Xiao Ding, Bing Qin, Ting Liu, 2015. BUEES: a bottom-up event extraction system. *Frontiers of Information Technology & Electronic Engineering*, **16**(7):541-552. [doi:10.1631/FITEE.1400405]

BUEES: a bottom-up event extraction system

Key words: Event extraction, Unsupervised learning, Bottom-up

Contact: Xiao Ding
E-mail: xding@ir.hit.edu.cn

ORCID: http://orcid.org/0000-0002-5838-0320

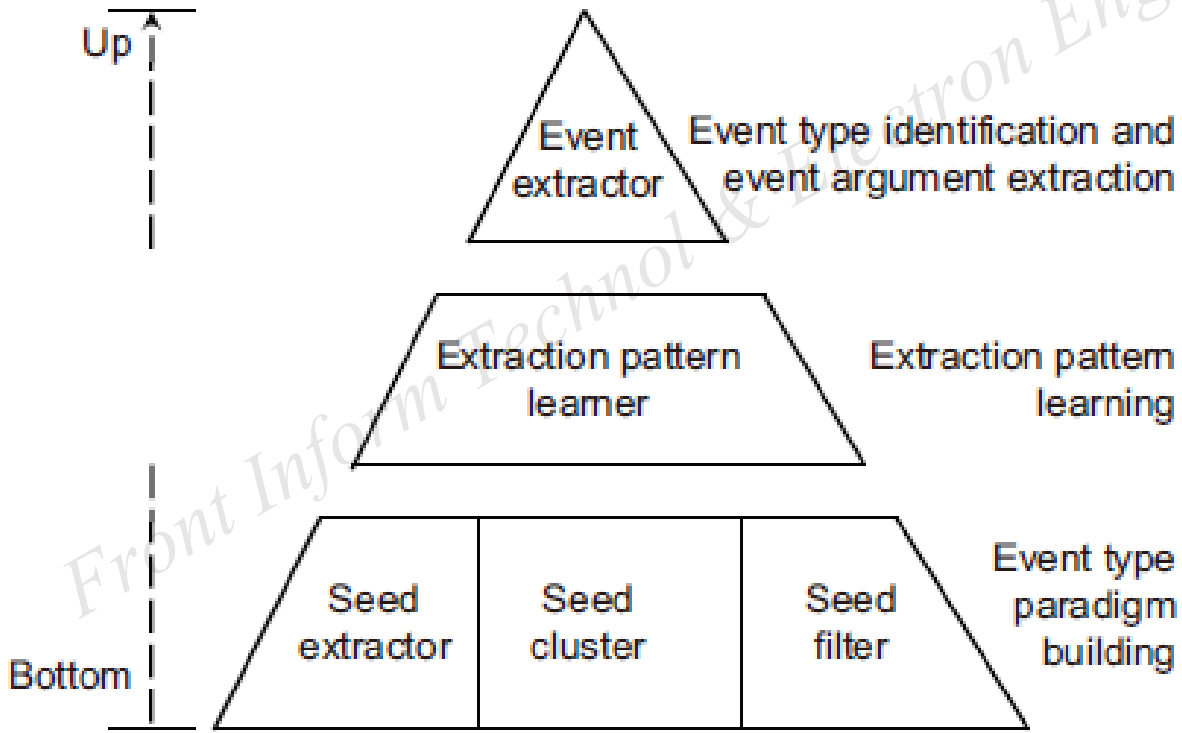
Introduction

- Traditional event extraction approaches require substantial human effort, and hence prevent event extraction systems from being domain adaptive and more widely applicable.

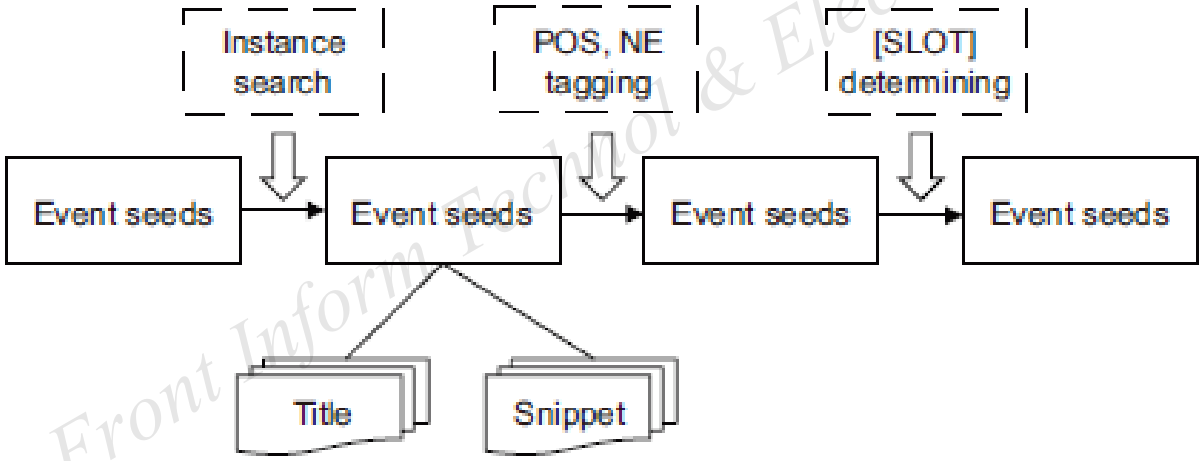
- To address the above challenges, in this study, we have designed and developed a bootstrapping-based bottom-up event extraction system, BUEES.

- The system automatically builds an event type paradigm in the input corpus, and then proceeds to extract a large number of instance patterns of these events. Subsequently, the system extracts event arguments according to these patterns.

The architecture of BUEES



The procedure of pattern learning



Measurement results

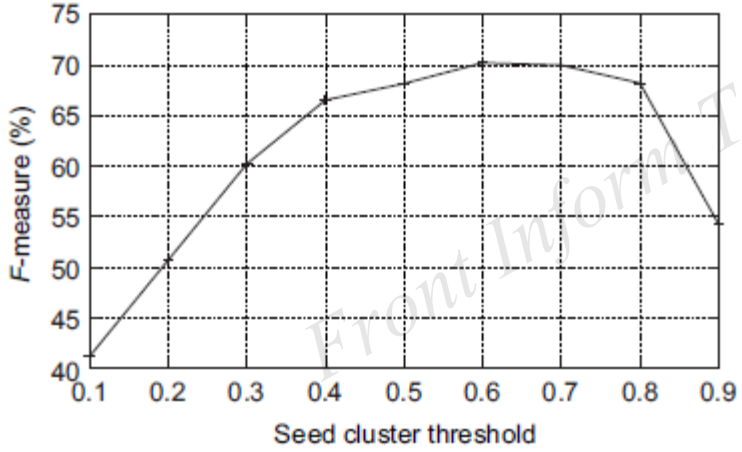


Fig. 4 Experimental results under different values of the event seed clustering threshold

Table 4 Experimental results of event type paradigm building

Corpus	F-measure		Purity	
	Baseline	Our method	Baseline	Our method
ACE	63.21%	69.57%	60.17%	70.24%
FN	71.52%	74.42%	74.81%	76.18%
MN	63.21%	75.08%	68.17%	80.28%

FN: Financial News; MN: Musical News

Table 5 Experimental error distributions of event type paradigm building

Error type	Proportion	Error type	Proportion
Trigger extraction	33.0%	Trigger filter	19.5%
Trigger ambiguity	28.3%	Others	19.2%

Overall experimental results

Table 6 Overall experimental results

System/ Human	Type identification			Argument recognition		
	P	R	F	P	R	F
Baseline	65.7%	50.9%	57.4%	53.1%	36.2%	43.1%
BUEES	72.7%	50.7%	59.7%	51.3%	45.9%	48.5%
HA1	75.2%	74.6%	74.9%	58.6%	60.9%	59.7%
HA2	82.7%	80.3%	81.5%	66.8%	69.6%	68.2%

HA1: human annotator 1; HA2: human annotator 2

Conclusions

- We have presented the BUEES system which discovers interesting events, learns extraction patterns, and extracts the event instances from the Web. BUEES relies neither on manually produced extraction patterns nor on a manually annotated training corpus.
- BUEES performs by clustering event seeds, generating seed instances, and bootstrapping its pattern. Based on a general sequential pattern, we propose a soft-pattern learning algorithm. Soft-pattern has a much greater generalizing ability and can reach a high performance for successful bootstrapping.