

# Guanxi in the Chinese Web

❖ Louis Yu

❖ Valerie King

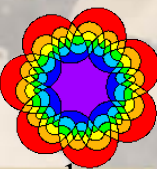
❖ Yan Zhuang

University of Victoria  
Victoria, BC, Canada  
*{yul, val, yzhuang}@uvic.ca*



University  
of Victoria

Department of Computer Science



# Outline

- ❖ Introduction and related work
  - ⌘ Link analysis and web graph measurement
  - ⌘ Random graph modelling
  - ⌘ Social network analysis
- ❖ The concept of *guanxi*
- ❖ *Guanxi* in the web
- ❖ Empirical study, analysis and modelling
- ❖ Identifying *guanxi* web sites
- ❖ Conclusion

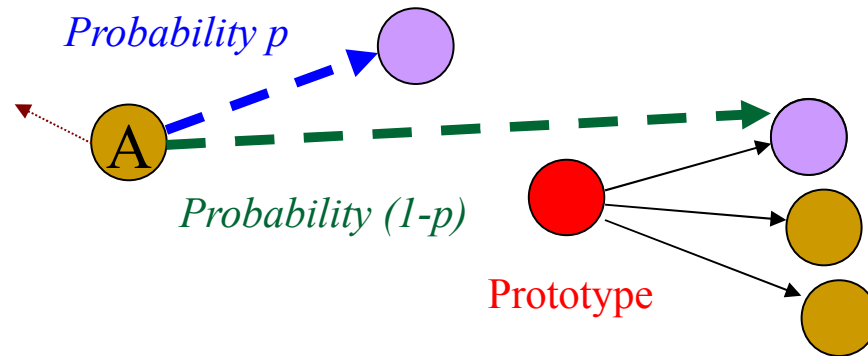
---

# Related Work

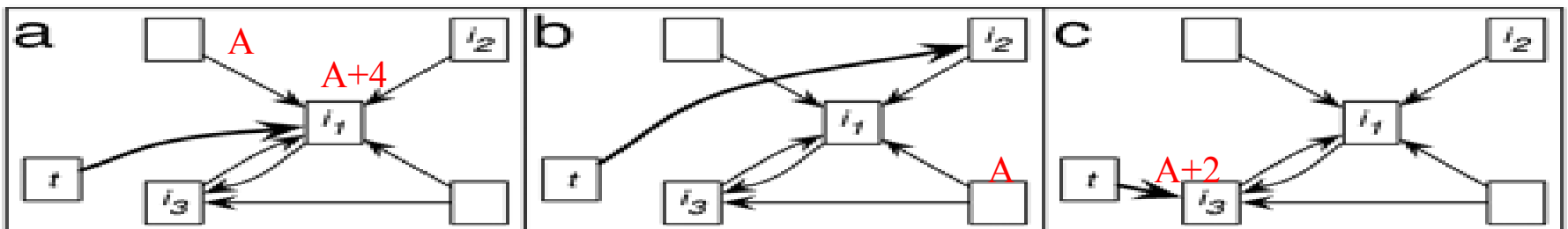
- **Link Analysis and Web Measurement**
  - PageRank [Brin and Page, 98] – peer endorsement
  - Hits [Kleinberg, 99] - authority and hub
  - Power law distribution - the fraction of nodes with degree  $i$  is proportional to  $\frac{1}{i^\alpha}$  for some constant  $\alpha$ 
    - The in-degree, out-degree distribution of the web follows the power law
    - Scale free network
- **Social Network Analysis**
  - Structural and mathematical properties of social networks
    - Cliques, dyads, components and circles
    - The significance of positions in these networks
    - Empirical study of the structure of online social networks
      - Youtube, Flickr, Twitter, Facebook, blogs ...
      - Scale free, small-world

# Random Graph Models

- Hostgraph Model (the copying mechanism)
  - Authors of web pages will randomly find a page and then copy some portion of the links to their own page



- Generalized Preferential Attachment Model
  - New nodes are more likely to link to existing nodes with high indegree.
  - Each site is assigned with an initial attractiveness score  $A$ . Over time, the attractiveness score of a site  $S = A + \text{indegree of } S$



# 关系

## guānxi

“An informal, particularistic personal connection between two individuals who are bounded [sic] by an implicit psychological contract to follow the social norm of guanxi such as maintaining a long term relationship, mutual commitment, loyalty, and obligation. A quality guanxi is also characterized by the mutual trust and feeling developed between the two parties through numerous interactions following the self-disclosure, dynamic reciprocity, long term equity principles.”

- A term in sociology that refers to connections/relationships
  - Varies in different Chinese societies and changes over time
- It had been studied by sociologists, political scientists, economists and anthropologists
- One can find a similar notion of *guanxi* in the ancient writings of Confucius

# Types of *Guanxi*

- Most Western network theories focus on network structures and the significance of positions in the network
- Dyadic (two-party) relationships are the fundamental units of *guanxi* networks.
- The classification of *guanxi*
  - (1) Socio-affective
  - (2) Instrumental
  - (3) Mix *guanxi*



# The Establishment of *Guanxi*

## ■ *Guanxi* base

- Same birth place
- Same school
- Friends, family members, partners...

老乡

同事

同学

朋友

亲人

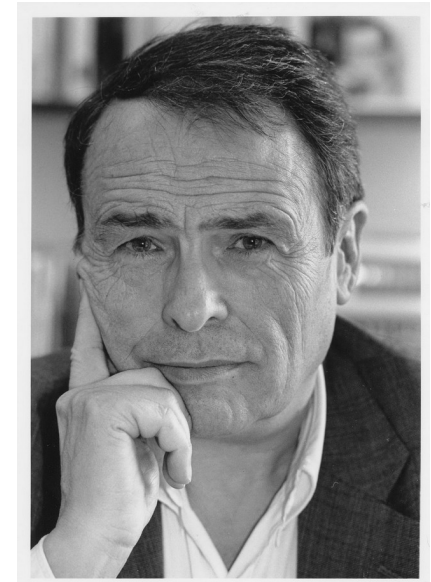
## ■ The exchange of resources/capital

- “*Qing*” = feeling 情
- “*Xin*” = trust 信 / 信任

## ■ Social ties

## ■ The exchange of capital

- Economic capital
- Social capital
- Culture capital



**Pierre Bourdieu**

# Guanxi Applied to the Web

- *Guanxi* between web sites
  - Regarding a web site as representing a company, a person, a news source ...
  - *Guanxi* base can be established between two web sites
    - Web sites that belong to family members, friends, collaborating partners
    - Web sites that originated in the same city, or have the same purpose
  - Two web sites may establish *guanxi*
  - The exchange of text or banner links

Collaborating partners

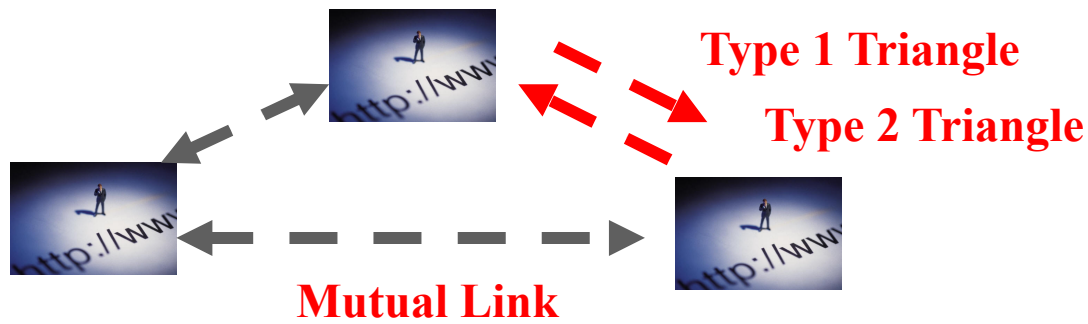
Same purpose





# Guanxi Applied to the Web

- Two web sites can also have *guanxi* due to an acquaintance through a third party web site with whom they both have *guanxi* with
- Web sites can establish *guanxi* due to an acquaintance through a link exchange platform
- Strong *guanxi* versus cheap *guanxi*
  - *Triangles are indications of cheap guanxi*



# Textual Indication of *Guanxi*

The image displays two web browser screenshots side-by-side, illustrating textual indications of *Guanxi* (connections) between the CCTV.com and Microsoft China websites.

**Left Screenshot (CCTV.com):**

- Browser: Windows Internet Explorer, URL: <http://www.cctv.com/program/C15676/14/12/index.shtml>
- Navigation: 关于cctv.com, 法律顾问, 版权管理, **友情链接** (highlighted with a red circle), 联系我们, 使用帮助
- Content: 公司介绍 | 管理团队
- Media: 媒体: 人民网 | 新华网 | 中国网 | 中国日报网 | 巨方网 | 人民日报 | 求是 | 光明日报 | 解放军工商时报 | 新民网 | 京华网 | 北京电视台宽频
- Government: 政务: 全国人民代表大会 | 中国人民政治协商会议 | 科技部 | 工业和信息化部 | 国家民委 | 公安 | 环境保护部 | 住房和城乡建设部 | 交通运输部 | 审计署 | 国资委 | 海关总署 | 税务总局 | 总局 | 统计局 | 林业局 | 知识产权局 | 旅游局 | 台办 | 新闻办 | 档案局 | 新华社 | 中科院 | 监会 | 保监会 | 电监会 | 社保基金会 | 自然 | 局 | 测绘局 | 民航局 | 邮政局 | 文物局 | | 国家语委 | 国务院扶贫办 | 国务院三峡办
- TV: 电视: 北京 | 天津 | 河北 | 石家庄 | 秦皇岛 | 山 | 南 | 浙江 | 杭州 | 绍兴 | 安徽 | 福建 | 广东 | 南方 | 广州 | 深圳 | 海南 | 珠海 |

**Right Screenshot (Microsoft China):**

- Browser: Microsoft, URL: Microsoft.com
- Navigation: 微软中国社区, 微软全球社区
- Content: 微软最有价值专家 - MVP, 微软技术社区联盟, 微软中文技术论坛, 技术聊天室, 社区资源网站集, 用户组, Webcasts, 更多资源
- Partnership: 微软中国社区合作伙伴 (highlighted with a red circle)
- Advertisement: IASK爱问 www.iAsk.com 爱问知识人 (zhishi.iask.com) <http://zhishi.iask.com>
- Text: 在新浪爱问知识人首页点击“电脑、通讯、互联网”，进入后即可看到该分类的问题列表。我们的MVP (微软最有价值专家)将为爱问提供的技术支持包括：汇编语言、C/C++、VB、VC++、Java、Delphi、Web、.Net、CMM、Windows、网络安全、办公软件等。另外，爱问知识人里设有“MVP专区”，用户可以随时进入并提交最有难度的问题，我们会动员微软全球MVP为用户解答。
- Product Technology Community: Windows, Office, 开发工具, 服务器, Mobile
- News: KAKU (卡酷) 卡酷出品《快乐奔跑》8月21日北京首映, 樱桃小丸子与“卡酷派”一起过暑假, 卡酷授权动漫元素文具用品首发上市, 《宝贝星计划》：关注宝宝的吃喝难, 看卡酷“动漫之王” 福迪斯尼之旅大奖
- 综合信息: 《故事汇》8月17日至8月21日节目预告, 文体群星聚北京 唱响精彩八月回眸奥运, BTV体育8月13-16日主要赛事及节目变更, 《手机江湖》液体电池生态手机, 《我的父亲母亲》中西医双博士教您自诊
- Partners: 北京电视台, 京视汽车, BTV 京视传媒
- Footer: 央视网 新华网 千龙网 北方网 京报网 北京广播网, 天津 河北 石家庄 秦皇岛 山西 太原 内蒙古 辽宁 沈阳 大连 吉林 长春 黑龙江 安徽 福建 厦门 江西 山东 济南 青岛 烟台 河南 湖北 武汉 湖南 广东 湛江 广西 重庆 四川 成都 云南 陕西 西安 甘肃 宁夏 新疆 贵州 宁波

**Annotations:**

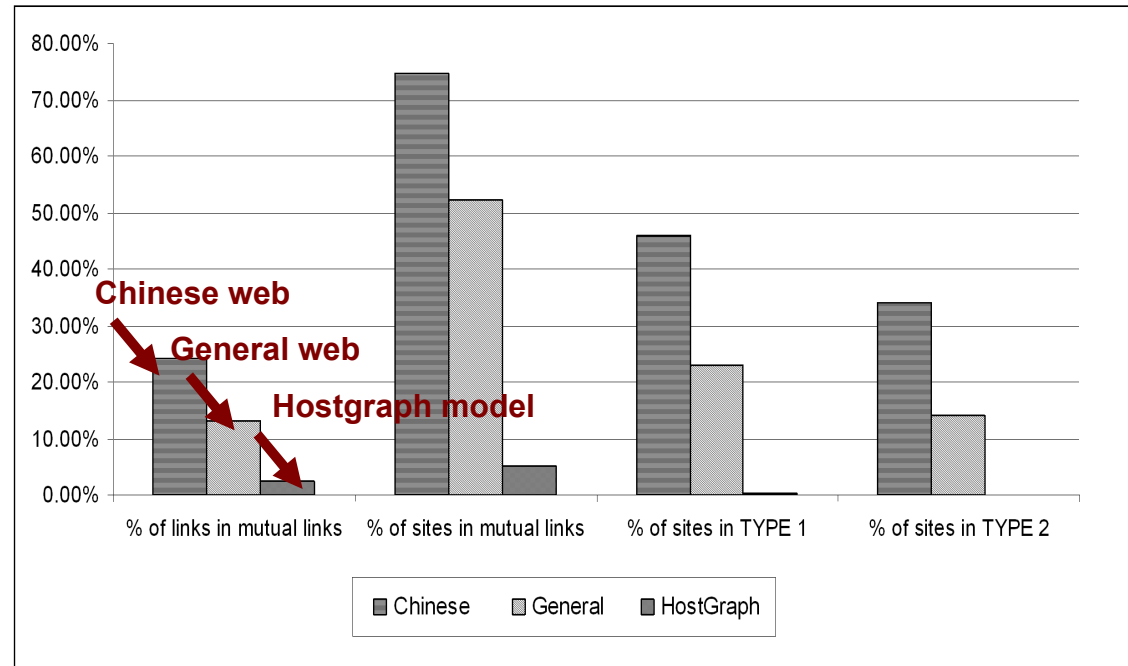
- Red arrows point from the "友情链接" button on CCTV.com to the "微软中国社区合作伙伴" link on Microsoft China.
- Red arrows point from the "媒体:" and "政务:" sections on CCTV.com to the "IASK爱问" advertisement on Microsoft China.
- A red circle highlights "北京" (Beijing) in the "电视:" list on CCTV.com, with a red dashed arrow pointing to the "北京电视台" logo on Microsoft China.
- A purple dashed arrow points from the "MVP专区" text in the IASK advertisement to the "我们的MVP" text in the same advertisement.

# Empirical Study of the Chinese web and the General Web

- The sample Chinese web - 10,000 web sites randomly selected from a crawl of the Chinese web conducted by Peking University
- The sample general web - 20,000 web sites randomly selected from the Open Directory Project ([www.dmoz.com](http://www.dmoz.com))
- Mutual Links, Type 1 triangles and Type 2 triangles

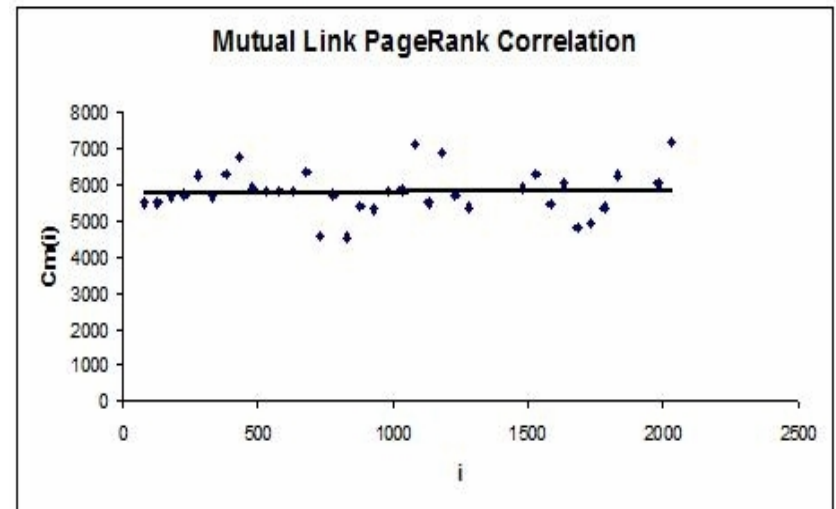
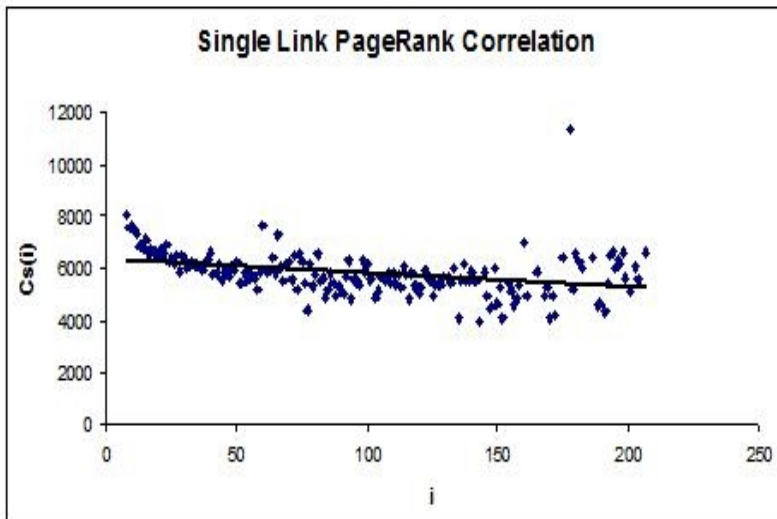
The Chinese web is denser relative to the general web and has been historically smaller

We use the hostgraph model to generate a random graph with size and density similar to the Chinese web



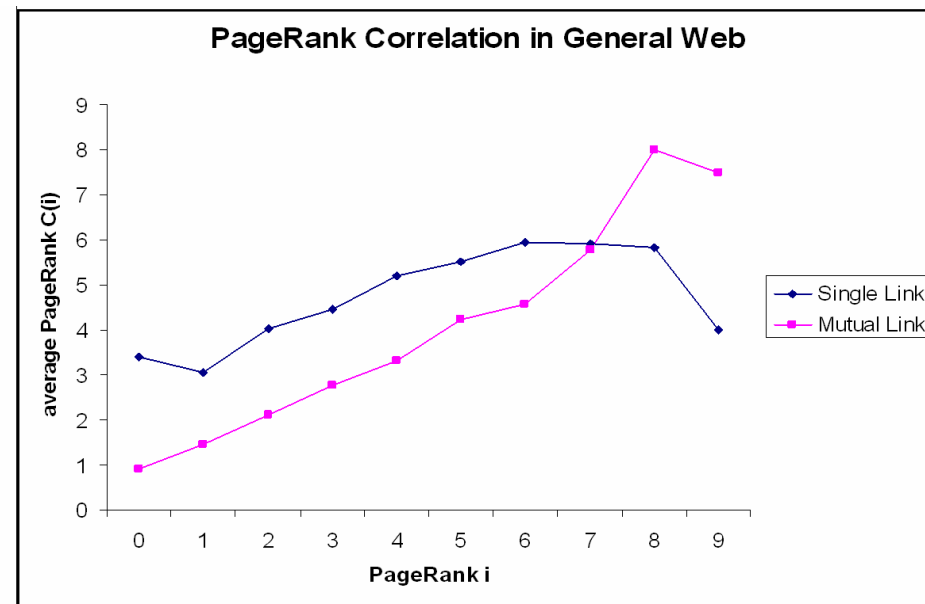
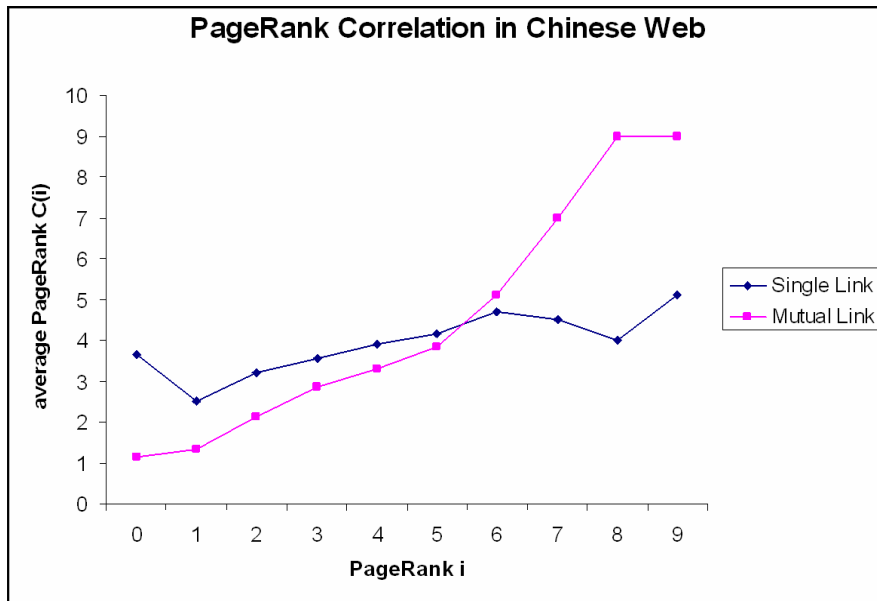
# PageRank Correlation

- We examine correlations between the sites' PageRank scores
- $C_S(i)$  ( $C_M(i)$ ) is the average PageRank of nodes with single (mutual) links from nodes with PageRank equal to  $i$
- *Under the preferential attachment model and the hostgraph model, the destinations of all links are determined independently of the PageRank of the source nodes*



# PageRank Correlation

- In the case of single links
  - The plot of  $C_S(i)$  as a function of PageRank  $i$  is almost flat
- In the case of mutual links
  - Sites with high PageRank  $\leftrightarrow$  sites with high PageRank
  - Sites with low PageRank  $\leftrightarrow$  sites with low PageRank



# A *Guanxi* Model of the Web

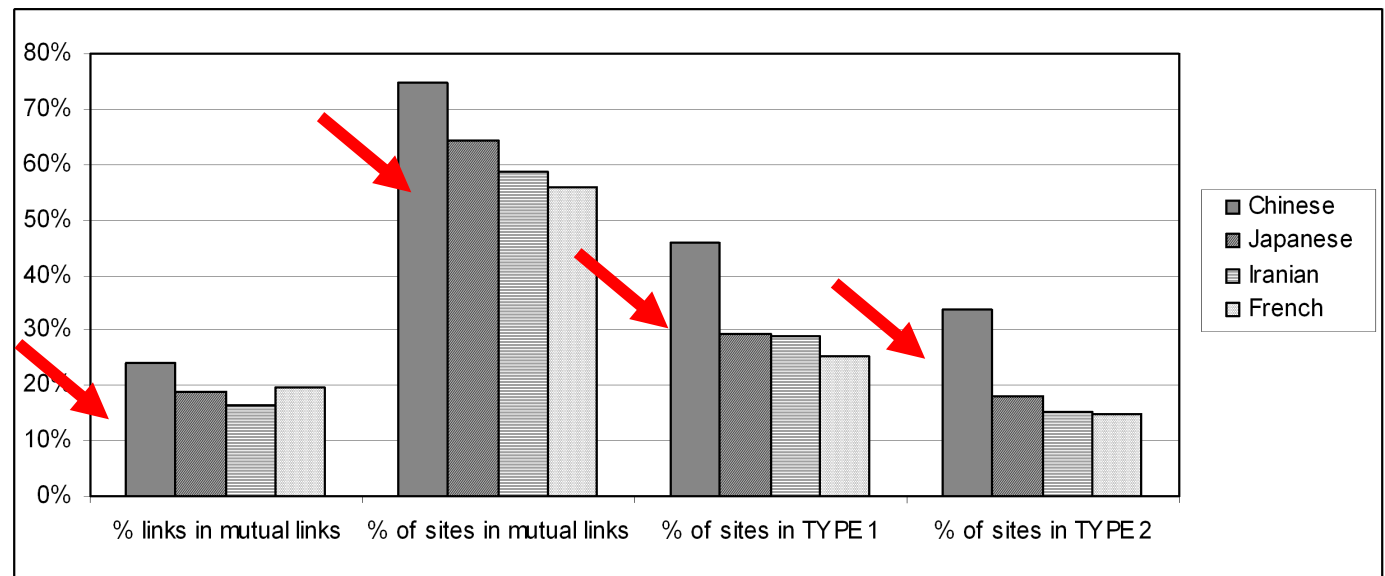
- The *guanxi* mechanism:
  - At each time step, we add  $k$  *guanxi* edges to a node  $A$ 
    - Choose a prototype and a destination node with a method similar to the hostgraph model.
      - (1) With probability  $q$ , the destination node links back
      - (2) With probability  $1-q$ , both the prototype and the destination node link back
- We Combine the *guanxi* mechanism with the hostgraph model

Parameter	Chinese	General
$\alpha$	0.37	0.61
$q$	0.44	0.71
$d_1$	0.87	0.82
$d_2$	0.83	0.74

*Table 1: Setting of Simulation*

# Comparison Against Web Sites in Different Countries

- We look at dyadic relationships between web sites in different countries.
  - Japan, France and Iran (.jp, .ir, .fr)
  - For each country, we randomly selected 5000 web sites



# Identifying *Guanxi* Web Sites

- We identified some characteristics that can be used to separate web sites that establish cheap *guanxi*, strong *guanxi* and no *guanxi*.
- We separated 17461 randomly selected web sites into three groups based on textual clues indicating that they are establishing different types of *guanxi*
  - Textual strong *guanxi* web sites
  - Textual cheap *guanxi* web sites
  - Textual no *guanxi* web sites
- After some experimentation, we formulated some rules to separate web sites based on characteristics in linking patterns and PageRank correlation
  - Structural cheap *guanxi* web sites
  - Structural strong *guanxi* web sites
  - Structural no *guanxi* web sites

Types of Guanxi	Cheap	Strong	No
Textual/Structural	0.774	0.795	0.909



---

# Conclusion

- We defined *guanxi* in the web
  - particular link patterns
  - supporting textual evidence in web pages
- Empirical study of the web sites in different countries
  - The Chinese web has a higher percentage of web sites in mutual links and triangles
  - PageRank correlation
- We present a mechanism to model the *guanxi* structure in the web
- We classify web sites that develop types of *guanxi* based on characteristics in linking patterns and PageRank correlation
- Applications:
  - Producing personally tailored recommendations
  - Filtering out web spam
  - Understanding social networks

# Questions?

THE WALL STREET JOURNAL. Digital Network | WSJ.com | MarketWatch | BARRON'S | All Things Digital.

Thursday, August 27, 2009 As of 3:56 PM PDT

THE WALL STREET JOURNAL. | BLOGS

U.S. Edition ▾ | Today's Paper | Video | Columns |

Home | World | U.S. | Business | Markets

WSJ Blogs >

## China Journal

How China is changing - and changing the v

A daily look at what we're reading.



- Is China America's bank, or should it be the other way around? An attempt to untangle the complex economic relations between the two countries. [Richard Spencer]

- Liu Xiaobo, one of the signers of the Charter 08, has reportedly been detained for the last week. Two human rights experts comment on some of the legal issues at stake. [China Law Prof Blog]

- A sign of hard economic times, or just a great moneymaking opportunity: Over a thousand college graduates compete to become pork sellers at unusually high salaries. [ChinaSMACK!]

- Guanxi (aka relationships, connections, social networks) is everywhere in China, including online. Here's an introduction to some recent research on the workings of guanxi on the Chinese Internet. [56minus1]

- Canceled: The NFL China Bowl, again, and this time it looks like for good. [China Sports Review]

- Director Jia Zhangke made this short film for the 10th anniversary of Modern Weekly magazine and discussed his work with the publication. [Danwei]

Our Work in the WSJ