Title	
Author(s)	Le, Hoang, Minh
Citation	
Issue Date	2006-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/820
Rights	
Description	Supervisor:　　　　　　,　　　　　　,

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Abstract

The rapid increase in volume of scientific literature has led to researchers overload in their pursuit of knowledge. Staying up-to-date with recently published literature - and actually finding relevant sources - is becoming increasingly difficult, time-consuming, and impossible. Experience varies widely, but the time when every essential journal was held in all major academic libraries has passed.

Emerging trend detection is a new challenge and an attractive topic in text mining. With the continued increases in performance of computational technologies, more complex implementations of text-processing techniques are become possible, this has spurred research into the development of more sophisticated methodologies for developing emerging trend detection methods. However, there is no model that is particularly constructed for scientific corpora while existing models do not appear to be appropriate for this especially important kind of text databases. Previous work lack of an appropriate represent scheme for research topics and an effective method to identify emerging trends.

Building a model for emerging trend detection in scientific corpora is our major research objective. To this end, we have made the following contributions:

1. A model for emerging trend detection in scientific corpora, which presents various advantages in comparing to existing models.

2. A scheme for topic representation based on the rich information commonly provided in scientific papers, which can adapt to different kinds of scientific corpora and also can be efficiently modified.

3. Methods for topic identification, which are used to extract temporal features from documents. They appear to be more accurate and powerful than others in our experiments.

4. A topic verification method based on the interest and utility functions. This can be used for classification of emerging trends.

5. A prototype system to evaluate the model, that is used to evaluate if our model can achieve significant results in emerging trend detection.