

Leveraging Writing Systems Change for Deep Learning Based Chinese Emotion Analysis

Rong Xiang¹, Yunfei Long¹, Qin Lu¹, Dan Xiong¹, I-Hsuan Chen²

¹Department of Computing, The Hong Kong Polytechnic University
csrxiang,csyulong,csluqin,csdxiong@comp.polyu.edu.hk

²Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University
ihcucb@gmail.com

Abstract

Social media text written in Chinese communities contains mixed scripts including major text written in Chinese, an ideograph-based writing system, and some minor text using Latin letters, an alphabet-based writing system. This phenomenon is called writing systems changes (WSCs). Past studies have shown that WSCs can be used to express emotions, particularly where the social and political environment is more conservative. However, because WSCs can break the syntax of the major text, it poses more challenges in Natural Language Processing (NLP) tasks like emotion classification. In this work, we present a novel deep learning based method to include WSCs as an effective feature for emotion analysis. The method first identifies all WSCs points. Then representation of the major text is learned through an LSTM model whereas the minor text is learned by a separate CNN model. Emotions in the minor text are further highlighted through an attention mechanism before emotion classification. Performance evaluation shows that incorporating WSCs features using deep learning models can improve performance measured by F1-scores compared to the state-of-the-art model.

1 Introduction

Emotion analysis has been studied using different NLP methods from a variety of linguistic perspectives such as semantic, syntactic, and cognitive properties (Barbosa and Feng, 2010; Balamurali et al., 2011; Liu and Zhang, 2012; Wilson et al., 2013; Joshi and Itkat, 2014; Long et al., 2017). In many areas, such as Hong Kong and the Chinese Mainland, social media text is often written in mixed text with major text written in Chinese characters, an ideograph-based writing system. The minor text can be written in En-

glish, emoji, Pinyin¹ (phonetic denotation for Chinese), or other new Internet shorthand notations using Roman characters of some Latin-based writing systems. Using mixed characters in different writing systems is known as WSCs.

Generally speaking, WSCs refers to the use of mixed text which switches between two or more writing systems (Clyne, 2000; Lee and Liu, 2012). A narrower definition, often referred to as code-switching, is the use of more than one linguistic variety in a manner consistent with the syntax and phonology of each variety². The use of alteration of different systems or languages of symbols is rooted in pragmatic and socio-linguistic motivations (Cromdal, 2001; Musk, 2012). The use of WSCs is a case of Economy principle in language (Vicentini, 2003) which is pursued by human being in various activities due to the innate indolence. It aims at the maximum effect with the least input. For instance, 'Good luck' becomes more popular than the Chinese version of '祝你好运' (*Good luck*) because inputting the English version takes shorter time in expressing the same emotion.

Studies in social psychology (Bond and Lai, 1986; Heredia and Altarriba, 2001) also show that WSCs is an effective and commonly used strategy to express emotion or mark emotion change especially in some society where social and political environment is more conservative (Wei, 2003). For instance, a new-born swear word 'zz' is often used in place of the Chinese version of 'moron'. This is because 'zz', which is the acronym of the Pinyin 'zhi zhang (moron)', looks less disrespectful lexically and more acceptable in social networks. With the rapid growth of internationalization, Chinese youngsters like to use English acronyms such as 'wtf' (*what the fuck*) 'stfu' (*shut*

¹<https://en.wikipedia.org/wiki/Pinyin>

²<https://en.wikipedia.org/wiki/Code-switching>

the fuck up). People also use WSCs to express idiosyncrasies written in English or other languages because text in other writing systems is much more difficult to be censored. For example, the sensitive term of democracy in Chinese (民主'min zhu') is often written as an intended misspelled Pinyin *minzu* or the English word *democracy*. This paper studies WSCs related textual features from the orthography perspective to explore their effectiveness as emotion indicators.

Previous studies in emotion analysis mostly rely on emotion lexicon, context information, or semantic knowledge to improve sentence level classification tasks. This linguistic knowledge is often used to transform raw data into feature vector, called feature engineering (Kanter and Veeramachaneni, 2015). However, WSCs can break the syntax of the major text and the switched minor text also lacks linguistic cues in this type of social media data (Dos Santos and Gatti, 2014). This makes feature engineering-based methods difficult to work. Neologism in the Internet forums increases the difficulty for both syntactic and semantic analysis. In particular, newly coined phrases tend to contain different types of symbols. Despite the challenges, this type of datasets is rich in shifts of writing systems orthographically. This characteristic offers reliable clues for emotion classification. Since WSCs is relatively common in real-time on-line platforms like microblog in China³. This work adopts a broader scope of WSCs to include switching between two languages, and change of writing systems in the same language such as Chinese characters to Pinyin notations. Notably, the accessibility of different character sets and symbols, as well as the frequent exposures to other languages and cultures characterize the nature of such short and informal text.

This paper presents our work in progress which uses a novel deep learning based method to incorporate textual features associated with WSCs via an attention mechanism. More specifically, the proposed Hybrid Attention Network (HAN) method first identifies all WSCs points. The representation of the major text is learned through a Long-Short Term Memory (LSTM) model whereas the representation of the minority text is learned by a separate Convolution Neural Network (CNN). Emotions expressed in the minor text are further highlighted through an attention mechanism

before emotion classification. The attention mechanism is achieved by projecting the major text representation into attention vectors while aggregating the representation of the informative words from WSCs context.

2 The Hybrid Attention Network Model

Let D be the dataset as a collection of documents for emotion classification. Each document d_i is an instance in D . The goal of an emotion analysis is to predict the emotion label for each d_i . The set of emotion labels includes $\{Happiness, Sadness, Anger, Fear, Surprise\}$. Let us use the term WSC segments to refer to the minor WSC text pieces. WSC segments can be easily marked in a pre-processing step using code ranges of Chinese characters and Romanized Pinyin or English text.

To make better use of WSCs scripts, a deep learning based HAN model is proposed to explicitly assemble WSCs information in an attention mechanism. Figure 1 shows the framework of HAN. The LSTM model on the left side is used to learn the representation of a document including the WSC segments. This is because documents with WSCs are generally coherent and intact despite few WSC segments that may break the syntax. The CNN model on the right is used to learn the representation of WSCs segments extracted from the sentence because they often occur discontinuously without syntactic structure. The outputs of both models are integrated into a hybrid attention layer before classification is carried out.

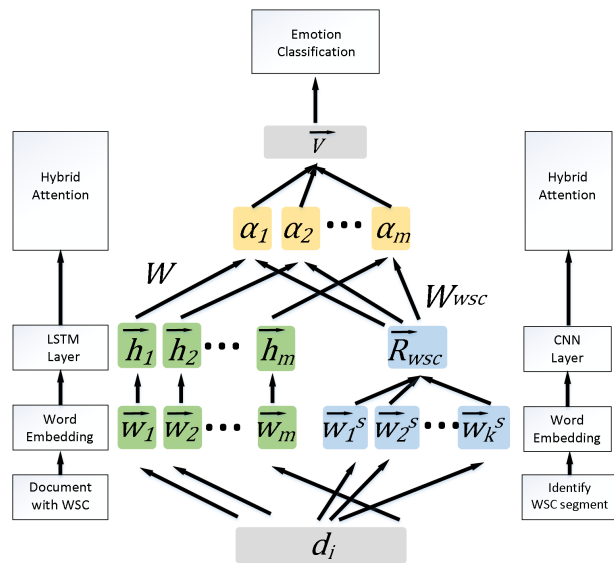


Figure 1: Hybrid attention network framework

³<https://en.wikipedia.org/wiki/Microblogging>

Using deep learning methods, the word representation in $d_i = \vec{w}_1, \dots, \vec{w}_m$, d_i , is learned using two networks. To distinguish the WSC units, they are given designed switch labels $\vec{w}_j^s (\vec{w}_j^s \subset d_i, j = 1 \dots k)$ and are extracted to be fed into the CNN as an extra feature. d_i is fed into LSTM to generate the hidden vector $\vec{h}_1, \vec{h}_2 \dots \vec{h}_m$ from d_i . In Chinese social media, WSCs segments are generally dispersed sporadically. So, for d_i with k WSC segments, the convolution is calculated using a sliding window of size $2n + 1$:

$$\overrightarrow{conv}_p = \sum_{p=k-n}^{k+n} \vec{w}_p^s, \quad (1)$$

and

$$\vec{R}_{wsc} = \frac{\sum_{p=1}^k \overrightarrow{conv}_p}{k}. \quad (2)$$

The WSC feature vector R_{wsc} is generated by average pooling.

Attention model was introduced by Yang et al.(2016) to show different contribution of different words semantically. To include both the information learned from LSTM and CNN, the consolidated representation, \vec{u}_p , includes the representations of both \vec{h}_p , and the WSC representation vector \vec{R}_{wsc} into a perceptron defined below:

$$\vec{u}_p = \tanh(W\vec{h}_p + W_{wsc}\vec{R}_{wsc} + b). \quad (3)$$

In order to re-evaluate the significance of word \vec{w}_p , a coefficient vector \vec{U} is introduced as an informative representation of the words in a network memory. The representation of a word \vec{u}_p and the corresponding word-level context vector \vec{U} is integrated to obtain a normalized attention weight:

$$\alpha_p = \frac{\exp(\vec{u}_p \cdot \vec{U})}{\sum_p \exp(\vec{u}_p \cdot \vec{U})}. \quad (4)$$

The updated document representation \vec{v} can be generated as a weighted sum of the word vectors given below:

$$\vec{v} = \sum_p (\alpha_p \vec{h}_p), \quad (5)$$

where \vec{v} contains both document information and WSCs representation with attention shall be used in the final SoftMax method, producing the output vector. Lastly, an argmax classifier is used to predict the class label.

3 Performance Evaluation

A Chinese microblog dataset is used for performance evaluation (Lee and Wang, 2015). We first present the dataset with some analysis first and then proceed to make performance comparison to baseline systems on emotion classification.

3.1 Dataset and Statistics

The dataset for WSCs is collected from Chinese microblog by Lee's group (2015). It contains 8,728 instances with an average length of 48.8. Every instance contains at least one WSCs script. In previous studies, half is used as the training set and the rest serves as the testing set.

The major text is written in Chinese characters. The WSC segments contain English words and Pinyin scripts, acronyms of Pinyin or other scripts. The annotation of emotions in each instance allows more than one class label. Each instance is labeled independently by the five emotion classes, happiness, sadness, anger, fear and surprise, based on the Ekman model (1992) except for Disgust. The emotion label can be contributed by Chinese text (E1), WSCs (E2) or Both (E3). Some of the instances have NULL emotion labels (E4). Out of the 6 labels, 25% of all instances has the happiness label, which is the most significant emotion. 16% has the sadness label, the percentages of anger, fear, surprise and NULL labels are 9%, 9%, 11% and 30%, respectively. Below shows four example instances in the three WSC types:

E1 Emotion: Happiness

这个年每天都吃好饱! 初三来点小朋友的最爱 麦当劳和pizza! ! (We are so full for every meal during Spring Festival! Will take kids to their favorite, MacDonald and pizza!)

E2 Emotion: Anger

我们会因为金希澈开了微博而看到许多无下限的nc发言。(We will see a lot of morons ("nc") comments once Jin Xicheu opens her Weibo.)("nc" is short for Pinyin "nao-can", which means moron.)

E3 Emotion: Happiness

真是速度啊, 收到快递了。 。 happy! (What a fast delivery, I got the parcel already, happy!)

E4 Emotion: None

我在bean bar (I am in the bean bar.)

	Hap	Sad	Anger	Fear	Surprise	Avg. F1	Wgt. F1
SVM	0.693	0.560	0.640	0.549	0.593	0.607	0.623
CNN	0.675	0.618	0.671	0.596	0.603	0.633	0.641
LSTM	0.717	0.642	0.704	0.606	<u>0.628</u>	0.659	0.671
BAN	<u>0.724</u>	<u>0.649</u>	<u>0.712</u>	0.627	<u>0.628</u>	<u>0.668</u>	<u>0.678</u>
HAN	0.729	0.658	0.729	<u>0.625</u>	0.641	0.676	0.688

Table 1: Performance evaluation (best result: **marked bold**; 2nd best: underlined.)

3.2 Analysis of WSCs Linked to Emotions

According to the work of Lee and Wang (2015), emotion words often serve as the cues of bilingual context. However, many WSCs segments which are not emotional words can also express emotion. To gain more insight on different types of WSC segments, we examine three types of WSCs: (1) English emotion words found from the NRC emotion lexicon (Mohammad and Turney, 2013), (2) Pinyin/acronym segments, and (3) others which include English words (no emotion), symbolic expressions, and emoji symbols, etc..

Figure 2 depicts their distribution in training dataset in different emotion labels. Note that English emotion words only serve about 1/3 of all WSCs for emotion. The largest group of WSC is in the category of 'Others'. This means non-emotion linked English words and symbols of other orthographic forms place an important part in emotion analysis of text with WSCs.

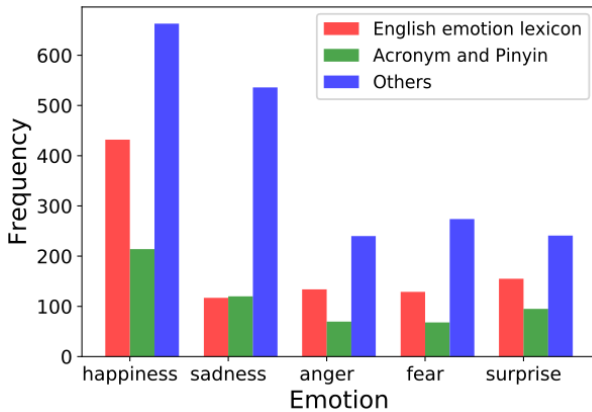


Figure 2: Distribution of different types of WSCs

3.3 Emotion Analysis

A group of experiments are implemented to examine the performance of different emotion classification methods evaluated using average F1-score

and weighted F1-score⁴. The baseline algorithms include BAN (Wang et al., 2016), the current state-of-the-art emotion classification algorithm. Others used in the comparison include SVM (Mullen and Collier, 2004), CNN and LSTM (Rosenthal et al., 2017). For all these baseline methods, WSC segments are included in the text. The difference compared to our HAN model is that we also separately extract WSCs segments and feed them into a separate CNN model.

Table 1 shows the performance evaluation result. From Table 1 we can draw a number of observations. Firstly, the performance of SVM is the worst since it lacks phrase level analytical capability because each word is considered independently in SVM. In other words, insufficient amount of information is learned in such a simple method. Secondly, the average weighted F1-score of CNN is lower than that of LSTM, indicating that the memory mechanism is effective in learning semantic information sequentially. The 3.0% gap of weighted F1-score shows that the order of words is valuable in emotion analysis. Thirdly, in addition to the improvement by BAN compared to CNN and LSTM, including WSCs in BAN can give performance gains 0.7% increase in weighted F1 measures. For the largest class Happy, the improvement is over 0.7% increase. Finally, compared to BAN, our proposed HAN which makes additional use of WSCs in a separate CNN gives another 1.0% performance gain.

3.4 Effects of Different Types of Text

In this set of experiments, we investigate the effect of three types of text, CN only (stands for Chinese), WSC segments, and CN+WSCs which are complete instances with both Chinese and WSC segments. We take LSTM and BAN to be compared to our proposed HAN.

Table 2 shows the performance evaluation of

⁴https://en.wikipedia.org/wiki/F1_score

	Hap	Sad	Anger	Fear	Surprise	Avg. F1	Wgt. F1
LSTM(WSCs)	0.631	0.546	0.682	0.589	0.529	0.595	0.598
LSTM(CN)	0.695	0.632	0.671	0.612	0.615	0.645	0.656
LSTM(CN+WSCs)	0.717	0.642	0.704	0.606	0.628	0.659	0.671
BAN(WSCs)	0.631	0.551	0.681	0.589	0.529	0.596	0.599
BAN(CN)	0.698	0.626	0.669	0.613	0.631	0.647	0.658
BAN(CN+WSCs)	<u>0.724</u>	<u>0.649</u>	<u>0.712</u>	0.627	0.628	<u>0.668</u>	<u>0.678</u>
HAN(CN+WSCs; WSCs)	0.729	0.658	0.729	<u>0.625</u>	0.641	0.676	0.688

Table 2: Performance using single writing system;best result is **marked bold**; second best is underlined.

	Hap	Sad	Anger	Fear	Surprise	Avg. F1	Wgt. F1
HAN(CN; WSCs)	0.629	0.613	0.682	0.588	0.531	0.606	0.613
HAN(CN+WSCs; CN)	<u>0.720</u>	<u>0.646</u>	<u>0.698</u>	<u>0.624</u>	<u>0.616</u>	<u>0.661</u>	<u>0.673</u>
HAN(CN+WSCs; WSCs)	0.729	0.658	0.729	0.625	0.641	0.676	0.688

Table 3: Performance by multiple writing systems;best result is **marked bold**; second best is underlined.

the three systems using different text types. Obviously, Chinese text carries more emotional information than WSCs text. The use of both Chinese and WSCs text gives better performance which shows that both WSCs and Chinese text contribute to the prediction task. However, by extracting WSCs into a separate CNN and re-merging into the attention layer, HAN still gives the best performance. This is because we are able to extract more emotion related information for the WSCs with a separate CNN. The attention model also gives dual consideration to WSCs. Nevertheless, note that BAN(CN+WSCs) can be competitive when both Chinese text and WSCs are learned.

To further analyze the effect of the two sub-models of LSTM and CNN in HAN, we examine the performance of HAN with different types of text to be taken by the two sub-models. In Table 3, the first type of text in the parenthesis denotes the text for the LSTM sub-model. The second text is used by the CNN sub-model. Note that in the first combination, only Chinese text without WSCs is used by the LSTM. Because this will break the syntax of the Chinese text, the result is the worst. In the second evaluation, the CNN sub-model is fed with Chinese text only. Still its performance is about 6% better than the first combination. Comparing HAN(CN+WSCs; CN) with the state-of-the-art method BAN(CN+WSCs), we can see that applying CNN to Chinese text will only introduce more noise and will not help to make performance improvement. Obviously, this gives more justification of using HAN(CN+WSCs; WSCs) as it gives the best performance gain.

4 Conclusion and Future Work

This paper presents a work in progress of an HAN model based on an LSTM model for emotion analysis in the context of WSCs in social media. We argue that WSCs text is potentially informative in emotion classification tasks such that they should be used as additional information contributing to deep learning based emotion classification models. Our proposed method offers a novel way to integrate multiple types of writing systems into an attention-based LSTM model. Along with WSCs text, the descriptive text of the major writing system is informatively valuable and semantic and affective information can be captured by LSTM effectively. Furthermore, WSCs texts indeed contain addition semantic and affective information which can be captured by a CNN model. After combining the representation of both the complete text and the WSCs, the two vectors are incorporated as the final feature.

Future works include two directions. One is to further evaluate the performance of HAN using larger corpora as currently only one public accessible corpus for writing systems for the Chinese communities can be studied. Another direction is to give more detailed study on how people use different types of WSCs to express emotions in censorship detection studies.

Acknowledgments

The work is partially funded by UGC under GRF projects(PolyU 152111/14E & U152006/16E).

References

- AR Balamurali, Aditya Joshi, and Pushpak Bhat-tacharyya. 2011. Harnessing wordnet senses for supervised sentiment classification. In *Proceedings of the conference on empirical methods in natural language processing*, pages 1081–1091. Association for Computational Linguistics.
- Luciano Barbosa and Junlan Feng. 2010. Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pages 36–44. Association for Computational Linguistics.
- Michael H Bond and Tat-Ming Lai. 1986. Embarrassment and code-switching into a second language. *Journal of Social Psychology*, 126(2):179–186.
- Michael Clyne. 2000. Constraints on code-switching: How universal are they. *The bilingualism reader*, pages 257–280.
- Jakob Cromdal. 2001. Overlap in bilingual play: Some implications of code-switching for overlap resolution. *Research on Language and Social Interaction*, 34(4):421–451.
- Cícero Nogueira Dos Santos and Maira Gatti. 2014. Deep convolutional neural networks for sentiment analysis of short texts. In *COLING*, pages 69–78.
- Paul Ekman. 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200.
- Roberto R Heredia and Jeanette Altarriba. 2001. Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science*, 10(5):164–168.
- Neha S Joshi and Suhasini A Itkat. 2014. A survey on feature level sentiment analysis. *International Journal of Computer Science and Information Technologies*, 5:5422–5425.
- James Max Kanter and Kalyan Veeramachaneni. 2015. Deep feature synthesis: Towards automating data science endeavors. In *Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on*, pages 1–10. IEEE.
- Jyh-An Lee and Ching-U Liu. 2012. Forbidden city enclosed by the great firewall: The law and power of internet filtering in china. *Minn. JL Sci. & Tech.*, 13:125.
- Sophia Lee and Zhongqing Wang. 2015. Emotion in code-switching texts: Corpus construction and analysis. In *Proceedings of the Eighth SIGHAN Workshop on Chinese Language Processing*, pages 91–99.
- Bing Liu and Lei Zhang. 2012. A survey of opinion mining and sentiment analysis. In *Mining text data*, pages 415–463. Springer.
- Yunfei Long, Lu Qin, Rong Xiang, Minglei Li, and Chu-Ren Huang. 2017. A cognition based attention model for sentiment analysis. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 473–482.
- Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a word-emotion association lexicon. 29(3):436–465.
- Tony Mullen and Nigel Collier. 2004. Sentiment analysis using support vector machines with diverse information sources. In *EMNLP*, volume 4, pages 412–418.
- Nigel Musk. 2012. Performing bilingualism in wales. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 22(4):651–669.
- Sara Rosenthal, Noura Farra, and Preslav Nakov. 2017. Semeval-2017 task 4: Sentiment analysis in twitter. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 502–518.
- Alessandra Vicentini. 2003. The economy principle in language. *Notes and Observations from early modern english grammars. Mots, Palabras, Words*, 3:37–57.
- Zhongqing Wang, Yue Zhang, Sophia Lee, Shoushan Li, and Guodong Zhou. 2016. A bilingual attention network for code-switched emotion prediction. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 1624–1634.
- Jennifer MY Wei. 2003. Codeswitching in campaigning discourse: The case of taiwanese president chen shui-bian. *Language and Linguistics*, 4(1):139–165.
- Theresa Wilson, Zornitsa Kozareva, Preslav Nakov, Sara Rosenthal, Veselin Stoyanov, and Alan Ritter. 2013. Sentiment analysis in twitter. In *Proceedings of the International Workshop on Semantic*.
- Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.