A Social Media Mining and Analysis Approach for Supporting Cyber Youth Work

W.M. Wang, Knowledge Management and Innovation Research Centre, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong
Benny C.F. Cheung, Knowledge Management and Innovation Research Centre, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong
Zeno C.S. Leung, Centre for Third Sector Studies, Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong
K.H. Chan, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong
Eric W.K. See-To, Knowledge Management and Innovation Research Centre, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

Cyber youth work is a pioneering and proactive approach used by non-government organizations to address the changing needs of the youth groups, particularly those at-risk and “hidden” young people. This paper describes the development process of a social media mining and analysis method, which is built to facilitate services for cyber youth work. The method incorporates an iterative method for collecting, selecting and extracting domain keywords in selected social media. A hybrid approach which combines heuristic rules and n-gram analysis for bilingual word segmentation has been developed. Then, social network analysis is used to analyze the extracted results. A pilot study has been done by using drug abuse as the study topic. It demonstrated high potential of the method to enhance cyber youth work. The weights deduced by the method have found to have a positive correlation with the benchmarked scores. It helps to understand the characteristic of the network, identify target clients, and provide data support for marking decisions.

KEYWORDS

Cyber Youth Work, Drug Abuse, Knowledge Mining, Social Media, Social Network Analysis

INTRODUCTION

Background

Outreaching youth work is an important type of social welfare service provided for the unattached, at-risk young people in Hong Kong (Central Guiding Committee on Outreaching Social Work, 1983). Conventionally, social workers approach young people in public and commercial venues such as soccer pitches and video game centres, befriend them and provide guidance and other services as needed. The rapid growth of the use of Internet in recent decades makes this traditional approach happened in the real world no longer sufficient. Young people tend to share their thoughts on social media such as forums and blogs (Mishna et al., 2015), many of them reveal their difficulties and needs too. Social workers therefore begin looking for those in need online, i.e. cyber outreaching.

Caritas Hong Kong is one of the pioneering organizations that has experimented cyber youth work since 2009. The “Life Architect” (hereafter as the “Project”) was an online project supported by the
Beat Drugs Fund of the Security Bureau, Hong Kong Special Administrative Region Government. The project aimed at serving the hidden youth in need via different internet tools. A team of four social workers would reach out the cyber world, identify and contact young people at-risk. The cyber outreaching process can generally be divided into 3 steps:

1. **Keyword Search**: Social workers make use of keyword search in popular online discussion forums to find posts that are related to their service scope.

2. **Private Contact with the Clients**: Social workers would then thoroughly read the searched results and identify potential targets. These young people may have expressed explicit or subtle messages calling for help. Common examples are “where can I find quit-drug service?” or “My friend was arrested!” Social workers send private messages to those users, showing their concern as well as giving advices, so as to build up trustful relationships with the latter.

3. **Face-to-Face Intervention**: By confirming that there are genuine needs of the clients, and that they are motivated to meet in the real world, social workers meet them face-to-face for providing further services.

In fact, cyber youth work in Hong Kong is still at the starting stage. Székely and Nagy (2011) found that even though there are a lot of benefits by the use of cyber youth work, but there are several limitations that are noticed.

1. **Random and Unsystematic Search**: The keywords used for online search are randomly proposed by the social workers based on their professional knowledge. The results would vary among different social workers. Moreover, keywords could be changed and new keywords (such as online slang) could be evolved from time to time (Leung et al., 2016). It is necessary to have a systematic method to retain the related keywords and to keep track on the trends of the keywords of the target scope.

2. **Massive Amount of Data**: The amount of data of social media is huge (Yaqoob et al., 2016). Social workers need to spend a lot of time and manpower to read and analyze all the collected data. There is a need to rank the search data based on their importance and relevance.

3. **Unstructured Format of Data and Lack of Quantitative Analysis**: The format of search results is unstructured. Due to the massive amount and unstructured format of the data, it is difficult for social workers to perform quantitative analysis. Most of the existing research of social media applications on social work remain “blackbox” (Chan & Holosko, 2016a). As a result, it is necessary to develop a method to convert the unstructured data into a more structured format, so that various quantitative analysis methods can be applied.

To overcome the above challenges, this paper develops a social media mining and analysis method to support cyber youth work. The method involves an iterative method for collecting, selecting and extracting keywords. It converts unstructured data into a structured format for easing data analysis. It ranks the results by assigning weights to the searched results. Moreover, it provides social network analysis for assisting social workers to understand the characteristic of the network, as well as helping them to identify the key persons from the collected data. Drug abuse is selected as the topic for pilot study. Preliminary results are demonstrated and discussed.

**LITERATURE REVIEW**

**Social Media Mining**

Social media mining is the process of representing, measuring, modeling, and extracting interested and meaningful contents and patterns from social media data. It is a rapidly growing interdisciplinary
Ontological Reflections on Peace and War
www.igi-global.com/chapter/ontological-reflections-peace-war/24191?camid=4v1a

Parameter Identification Using ANFIS for Magnetically Saturated Induction Motor
www.igi-global.com/article/parameter-identification-using-anfis-magnetically/66886?camid=4v1a