The Spanish-speaking Blogosphere: towards the powerlaw?

Fernando Tricas
(Depto Informática e Ing. de Sistemas, U. Zaragoza, Spain)
http://fernand0.blogalia.com/

Juan J. Merelo
(Depto. Arquitectura y Tecnología de Computadores, U. Granada, Spain)
http://atalaya.blogalia.com/
Outline

- Motivation
- Some definitions
- Tools
- The Spanish blogosphere
- Powerlaws and evolution
- Conclusions and future work
Motivation

- The Spanish blogosphere has a manageable size
- Desire of better understanding
- Detection of problems
- Improving self-consciousness
- We can think of it as a community

→ We are observing its evolution
What is a weblog?

- Blog, weblog, logbook of thoughts, links, histories on the web
  - Dave Winer’s
    http://www.scripting.com/
  - Cameron Barret’s
    http://www.camworld.com/
- Blog-hosting sites
  - Blogger
  - Userland
  - Blogia, Blogalia
What is a blog?

- Reverse-chronological order
- Comments.
- Alternative formats (Rich Site Summary, RSS).
- Other stuff: polls, blogroll, music, books, latest comments...
Tools

http://blogometro.blogalia.com/

- Each day a ranking of fresh links
- Open source
- Since the end of 2002
- Some purging
The Spanish blogosphere

First blogs: Mid of 1999

- Several individual people
- Barrapunto (Slashdot clone)
- Blogspot
- Other hosting sites
- Blogalia, Blogia, libreexpresion, zonalibre, bitacoras.com, ...
The Spanish blogosphere

Complex thing!

- Several countries
- Several languages (in Spain there are four official languages, but there is also people blogging in English, French, Dutch, ...)
- Of course, bilingual and even trilingual blogs
Just in case ...
http://oreneta.com/baldie/blog
- Definition subject to discussion!
Successful initiatives

- Ciberamigo Invisible
  http://www.awacate.com/amigo/

- Topic Exchange
  http://topicexchange.com/t/bitacoras/

- Directories
  http://bitacoras.net,
  http://bitacoras.com/

- Vecindario
  http://www.pisotrece.com/vecindario/

- ‘Mapa de la Blogosfera hispana’
  http://www.microsiervos.com/blogosfera/

- More ...
Recent initiatives

- bitacoras.com (pings, directory, ...)
- bitacoras.org (meta–weblog, there are others...)

[YO]

[trackback]
Some figures

- More than 10500 blogs
- More than 1000000 links were observed (723000 in the paper)
- Around 250000 links in any given day
- 3500 new posts a day (?)
- January 2004: 3060 blogs with at least, a link
  January (1299 blogs), and in November (943 blogs).
Spanish blogosphere and powerlaws

The global blogosphere

http://www.shirky.com/writings/powerlaw_weblog.html
http://homepage.mac.com/kevinmarks/powerlaws.html
Number of blogs

Number of weblogs

Total number

01/01 01/02 01/03 01/04 01/05 01/06 01/07 01/08 01/09 01/10 01/11

1500 2000 2500 3000 3500 4000 4500 5000 5500
Spanish blogosphere and powerlaws

Links vs blog distribution

# links vs rank

Data +
Fitted curve --

Web Based Communities 2004. Lisbon, Portugal– p.14/19
Spanish blogosphere and powerlaws

Links vs blog distribution

# links vs rank

Data vs fitted curve

Web Based Communities 2004. Lisbon, Portugal– p.15/19
Evolution

chi-square test

rms
Evolution

- The fit is better with time (and when the number of weblogs grows)
- This can show that the blogosphere is still young (but evolving)
- Maybe we are a bit different from the global one?
Conclusions

• No clear idea about size and extension (but ...)
• Is the Spanish blogosphere reaching a mature state?
• Not many links
Future Work

• Software improvement
• Results diffusion
• Other measurements (words, phrases, ...)
• Identification of in/out links
• Cluster formation (WBC 2004, first approach)
• Open access to our data
• Other ‘blogospheres’