

First ACM Conference on Web Search and Data Mining (WSDM 2008)

# On Ranking Controversies in Wikipedia: Models and Evaluation

Ba-Quy Vuong, Ee-Peng Lim, Aixin Sun,  
Minh-Tam Le, Hady Wirawan Lauw,  
Kuiyu Chang  
School of Computer Engineering



Funded by



# Outline

- **Motivation**
- **Wikipedia approach**
- **Basic model**
- **ControversyRank models**
- **Experiments and results**
- **Age-aware models and new results**
- **Conclusions**

# Wikipedia



- **Collaboratively edited online encyclopedia**
- **> 7M articles in > 250 languages**
  - > 1.9M articles in English
  - > 4M registered users
  - 12th most popular websites (Alexa.com)

# Controversy topics in Wikipedia

- **Controversy [Random House Unabridged Dictionary]**
  - prolonged public dispute, debate, or contention; disputation concerning a matter of opinion.
- **Similar to real life communities, controversies can arise in Wikipedia community**
- **Neutral Point of View (NPOV)**
- **Edit warring**
  - “Individual editors or groups of editors repeatedly revert content edits to a page or subject area.”

# Example

Disagreement in choice and rank order of sites

Request for rename/removal of article

Microsoft Internet Explorer window showing the Wikipedia article "Holiest sites in Islam". The browser title is "Holiest sites in Islam - Wikipedia, the free encyclopedia - Microsoft Internet Explorer". The address bar shows "http://en.wikipedia.org/wiki/Holiest\_sites\_in\_Islam". The article title is "Holiest sites in Islam" with a subtitle "From Wikipedia, the free encyclopedia". A notice box states: "The examples and perspective in this article or section may not represent a worldwide view of the subject. Please improve this article or discuss the issue on the talk page." The main text reads: "There are many holy sites in the various Islamic traditions. The Kaaba is considered the holiest shrine, while the mosques of Nabwi (the Prophet) and al-Aqsa (farthest) are widely regarded as the second and third holiest respectively. Other shrines include the tombs of the twelve imams which are considered holy in Shiite Islam, and shrines revered by locals." A "Contents" table of contents is visible below the text, listing: 1 Masjid al-Haram, Mecca; 2 Masjid-an-Nabawi, Medina; 3 Masjid Al-Aqsa, Jerusalem; 4 Shi'a shrines; 4.1 Tomb of Imam Ali, Najaf; 4.2 Tomb of Imam Husayn, Karbala; 4.3 Tomb of Imams al-Hadi and al-Askari, Samarra; 4.4 Al-Kadshim Mosque, Baghdad. The left sidebar contains navigation and interaction links.



# Controversial topics in Wikipedia

- **Controversial topics are reflection and documentation of real world**
- **Articles with controversial topics are likely to cause much discussions and disputes**
  - Need to manually judge degree of controversy
  - Need much moderation efforts

# Challenges in finding controversial articles

- **Large number of articles**
  - New articles get created
  - Long article history
- **Diverse content**
  - Lack of domain experts
- **Evolving content**
  - Non-controversial articles may become controversial
  - Controversial articles may turn non-controversial after consensus is reached.



# Current solution

- **Community efforts to identify controversial articles and to resolve controversy**
- **Manually assign controversy related tags to articles**
- **Maintain**
  - [Wikipedia:List of controversial issues](#)
  - [Category:NPOV disputes](#)

Tag	Meaning
<code>{{disputed}}</code>	The factual accuracy of this article is disputed.
<code>{{totallydisputed}}</code>	The neutrality and factual accuracy of this article are disputed.
<code>{{controversial}}</code>	This is a controversial topic, which may be under dispute.
<code>{{disputed-section}}</code>	Some section(s) has content whose accuracy or factual nature is in dispute.
<code>{{totallydisputed-section}}</code>	The neutrality and factual accuracy of some section are disputed.
<code>{{pov}}</code>	The neutrality of the article is disputed.



# Overview of our approach

- **Fully automated**
  - No training
- **Data driven**
  - Measure disputes among Wikipedia users
  - Determine controversial articles from disputes
- **Dispute is measured by word deletions**

# Article Tag Count (ATC)

- An estimated degree of controversy for an article

$$ATC_k = \sum_{i=1}^{N_k} c_{ki}$$

where  $N_k$  : # of revisions of article  $k$

$C_{ki}$  : # of controversy tags in revision  $i$  of article  $k$

- Very few (labeled) articles have >0 ATCs
  - In our dataset, 71 out of ~20K articles have ATC > 0

# Dataset

- 19,456 articles from the Religious Objects category and sub-categories on 12 June 2007.
- 174,338 contributors

**Table 4: Dataset Statistics**

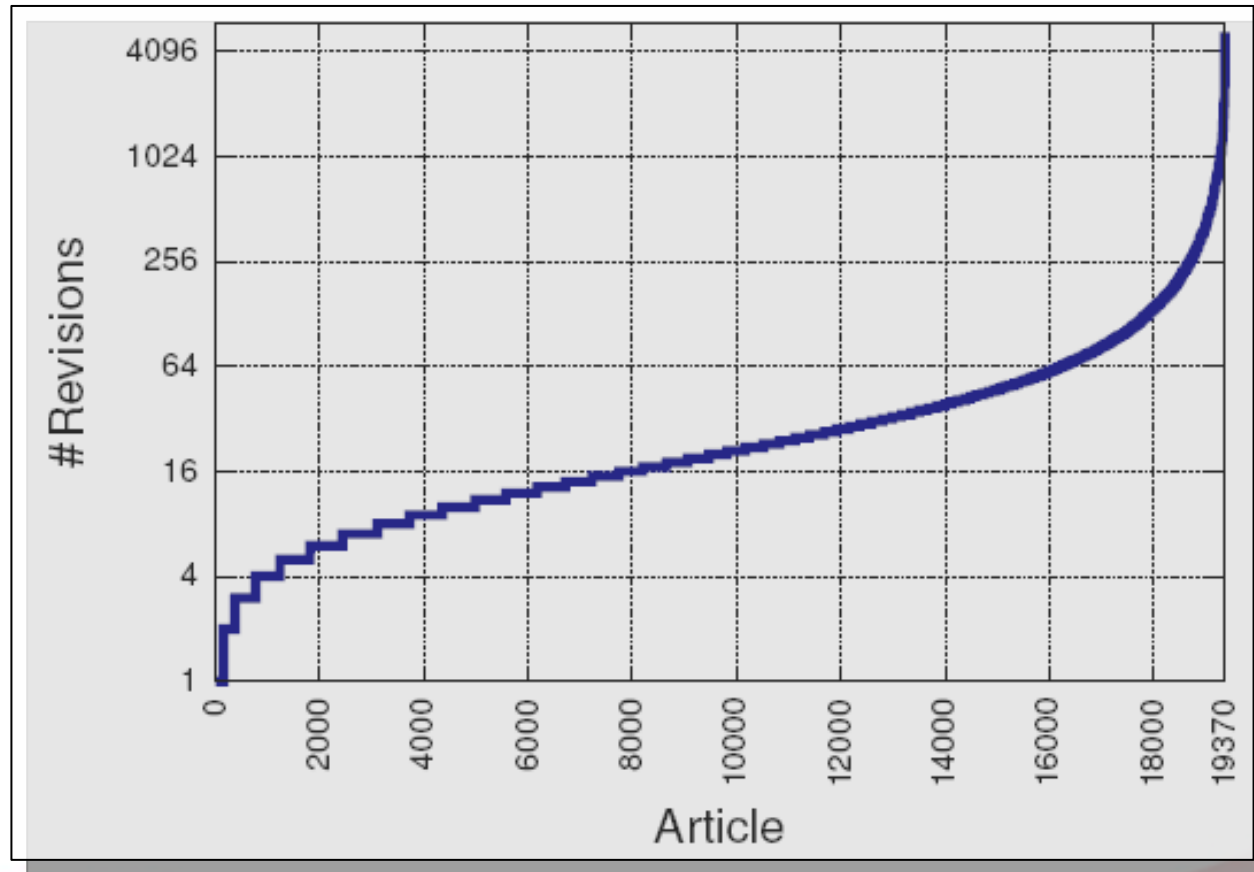
Count		<i>Min</i>	<i>Max</i>	Avg	Std Dev
# contributors	per article	1	3190	39.69	99.47
# articles	per contributor	1	937	2.71	23.71
Contributions	per article	3	360,929	1324.21	9118.68
	per contributor	0	347,727	140.22	3092.58
Disputes	per article	0	359,165	902.82	8863.65
	per contributor	0	348,800	95.60	3888.33

**Table 5: Distribution of *ATC* Values**

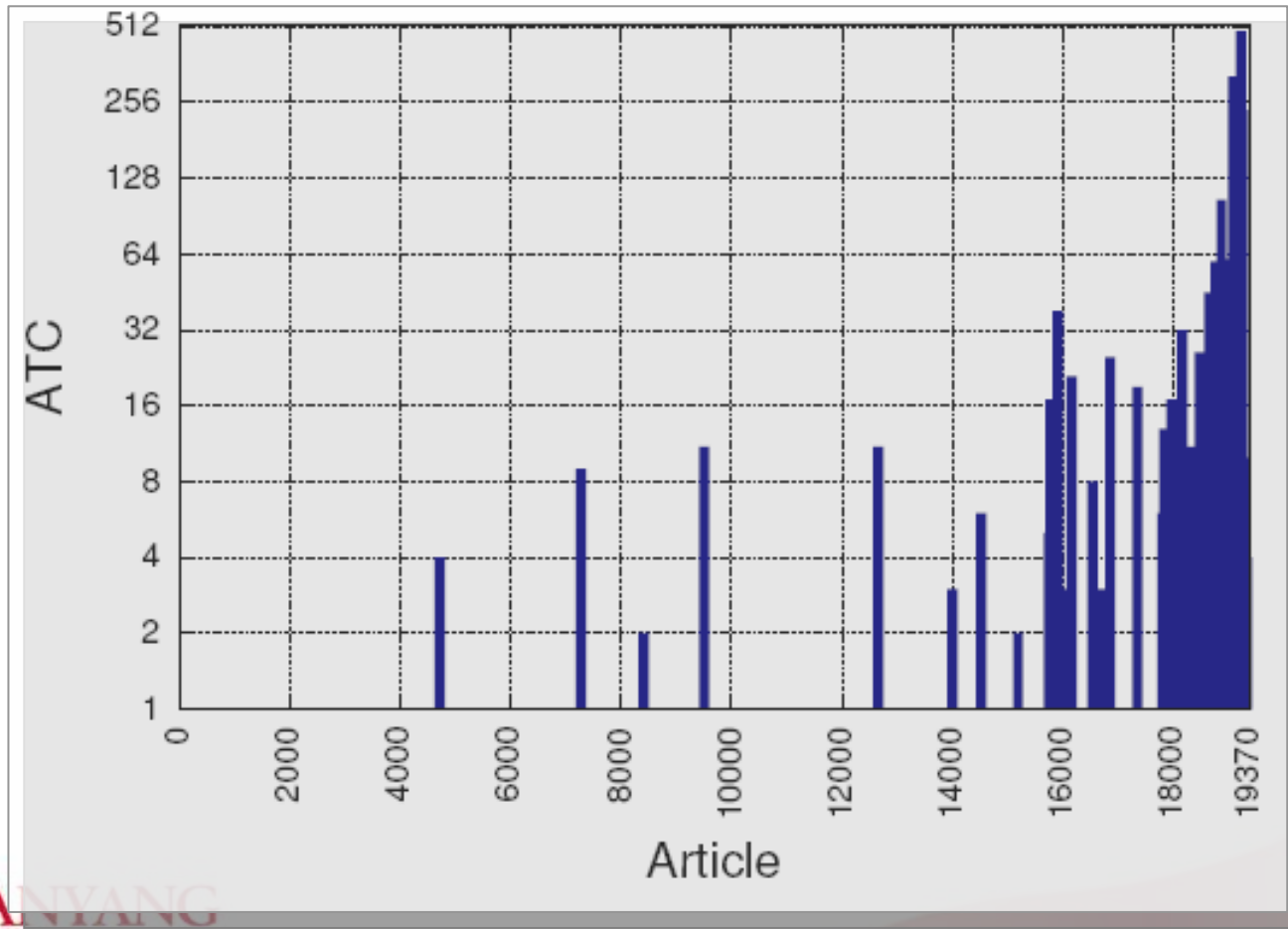
<i>ATC</i>	>500	101-500	21-100	5-20	1-4	0	Total
# articles	0	6	19	21	25	19,385	19,456
%	0.0	0.031	0.098	0.108	0.128	99.635	100



# Articles ordered by revision count



# Controversy labeled articles ordered by revision count



# Top 20 Articles (Revision Count)

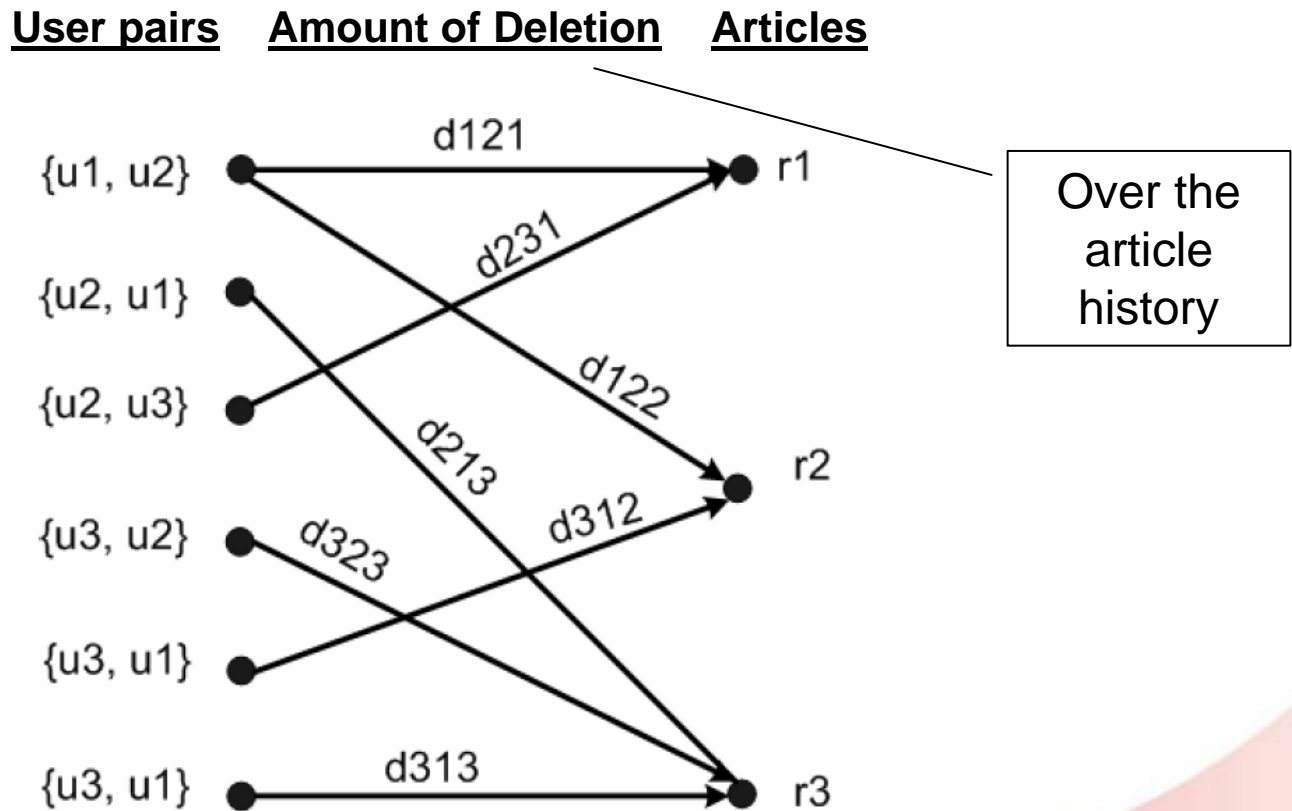
No	Article Name	#Rev	ATC	ATC Rank	No	Article Name	#Rev	ATC	ATC Rank
1	Podcast	5344	0	>71	11	Iain Lee	2384	0	>71
2	Emma Watson	4115	0	>71	12	Globe Theatre	2330	0	>71
3	Stephen Hawking	3682	0	>71	13	Grigori Rasputin	2211	0	>71
4	John McCain	3072	0	>71	14	John Howard	2200	0	>71
5	George Orwell	2948	0	>71	15	Keira Knightley	2177	0	>71
6	David Cameron	2692	0	>71	16	Salem witch trials	2176	0	>71
7	Dr.Seuss	2625	0	>71	17	Easter	2146	0	>71
8	James Madison	2623	0	>71	17	Constantine I	2111	0	>71
9	John Locke	2477	0	>71	19	Winston Churchill	2100	0	>71
10	Oscar Wilde	2432	0	>71	20	Rupert Murdoch	2003	0	>71



# Top 20 Articles (Contributor Count)

No	Article Name	#Cont	ATC	ATC Rank	No	Article Name	#Cont	ATC	ATC Rank
1	Podcast	2146	0	>71	11	Easter	1041	0	>71
2	Stephen Hawking	1933	0	>71	12	Keira Knightley	1030	0	>71
3	Emma Watson	1619	0	>71	13	Rupert Murdoch	1028	0	>71
4	John McCain	1459	0	>71	14	Globe Theatre	974	0	>71
5	George Orwell	1342	0	>71	15	Winston Churchill	948	0	>71
6	Dr. Seuss	1186	0	>71	16	David Cameron	912	0	>71
7	John Locke	1174	0	>71	17	Salem witch trials	899	0	>71
8	Grigori Rasputin	1110	0	>71	18	Jefferson Davis	879	0	>71
9	Oscar Wilde	1093	0	>71	19	Constantine I	871	0	>71
10	James Madison	1085	0	>71	20	Robert Frost	848	0	>71

# Disputes in a bipartite graph

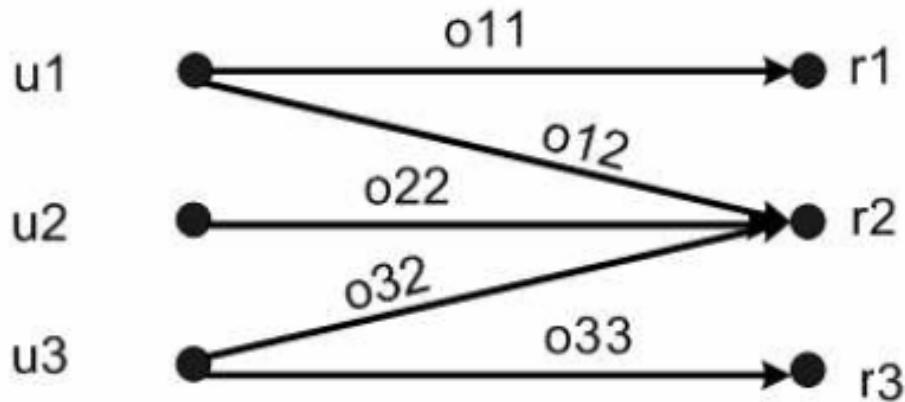


$d_{ijk}$  = # of words contributed by user  $j$  to article  $k$  but deleted by user  $i$



# Contribution by Users

Users      Amount of Added Content      Articles



Over the  
article  
history

$o_{ik}$  = # of words contributed by user  $i$  to article  $k$



# Basic model

- Controversial score of article

$$C_k^r = \frac{\sum_{i,j} d_{ijk}}{\sum_j o_{jk}}$$

- $d_{ijk}$  = # of words contributed by user  $j$  to article  $k$  but deleted by user  $i$
- $o_{ik}$  = # of words contributed by user  $i$  to article  $k$

But, are all users the same?

# ControversyRank (CR) model

- **Main idea (mutual dependency principle):**
  - Article Controversy  
An article is controversial if it contains more disputes among non-controversial contributors.
  - Contributor Controversy  
A contributor is controversial if s/he is engaged in more disputes in non-controversial articles.

# ControversyRank (CR) model

- **Controversial score of article**

$$C_k^r = \frac{\sum_{i,j} \text{agg}[(1 - C_i^u), (1 - C_j^u)] \times d_{ijk}}{\sum_j o_{jk}}$$

- **Controversial score of user**

$$C_i^u = \frac{\sum_{j,k} (1 - C_k^r) \times (d_{ijk} + d_{jik})}{\sum_{j,k} o_{jk} \times I(i, j, k) + \sum_k o_{ik}}$$

# Choice of Agg()

- **CR Average Model**

$$\text{agg}[(1 - C_i^u), (1 - C_j^u)] = \frac{1 - C_i^u + 1 - C_j^u}{2}$$

- **CR Product Model**

$$\text{agg}[(1 - C_i^u), (1 - C_j^u)] = (1 - C_i^u) \times (1 - C_j^u)$$

# Evaluation metrics

$$\text{Precision@}k = \frac{\text{\#relevant articles in top } k \text{ articles}}{k}$$

$$\text{Recall@}k = \frac{\text{\#relevant articles in top } k \text{ articles}}{\text{\#relevant articles in the dataset}}$$

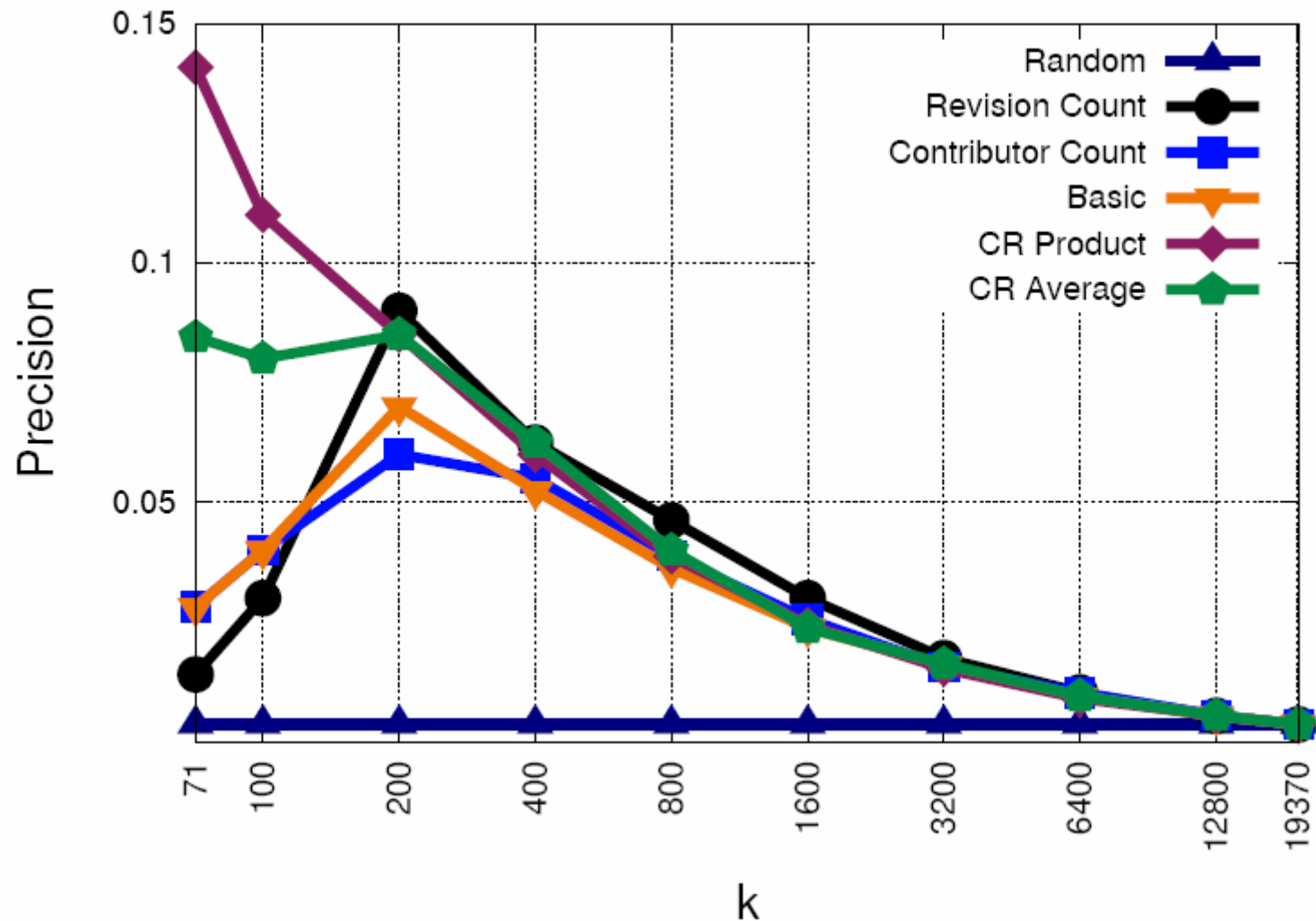
$$F1@k = \frac{2 \times \text{Precision@}k \times \text{Recall@}k}{\text{Precision@}k + \text{Recall@}k}$$

$$NDCG = \frac{1}{Z} \sum_{p=1}^k \frac{2^{s(p)} - 1}{\log(1 + p)}$$

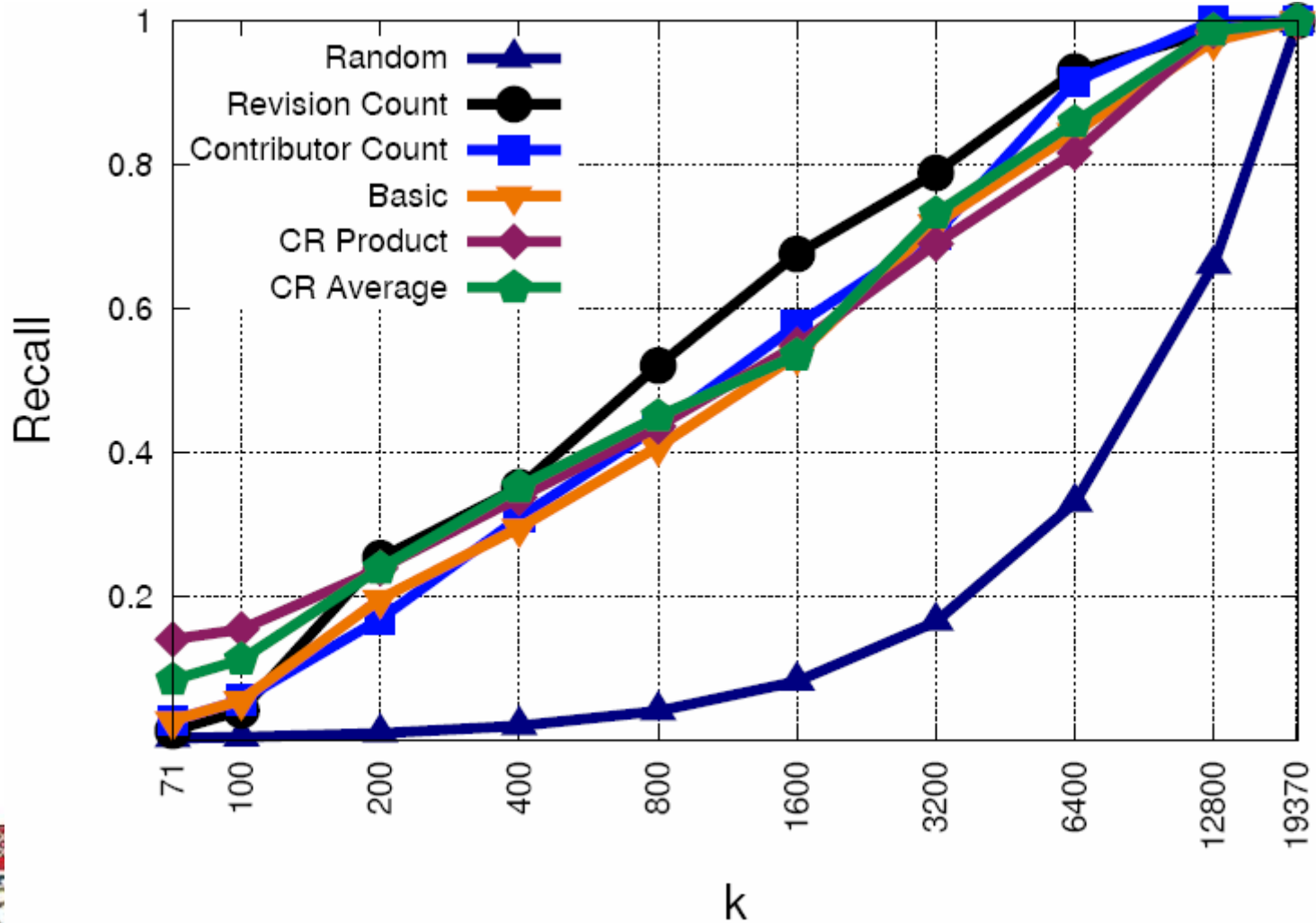
$$s(p) = \log(ATC_p + 1)$$



# Result : Precision

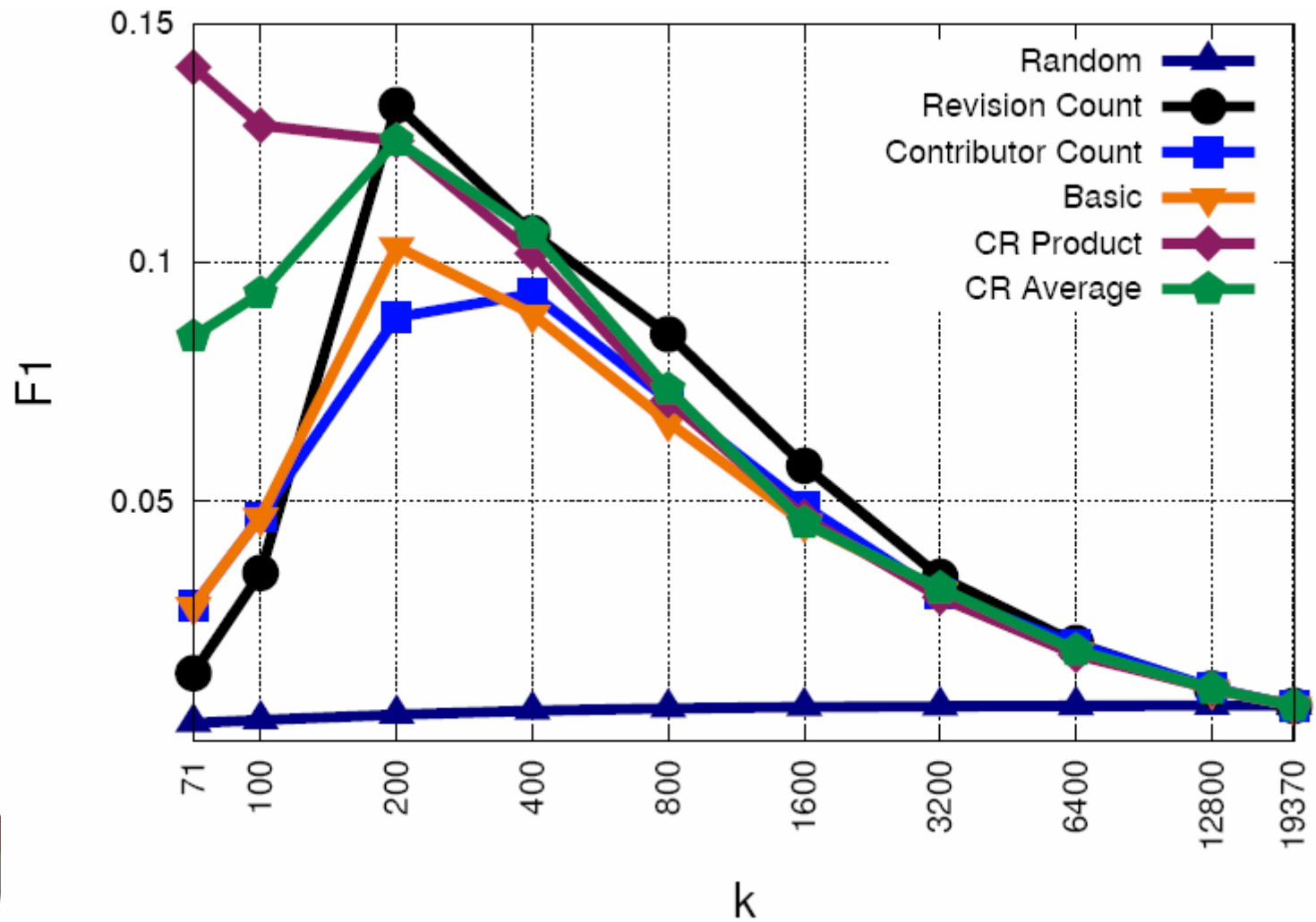


# Result : Recall

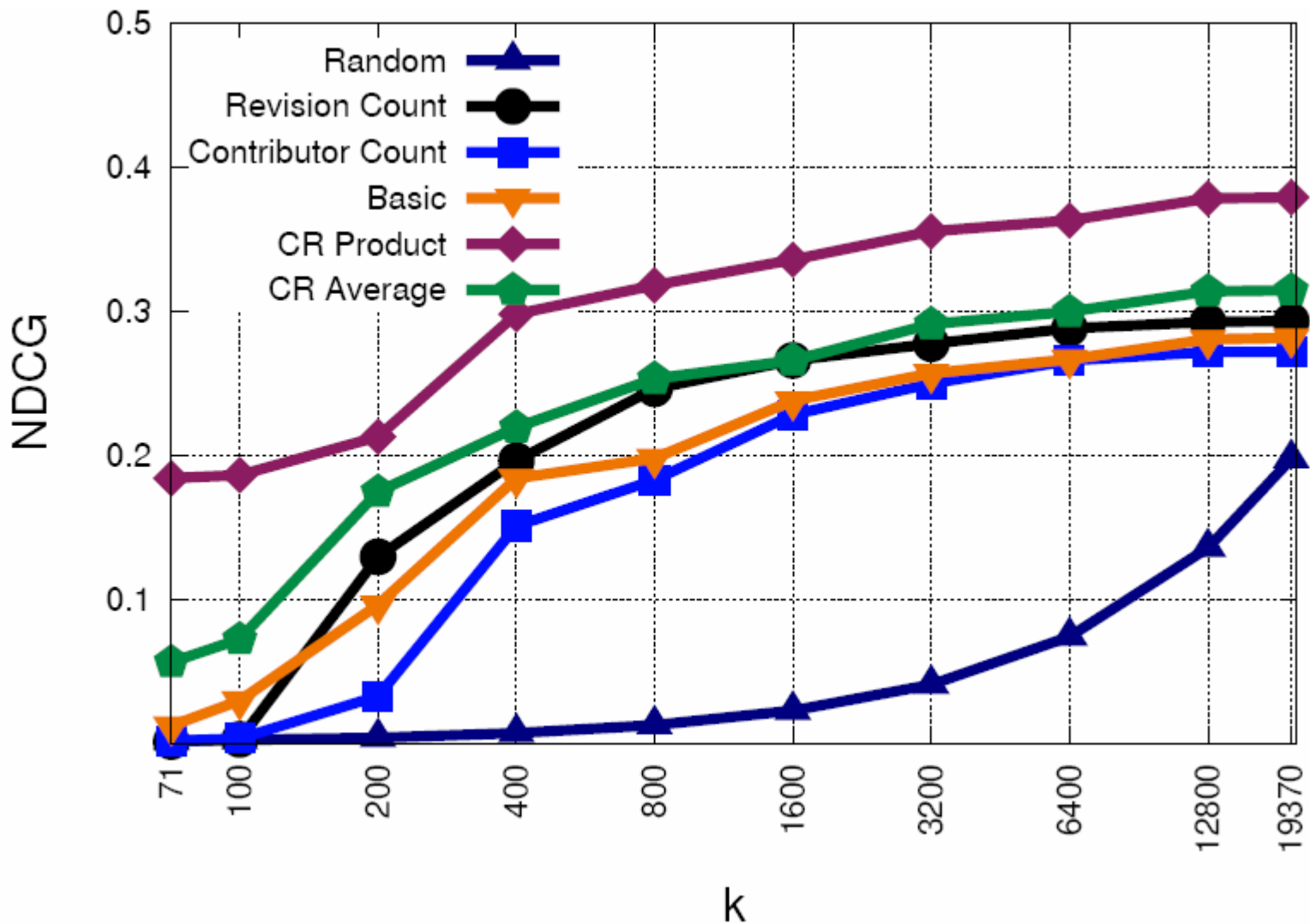




# Result : F1



# Result : NDCG



# Result: Top 20 Articles

Table 6: Basic Model Top 20 Articles

Rank	Article Name	ATC	ATC Rank	Rank	Article Name	ATC	ATC Rank
1	Dominus Illuminatio Mea	0	>71	11	John Howard	0	>71
2	North Marston	0	>71	12	Emma Watson	0	>71
3	Will Young	0	>71	13	Rozen Maiden	0	>71
4	Abingdon School	0	>71	14	Saint Sophia...	0	>71
5	Kamakhya	0	>71	15	Stephen Hawking	0	>71
6	John Dalton	0	>71	16	Queen Elizabeth II Bridge	0	>71
7	Christ Church...	0	>71	17	Aaron	0	>71
8	Podcast	0	>71	18	Kevin Rudd	0	>71
9	Jyotiba	0	>71	19	George Orwell	0	>71
10	Iain Lee	0	>71	20	Our Lady of the...	0	>71

Table 7: CR Average Top 20 Articles

Rank	Article Name	ATC	ATC Rank	Rank	Article Name	ATC	ATC Rank
1	John Howard	0	>71	11	Globe Theatre	0	>71
2	Podcast	0	>71	12	Aaron	0	>71
3	Iain Lee	0	>71	13	Anton Chekhov	0	>71
4	Zt <sup>nl</sup>	0	>71	14	Pro-Test	58	10
5	Stephen Hawking	0	>71	15	Myron Evans	60	9
6	George Orwell	0	>71	16	Our Lady of the...	0	>71
7	Emma Watson	0	>71	17	Robert Hooke	0	>71
8	11-Sep	0	>71	18	St. Dr. Seuss	0	>71
9	Jyotiba	0	>71	19	John Dalton	0	>71
10	Andrew Adonis...	0	>71	20	John Locke	0	>71



# Result: Top 20 Articles

Table 8: CR Product Top 20 Articles

Rank	Article Name	<i>ATC</i>	<i>ATC</i> Rank	Rank	Article Name	<i>ATC</i>	<i>ATC</i> Rank
1	Zt'l	0	>71	11	Bishop of Salisbury	0	>71
2	Myron Evans	60	9	12	First Baptist Church...	3	52
3	Solomon's Temple	0	>71	13	Holiest sites in Islam	490	1
4	University College Record	0	>71	14	San Lorenzo...	0	>71
5	Nassau Presbyterian Church	0	>71	15	Guy Davenport	0	>71
6	Shrine of St. Margaret...	0	>71	16	Bonn Minster	0	>71
7	Bishop of Worcester	0	>71	17	Temple Rodef Shalom	0	>71
8	Yell Group	0	>71	18	Ta Som	0	>71
9	St Volodymyr's Cathedral	0	>71	19	Romanian Orthodox...	0	>71
10	Ashtalakshmi Kovil	0	>71	20	Italo-Greek Orthodox...	0	>71

# Mis-“controversial” articles

- During the construction phase, edits are expected
- Disputes during this phase are rank and file
- **Age** should be considered

$$Age(r_k) = \frac{1}{1 + e^{-\frac{1}{2}(rev_k^r - 20)}}$$

# Controversial score of article

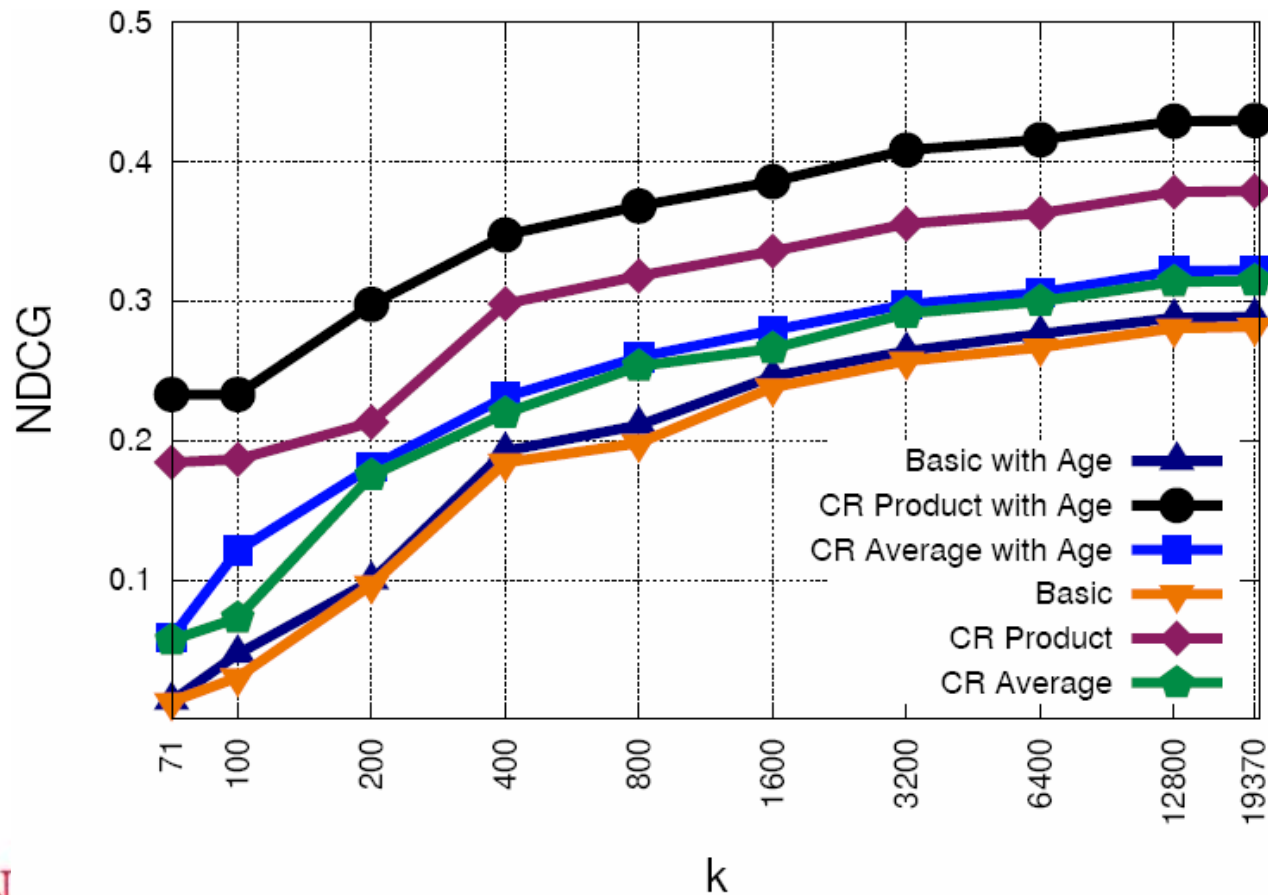
- **Age-aware Basic Model**

$$C_k^r = \frac{\sum_{i,j} d_{ijk}}{\sum_j o_{jk}} \times \text{Age}(r_k)$$

- **Age-aware CR Model**

$$C_k^r = \frac{\sum_{i,j} \text{agg}[(1 - C_i^u), (1 - C_j^u)] \times d_{ijk}}{\sum_j o_{jk}} \times \text{Age}(r_k)$$

# Results: NDCG



# Result: Top 20 articles

Table 11: Age-aware CR Product Top 20 Articles

Rank	Article Name	ATC	ATC Rank	Rank	Article Name	ATC	ATC Rank
1	Myron Evans	60	9	11	Temple Rodef Shalom	0	>71
2	Solomon's Temple	0	>71	12	Romanian Orthodox...	0	>71
3	Bishop of Worcester	0	>71	13	Ashtalakshmi Kovil	0	>71
4	Yell Group	0	>71	14	Italo-Greek Orthodox...	0	>71
5	St Volodymyr's Cathedral	0	>71	15	City Harvest Church	27	18
6	Bishop of Salisbury	0	>71	16	Macedonian Orthodox...	61	8
7	First Baptist Church...	0	>71	17	Oxford University Conservative...	1	64
8	Holiest sites in Islam	490	1	18	Church of Kish	21	24
9	Guy Davenport	0	>71	19	Iain Lee	0	>71
10	Bonn Minster	0	>71	20	Waldegrave School...	0	>71





# Conclusions

- **ControversyRank (CR) model for measuring degree of controversy in Wikipedia articles**
- **Degrees of controversy among Wikipedia users are considered**
- **CR models are more accurate than Basic and others**
- **CR Product model is the best**

# Future Work

- **Convergence of CR Models**
- **Visualization of controversy articles and disputes in their content**
- **Evolution of controversy**

# Thank you



**NANYANG**  
TECHNOLOGICAL  
UNIVERSITY