

Quantitative indicators for role behaviour in online discussion groups

Eleni Hichinson and Christian von Ferber

November 30, 2012

Source

Newsgroup discussions are organised in *threads* .

Re: poorly formed underbrace (s)

- 1 [Ilya Zakharevich](#) Dec 1 2011
- 2 [Donald Arseneau](#) Dec 2 2011
- 3 [Donald Arseneau](#) Dec 2 2011
- 4 [Ilya Zakharevich](#) Dec 2 2011
- 5 [Donald Arseneau](#) Dec 2 2011
- 6 [Ilya Zakharevich](#) Dec 4 2011
- 7 [Donald Arseneau](#) Dec 6 2011
- 8 [Ilya Zakharevich](#) Dec 7 2011

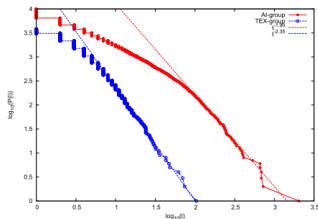
We analyse and compare the
groups

`comp.text.tex`

and

`comp.ai.philosophy`

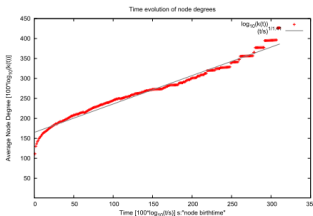
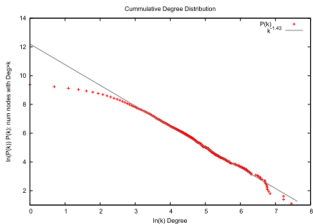
Threadlength distribution



Cummulative threadlength distributions:

$$P(\ell) \sim \ell^{1-\lambda}$$
$$\lambda = 3.35(\text{TeX})$$
$$\lambda = 2.90(\text{AI})$$

AI threads are $10 \times$ longer.



Common threads imply links:
Link any two posters sharing a thread.

Preferential attachment relates the degree distribution exponent γ :

$$P(k) \sim k^{1-\gamma}$$

to degree time evolution

$$k(s, t) \sim (t/s)^{1/(\gamma'-1)}$$

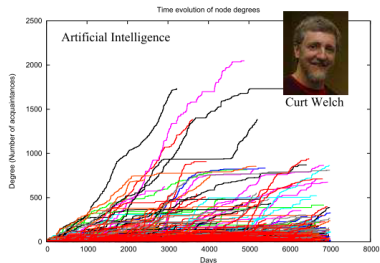
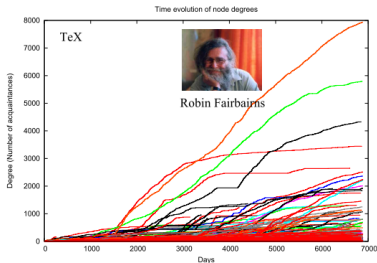
(s : 'birth time' of the node)

Here, we find

$$\gamma = 2.43 \approx 2.41 = \gamma'$$

- as expected for the preferential attachment mechanism.

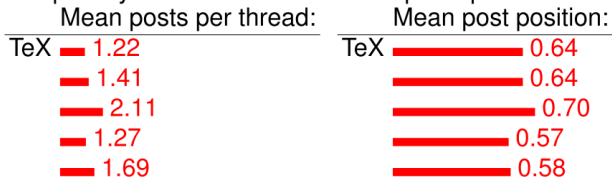
Evolution of individual degree



Time evolution of individual degrees (number of 'followers').
TeX high frequency posters post with almost constant frequency.
AI high frequency posters show 'on' - 'off' periods.

Roles of high frequency posters

The roles of high frequency posters differ between the two groups. We quantify the differences for the top five posters of each group:



TeX: High frequency posters in TeX answer and often close questions raised in the thread. This is quantified by a low number of posts per thread and a post position towards the end of the thread.



AI: High frequency posters in AI open threads and take part in the evolving discussion. This is quantified by a high number of posts per thread and a post position near the centre of the thread.

Other Roles and quantitative measures

90% of Reinhard Zierke's posts ('TeX' member) are not answered upon.

Reinhard is not a **troll**, he just regularly announces new TeX packages.

92% of George Bajszar's posts (AI member) remain idle, George is classified by fellow group members as a **Usenet Spammer**.

Comparing two online discussion groups, one on TeXnical issues the other one on philosophical (but computer related) issues we find statistical similarities in network growth, scalefree behaviour and appearance of high frequency posters. More work is needed to quantify the specific experts vs newbies division prevailing in technical discussions.