

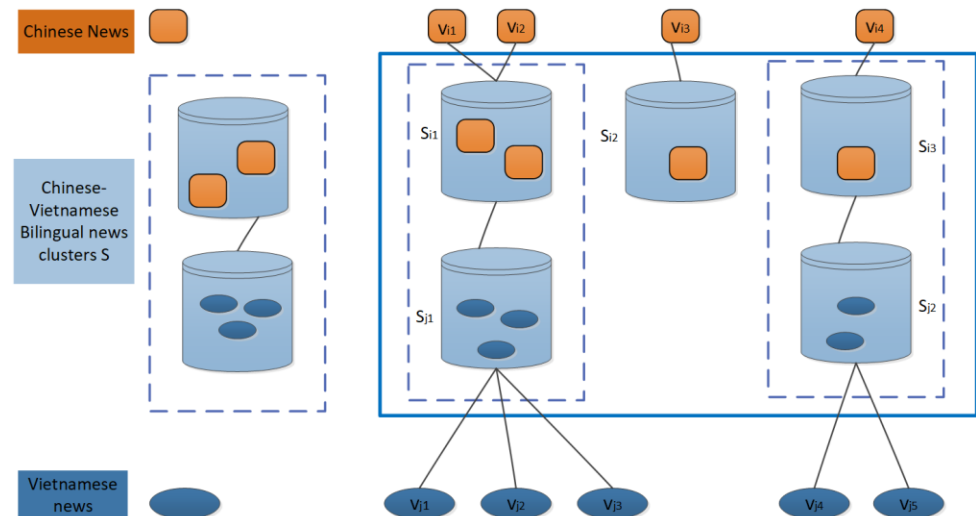
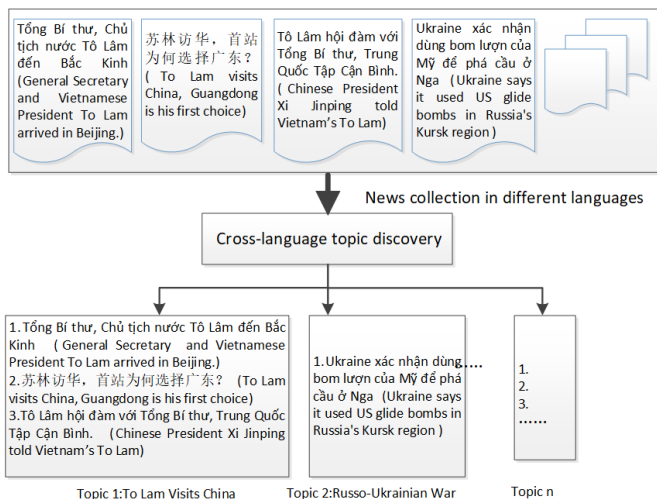
# Chinese and Vietnamese Bilingual News Topic Discovery via Association Graph Clustering

**Xiacong WANG, Peili TANG, Yuxin HUANG,  
Shengxiang GAO, Zhengtao YU**

Frontiers of Computer Science, DOI: [10.1007/s11704-025-41049-1](https://doi.org/10.1007/s11704-025-41049-1)

# Problems & Ideas

- Problems of conventional Chinese and Vietnamese Bilingual News Topic Discovery:
  - words may not be included in the bilingual resources, or there may be great variation.
  - only focused on the direct associations between events, ignoring the propagation of associations throughout event networks.
- Ideas: Use an association graph to represent the complex structure and relationships between events. Weaken the effect of language differences on clustering results.



# Main Contributions

- Contributions:
  - Using the features of element association, entity co-occurrence, and text content similarity between related news reports; Using that of similarity propagation between nodes in the graph to cluster Chinese-Vietnamese bilingual texts;
  - Using the propagation features between graph network nodes, weakening the clustering inaccuracy caused by differences between the two languages. Empirical results indicate that method substantially outperforms baseline models in low resource language.

---

Language type	Method of this paper			STSA			k-means			PromptTopic		
	P	R	F	P	R	F	P	R	F	P	R	F
Chinese	0.725	0.786	0.754	0.690	0.752	0.720	0.652	0.72	0.684	<b>0.73</b>	<b>0.80</b>	<b>0.78</b>
Vietnamese	<b>0.72</b>	<b>0.785</b>	<b>0.751</b>	0.683	0.798	0.735	0.642	0.70	0.671	0.70	0.773	0.724
Chinese- Vietnamese	<b>0.723</b>	<b>0.788</b>	<b>0.754</b>	0.681	0.750	0.710	0.643	0.70	0.670	0.718	0.768	0.721

---