

Blogosphere



- ✦ What is blogosphere?
- ✦ Why do we need to study Blog-space or Blogosphere?



✦ What is blogosphere?

The interaction between blogs can be viewed as a network of hyper-linked and time stamped posts, called “blogosphere”.

✦ Why do we need to study Blog-space or Blogosphere

There are at least two important reasons for the systematic study Blog-space.

Sociological reasons and Technical reasons.

On the Bursty Evolution of Blogspace

- ✦ **Purpose:** address the evolution of hyperlinked corpora.
- ✦ **Tools:** “time graph” and “time-dense community tracking” in the context of Blogspace.
- ✦ **Main contributions:** 1 introducing a combinatorial object we call a time graph for the study of graphs that evolve in continuous time. 2 defining a notion of communities in Blogspace and extend Kleinberg’s notion of temporal bursts. 3 conducting a series of experiments, showing the development of macroscopic and microscopic community structure, and the evolution of burstiness.

Burst

- ✦ An event might correspond to the appearance of an email containing particular keywords. The crucial step is to model such bursts so that they can be identified efficiently.
- ✦ Within a community of interacting bloggers, a given topic may become the subject of intense debate for a period of time, then fade away. These bursts of activity are typified by heightened hyperlinking amongst the blogs involved-within a time interval.(Bursty communities of blogs.)

Modeling Blog Dynamic



✦ **Problems:** How do blogs produce posts?

What local, underlying mechanisms lead to the bursty temporal behaviors observed in blog networks?

✦ **Method:** “Zero Crossing” model.

We propose model called ZC which uses simple and intuitive principle for each individual blog, and it is able to produce the temporal characteristics of the blogosphere together with global topological network patterns.

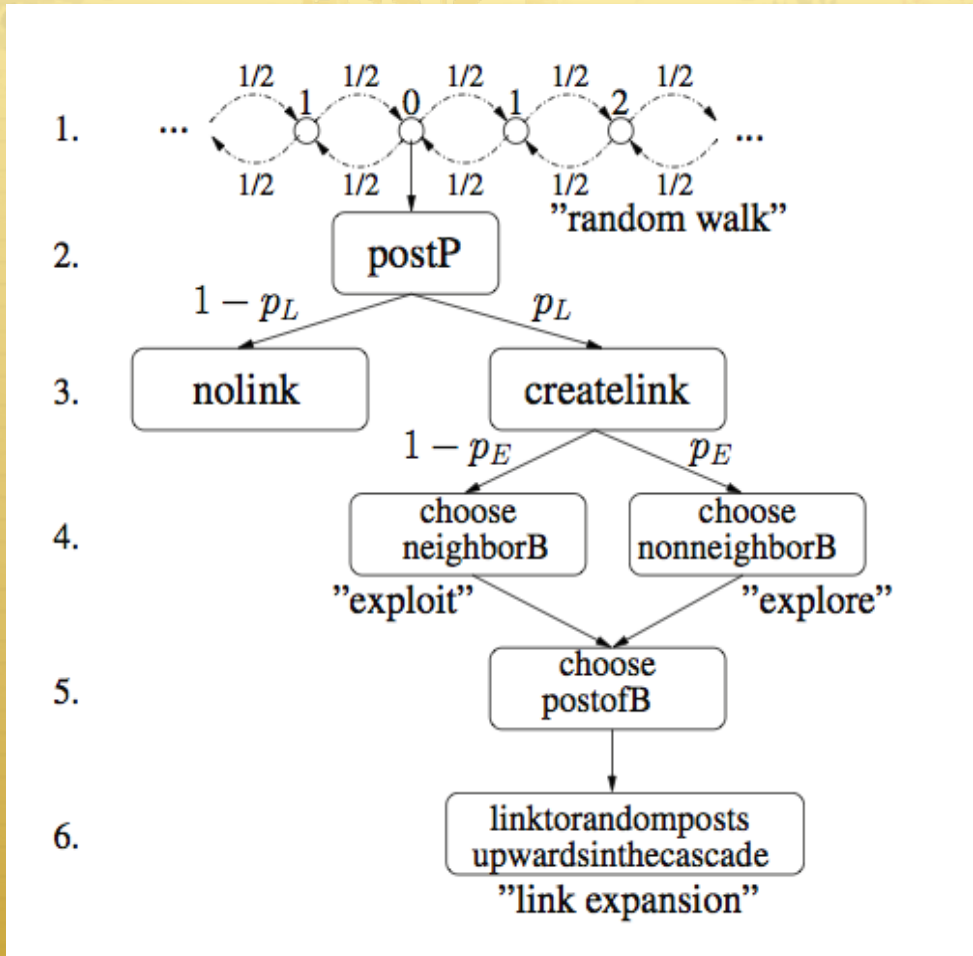


Figure 1: Our zero-crossing model ZC. Each blog behaves according to this model. Numbers correspond to the steps of our ZC generative model.

The model involves three major mechanisms,
each handling one aspect of the dynamics of
•◆• the blogosphere.◆•

✦ WHEN

When would a blogger write a post?

✦ WHAT

Once a blogger has decided to blog, which other blogs(if any) will he choose to read, and which posts inside those chosen blogs will he choose to cite?

✦ FOLLOW UP

Once a blogger decides to cite post Q, he may follow up on it, and also cite some of the posts that Q is citing.; the blogger may do that recursively. We will refer to this mechanism as link expansion

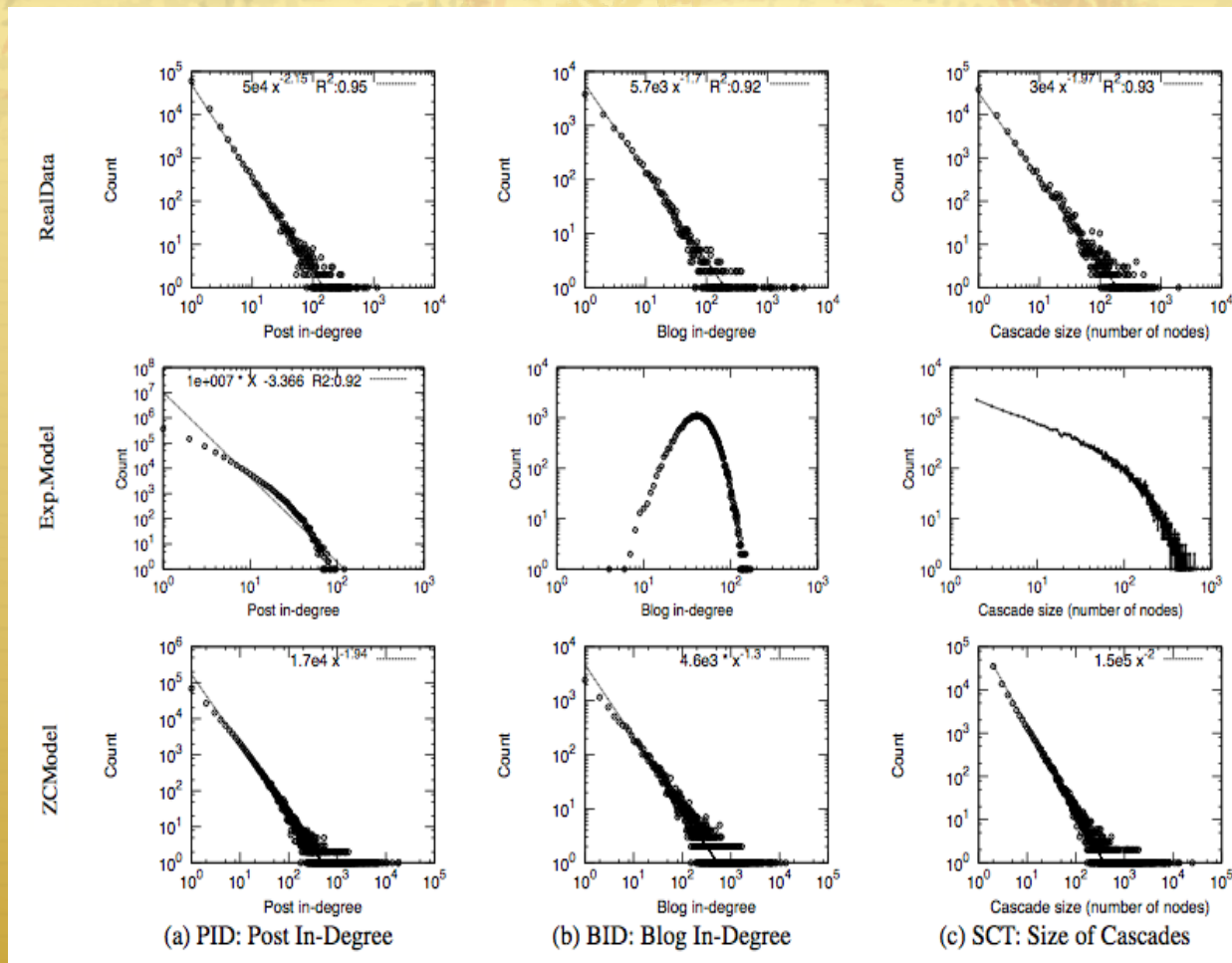
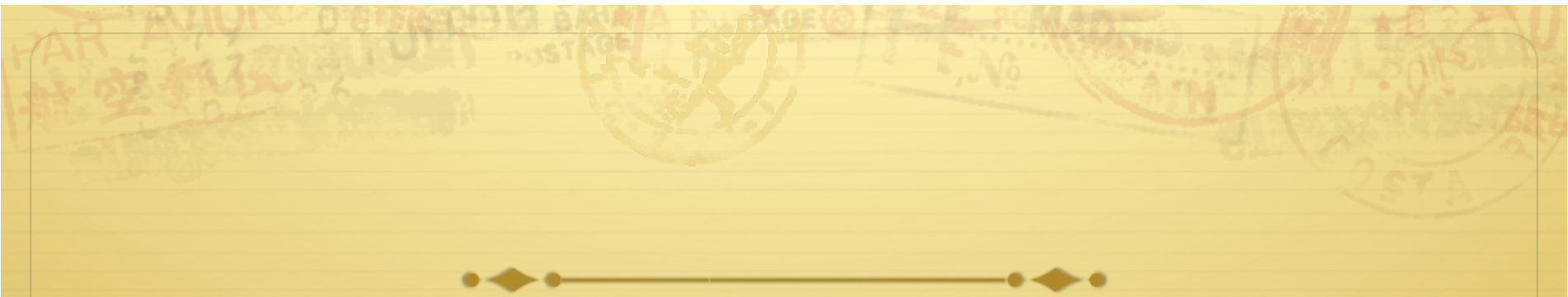


Figure 2: Topological patterns of the blogosphere. Top: real blogosphere; Middle: Experimental model; Bottom: blogosphere as modeled by the ZC model. Notice ZC model outperforms EXP model and matches the properties of real blogosphere



✦ **Contribution:** This model can naturally be used to generate synthetic blogospheres for what-if scenarios, to explore and model blog dynamics for the purpose of information propagation, marketing, and advertising.


Meme-tracking and the Dynamics of the News Cycle



- ✦ **Problem:** How to track new topics, ideas , and memes across the web ?

Prior work has identified two main approaches to this problem: probabilistic term mixtures and entities which however lies much of the temporal and textual range over which propagation on the web and between people typically occurs, through the continuous interaction of news, blogs, and websites on a daily basis.

- ✦ **New problem:** How to track short unit of text, short phrases, and memes that act as signature of topics and events propagate and diffuse over the web, from mainstream media to blogs, and vice versa

- 
- ✦ **Method:** Here we propose a approach named meme-tracking . And our approach to meme-tracking with applications to the news cycle.
 - ✦ **Contribution:** Developed a framework for tracking short, distinctive phrase that travel relatively intact through on-line text and presented scalable algorithm for identifying and clustering textual variants of such phrases that scale to a collections of 90 million articles.

Turning Down the Noise in the Blogosphere

Problem: Users have to cope with overload since the blogosphere has experienced a substantial increase in the number of posts published daily.

Approach: present a principled approach for picking set of posts that best covers the important stories in the blogosphere.

Result: derive two different algorithms based on general framework. One efficient algorithm to select the top stories in the blogshpere. Another achieves no-regret personalization which based on his personal preferences.