

Date of acceptance

Grade

Instructor

Advance in news event aggregation methodology

Phuong Nguyen

Helsinki October 4, 2017

UNIVERSITY OF HELSINKI

Department of Computer Science

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Faculty of Science		Department of Computer Science	
Tekijä — Författare — Author			
Phuong Nguyen			
Työn nimi — Arbetets titel — Title			
Advance in news event aggregation methodology			
Oppiaine — Läroämne — Subject			
Computer Science			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
		October 4, 2017	6 pages + 0 appendices
Tiivistelmä — Referat — Abstract			
<p>News event grouping methods have been developed constantly over the year to aid news consumption, especially in the age of information. This report describes the progress and development in response to the modern context of digital information explosion.</p>			
Avainsanat — Nyckelord — Keywords			
news event, aggregation, named entity, keyword			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Contents

1	Introduction	1
2	Published work	2
2.1	Early work	2
2.2	Advance in aggregation method	3
2.3	Recent work	3
3	Conclusion	5
	References	6

1 Introduction

During the vigorous growth of information in the 21st century, we have witnessed many challenges that require our attention. Questions have risen concerning factual accuracy, relevance, privacy, etc. . . While the abundance of information is useful in many cases to human society, the excess is very often the seed that create struggle in finding beneficial information that satisfy our demands.

In the case of news reporting, the evolution of news delivering has gone far from the traditional physical newspaper to embrace the new mainstream called digital news. However, to properly dispatch news in a timely manner to the right audience, with first-class correctness and without redundancy, various strategies have been developed over the time.

To illustrate the challenges we are facing, imagine what happens everytime a new event occurs, regardless of where in the world. There is almost always a target group of people who need access to the story regarding the event. The story needs to be classified in accordance to target's interests. However, some events might generate a significant amount of interest, which in turn causing a huge corresponding amount of different stories that refer to the same event. This inevitably leads to the problem where audiences have to deal with a mass amount of similar information, while the truthfulness of the stories related to the event become questionable.

2 Published work

In order to understand how the aforementioned problems are being solved, several works in the field will be reviewed as a brief introduction to the development of news aggregation methodology.

The report will conclude with a summary of the general work that have been done so far, contemporary limitations and a brief discussion on the prospect in the coming years.

2.1 Early work

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris non felis quam. Curabitur molestie ultrices nibh vel faucibus. Praesent vitae dignissim leo. Cras finibus ornare odio, ac facilisis lacus egestas vitae. Aliquam erat volutpat. Ut in consequat odio. Aenean tincidunt, tortor vitae facilisis fermentum, eros nulla mattis nisi, id bibendum augue ante eget nunc. Nullam facilisis nunc eu placerat laoreet. Vestibulum vestibulum arcu lorem, in rutrum odio laoreet eget. Maecenas id massa sapien. Vivamus imperdiet nibh vel elit gravida aliquet. Donec nunc diam, ullamcorper quis nibh eget, vulputate vulputate lacus. In nunc urna, ultricies vel eleifend id, mattis at libero. Phasellus placerat nisl vitae consequat aliquet. In dignissim vestibulum libero, a dignissim odio consequat a. Morbi quis vehicula eros.[YEP⁺17, LMWY01]

Praesent eu suscipit ligula, non venenatis sapien. Cras tempor magna ut nisi pharetra semper. In ac arcu sapien. Pellentesque fermentum nec erat sed interdum. Praesent vitae tincidunt metus. Phasellus vel imperdiet justo. Morbi diam tortor, posuere id dolor quis, placerat lacinia lacus. Duis nec orci fringilla leo lacinia rutrum eget quis arcu. Aenean bibendum metus at mi aliquet, at luctus massa euismod.

Duis fermentum turpis eget pellentesque scelerisque. Maecenas elementum suscipit lectus, id ullamcorper lectus molestie vitae. Nulla et suscipit libero. Nulla ac mi eget libero commodo volutpat. Nunc malesuada quis eros a suscipit. Nullam faucibus volutpat pulvinar. Praesent in suscipit ante. Quisque eget aliquet lacus. Maecenas viverra libero nunc, sit amet ullamcorper metus egestas et. Sed at sapien sed odio congue porta. Nunc molestie lectus nec nulla tristique, eu suscipit risus varius. Aliquam mollis ipsum velit, at interdum elit porttitor sed. Sed eget tincidunt nisi, vel condimentum mauris. Nam at urna leo. Nulla sit amet ultricies sapien. Aliquam

neque arcu, auctor et commodo a, hendrerit pellentesque velit.

2.2 Advance in aggregation method

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris non felis quam. Curabitur molestie ultrices nibh vel faucibus. Praesent vitae dignissim leo. Cras finibus ornare odio, ac facilisis lacus egestas vitae. Aliquam erat volutpat. Ut in consequat odio. Aenean tincidunt, tortor vitae facilisis fermentum, eros nulla mattis nisi, id bibendum augue ante eget nunc. Nullam facilisis nunc eu placerat laoreet. Vestibulum vestibulum arcu lorem, in rutrum odio laoreet eget. Maecenas id massa sapien. Vivamus imperdiet nibh vel elit gravida aliquet. Donec nunc diam, ullamcorper quis nibh eget, vulputate vulputate lacus. In nunc urna, ultricies vel eleifend id, mattis at libero. Phasellus placerat nisl vitae consequat aliquet. In dignissim vestibulum libero, a dignissim odio consequat a. Morbi quis vehicula eros.

Praesent eu suscipit ligula, non venenatis sapien. Cras tempor magna ut nisi pharetra semper. In ac arcu sapien. Pellentesque fermentum nec erat sed interdum. Praesent vitae tincidunt metus. Phasellus vel imperdiet justo. Morbi diam tortor, posuere id dolor quis, placerat lacinia lacus. Duis nec orci fringilla leo lacinia rutrum eget quis arcu. Aenean bibendum metus at mi aliquet, at luctus massa euismod.

Duis fermentum turpis eget pellentesque scelerisque. Maecenas elementum suscipit lectus, id ullamcorper lectus molestie vitae. Nulla et suscipit libero. Nulla ac mi eget libero commodo volutpat. Nunc malesuada quis eros a suscipit. Nullam faucibus volutpat pulvinar. Praesent in suscipit ante. Quisque eget aliquet lacus. Maecenas viverra libero nunc, sit amet ullamcorper metus egestas et. Sed at sapien sed odio congue porta. Nunc molestie lectus nec nulla tristique, eu suscipit risus varius. Aliquam mollis ipsum velit, at interdum elit porttitor sed. Sed eget tincidunt nisi, vel condimentum mauris. Nam at urna leo. Nulla sit amet ultricies sapien. Aliquam neque arcu, auctor et commodo a, hendrerit pellentesque velit.

2.3 Recent work

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris non felis quam. Curabitur molestie ultrices nibh vel faucibus. Praesent vitae dignissim leo. Cras finibus ornare odio, ac facilisis lacus egestas vitae. Aliquam erat volutpat. Ut in consequat odio. Aenean tincidunt, tortor vitae facilisis fermentum, eros nulla

mattis nisi, id bibendum augue ante eget nunc. Nullam facilisis nunc eu placerat laoreet. Vestibulum vestibulum arcu lorem, in rutrum odio laoreet eget. Maecenas id massa sapien. Vivamus imperdiet nibh vel elit gravida aliquet. Donec nunc diam, ullamcorper quis nibh eget, vulputate vulputate lacus. In nunc urna, ultricies vel eleifend id, mattis at libero. Phasellus placerat nisl vitae consequat aliquet. In dignissim vestibulum libero, a dignissim odio consequat a. Morbi quis vehicula eros.

Praesent eu suscipit ligula, non venenatis sapien. Cras tempor magna ut nisi pharetra semper. In ac arcu sapien. Pellentesque fermentum nec erat sed interdum. Praesent vitae tincidunt metus. Phasellus vel imperdiet justo. Morbi diam tortor, posuere id dolor quis, placerat lacinia lacus. Duis nec orci fringilla leo lacinia rutrum eget quis arcu. Aenean bibendum metus at mi aliquet, at luctus massa euismod.

Duis fermentum turpis eget pellentesque scelerisque. Maecenas elementum suscipit lectus, id ullamcorper lectus molestie vitae. Nulla et suscipit libero. Nulla ac mi eget libero commodo volutpat. Nunc malesuada quis eros a suscipit. Nullam faucibus volutpat pulvinar. Praesent in suscipit ante. Quisque eget aliquet lacus. Maecenas viverra libero nunc, sit amet ullamcorper metus egestas et. Sed at sapien sed odio congue porta. Nunc molestie lectus nec nulla tristique, eu suscipit risus varius. Aliquam mollis ipsum velit, at interdum elit porttitor sed. Sed eget tincidunt nisi, vel condimentum mauris. Nam at urna leo. Nulla sit amet ultricies sapien. Aliquam neque arcu, auctor et commodo a, hendrerit pellentesque velit.

3 Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris non felis quam. Curabitur molestie ultrices nibh vel faucibus. Praesent vitae dignissim leo. Cras finibus ornare odio, ac facilisis lacus egetas vitae. Aliquam erat volutpat. Ut in consequat odio. Aenean tincidunt, tortor vitae facilisis fermentum, eros nulla mattis nisi, id bibendum augue ante eget nunc. Nullam facilisis nunc eu placerat laoreet. Vestibulum vestibulum arcu lorem, in rutrum odio laoreet eget. Maecenas id massa sapien. Vivamus imperdiet nibh vel elit gravida aliquet. Donec nunc diam, ullamcorper quis nibh eget, vulputate vulputate lacus. In nunc urna, ultricies vel eleifend id, mattis at libero. Phasellus placerat nisl vitae consequat aliquet. In dignissim vestibulum libero, a dignissim odio consequat a. Morbi quis vehicula eros.

Praesent eu suscipit ligula, non venenatis sapien. Cras tempor magna ut nisi pharetra semper. In ac arcu sapien. Pellentesque fermentum nec erat sed interdum. Praesent vitae tincidunt metus. Phasellus vel imperdiet justo. Morbi diam tortor, posuere id dolor quis, placerat lacinia lacus. Duis nec orci fringilla leo lacinia rutrum eget quis arcu. Aenean bibendum metus at mi aliquet, at luctus massa euismod.

Duis fermentum turpis eget pellentesque scelerisque. Maecenas elementum suscipit lectus, id ullamcorper lectus molestie vitae. Nulla et suscipit libero. Nulla ac mi eget libero commodo volutpat. Nunc malesuada quis eros a suscipit. Nullam faucibus volutpat pulvinar. Praesent in suscipit ante. Quisque eget aliquet lacus. Maecenas viverra libero nunc, sit amet ullamcorper metus egetas et. Sed at sapien sed odio congue porta. Nunc molestie lectus nec nulla tristique, eu suscipit risus varius. Aliquam mollis ipsum velit, at interdum elit porttitor sed. Sed eget tincidunt nisi, vel condimentum mauris. Nam at urna leo. Nulla sit amet ultricies sapien. Aliquam neque arcu, auctor et commodo a, hendrerit pellentesque velit.

References

- LMWY01 Lam, W., Meng, H. M. L., Wong, K. L. and Yen, J. C. H., Using contextual analysis for news event detection. *International Journal of Intelligent Systems*, 16,4(2001), pages 525–546. URL <http://dx.doi.org/10.1002/int.1022>.
- YEP⁺17 Yangarber, R., Escoter, L., Pivovarova, L., Du, M. and Katinskaia, A., Grouping business news stories based on salience of named entities. *EACL*, 2017.