# Detecting Expressional Anomie in Social Media via Fine-grained Content Mining

Qingqing Zhou, Nanjing Normal University, Nanjing, Jiangsu Province, China

Ming Jing, Nanjing Normal University, Jiangsu Province, China

## ABSTRACT

Expression plays an important role in language inheritance, interpersonal communication, and social stability. With the rapid development of the Internet, people are becoming frequently interested in expressing themselves on social media. Meanwhile, massive anomic expressions are generated, which pollute network environments and even hinder social development. Hence, the purpose of this article is detecting anomic expressions in social media automatically, so as to reveal fine-grained status of online expressional anomie. Specifically, the authors used machine learning to detect anomic expressions and identify anomic types. Then, impacts of different factors (e.g. gender, region, time) on expressional anomie were analyzed. Finally, distributions and characteristics of expressional anomie about online contents were obtained. Empirical results indicate that the current situation of expressional anomie is severe, and scientific and effective treatments for anomic expression are necessary and urgently. Meanwhile, gender, region, and time should be taken into consideration in the formulation of treatments.

## KEYWORDS

Anomie Classification, Expressional Anomie, Machine Learning, Social Media, User Generated Content

## INTRODUCTION

With the rapid development of web 2.0, people are more accustomed to express their opinions and attitudes towards policies, events and products on the Internet, which generates massive online expressions (Sharma & Lbansal, 2015). These abundant online expressions have obvious advantages and disadvantages (Nah et al., 2010). It can help policy promulgators understand the scientificity of policies, event respondents develop solutions, and product designers optimize products. However, some online expressions are full of abuses, attacks or slanders, which may lead to the deterioration of events or hinder product development. Moreover, it may pollute Internet environment and destroy norms of language, thus affecting social stability, especially affect teenagers' Internet use and expression. Such expressions are regards as anomic expressions. Hence, in this paper, expressional anomie is identified as improper expressions with irregular words (offensive words, insults, slanders, etc.) based on Merton's theory (Merton, 1938). As most netizens are easily influenced by extreme or irrational emotions, and lack calm analysis or judgment, anomic expressions are often spread rapidly and widely, thus resulting in large-scale adverse effects, such as frequent cyber bullying, vicious political events. Therefore, research on expressional anomie are necessary and urgently.

Traditional survey methods for analyzing anomic expressions such as questionnaires, observation and interviews have been widely employed (Cooke & Jane, 2005). Although these methods are mature in theory and easy to operate, the cost of resources and time are relatively consuming. Using

artificial intelligence technology (e.g. machine learning) to conduct expressional anomie analysis based on large-scale corpus in social media may overcome shortcomings of traditional methods in data size and data acquisition cost (Zhou et al., 2016). Weibo (http://weibo.com/), one of the most popular social media, has attracted 430 million active users and generated mass online expressions (i.e. microblogs), thus providing massive authentic corpus for anomie research (Wang et al., 2017; Zhou & Zhang, 2018). Hence, this paper analyzed users' expressional anomie with microblogs in Sina Weibo. Specifically, we first distinguished anomic expression from microblog corpus, and then classified different types of expressional anomie automatically. Secondly, we mined the impacts of users' genders and regions on expressional anomie which were found to have obvious impacts. Finally, we analyzed anomie distributions of different time spans. Empirical results on microblogs indicate that the proportion of anomic expression is quite high in online contents, and current expressional anomie needs effective treatments urgently. Meanwhile, different genders, regions and time spans represent different anomie characteristics.

## RELATED WORK

In this paper, we use social media to analyze users' expressional anomie. There are two aspects of related works, namely objects of anomie research and methods of anomie research.

### Objects of Anomie Research

Durkheim held that anomie was destruction of social norms by individual behaviors, which led to various conflicts and confusion (Durkheim, 1933). Merton regarded anomie as the result of a conflict between socially agreed goals and socially agreed ways of achieving them (Merton, 1938). Behaviors that share common goals, but lack socially agreed approaches may be anomic. Both the two classic anomie theory held that anomie was caused by the weakening of social control, which leads to destruction of role of norms. Anomie research can be oriented to multiple objects, such as expressional anomie, institutional anomie, organizational anomie, and business anomie.

Specific to the field of expressional anomie, we can consider anomic expression as the destruction of expression norms. Current researchers analyzed expressional anomie in television programs, social media and daily communication. Sapolsky (2009) investigated anomic language on prime-time broadcast and cable programs, and found that 9 out of 10 programs contained at least one incident of anomie, and viewers were exposed to 12.58 cuss words per hour in 2005. Mccreary (2014) analyzed anomic language therein primetime television programming, and found that cable programming also received supervision from advertisers and viewers. Rost et al. (2016) introduced social norm theory to identify online expressional anomie in a social-political online setting. Results showed that in the context of online firestorms, non-anonymous individuals were more aggressive compared to anonymous individuals. Silva et al. (2016) collected expression from social media, so as to understand nature of expressional anomie in social media. Wong (2016) mined Cantonese-speakers' opinions towards tabooed words in daily communications, and found that Internet has changed speakers' attitudes towards tabooed words. Chetty and Alathur (2018) conducted a review on expressional anomie with different classes and terrorism with cyber use in the framework of online social networks. They held that with the help of combined effort from the government, the Internet Service Providers and online social networks, the proper policies can be framed to counter expressional anomie efficiently and effectively. The above analysis suggests that expressional anomie exists in many scenarios and occurs frequently in these scenarios. The expressional anomie involves almost all channels of human expression and affects human expression imperceptibly. It has a great impact on the development of human expression, especially the healthy development of adolescent expression. Therefore, research on expressional anomie is imminent and necessary.

Regarding institutional anomie, institutional anomie theory by Messner and Rosenfeld was mined and tested, aimed at explaining contemporary crime from the perspective of cultural goals

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/detecting-expressional-anomie-in-social-media-via-fine-grained-content-mining/245297

## Related Content

### Regression Test Selection for Database Applications
Ramzi A. Haraty, Nashat Mansourand Bassel A. Daou (2004). *Advanced Topics in Database Research, Volume 3 (pp. 141-165).*
www.irma-international.org/chapter/regression-test-selection-database-applications/4358

### Multilevel Databases
Alban Gabillon (2005). *Encyclopedia of Database Technologies and Applications (pp. 383-389).*
www.irma-international.org/chapter/multilevel-databases/11177

### QGLG Automatic Energy Gear-Shifting Mechanism with Flexible QoS Constraint in Cyber-Physical Systems: Designing, Analysis, and Evaluation
Xindong You, Yeli Li, Zhenyang Zhu, Lifeng Yuand Dawei Sun (2018). *Journal of Database Management (pp. 43-65).*
www.irma-international.org/article/qglg-automatic-energy-gear-shifting-mechanism-with-flexible-qos-constraint-in-cyber-physical-systems/201042

### Mobile Collaboration Support for Virtual Teams: The Case of Virtual Information Systems Development Teams
Keng Siauand Min Ling (2017). *Journal of Database Management (pp. 48-69).*
www.irma-international.org/article/mobile-collaboration-support-for-virtual-teams/189138

### Evolutionary Algorithms in Supervision of Error-Free Control
Bohumil Sulcand David Klimanek (2010). *Soft Computing Applications for Database Technologies: Techniques and Issues  (pp. 39-48).*
www.irma-international.org/chapter/evolutionary-algorithms-supervision-error-free/44381