assigned with token labels in each Tweet, as shown in **Table 1-2**.

Table 1 Examples of positive emoticon sets.

Positive Emoticons		
Emoticons	Meaning	Token Label
:-) , :) , :D , :o) , :]	Happy	ehappyw
(^v^)^ , (^u^)^ , (^o^)^ , ^-^	Happy	ehappye
:-D, 8-D, XD, =3, B^D	Laugh	elaughw

Table 2 Examples of negative emoticon sets.

Negative Emoticons		
Emoticons	Meaning	Token Label
:- (, :(, >:[, :<	Sad	esadw
T-T , T^T , '_ ' , =_ =	Sad	esade
:'(, :(Cry	ecryw

3) For Thai word segmentation, we used the LongLexTo library developed by the National Electronics and Computer Technology Center (NECTEC), Thailand [11]. For the Thai word segmentation process, we modified the LongLexTo java library with a total of 42,983 words: 42,221 words from the Lexitron data dictionary [11] and 762 words from related words in the utility and sentiment domains. Texts passed on from the text pre-processing step were automatically split into word tokens. For any English words (often included in Thai Tweets), conversion to lowercase was carried out. Lastly, other symbols and numbers were removed.

Classification Step

For the classification process, we used a lexicon-based technique to filter Thai opinions and classified them into positive, negative or neutral sentiment. In order to identify the polarity of an opinion, we use two sub-processes for this step by using the main idea from ASTS [9]. The details of our method for opinion polarity identification are shown in **Figure 2**.

```

Procedure Classification
    SentimentScore = 0
    // checking each emoticon token word in text using emoticon classifier
    For each emoticon token Do
        If word token is in a positive emoticon token list
            SentimentScore +1
        Else If word token is in a negative emoticon token list
            SentimentScore - 1
        End If
    End For
    If SentimentScore > 0 Then classified POSITIVE
    Else If SentimentScore < 0 Then classified NEGATIVE
    Else If SentimentScore = 0 Then Go To the polarity lexicon classifier
    End If
    // checking each word in text using the polarity lexicon classifier
    For each word token Do
        If word token is in a positive word list
            SentimentScore +1
        Else If word token is in a negative word list
            SentimentScore - 1
        End If
    End For
    If SentimentScore > 0 Then classified POSITIVE
    Else If SentimentScore < 0 Then classified NEGATIVE
    Else If classified NEUTRAL
    End If
End
    
```

Figure 2 The classification algorithm

Example 1:

Input Message: รถบรรทุกพุ่งชนเสาไฟฟ้าหัก 2 ต้น บนถนนรามคำแหง ที่ปากซอยพวงศิริ การจราจรเคลื่อนตัวช้า (A truck hit 2 concrete power poles and they were smashed off at ground level on Ramkhamhaeng rd. at Soi Phasiri and traffic was slow.)

Output Message: รถบรรทุก พุ่งชน เสาไฟฟ้า หัก 2 ต้น บนถนนรามคำแหง ที่ ปากซอย พวงศิริ การจราจร เคลื่อนตัวช้า

Output Sentiment: Negative

Example 2:

Input Message: สอบถามการลงทะเบียนรับสิทธิ์ค่าไฟฟ้าบัตรสวัสดิการแห่งรัฐ (How to register for getting free electricity tariff for those who have Thailand welfare cards?)

Output Message: สอบถาม การลงทะเบียน รับสิทธิ์ ค่าไฟฟ้า บัตรสวัสดิการแห่งรัฐ

Output Sentiment: Neutral

Example 3:

Input Message: ชำระค่าไฟฟ้าด้วยบัตรเครดิตได้มั๊ย (Can I pay electricity bills by credit card?)

Output Message: ชำระ ค่าไฟฟ้า ด้วย บัตรเครดิต ได้ มั๊ย

Output Sentiment: Neutral

IMPLEMENTATION PLAN

VOCMS was launched in January 2018 and the information from structured data has been used in the strategic planning process. It was used in three dimensions: product leadership, operational excellence and customer intimacy [7]. Moreover, the analysed unstructured data (social media content) is part of an initiative project for MEA social media management and marketing.

The first implementation was to solve customer dissatisfaction from messages assigned as negative sentiment. Firstly, they were sent to the customer complaint management team to analyse, manage and resolve these issues, after which they were collected and summarized to help create long-term solutions.

The second implementation was aimed at increasing customer satisfaction. In 2018, the statistics in the Introduction show that Thailand ranks in the World's top 10 for social media usage and Thais also spend more than 9 hours each day using the Internet. In fact, since 2012, MEA has opened an official account on Facebook (Metropolitan.Electricity.Authority), Youtube (MEA Multimedia) and Twitter (@mea_news), and an MEA Line account (@meanews) was opened in 2015. These channels were purposely used for posting news on MEA. Based upon an MEA survey on customer behaviour research conducted in 2018, the results show that customers prefer to communicate with the company through social media channels. Therefore, in April 2018, MEA opened two-way communication on their official social media accounts operated by the call centre team.

The third implementation, which is the long-term plan, is to integrate social media with CRM data. This plan is aimed at merging digital and traditional communication channels to engage customers more effectively. Integrating the CRM system and social media content should be about converting conversations into transactions [12]. However, to map and integrate customer accounts with social media accounts is a challenge that is currently being implemented mainly because MEA still needs to upgrade and modify the workflow, processes, rules, data structure, training, call centre, people, etc. with the existing CRM system. Therefore, in 2019, MEA plans (1) to create a digital marketing campaign for mapping

accounts, (2) to create the new services and (3) to upgrade the CRM system. In addition, a specialist team is needed to establish a social media roadmap to consider three levers [12]:

People: culture, skills, training, planning policies, etc.

Processes: organisational processes such as marketing, customer service, customer data management, etc.

Technology: content management, integration, infrastructure, analytics, software, use cases, etc.

CONCLUSION

This results from this initiative project clearly show that enhancing the VOCMS with social media analytics has been beneficial for helping MEA to improve its customer services and satisfaction by enabling quick responses to customer complaints by obtaining more details to help solve problems. Moreover, the information from analysing social media information could give customers knowledge, especially on the experience of using the company's services, thereby helping MEA to understand its customers and to deliver a great customer experience to them. In addition, it could be helpful for both long-term and short-term planning, which will definitely increase the efficiency and effectiveness of planning activities related to the customer. However, with the increasing usage of social media in Thailand and customer behaviour changing rapidly, MEA needs a specialist team to establish a social media roadmap to consider people, processes and technology.

Acknowledgments

This project was supported by the MEA Information Technology Development Department.

REFERENCES

- [1] V. A. Zeithaml; M. J. Bitner and D. D. Gremler, 2009, *Services Marketing: Integrating Customer Focus Across the Firm*, McGraw-Hill Education, New York, 80-81.
- [2] J. A. Goodman, 2009, *Strategic Customer Service: Managing the Customer Experience to Increase Positive Word of Mouth, Build Loyalty, and Maximize Profits*. AMACOM, New York, 17-23.
- [3] ETDA, 2018, *Thailand Internet User Profile 2018 (In Thai)*, Available at <https://www.etcha.or.th/content/etcha-reveals-thailand-internet-user-profile-2018.html>, accessed 1 Dec 2018.

- [4] L. Suchit, 2018, *Thailand Makes Top 10 in Social Media Use*, Available at <https://www.bangkokpost.com/business/news/1420086/thailand-makes-top-10-in-social-media-use>, accessed 1 Dec 2018.
- [5] C. Zinner and C. Zhou, 2011, Social Media and the Voice of the Customer, In: N Smith, R Wollan and C Zhou (eds.). *The Social Media Management Handbook: Everything You Need to Know to Get Social Media Working in Your Business*. John Wiley & Sons, New Jersey, 67-70.
- [6] L. Dey; S. M. Haque; A. Khurdiya and G. Shroff, 2011, "Acquiring competitive intelligence from social media", *Proceedings of the 2011 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data*, Beijing, China.
- [7] R. Jitrlada and E. Chingchai, 2018, "Customer Analytics for Enhancing the Customer Experience", *Proceedings CEPSI Conference*. TNB Malaysia, G.4:1-6.
- [8] Wikipedia, 2016, *Thai alphabet (in Thai)*, Available at https://en.wikipedia.org/wiki/Thai_alphabet, accessed January 2016.
- [9] R. Jitrlada; V. Preecha and P. Sukanya, 2018, "Acquiring Sentiment from Twitter using Supervised Learning and Lexicon-based Techniques", *Walailak Journal of Science & Technology*. Vol. 15 Issue 1, 63-80.
- [10] C Goncalves. GitHub Inc, 2012, *Twitter-text Library*, Available at <https://github.com/twitter/twitter-text>, accessed January 2015.
- [11] NECTEC, 1994, *LexTo - Thai Lexeme Tokenizer (in Thai)*, Available at <http://www.sansarn.com/lexto>, accessed August 2014.
- [12] COGNIZANT, 2013, *Social Media for Utilities: Developing a Satisfying Customer Experience*, Available at <https://www.cognizant.com>, accessed October 2018.