

[DOI: 10.20472/IAC.2018.040.027](https://doi.org/10.20472/IAC.2018.040.027)

**CHIH LI HUNG**

Chung Yuan Christian University, Taiwan

## **USING PREFERENCE VECTOR MODELING TO POLARITY SHIFT FOR IMPROVEMENT OF OPINION MINING**

### **Abstract:**

This research proposes the preference vector modeling (PVM) to deal with polarity shifts for improvement of sentiment classification for word of mouth (WOM). WOM has become a main information resource of consumers while making business or buying strategies. A polarity shift happens when the sentiment polarity of a term is different from that of its associated WOM document, which is one of the most difficult issues in the field of opinion mining. Traditional opinion mining approaches depend on predefined sentiment polarities of terms to be accumulated as the WOM's sentiment polarity or to be trained based on machine learning techniques, but ignore the significance of polarity shift due to some specific usage of terms. There are two kinds of approaches used for detection of polarity shifts in the literature, which are rule-based approaches and machine learning approaches. However, it is hard for a rule-based approach to manually define a complete rule set. The machine learning approach, which is based on the vector space model (VSM), suffers from the curse of dimensionality. Therefore, this research proposes a novel approach to deal with polarity shifts for sentiment analysis because of the weakness of existing research in the literature. Firstly, this research proposes PVM based on an integration of opinionated documents and a star ranking system. The dimensionality of preference vectors equals the number of the star ranking system. Thus, the proposed PVM overcomes the curse of dimensionality as the number of dimensionality of the star ranking system is much fewer than that of the document vector based on VSM. Then, the automatic approach for polarity shift detection is proposed. The document preference vector is represented based on the average vector of term preference vectors. This way is able to deal with opinionated documents if they are extracted from the same scale of the star ranking systems and the same domain. Finally, the integrated approach of PVM and some classification techniques is used for improvement of sentiment classification for word of mouth.

### **Keywords:**

Polarity Shift; Preference Vector Modeling; Opinionated Text; Sentiment Analysis; Opinion Mining

**JEL Classification:** D80, L86