

Date of acceptance

Grade

Instructor

My Topic

Behrouz Derakhshan

Helsinki June 3, 2014

UNIVERSITY OF HELSINKI

Department of Computer Science

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Faculty of Science		Department of Computer Science	
Tekijä — Författare — Author			
Behrouz Derakhshan			
Työn nimi — Arbetets titel — Title			
My Topic			
Oppiaine — Läroämne — Subject			
Computer Science			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
		June 3, 2014	4 pages + 0 appendices
Tiivistelmä — Referat — Abstract			
<p>Network influence maximization. Something about graphs, about network influence, social network, other related areas, big data, hadoop and other technologies, graph processing</p> <p>ACM Computing Classification System (CCS):</p> <p>A.1 [Introductory and Survey],</p> <p>I.7.m [Document and text processing]</p>			
Avainsanat — Nyckelord — Keywords			
layout, summary, list of references			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Contents

1	Introduction	1
2	Literature Review	1
2.1	Graph	1
2.2	Social Graphs	1
2.3	Influence Maximization	1
2.4	Big Data and hadoop	2
2.5	Use cases reasearch and industry	2
3	Big data and Network influence maximization	2
3.1	Tools	2
3.2	Data set	2
3.3	Algorithm	2
3.4	Other challenges	2
4	Conclusion	2
	References	2

1 Introduction

In [KKT03] they have studied models by which influence propagates through social networks. They have discussed two diffusion models

Linear Threshold Model (add reference) and Independent Cascade Models (add reference). Describe the two models briefly. The problem of influence maximization which is finding an initial set that will maximize the profit (target function) is NP-hard. Formal definition: it asks for a parameter k , to find a k -node set of maximum influence. They used approximation with a natural greedy hill-climbing strategy which will guarantee a 63 percent of the optimal solution performance (reference). For the approximation to work the target function should be submodular, they have proved that for both models the target function is indeed submodular. They have run experiments on co-authorships in physics section of arXiv (www.arxiv.org). The results show that their greedy algorithm performs better than other methods such as random selection, high-degree or central nodes discussed in book Social Network Analysis by S. Wasserman (find reference)

2 Literature Review

Let us start by looking at the sections expected to be in a scientific text. Keep in mind that the same expectations go for all kinds of technical writing.

2.1 Graph

2.2 Social Graphs

2.3 Influence Maximization

The nature of the matter at hand determines how the topic chapters are disposed.

In order to guide the reader, it is a good idea to start each main chapter with a short paragraph on what the main topic of the chapter is and how it progresses from one sub-chapter to the next.

Texts with only one sub-chapter, or with more than two chapter levels (main and sub-chapters) are a sign of a problem with the disposition of the text. There may be justifiable reasons to use three-level headings in some technical documents, but

they are an exception to the rule.

2.4 Big Data and hadoop

2.5 Use cases reasearch and industry

So-called mnemonic references are used for referring to sources; they are constructed as described in the section on the list of references. The page numbers should be added to the reference if it would be too laborious for the reader to find the reference in the source without them.

References are always placed inside sentences. This means that e.g. a separate reference at the end of a paragraph would be inappropriate.

The structure of the text must clearly show what the reference relates to. At the same time, it shows how long a piece of the text that the reference relates to.

3 Big data and Network influence maximization

TDescription of what I'm going to do, big data/hadoop/mapreduce implementations for network influence

3.1 Tools

3.2 Data set

3.3 Algorithm

3.4 Other challenges

4 Conclusion

References

- AQM97 Abiteboul, S., Quass, D., McHugh, J., Widom, J. and Wiener, J., The lorel query language for semistructured data. *In-*

- ternational Journal on Digital Libraries*, 1,1(1997), pages 68–88. [Myös <http://link.springer.de/link/service/journals/00799/bibs/7001001/70010068.htm>, 18.1.2000].
- BPS98 Bray, T., Paoli, J. and Sperberg-McQueen, C., Extensible Markup Language (XML) 1.0. W3C Recommendation 10-February-1998. <http://www.w3.org/TR/1998/REC-xml-19980210>. [18.1.2000]
- Die99 Dietinger, T. et al., Dynamic Background Libraries - New Developments in Distance Education Using HIKS (Hierarchical Interactive Knowledge System). *Journal of Universal Computer Science*, 5,1(1999). [Myös http://www.iicm.edu/jucs_5_1/dynamic_background_libraries_new, 18.1.2000].
- ErM96 Erkiö, H. and Mäkelä, M., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelytieteen laitos, Helsinki, 1996.
- Erk94 Erkiö, H., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelyopin laitos, Helsinki, 1994.
- FHS89 Fogelberg, P., Herranen, M. and Sinikara, K., *Tuumasta toimeen, tutkielman tekijän opas*. Yliopistopaino, Helsinki, 1989.
- Gan89 Gannon, D. et al., Programming environments for parallel algorithms. In *Parallel and Distributed Algorithms*, Cosnard, M. et al., editors, North-Holland, 1989, pages 101–108.
- Gri87 Grimm, S. S., *How to write computer documentation for users*. Van Nostrand Reinhold Co., New York, 1987.
- HaP82 Harkins, C. and Plung, D. L., editors, *A guide for writing better technical papers*. IEEE Press, 1982.
- Jul81 Julkaisusarjoja ja opinnäytteiden tiivistelmiä koskevat ohjeet ja suositukset. Helsingin yliopiston kirjastolaitoksen julkaisu A 3, Helsinki, 1981.
- Kil00 Kilpeläinen, P., WWW-lähteisiin viittaaminen tutkielmatekstissä, 2000. <http://www.cs.helsinki.fi/u/kilpelai/TiKi/wwwlahteet.html>. [19.1.2000]

- EMN01 Kempe, D., Kleinberg, J. and Tardos, E., Maximizing the spread of influence through a social network. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelyopin laitos, Helsinki, 2001.
- KKT03 Kempe, D., Kleinberg, J. and Tardos, E., Maximizing the spread of influence through a social network. *PKDD '03 Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA, 2003, pages 137–146.
- Smi78b Smith, A. J., Sequential program prefetching in memory hierarchies. *Computer*, 11,11(1978), pages 7–21.
- Smi78a Smith, A. J., Sequentiality and prefetching in database systems. *ACM Transactions on Database Systems*, 3,3(1978), pages 223–247.
- Ver92 Verkamo, A. I., Opinnäytetyön ulkoasun malli. Tieteellisen kirjoittamisen kurssiin liittyvä julkaisematon moniste, Tietojenkäsittelyopin laitos, Helsinki, 1992.