Estimating Emotional States of Internet Users in Real-World Situation

Wataru Sasaki[1,a]   Tadashi Okoshi[1]   Keisuke Uehara[1]   Yasuto Nakanishi[1]   Jin Nakazwa[1]

Abstract: With the advances in computer technology and the pervasiveness of the Internet, people's lives continue to change significantly. Most have multiple smart devices with a wide variety of applications. They communicate remotely via the Internet, and purchase diverse items, not limited to daily necessities, from online shops. In an era where people around the world are connected and share their circumstances via the Internet, detecting the emotional state of users has considerable potential to enable the provision of user-centric services tailored to specific user statuses and needs. However, considering the user's convenience, estimating Internet users' emotional states is difficult in real-world situations. To address the problem, due to almost all Internet users regularly using search engines, this study targets users' queries input to web search engines as an easily obtainable and non-invasive proxy feature to elucidate users' emotional states. Because web services typically store the historical logs of users' search queries on the server-side, they can be used immediately without awaiting the widespread adoption of new types of sensors. This study predicts users' online web behavior based on their emotional states estimated from large-scale real-world data. In addition, the weekly rhythm of the emotional state and the relationship between the number of new COVID-19 infected people and the ups and downs of the emotional state were clarified.

Keywords: affective computing, emotion estimation, human computer interaction, web search, mobile sensing

1. Introduction

Emotions have various effects on our daily lives. Darwin explained the role of emotion as an adaptive function from an evolutionary perspective [1]. Fear is a response to imminent danger and encourages us to escape, while anger raises the heart rate and prepares us to keep the adversary at bay. It was also observed that people who were deprived of emotional awareness due to brain damage were unable to make logical decisions [2]. This indicates that an emotional state affects decision-making. In addition, emotions contribute to other consciousness functions besides decision-making, especially attention and memory. It has been suggested that positive moods broaden and negative moods narrow the breadth of attention [3]. Emotions are used to discern what is relevant to us and what is not, and facilitate the memory easily to recall the emotional experiences that have affected us [4]. Emotions are an important element of social activities. Long-term relationships are cultivated by sharing emotional states and empathy [5].

Although the advent of digital information technology and the Internet have significantly changed our lives, emotions remain important. With the development of the Internet, we are now in an era where we can communicate without physically facing each other and, using smart devices, can connect with anyone in the world. In the near future, as the transition to virtual space progresses due to the development of virtual reality technology, a new world will be created that is different from real space and fur-ther expands the range of communication. Even in such a future, as long as we are human, emotions will continue to influence our decision-making, memory, and communication.

2. The Problem

The ability to estimate the current emotional states of users who are remotely connected by the Internet is useful for realizing services tailored to the receptivity of individual users and for analyzing people's current emotional states at a higher level, e.g., for an entire web service, or people in different cities, regions, or demographics. Emotion estimation also helps Internet users to communicate with each other. Since the advocacy of affective computing, research has been actively conducted on how systems estimate human emotional states [6]. However, estimating the emotional states of Internet users in real-world situations is difficult outside the laboratory. Typically, sensing and determining the emotional states of a person requires psychophysiological data such as heart rate (HR), heart rate variability (HRV) from electrocardiograms, and the results of electroencephalograms. However, obtaining such data from Internet users in real-world conditions is not feasible owing to the low general availability of such sensors in society, the additional burden on users of using such devices, and the lack of social acceptance for the collection of such data. In addition, when estimating emotion from other than physiological data, it is also difficult to collect modalities such as facial-expression and voice data from Internet users in real-world situations, because of privacy issues and inconvenience to the users.

1   Keio University
a)   wataruew@sfc.keio.ac.jp

3. Research Approach

To resolve the present research problem, this study proposes the idea of using the keywords input to the search engine as a proxy feature possibly to explain the emotional states of Internet users. Users' query inputs to the web search engine are easily obtainable and noninvasive proxy features to define their emotional states, focusing on the proposition that the majority of Internet users frequently use search engines in normal life. Because web services typically store a log of users' search queries on the server-side, this methodology has an advantage, since it can be deployed immediately without awaiting the widespread adoption of new types of sensors. This study focuses on multiple emotional states such as mind-wandering state and mood state.

First, this study focuses on the **wandering state** of Internet users, which appears in their web search behavior (not in the individual queries). This study estimates without using the mental state itself but the wandering aspect of emotional states. It is hypothesized that the wandering of interest manifested in a time series of users' search keyword histories, somewhat reflects the fluctuations in their emotional states. Using this idea and analyzing large-scale data containing shopping and search logs produced by approximately 200,000 users of a real web portal for over a year, we classified online store users into three types, and showed that it was possible to predict the timing of purchasing with high accuracy by analyzing the data. We showed that it is possible to predict purchase timing with high accuracy by classifying online shoppers into three types and analyzing the data.

Next, with a finer time resolution compared to the wandering state, this study focuses on the **mood state** of Internet users. This study proposes a novel two-step mood estimation using different types of models, namely the "Sensor Mood Model (SMM)" and the "Search-Query Mood Model (QMM)" to generate more mood status labels for building the model and boosting the overall classification performance as shown in Figure 1. The system was deployed in this product's server stack, and large-scale data analysis was conducted on more than 11,000,000 users. Every 3 hours this system calculates the mood scores of users registered on Yahoo! Japan (more than 11 million users) by processing their search queries with QMM. The average of all users' scores is calculated as the "Nation-wide Mood Score." This study evaluated the change in nation-wide mood scores over time and discussed how it matches events and occurrences in society. Interestingly, this research found that (i) the score obtained based on the proposed algorithm shows the weekly rhythm of people's mood (which drops every 1st working day of the week and increases again every weekend) and (ii) the longer-term trace of people's mood in the COVID-19 pandemic period in 2020 is inversely synchronized with the daily number of COVID-19 new cases as shown in Figure 2. This research also found that prefectures with more COVID-19 cases experienced a deeper drop in their mood, according to a per-prefecture analysis. Furthermore, (iii) this research also found that the nationwide mood score successfully captures mood changes influenced by significant news that simultaneously affects many users' mood states.
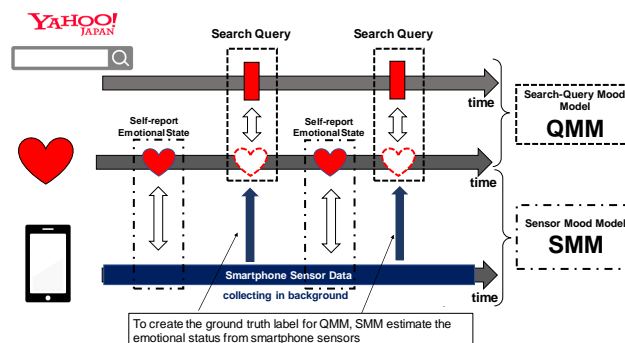

Fig. 1 The Approach of Building Estimation Model by Combinational Use of Search Queries and Smartphone Sensors
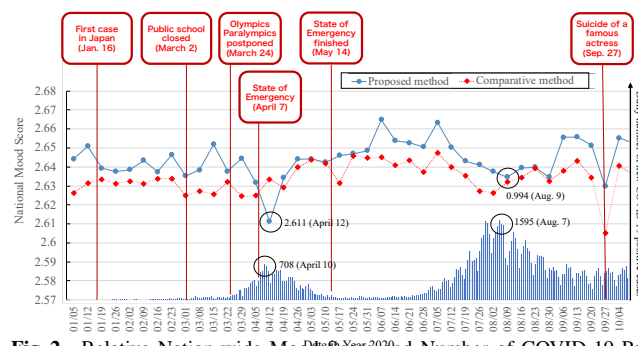

Fig. 2 Relative Nation-wide Mood Score and Number of COVID-19 Patients in Year 2020

4. Conclusion

We focus on the search keywords to estimate the emotional states of Internet users in the real world, taking into account the burden of users. As a result, we predicted the timing of online purchases and clarified the relationship between the national mood and the newly infected COVID-19 users. This study also summarized the research concepts based on the study of recognizing and estimating emotional states from various modalities through multiple studies, organized along three axes: invasiveness, time resolution, and reliability of data. Therefore, this study can be used as a guideline for research on the estimation of emotional states from various modalities, and has made a significant contribution to emotion estimation research.

References

[1] Charles Darwin. The expression of the emotions in man and animal. 1872.

[2] Antoine Bechara, Hanna Damasio, and Antonio R Damasio. Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, 10(3):295–307, 2000.

[3] Barbara L Fredrickson. The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions. *American psychologist*, 56(3):218, 2001.

[4] Kevin N Ochsner. Are affective events richly recollected or simply familiar? the experience and process of recognizing feelings past. *Journal of experimental psychology: general*, 129(2):242, 2000.

[5] Bernard Rimé. Interpersonal emotion regulation. *Handbook of emotion regulation*, 1:466–468, 2007.

[6] Rosalind W Picard and Roalind Picard. *Affective computing*, volume 252. MIT press Cambridge, 1997.