

网络论坛的一些统计规律及其 数学模型

胡延庆

2010年7月26日

人类动力学:

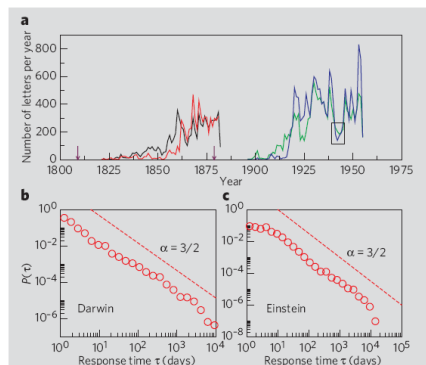
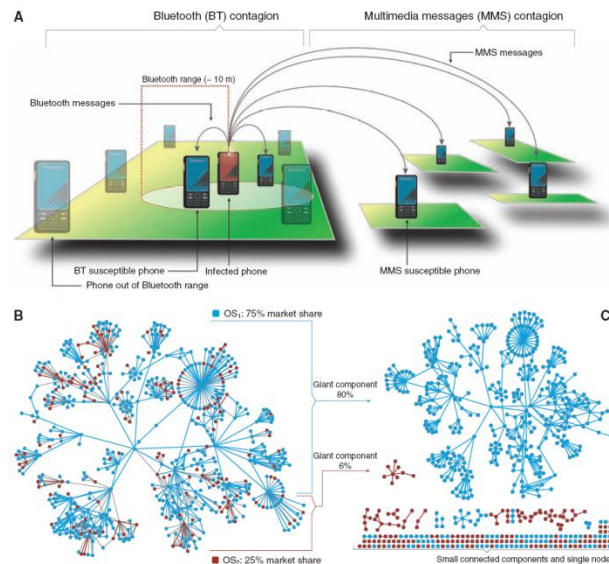
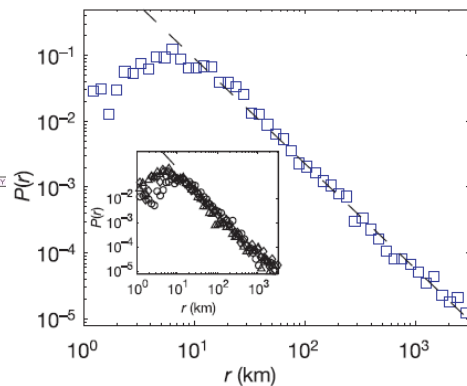
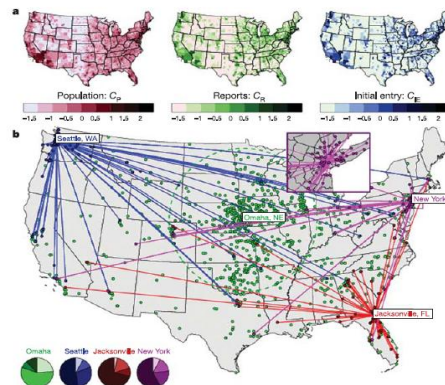
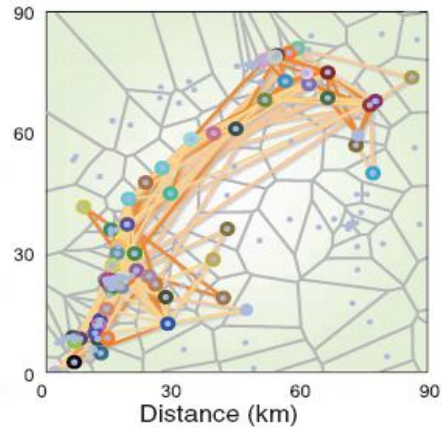
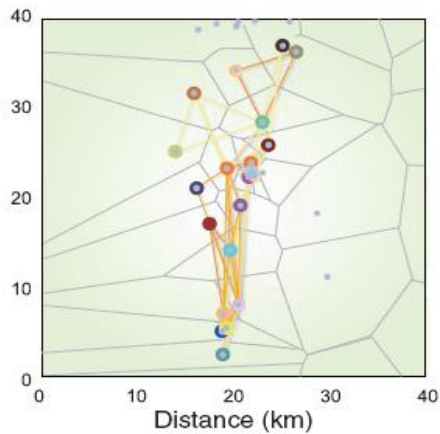
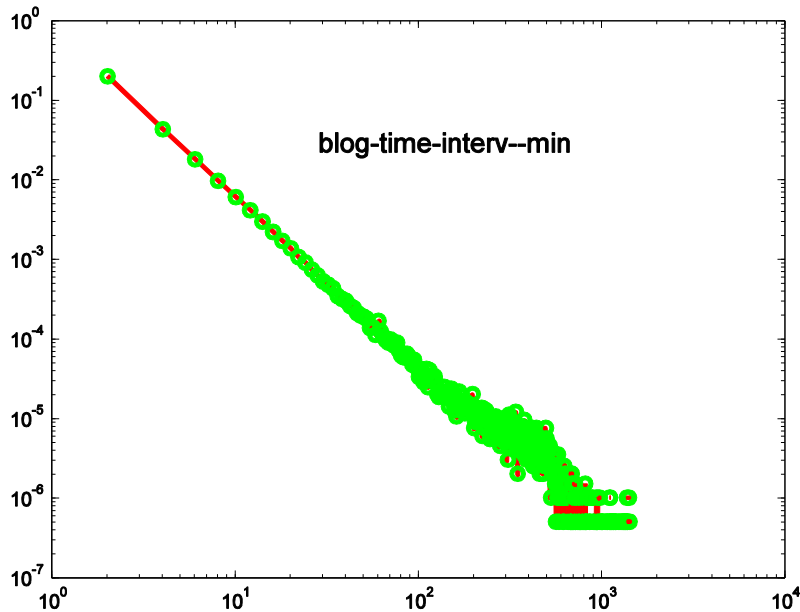


Figure 1 | The correspondence patterns of Darwin and Einstein.

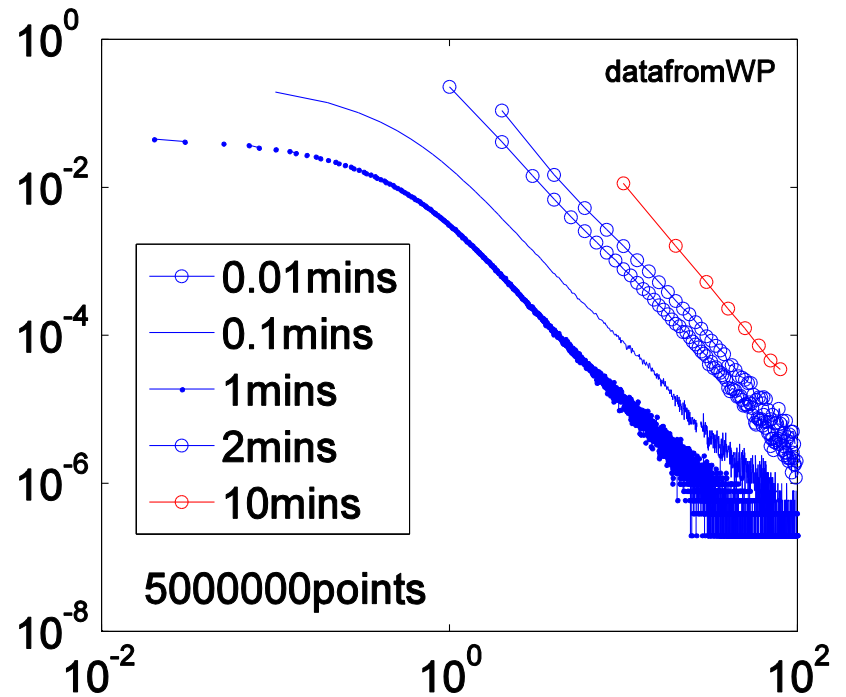


在信息社会下人的行为研究具有很大潜力!

BBS、论坛、博客上的一些统计特征



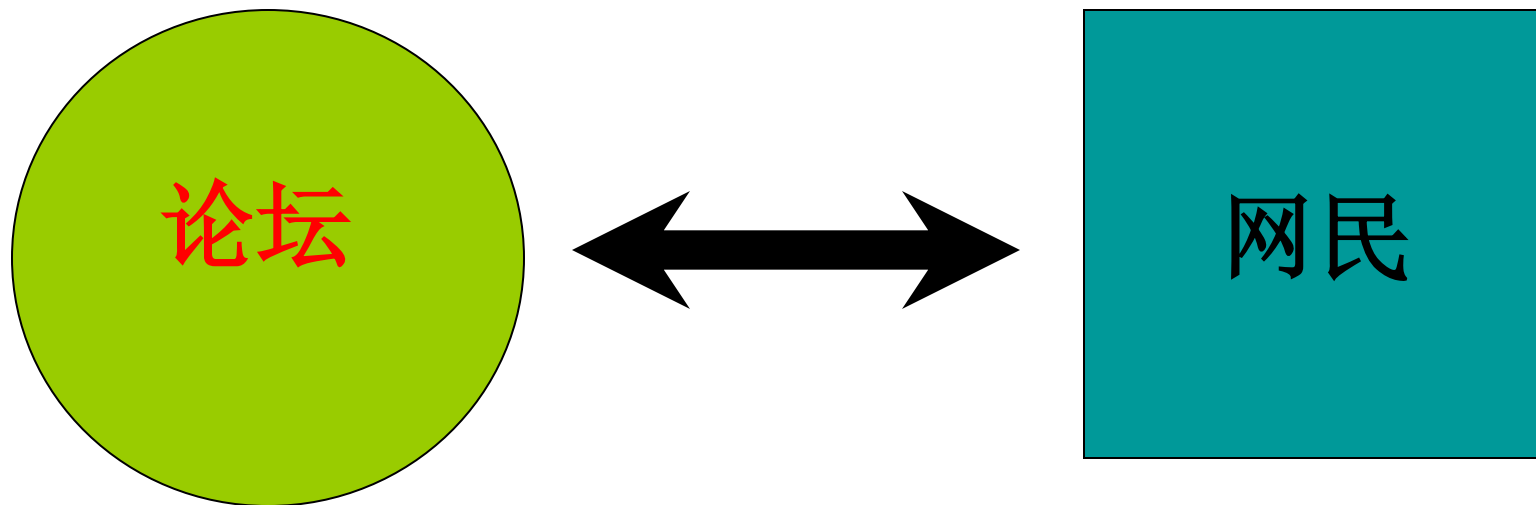
Openlab数据



南京大学BBS

当把**BBS**或者一个论坛看成一个系统，把人们它们上面留下记录的时候看成是系统的活跃点，实证研究显示，在**BBS**或者论坛上面这些活跃点之间的时间间隔会呈现出明显的无标度特征。

模型



规则:

1. 网上信息愈多，人们越倾向于参与发帖。
2. 人们过一段时间总有些东西想发到网上。

模型的数学描述

$$\begin{cases} S(t) = N(t) - N(t-T) + 1 \\ M(t) = t * c - N(t) + b \\ r(t) = 1 - (1 - S(t) * tc)^{M(t)} \approx S(t) * M(t) * tc \end{cases}$$

这里：

$S(t)$ 表示当前 t 时刻人们能够看见有效帖子的数量

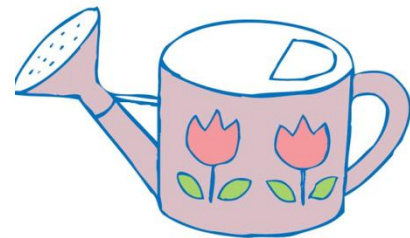
$N(t)$ 表示当前 t 时刻论坛上面帖子的数量

$M(t)$ 表示到目前 t 为止前人们潜在发帖的数量

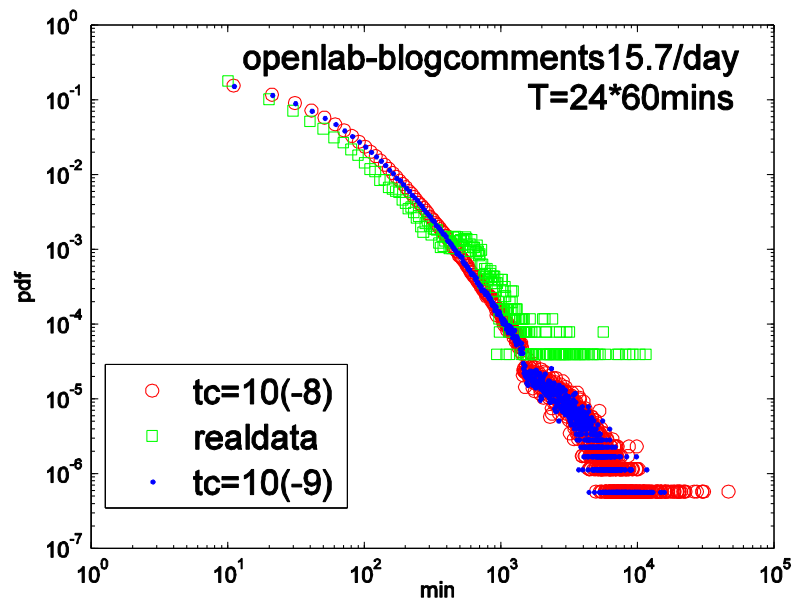
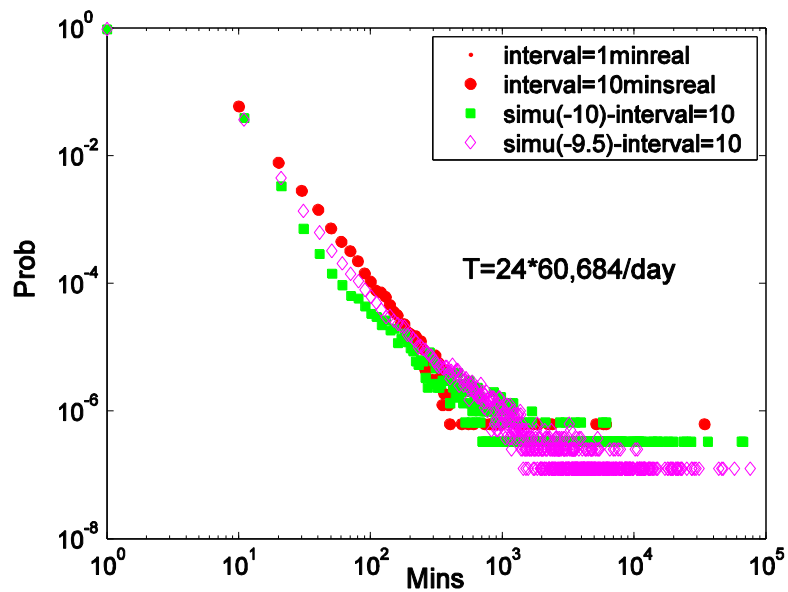
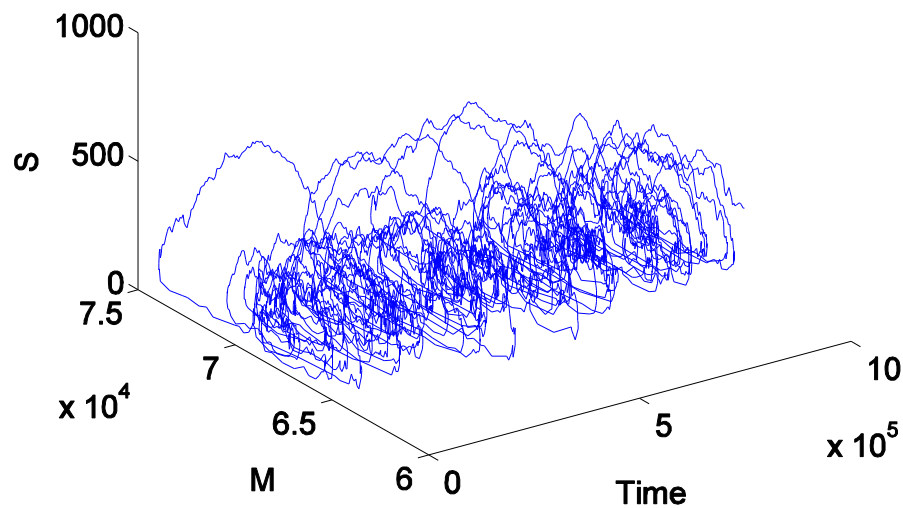
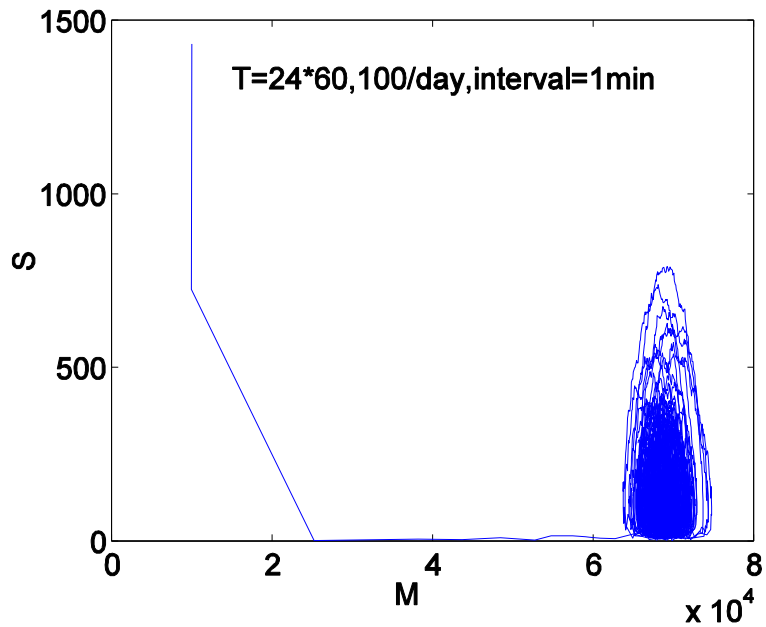
$r(t)$ 表示当前 t 时刻发一个帖子的概率

c, b 是常数， c 可以通过网络算出来。

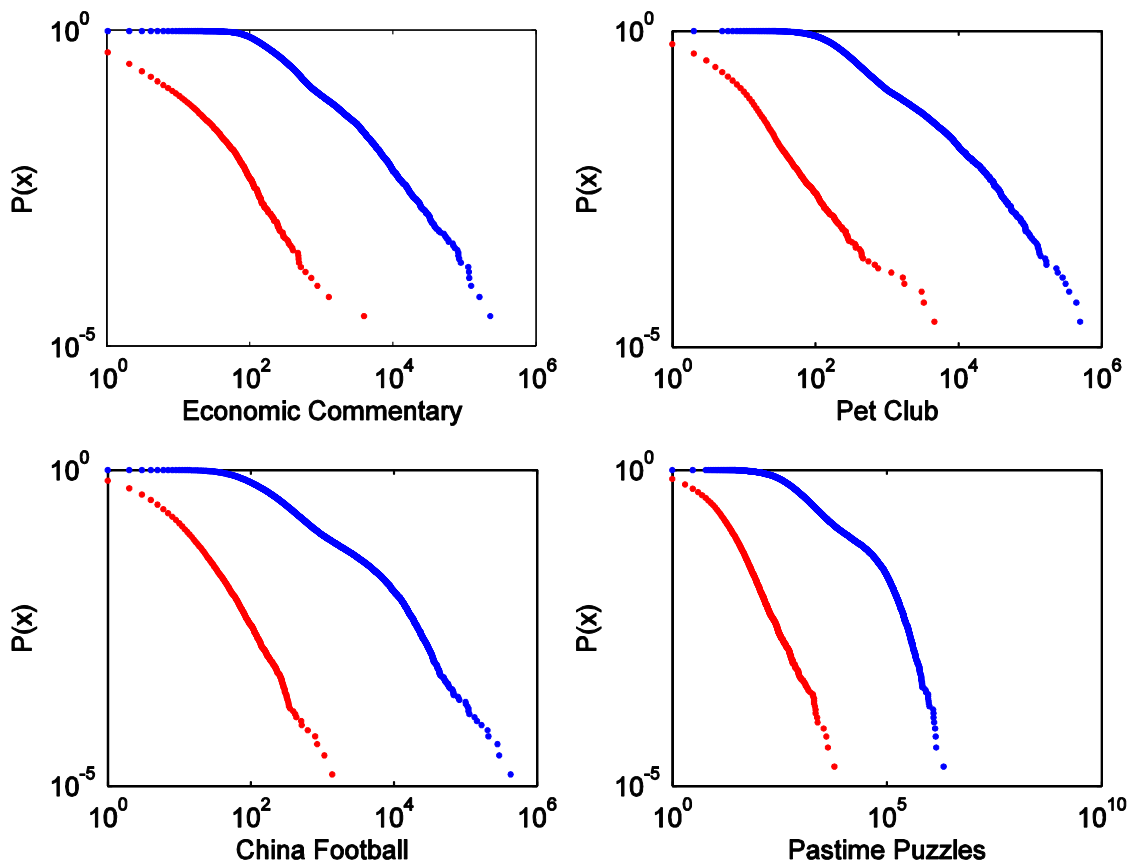
tc 可调节参数



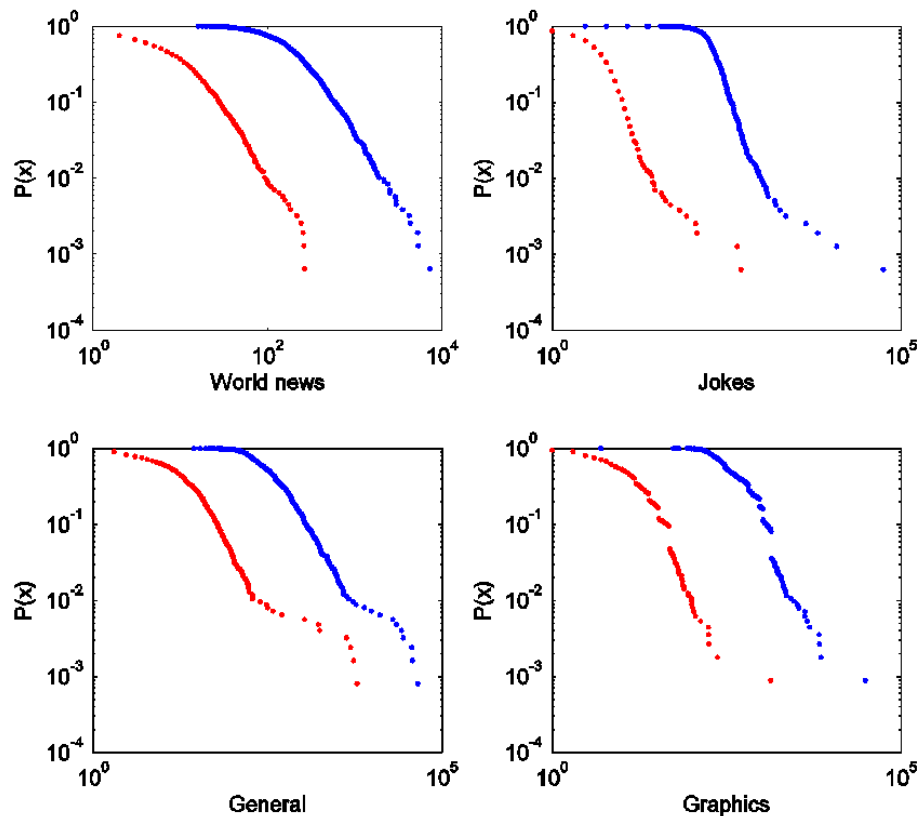
结果



网络论坛帖子关注程度特征



Here we use the number of views and replies in entertainment, sports, life, economics themes, including Economic Commentary, there are 32,480 threads from April 17, 2001 to November 11, 2008; Pet Club, there are 39,950 threads from December 1, 2005 to November 17, 2008; China Football, there were 95,891 threads from June 14, 2001 to July 13, 2008; Pastime Puzzles, there were 49,346 threads from November 19, 2004 to July 8, 2008.

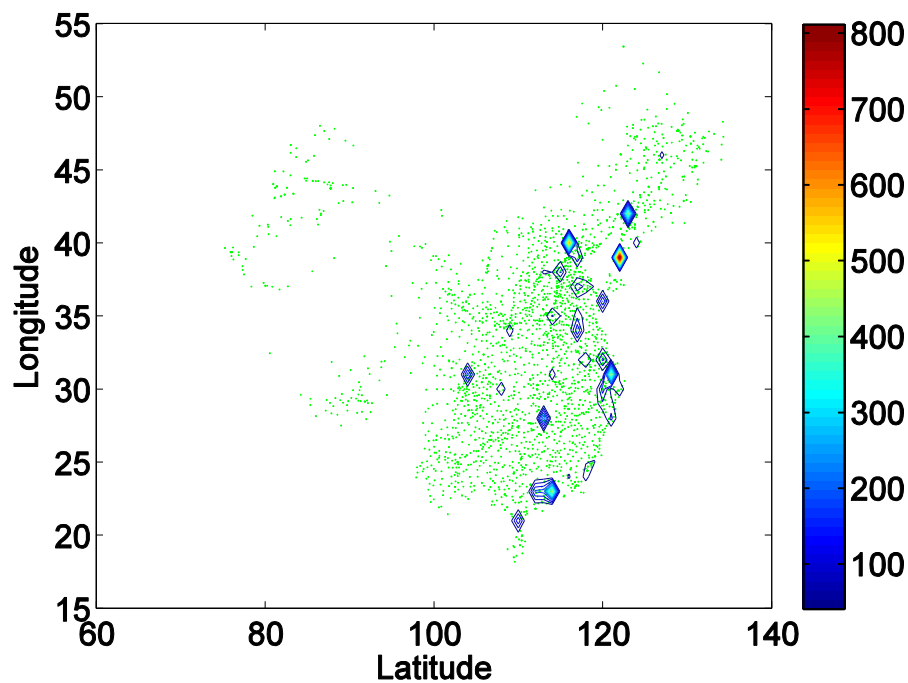


Chat Avenue is the current leader in western free online chat rooms. Most of its users are from the USA, Canada, Australia and the United Kingdom. There are thirteen chat rooms to choose from with a specific demographic for each and every user [16]. We mainly get the views and replies from World Events Politics and News, Jokes, General, and Graphics from the Chat Avenue forum, including World Events Politics and News, there were 1521 threads from June 10, 2006 to February 23, 2009; Jokes, there were 1463 threads from June 10, 2006 to February 25, 2009; General, there were 1212 threads from June 10, 2006 to March 3, 2009; Graphics, there were 1131 threads from June 10, 2006 to March 13, 2009.

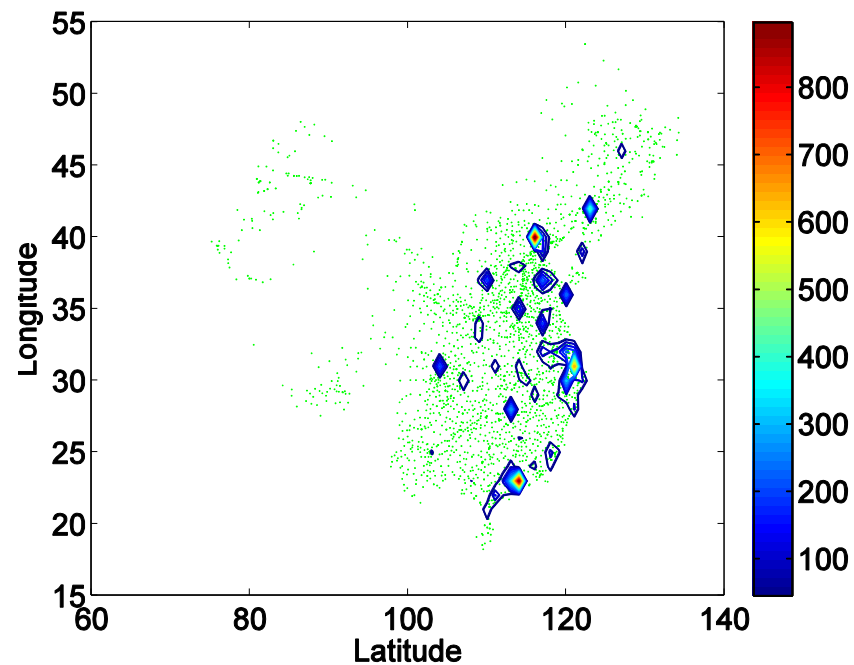
Forum	Theme	Views			Replies		
		α	x_{\min}	p	α	x_{\min}	p
<i>Sina</i>	Economic Commentary	2.05	256	0.37	2.91	57	0.438
	Pet Club	1.94	207	0	2.66	13	0.73
	China Football	1.87	119	0	3.06	60	0.43
	Pastime Puzzles	1.62	92	0	2.51	46	0.58
<i>Chat Avenue</i>	World News	2.69	358	0.303	2.81	26	0.469
	Jokes	3.44	182	0.185	3.31	7	0.042
	General	2.45	559	0.205	2.5	31	0.118
	Graphics	2.73	534	0.488	2.4	14	0.038

Forum	Theme	Views			Replies		
		α	x_{\min}	p	α	x_{\min}	p
<i>Sina</i>	Synthesis Competition	2.16	24 109	0.33	2.39	53	0.02
	Stocks and Capital	1.89	550	0.55	2.29	23	0.95
	Postgraduate Communication	2.04	2 372	0.01	2.13	5	0.27
	Words on World	2.09	13 689	0.2	2.18	8	0.001
	08European Cup	2.05	162	0	2.55	35	0.087
<i>Chat Avenue</i>	Music	2.34	194	0	2.04	10	0
	Member Pics	2.62	1 254	0.57	3.22	122	0.645
	Literature	2.18	764	0.71	2.31	9	0.12
	Movies	2.22	493	0.5	2.24	14	0.13
	Debate	2.63	2 104	0.87	2.67	63	0.12

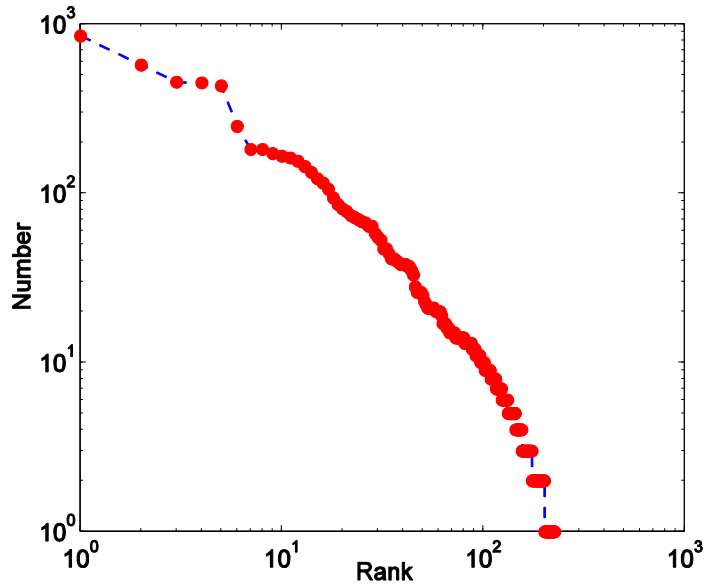
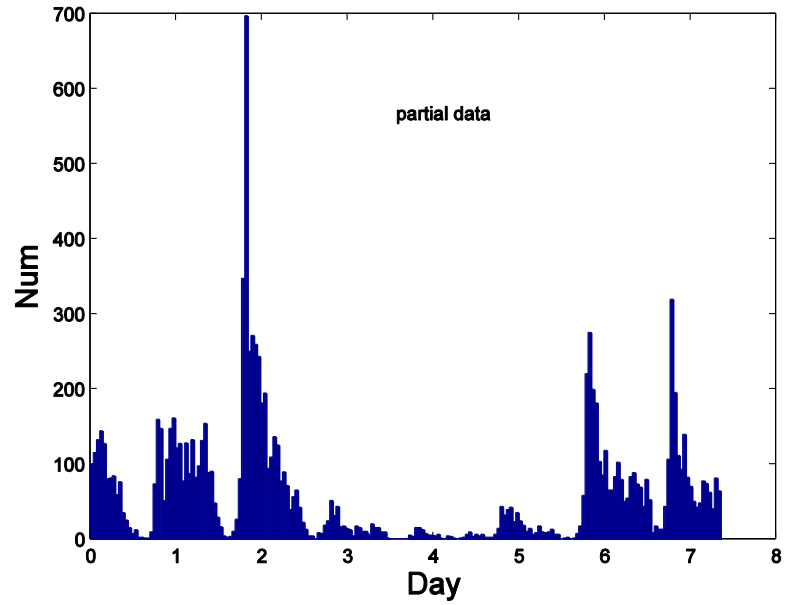
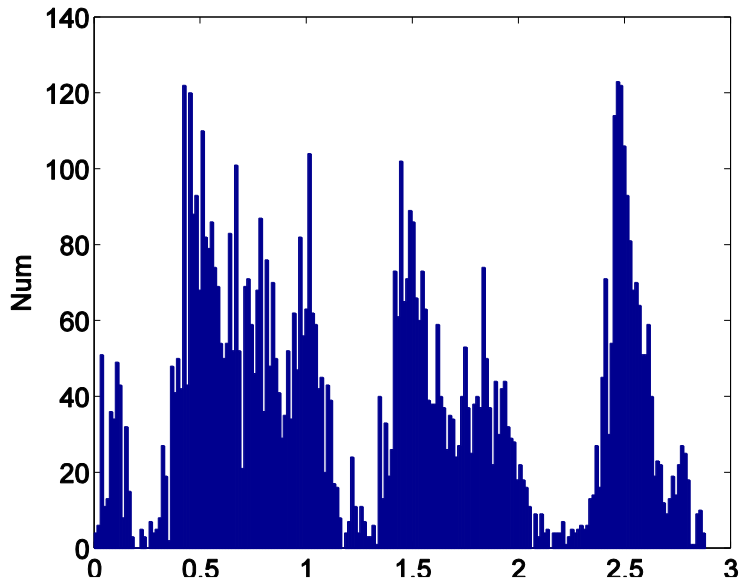
中国网民的地理分布特征



大连湾石油事件



南方暴雨事件



To be continued....

Thank you!