MEASURING POLITICAL DEBATE ON THE CHINESE INTERNET FORUM WU MEI
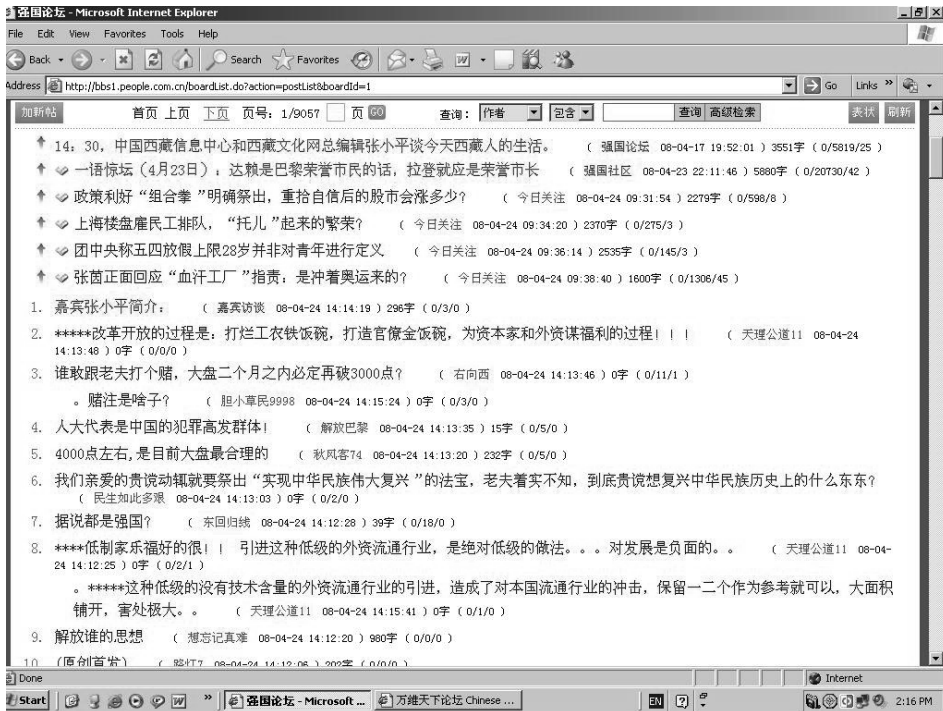
ABSTRACT

This paper addresses the methodological issue of how to quantify and measure seemingly chaotic, crude and impulsive online interactions in Chinese Internet forums, which have played an increasingly important role in shaping the country's political spectrum. It introduces an integrated research methodology, designed and developed specifically in a four-year project conducted by this author on the Chinese Internet forums, the first of its kind which scrutinises the interactive content of 14 major political forums in China and overseas over a one-year period, consisting of over 450,000 postings. Using commercial and self-developed online content analysis tools, this methodology features a series of computer-assisted data collection, quantitative analysis, and qualitative approaches. It investigates first from a wide range of mega data extracted online to capture the general dynamics of the Chinese Forum. It then focuses on the selected categories of text data to delineate the linkage of global connectivity among forum networks and reveal the spectrum of discourse and pattern of deliberation. The paper, which conceptualises Chinese Internet forums as discursive public spaces mediated by the new online conversational media, aims to establish a series of applicable methods to study the publicly accessed interactive content of online political talk in the global network of the Chinese Internet.

Wu Mei is Assistant Professor, Department of Communication, University of Macau, China, e-mail: meiwu@umac.mo.

Introduction

The Chinese Internet forum (*lutan* in Chinese, also known as the Chinese bulletin board system or BBS) features one of the most distinguished types of online political activism in the global Chinese language Internet, as about one quarter of the total online population in China are engaged in various kinds of Internet forum discussions (CNNIC 2006). Different from Newsgroups and message boards popular on the English Internet, the Chinese discussion boards have evolved into a unique public medium for information dissemination and public debate in global Chinese communities. The Chinese Internet Forum is defined in this paper as a particular kind of online discussion prototype, which in essence has become a generic term in the Chinese Internet culture referring to *a publicly accessible threaded discussion platform* on the Internet.[1]

Figure 1: A Typical Screen Page of the Chinese Internet Forum

The Chinese Internet Forum has been the subject of a large number of studies about politics, public participation and the Internet in Chinese society and diaspora. There have been many assertions about the potential of online discussion in the transformation of political and media systems in China. One of the theoretical constructs often referred to is the concept the public sphere. However, much of the research tends to use a qualitative approach; or if quantitative methods (for example, content analysis) are employed, they often focus on a particular Internet forum or a particular issue in a short period of time (Li, Qin & Kluver 2003; Qiu 2001; Tang & Shi 2001; Chen & Deng 2002; Guo 2002, Tai 2006). These studies have

drawn attention to the emergent phenomenon of active political communication on the Chinese Internet. However, they are remarkably limited in data selection as compared with the magnitude of the entire social act of political debate over the Internet. Therefore, the generalisation derived from these case-based inquiries is not sufficiently validated due to the methodological weakness of data limitation and subjective observation. Qualitative studies are necessary to determine why events occur (Black 2002), but they lack empirical reasoning and operationalised measurements to demonstrate what is happening on a large scale or over a long period of time. Many essential qualities of the e-publics (as the political Internet forum is defined in this paper), such as the size of contributing population, the patterns of participation, the quantity and sources of information circulated, and the degree of global interconnectivity, cannot be fully substantiated without a large-scale quantitative analysis, specifically an integrated research methodology which combines both quantitative and qualitative methods and complementarily explores the dynamic nature of computer-mediated public discussion.

Drawing from previous online political discussion studies, mainly newsgroup analyses (Schneider 1997; Bentivegna 1998; Smith & Kollock 1999; Jankowski & Selm 2000; Sack, 2000; Nonnecke & Preece 2003; Smith 2003; Whittaker Terveen, Hill, & Cherny 2003, Kelly, Fisher, & Smith 2005), the Complementary Explorative Date Analysis (CEDA) of online communities (Sudweeks & Simoff 1999), and various other studies, I introduce an integrated research methodology designed and developed specifically to analyse the Chinese Internet forums termed e-publics, discursive public spaces mediated by the new online conversational media.[2] This methodology, using commercial and self-developed online content analysis tools, features a series of computer-assisted data collection, quantitative analysis, and qualitative approaches. It investigates the online political discussion first from a wide range of mega data extracted to capture the general dynamics of the Chinese Forum. It then focuses on the selected categories of text data to delineate the linkage of global connectivity among forum networks and reveal the spectrum of discourse and pattern of deliberation. In essence, this study adopts an integrated research schema merging a series of quantitative and qualitative approaches to analyse Chinese Internet forums which are conceived as a networked sphere of micro e-publics of political discourse and participation. It aims at establishing a series of applicable methods to study the publicly accessed interactive content (postings in a threaded discussion format) of online political talk in the global network of the Chinese Internet.

Chinese Internet Forums and Quantitative Inquiry

Different from newsgroups and message boards popular on English-language Internet sites, the Chinese discussion boards have evolved into a unique platform for public debate and political activism in global Chinese communities. Technically, there are several types of BBS, such as the PC-based BBS, the web-based BBS, and the BBS on the UNIX system. In the Chinese Internet culture, the BBS has become a generic term, which refers to a threaded discussion platform accessible through the Web, Telnet or other means. A report released in 2006 suggests that 43.2% of the Chinese Internet population used BBSs and other discussion forums (CNNIC 2006).[3] All major Chinese portal sites and web sites host multiple Internet forums

with topics ranging from politics, finance, the military, sports, computers, health, books, travel, lifestyles, relations, hobbies and animals. The research presented here concentrates specifically on Internet forums discussing politics and current affairs.

The development of the Chinese Internet Forum has been closely linked to international political events and cyber activism in global Chinese societies. The Indonesian riots of 1998 were documented as a major case in the activation of an online gathering place for the Chinese diaspora (Tay 2000). The establishment of the Qiangguo Forum (Strong Nation Forum) in May 1999, the first Internet forum endorsed by the state media, marked the beginning of a new era of digital participation and political activism in China. Since then, the Chinese online forums, which originated from Taiwan and overseas Chinese communities in North America, entered the country's political-media spectrum as part of the "fourth medium" (Min 2005). The influence of the BBS and other new medium applications is evidenced in a series of international and domestic political events including the mass demonstration against Japan's bid to become a permanent member of the UN Security Council in 2005 (Tai 2006). There are four types of BBSs in China depending on their affiliations: the mainstream BBS (those hosted by mainstream media and governments), the commercial BBS (sponsored by commercial businesses), the citizens BBS (those have no governmental or commercial affiliation) and the campus BBS (those operated by universities).

Apart from online or offline Internet user surveys, research efforts on the content of Chinese forums are mainly qualitative in nature and they tend to examine discussion in a single forum (Li et al. 2003; Tang & Shi 2001; Guo 2002; Chen & Deng 2002; Wang 2003) or around a specific issue (Qiu 2001; Chen 2005; Xiao & Polumbaum 2006; Tai 2006). Most of them cover only a certain period of time. There is an obvious lack of comprehensive measurements and descriptions of the Chinese Forum in its general state that are derived from analyses of statistically representative data.

There are several concerns in the study of online political discourse:
- quantity of political communication over the Chinese Internet Forum;
- size of the participating population and the patterns of participation, both as active and silent participants;
- sources of information circulated in the networks of forums in relation to the mainstream media;
- interconnectivity of the global Chinese Forum network;
- spectrum of ideological argumentations;
- pattern of online deliberation.

None of these issues can be fully elaborated by forum-specific or issue-specific qualitative investigations alone. They should be scrutinised simultaneously by large-scale quantitative projects which cover numerous forums over an extended period of time.

There are many obstacles to conducting a large-scale quantitative inquiry of the Chinese Internet forums. First, Chinese forums, unlike newsgroups, are not based wholly on the Usenet. They are configured and operated in varied technological designs, thus it is extremely difficult and time-consuming to download content and archive data on a large scale and over an extended period of time. Second, there

is a large quantity of information (postings) generated by Chinese forums, which makes it impossible to process such data without proper computer-assisted tools. Third, there is a large number of Internet forums; each having its own interactive culture. It is a formidable task to develop standard criteria to measure activity and interactivity of these online publics. Finally, online discussion is ephemeral in nature; its content is updated so swiftly that studies of these talks, which require considerable labour, may not become meaningful as these talks disappear quickly on the Internet.

In addition to these obstacles, there is also a lack of recognised scopes of inquiry and units of analysis for quantitative investigations of online political discussion in general and Chinese Internet forums in particular. As the Chinese BBS represents a unique and evolving form of digital political activism in Chinese societies, there has been limited literature and established approaches in the West that can be readily adopted into a large-scale complementary exploration of the Chinese BBS. In the sections below, I will first examine the methodological concerns in the study of online political discussion, specifically in connection with the Chinese cyber interactivity. This will be followed by an elaboration of an integrated research schema which I developed during a four-year project of the Chinese Internet Forum.

Methodologies of Online Political Discussion Research

Online political discussion is a particular type of public commons created by the instantaneously and globally accessible conversational media of the Internet, bulletin board systems (BBS), Usenet, e-mail, chat rooms and WWW. Thanks to these online talk applications, a new medium and a locale in political communication has emerged: the medium of micro-publics and a computer-mediated informal zone of public debate.

The essential nature of this e-public is that it is an individually initiated network of activity and interactivity of political participation. Therefore, measurement of the content of online discussion boards has to take this feature into consideration. The online discussion content differs from the content of traditional mass media; it is individual, transitional and dynamic. In contrast, the traditional mass media content is static in the sense that it is generated by recognised institutions with a range of commonly accepted and comprehended codes of meanings and norms. Once it is created, its inference is relatively controlled (Früh in Rössler 2002). As a consequence, the content study of online discussion should not only study the text (postings) of discussion, but also, and more importantly, measure the activity and interactivity of the discussion. Essentially, within the online discussion process, the meaning of individual postings may not be as significant as the interactivity among these posters. As Patrick Rössler says:

> Going online with a web homepage or a posting in a newsgroup does not mean that the respective content contributes automatically to the online communication process. It is part of this process in so far as the content is published somewhere; but it is not a relevant part of this process unless it is found and consumed by at least one other user (Rössler 2002, 294).

Several methodologies have been developed to measure the activity and interactivity of the online content. They include:

(1) Newsgroup studies (Schneider 1997; Bentivegna 1998; Smith & Kollock 1999; Jankowski & Selm 2000; Sack 2000; Nonnecke & Preece 2003; Smith 2003; Whittaker, Terveen, Hill & Cherny 2003; Kelly, Fisher & Smith 2005). These studies provide a basic framework of quantifying the digital talk in threaded discussion format. Based on a large amount of online data, they measure various types of relationships among posts, authors, topics and perspectives and thus detail a quantitative description of online discourse of politics;

(2) The Complementary Explorative Date Analysis (CEDA) of online communities (Rafaeli & Sudweeks1998; Sudweeks & Simoff 1999). This method is proposed to analyse computer-mediated communication (threaded discussion) in online communities. The method collects two sets of data, one for quantitative analysis of interactivity (reference depth, width and height and so on), the other for qualitative analysis of leadership and communication management. A qualitative refinement is required to modify both analyses and incorporate quantitative and categorical models (Sudweeks & Simoff 1999);

(3) Network analysis. There is considerable literature concerning this scope of methodological applications, which covers such a wide range of academic fields as physics, computer science, bio-science, sociology, information science and CMC studies. The analysis attempts to map relationships between individual nodes and describe the structure within a networked system. Applying this graph methodology, a number of studies provide a variety of structural models of online information exchange, which include Krebs' model of purchasing patterns of political books from Amazon.com (Krebs 2004), Adamic and Grlance's (2005) model of political blogsphere, and Kelly, Fisher and Smith's (2005) model newsgroup discussion.

The network analysis is also used in investigating the global transaction of online information. By surveying the link structure of 4,000 web sites, Halavais (2000) discovered that the national boundaries are inscribed in the online networks. Park and Thelwall's (2003) inquiry of the hyperlink network of World Wide Web indicates a centrality structure of hyperlinks. An analysis by Naewon Kang and Junho H. Choi on the crossposting structure among newsgroups in Clarinet also confirms a center-peripheral pattern of information redistribution (Kang & Choi 1999).

All these methods, particularly the newsgroup studies, serve as reference for this project. However, they require a completely new design of integrated methodology so as to consider the specific nature of the Chinese Internet Forum, which is deeply embedded within the Chinese political and socio-cultural environment. There are basically four distinctive features which make the Chinese Forum different from a newsgroup discussion.

(1) Difference in online discussion format. The Chinese Internet Forum, although having varied designs of discussion format, features net names as posters as opposed to authors' names and email addresses in the Newsgroup. A poster usually has to register a net name before he/she participates in discussion. The concealing of posters' addresses protects the privacy of the discussants, thus may encourage more candid comments. However, this feature makes it more difficult to evaluate the relationship between posts and authors as an author can actually register numerous net names and post a large number of messages under various names.

Another design feature is that many Chinese forums provide a "hit count" on each post. This information serves to encourage authors to contribute quality and

popular postings. It is also very crucial for observers and scholars to have some sense of the readership (the extent of lurking) of the forums.

(2) Difference in forum management. The Chinese Forum is monitored and managed by a *banzhu* (a forum host), whose role is more or less similar to a discussion moderator. The *banzhu* sets rules, monitors etiquette, cleans spams, uploads or deletes posts, communicates with authors and sometimes exerts authority by blocking IPs of unwelcome posters. The *banzhus* style of moderation and management varies from one forum to another, ranging from a very participatory *banzhu* to a virtually invisible one. Consequently, this influences the specific discussion culture of individual forums.

(3) Difference in the political-media landscape: The Chinese online discussion has evolved in the country where the state media system still dominates. The government openly recognises the validity of information control and has installed a sophisticated "filtering system" to block sensitive information. It is therefore very important to set the scope of the investigation within the context of the state media domination in an attempt to detect the contested and negotiated space for political articulation in China's e-publics.

(4) Difference in global connectivity: Chinese Internet forums are a significant component of the global Chinese Internet. They are not hosted only in China, but also dispersed in many countries, mainly in North America. They form a transnational network of information exchange and political debate. This globalist nature contributes to the transnational intertextuality of the online environment. As Mitra and Cohen (1999) pointed out in a study of web communication, the meaning of a text is conditioned by how it is produced.

> *…a text produced within the "free speech" environment of America could be quite different in meaning from a text produced under a more oppressive system of government. However, to the reader, these WWW texts can appear next to each other and in their juxtaposition can produce meanings that neither of the texts could have created individually. The global linkages offered by the WWW produce the possibility of combining texts from different national sources and create different potentials of meaning* (Mitra & Cohen 1999, 190).

Because of all these differences, investigations of the Chinese discussion boards require a unique set of methodological approaches to address specific research concerns. In the following section, I elaborate on an integrated methodology adopted to my study of Chinese online discussions.

Integrated Methodology to Study Online Discussion in the Chinese Context

This study on the Chinese Forum represents a merging of quantitative and qualitative methodologies, which addresses the uniqueness of the Chinese Forum. Set within the conceptual framework of the public sphere and political discourse, it integrates several dimensions of publicness, interactivity, diversity, discourse and deliberation in computer-mediated public discussion. Originating from a preliminary small-scale study of eight Internet forums (Wu 2004), it draws on a variety of the above-mentioned online discussion and community investigations,

including quantitative inquiries of newsgroup political discussion, online network analyses, and political discourse. In addition, it also develops several new empirical approaches, such as how to measure the size and interest of peripheral participation (lurkers) in relation to the active participation, and how to calculate the flow pattern and global interconnectivity of discussion forums. Furthermore, new qualitative categorisations of an ideological spectrum, deliberative discourse and intertextuality are also developed to interpret online discussion in the Chinese context. Figure 2 below summarises this integrated methodology. Basic elements for analysis include all those that are openly exhibited by the forums: postings (thread postings and replies), net names, hits, and bytes.

In the "methods" column, methods used in each task are listed. The "data type" column indicates the types of data used in analysis. There are three types of data (postings) which are defined as the "primary data", "secondary data" and the "third level data". Detailed explanation about the definition and selection of these data are provided in a later section. The column of "specifics of analysis" describes a set of particular tasks each analysis intended to achieve through quantifying and measuring various variables. The "categories of inquiry" column illustrates more generalised theoretical issues concerned with the e-publics which are attempted to construe through inductive inference of the findings.

Figure 2: Integrated Methodology to Study Chinese Internet Forums

	Methods	Data type	Specifics of analysis	Categories of inquiry
Quantitative	Statistical sampling		Primary data; Secondary data	Data collection
	Computer-assisted numeric analysis	Primary data	Traffic	Publicness
			Population	Active participation
			Lurking ratio, Lurking index	Peripheral participation
			Reciprocity	Interactivity
	Computer-assisted network analysis	Primary data	Crossposting linkage	Global connectivity
	Combined content analysis: human coders and computer-assisted analysis	Secondary data	Information types	Diversity in information
			Information sources	
		Third level data	Discussion agendas	
Qualitative	Categorisation of dimensions	Secondary data	Posting categories	Diversity in opinion
			Ideological categories: Maoism, Liberalism, Commercialism, Nationalism, Globalism, Traditionalism	Ideological map
	Categorisation of dimensions	Issue-specific data selected from secondary data and third level data	Mode of speech, Justification, Respect, Constructive politics	Discourse quality
			Locality-based text, Locality-based perspective, Mixed perspective	Intertextuality

Data Selection

Selection of Sample Forums. There is a multitude of political Internet forums in the global Chinese Internet. However, the traffic and participation among them vary tremendously. There is a significant "power law distribution" evidenced in the Internet forums which means a very small number of forums attract much traffic and generate a large quantity of content, whereas a majority of forums have a very minimal content and participation. A few very active Internet forums can have over 300 posts per hour while most other forums generate only a few posts per day. For my project, a "purposeful" sampling strategy is adopted: 14 Internet forums of politics and current affairs were selected with consideration of popularity, geographic location, ownership and affiliation, political inclination and influence in the development of political Chinese BBS culture, rather than a random sampling. Seven are based in China with one hosted by the state media (People's Daily), three hosted by commercial media (China.com, Netease.com, cat898.com), one by a local government (EastDay, Shanghai) and two by universities (Tsinghua University and Peking University). The seven overseas forums are all hosted by diasporic Chinese in the United States except one associated with a university in Hong Kong.[4] This "purposeful sampling" is decided mainly from the consideration to choose the most influential forums which can be representative to reflect a spectrum of social groups, interests and political agendas, and be meaningful to identify the global connectivity under the filtering system of the Chinese authorities.

Selection of Sample Dates. Totalling 42 days between September 22, 2003 and September 21, 2004 were statistically selected as sample dates to represent a one-year content of the Internet discussion so as to reduce a huge workload of processing a massive number of postings. This selection design of sample dates and sample postings itself is a pioneer project which, commissioned by this author, was conducted by the Statistical Studies Institute (SSI) of China Communication University in 2003. The research group, led by Prof Ke Huixin, a leading scholar in China's statistical studies and communication, applied a range of statistical methods to finalise the design of 42 sample dates and 67 thread postings for each sample date. The following list summarises the method employed:

- Sampling population. The entire number of thread postings in the 11 selected current affairs forums inside China (provided by this author) between September 22, 2003 to September 21, 2004. It is composed of 11 sets of sub-sampling population. The sampling population did not cover overseas forums as most of them were not accessible from China.
- The construction of the sampling frame. It was designed through an analysis of statistics based on results of a close observation of two-week postings of the 11 forums between April 4-20, 2003.
- Sampling unit. The "week" serves as the primary sampling unit.
- Sampling size. The size of the sub-sampling population is 1900, the design of *deff* is assumed to be 2, and the total sampling population is 22,000.
- Accuracy: Under the 95% confidence level, the maximum margin of error is below 0.58%.
- Sampling method. A systematic sampling was adopted; six "combined weeks" were selected. All the postings in the selected "combined week" were down-

loaded to serve as the first stage sampling for the estimation of population parameters. Another systematic sampling method was employed to select the second stage sampling from the first-stage sampling. The second-stage sampling was used for in-depth content analysis.[5]

- Primary data: All postings (about 450,000) in the 42 days formed the "primary data" for the quantitatively computer-assisted analysis without human raters to review the content of postings.
- Secondary data: 67 thread postings in each forum in each sample date were statistically selected in the same SSI project as the "secondary data" to represent a daily provision of information and discussion topics for combined content analysis with human raters to code the content of individual messages.
- Third level data: The most replied thread post (MRP) and most hit post (MHP) were selected from the "primary data" in each forum and in each sample date as representative of the most commented posting and most read posting for further textual analysis.[6]

Computer-assisted Quantitative Analysis of Online Activities

This analysis examines (and is based on) the primary data the discussion activity from the following set of dimensions: a) traffic; b) population; c) reciprocity; d) active participation; and, e) periphery participation. The traffic (measured by the total number of postings), the refreshing speed of a post, and the mean length of a post, indicate essentially the quantity of information exchanged over an online forum. The population, measured by the number of net names, represents to some extent the size of each public forum in terms of visible participants, although the number of net names does not equal that of real participants since a user may have multiple nicknames.

The reciprocity is defined as the interactive relationship in the Internet discursive environment. It is examined by two sets of measuring categories: the first, message reciprocity, is measured by the analysis of postings, which analyses how conversational topics are initiated, replied to, and discussed by calculating the percentages of the lead postings (the initiated thread) and replies in relation to the total number of postings and the ratio of lead postings and replies (the thread depth). The second, author reciprocity, is measured by the analysis of the posting track of authors (net names). It probes how authors respond to and are followed by other authors in discussions.

For the investigation of participation in the digital public, I argue that participation in the online forum public consists of two modes. One is visible and active participation, which involves posters who contribute to the content (postings) of the Internet public. However, the online public is also characterised by having a large number of lurkers who act as the audience of the discussion content. The phenomenon of lurkers in the Chinese Internet Forum can be considered as periphery participation of the digital public.[7] The active participation analysis provides an examination of the nature of digital political talk, and to what degree the online discussion is considered equal and participatory. It uses the Gini coefficient of postings/posters to indicate the pattern of active participation. The participation distribution is calculated by measuring the concentration of authorship (the relation between the percentage of total messages per author and the rank of authors in terms
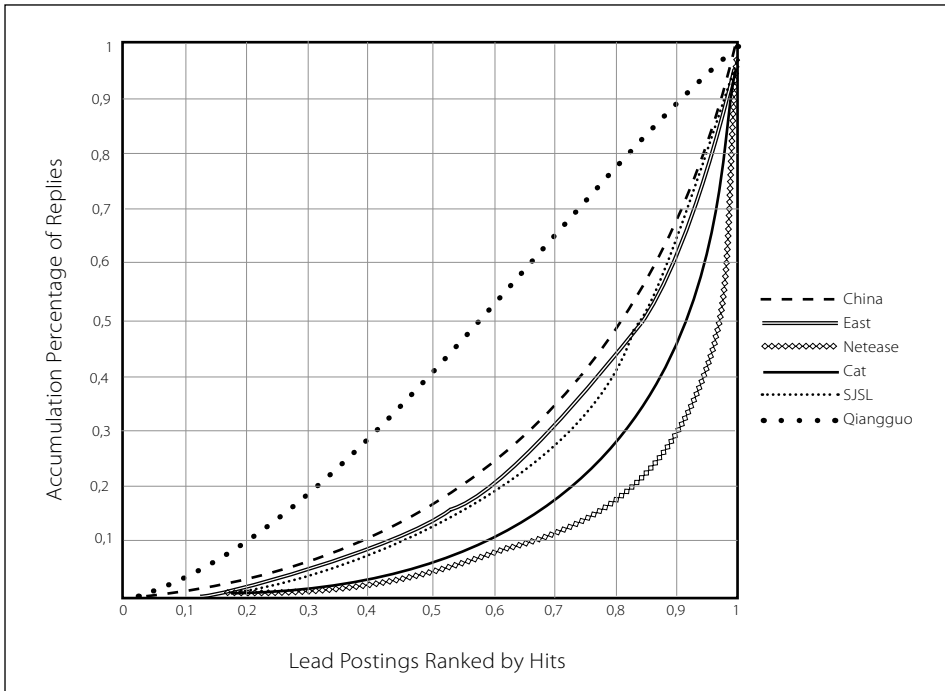
of frequency of messages). This figure is reflected by the Gini coefficient (the closer the Gini coefficient is toward 0, the more it indicates an equal distribution).[8]

Analysis of Periphery Participation. The analysis of the periphery participation looks at the ratio of hits and replies of postings and the relationship between replies per post and hits per post. It adopts two sets of measurement: lurking rate and lurking interest. Lurking rate defines the scale of lurking practice by measuring the relationship of hits and replies in thread postings. Lurking interest reflects the interest of lurkers in relation to that of active participants in the respective Internet forums using the Gini coefficient.

(1) Lurking Rate. The findings demonstrate that for each thread post, if it has 100 hits, about 86-99 hits are lurking; or in other words, every 100 hits of a thread message generate 1-3 replies on average. This shows that lurking is a normal behaviour of online participation. This finding may also offer some estimate about the scale of invisible participation of a forum in relation to the visible participation (posters and repliers). The number of lurkers in any given digital public is no more than the 100 times of the number of reply messages if it is assumed that one hit equals one reader. As most readers make more than one hit, the actual number of lurkers is much smaller.

(2) Lurking Interest. The lurking interest index indicates whether the interest of peripheral participants correlates with the interest of active participants by calculating whether the posts with a higher number of hits also generate a higher number of replies. The findings are indicated by a lurking interest index chart; see Figure 3.[9]

Figure 3: Lurking and Reply Interest Index

The chart describes the distribution of all the replies among hits (the relation between the accumulative percentage of replies and the total lead postings ranked by the number of hits). This figure is reflected by the Gini coefficient (the closer the Gini coefficient is toward 1, the more it indicates a positive correlation). If the interest distribution is even (meaning the number of replies is related more negatively with the number of hits), the value of the Gini coefficient is 0. If the interest distribution is highly skewed, which means that the number of replies correlates more positively to the number of hits, the value would go toward 1.

The findings indicate that in most forums, the interest of lurkers corresponds positively to the interest of repliers, meaning there is a sort of common interest among the active and peripheral participants. However, a closer examination of the data reveals that the difference in this index may be more to do with the display format of a forum than a disparity of interest among posters and lurkers.

Network Analysis of Global Connectivity

Global Network of Forums. There is a multiplicity of digital micro-publics which are frequently and globally intertwined. Through crosspostings of individual netters, there emerges an information flow pattern that has never been seen before – a wildfire-like spread of news over the global networks of BBSs. This study explores the issue by tracking and calculating the recurrence or multi-occurrence of postings (primary data) in various forums.

The global connectivity chart (Figure 4) illustrates the network connections of these forums based on computer-assisted subject line search (equal and contain the same subject line) among all the thread postings. Altogether, there are 14,034 crossposting connections among all the forums. The distribution of these connections exhibits a sort of power law structure as well.

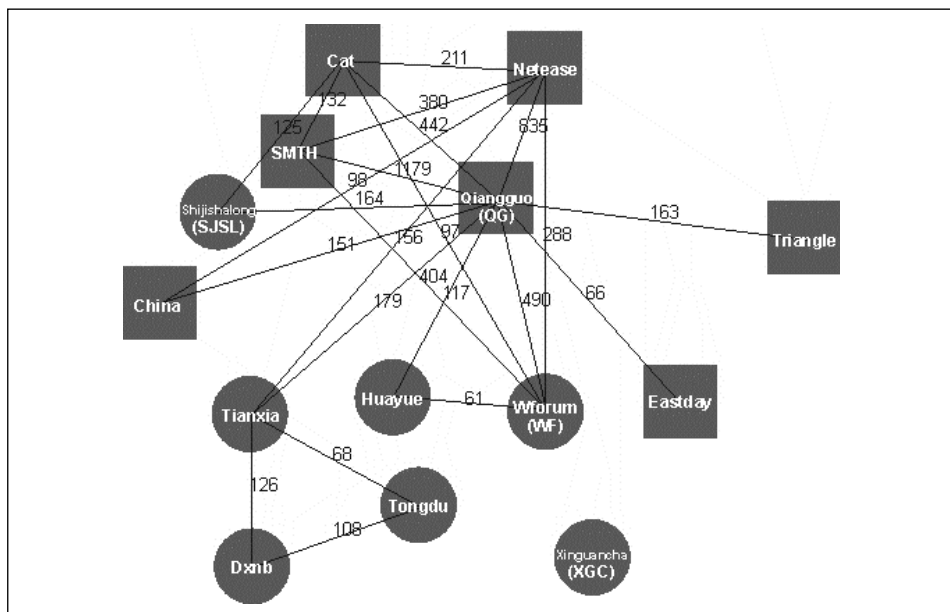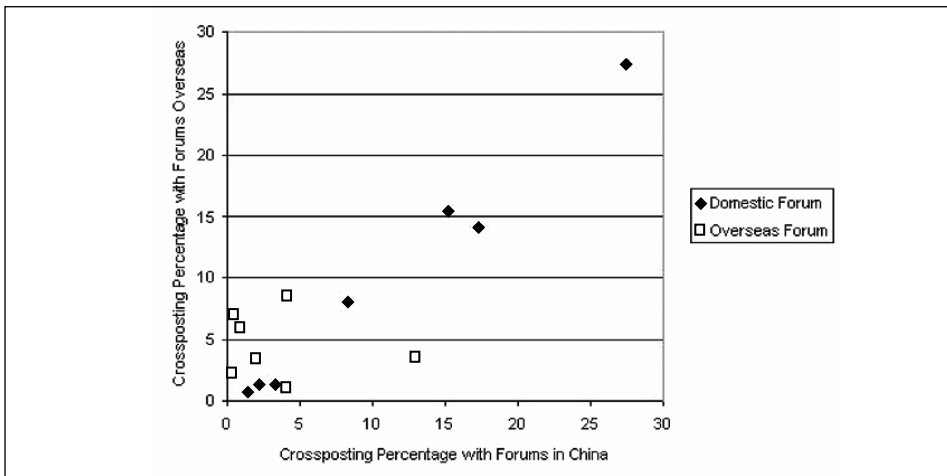Figure 4: Crossposting Connections Among Forums

Figure 5 demonstrates a posting flow pattern between domestic and overseas forums. Axis X indicates the crossposting percentage with the domestic forums, while Axis Y indicates the crossposting percentage with overseas forums. The figure shows that domestic forums have a similar crossposting pattern between the domestic and overseas posting flow. However, for overseas forums, they are connecting mostly with either domestic forums or with overseas forums. This finding signifies that overseas forums are divided into two camps, one has a closer tie of posting exchange with domestic forums, and the other has a weak tie or even no tie with forums in China. The former contains those with political inclinations of nationalism or Maoism. And the latter includes those which openly advocate the overthrow of the Chinese regime. Forums with modest democratic position are somewhere in the middle.

Figure 5: Flow Pattern of Forums

Combined Content Analysis of Information Re-distribution

This analysis is conducted based on the "secondary data" – 67 postings in each forum in 42 days and the "third level data." It concentrates on how forums are functional as an alternative channel of information in relation to the state media system in China. There are four parameters for analysis: forum-generated information, forum-disseminated information, mainstream media information in the forum sphere, and topic and agenda of the BBS sphere.

Key units of analysis include: original postings, crosspostings, most-hit postings (MHPs) and most-replied postings (MRPs). Throughout the entire process of qualitative coding, only one human coder was employed to categorise postings with the aid of an in-house software, which was designed by the same coder based on the coding specifics formulated by the project supervisor.

Forum Agenda. MHPs and MRPs in each forum and in each sampling date are defined as indicators of forum agendas. In both categories, original postings (written by forum authors) account for a higher percentage. Among the total number of MRPs studied, nearly two-thirds of them are original postings. For MHPs, crosspost-

ings account for 40%, indicating that crosspostings tend to provide information for reading, whilst original postings are more stimulating to prompt a discussion.

Sources of Crosspostings. Crosspostings reflect a voluntary redistribution of news from other sources, mostly mainstream media, by BBS participants. This analysis, intended to assess the traditional media's influence on online political talk, analyses the source of crosspostings by two categories: media types—traditional media, Internet-only media, and BBSs; and geographic areas where these media organisations and operations are located.

Combined content analysis on posting categories. This analysis is also based on the "secondary data." It attempts to define how much online discussants are devoted to opinion assertion and call for social activism.

Qualitative Analysis of Ideological Spectrum

Ideology-scape in the Chinese Forum. One of critical views contending the potential of Chinese online political debate is to what extend the freedom of expression is permitted. Contrary to a commonly accepted view of the monolithic ideology of communism in the country, the Chinese ideology-scape reflects a hybrid ideology of the current state establishment, embodying such inherently contradictory elements as Maoism, socialism, capitalism, modernism and globalism (Liu 2004). This Chinese context of political views and ideology is conceptualised in this study as a site of constant process of discursive constructs, re-constructs and negotiations within a pool of ideological resources from Maoism to liberalism, commercialism, nationalism, globalism to Confucianism.

Ideological Compass. A set of five scales – Political, Economic, Social, Cultural and International Relation – is used to illustrate the dimensional change of discursive formation in China from the rigid Communist era of Mao Zedong (1949-1976), to the reform-era of Deng Xiaoping (1978-1997) which introduced a free market economy into the country, and the post-Deng era (1997-present). This set of scales also attempts to illustrate the discursive contestation and reconciliation in the BBS sphere.

The official ideology is a commonly recognised paradigmatic set of ideas and viewpoints consistently propagated by the mainstream media of mass communication in China of the respective time period. It is difficult to identify the precise "fixing" of meanings and semiotic systems connected with the official discourse as it, on one hand, is subjected to the successive transformation of social relations and historical developments in Chinese society. On the other hand, it is ascribed to a milieu where ideological struggles between ideas and discourses are constantly forged and social resources of making sense are formed, contested, institutionalised and re-mobilised.

Future work involves establishing a set of coding criteria to analyse the online political debate in the Internet forums, with a particular concern about whether the threaded pattern of interaction in an instantaneous and global environment would introduce anything new into the pattern of political debate and deliberation.

Conclusion

This research project is primarily concerned with the development of an integrated methodology, exclusively based on online data, to examine digital politi-

cal talk in the Chinese Internet. It signifies an effort to establish a new method of content analysis to scrutinise the interactive content of online discussion. This user-generated content differs considerably from that found in traditional media. Compared with the "static" nature of the traditional media, the online content is dynamic; it is individually initiated, network connected, and globally transitional. The integrated research schema introduced in this article reflects one of the initial attempts to quantify and measure the dynamics of online discussion, specifically in the Chinese cyber publics. It merges quantitative and qualitative methodologies to comprehensively inquire on the online debate from several dimensions of the public sphere, such as publicness, interactivity, diversity, discourse and deliberation. Several novel empirical approaches have been invented to discover some of the distinguished attributes of the online discussion, such as how to measure the size and interests of peripheral participation (lurkers) in relation to the active participation, how to map the flow trajectory and global interconnectivity of discussion forums, and so on. Furthermore, new qualitative categorisations of an ideological spectrum, deliberative discourse and intertextuality are also developed to provide a more in-depth interpretation of online publics in the Chinese context. All these approaches enable not only a deeper comprehension of the dynamics in Chinese Internet which engages 210 million netizens, but also render a frame of reference to studies of online political discussion in other countries.

It is worth noting that the integrated research methodology adopted in this study offers only one of the possible approaches to analysing interactive content online. It may not be commonly applicable without adaptations to other domains of CMC research or in other socio-cultural contexts. Limitations in this schema are related to two aspects. On one hand, understanding of the emergent phenomenon of online commons is still at an embryonic stage. There is a deficiency in qualitative reasoning and categorical models to define and conceptualise them within the existing dimensions of communication studies. On the other hand, there is a lack of useful computer-assisted tools which enable extensive and efficient quantitative analyses. This serves as both a challenge and an opportunity for Internet researchers.

Acknowlededgement

Notes:

1. This paper focuses attention on a specific domain of computer-mediated online political activism. The Internet forum studies, particularly the Internet forum methodologies reviewed in this paper are largely concerned with this narrow area of the field rather than the Chinese Internet in general.

2. For detailed conceptualization of the E-publics and Chinese Internet forums, please refer to several conference papers by this author (Wu 2005, Wu 2006a, 2006b).

3. Another survey conducted by Guo Liang's group indicates that 23.8% of netizens in China use BBSs (Guo 2005).

4. Originally 18 Internet forums were selected. However, some of the selected Internet forums like Fazhan Forum (Xinhuanet), Zaobao Forum, Dajia Forum (Chinesenewsnet.com) were not included due to technical problems in downloading the content.

5. For a detailed account about selection of sampling dates and postings, see Ke et al. 2004.

6. Eight forums did not carry "hit count" information and thus only six forums have the third level data of MHP posts.

7. Although lurkers are said to make up over 90% of online communities (Katz 1998; Mason 1999), their role in the online public is difficult to define. Some researchers describe lurkers as "free-riders," a sort of non-contributing, resource-taking people (Kollock et al. 1996). Others consider them "silent participants," whose viewing act is a valid form of participation (Nonnecke et al. 2003). I argue that lurkers, to a large extent, perform the similar role as readers/viewers to the traditional print and TV media. The size of the lurkers compared to an online discussion group defines the degree of publicness of the digital community.

8. This approach of using Gini coefficient and Lorenz curve to measure participation in online discussion was also first adopted by Schneider (1997). He discovered that the distribution of authorship in the newsgroup discussion was highly skewed (Gini coefficient: 0.44). About 0.05 percent of the authors (among 3,000 authors) contributed more than 40 percent of the 46,000 postings. Or in another measurement, about 5 percent of the posters supplied almost 80 percent of the messages. This relation between authors and postings is graphically represented in a Lorenz curve which relates to the percentage of total messages per author while ranking the authors in terms of frequency of messages.

9. The chart covers only six forums as the rest eight forums do not provide hit count of postings. It is common that a forum with high traffic tends to disply hit count of each post.

References:

Adamic, Lada A. and Natalie Glance. The Political Blogosphere and the 2004 U.S. Election: Divided They Blog. BlogPulse. <http://www.blogpulse.com/papers/2005/AdamicGlanceBlogWWW.pdf>

Bentivegna, Sara. 1998. Talking Politics on the Net. Research Paper. Cambridge, MA: The Joan Shorenstein Center, John F. Kennedy School of Government, Harvard University.

Black, Thomas. R. 2002. *Understanding Social Science Research*. London: Sage.

Chen, Hongmei. 2005. *Wangluo chuanbo yu gongzhong biaoda* [Internet Communication and Public Expression]. Lanzhou: Gansu People's Publishing House.

Chen, Tongxu, and Lifeng Deng. 2002. *BBS yiti de xingcheng yu shuaibian* [The formation and decay of BBS agenda]. *Journalism and Communication Studies*, 1.

CNNIC China Internet Network Information Center. 2006. The 18th Survey Report, Statistical Reports on the Internet Development in China. <http://www.cnnic.org.cn/en/index/0O/02/index.htm>

Dahlberg, Lincoln. 2001a. The Internet and Democratic Discourse: Exploring the Prospects of Online Deliberative Forums Extending the Public Sphere. *Information, Communication & Society* 4, 4, 615-634.

Dahlberg, Lincoln. 2001b. Computer-Mediated Communication and the Public Sphere: A Critical Analysis. *Journal of Computer-Mediated Communication* 7, 1.

Guo, Liang. 2002. *Qiangguo luntan: 911 kongbu xiji hou de 24 xiaoshi* [Qiangguo Luntan: 24 hours after 911 attack]. *Journalism and Communication Studies*, 4.

Guo, Liang. 2005. *2005 nian Beijing, Shanghai, Guangzhou, Chengdu, Changsha hulian wang shiyong zhuangkuang ji yixiang diaocha baogao* [2005 Survey Report on Internet Use and Impact in Beijing, Shanghai Guangzhou, Chengdu and Changsha.]

Habermas, Jürgen. 1996. *Between Facts and Norms: Contributions to a Discourse of Law and Democracy*. Cambridge, MA: MIT Press.

Hacker, Kenneth L. and Jan van Dijk. 2000. *Digital Democracy: Issues of Theory and Practice*. London: Sage.

Halavais, Alexander. 2000. National Borders on the World Wide Web. *New Media & Society* 2, 1, 7-28.

Jankowski, Nicholas and Martine van Selm. 2000. The Promise and Practice of Public Debate in Cyberspace. In K. L. Hacker and J. van Dijk (eds.), *Digital Democracy: Issues of Theory and Practice*, 149-65. London: Sage.

Kang, Naewon and Junho H. Choi. 1999. Structural Implications of the Crossposting Network of International News in Cyberspace. *Communication Research* 26, 4, 454-81.

Katz, Jon. 1998. Luring the Lurkers. Slashdot. <http://slashdot.org/features/98/12/28/1745252.shtml>

Ke, Huixin, Ke Huang, and Tingting Xie. 2004. *Zhongwen wangluo luntan yanjiu de chouyang sheji* [Sample design of Chinese BBS studies]. Paper presented at the Annual conference of China's Computer-Mediated Communication Association, Nanjing.

Kelly, John, Danyel Fisher, and Marc Smith. 2005. Debate, Division and Diversity: Political Discourse Networks in USENET Newsgroups. Paper presented at the Online Deliberation Conference 2005, Stanford University.

Kollock, Peter and Marc Smith. 1994. Managing the Virtual Commons: Cooperation and Conflict in Computer Communities. <http://www.sscnet.ucla.edu/soc/csoc/papers/virtcomm>

Krebs, Valdis. 1999. The Social Life of Books: Visualizing Communities of Interest via Purchase Patterns on the WWW. Orgnet.Com. <http://www.orgnet.com/booknet.html>

Li, Xiguang, Xuan Qin, and Randolph Kluver. 2003. Who is Setting the Chinese Agenda? The Impact of Online Chatrooms on Party Presses in China. In K. C. Ho, R. Kluver and C.C. Yang (eds.), *Asia. com*, 143-58. London: Routledge.

Liu, Kang. 2004. *Globalization and Cultural Trends in China*. Honolulu: University of Hawaii Press.

Mason, Bruce. 1999. Issues in Virtual Ethnography. In K. Buckner (ed.), *Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and Connected Communities, Proceedings of Esprit i3 Workshop on Ethnographic Studies*, 61-69. Edinburgh: Queen Margaret College.

Min, Dahong. 2005. *Zhongguo wangluo chuanbo xueke fazhan baogao* [Report on the development of China's CMC research]. Paper presented at the Annual conference of China's Computer-Mediated Communication Association, Nanjing.

Mitra, Ananda and Elisia Cohen. 1999. Analyzing the Web: Directions and Challenges. In S. Jones (ed.), *Doing Internet Research* , 179-202. Thousand Oaks, CA: Sage.

Nonnecke, Blair and Jenny Preece. 2003. Silent Participants: Getting to Know Lurkers Better. In C. Lueg and F. Danyel (eds.), *From Usenet to CoWebs*, 110-32. London: Springer.

Park, Han Woo and Mike Thelwall. 2003. Hyperlink Analyses of the World Wide Web: A Review. *Journal of Computer-Mediated Communication* 8, 4.

Qiu, Jack Linchuan. 2001. Chinese Nationalism on the Net: An Odd Myth with Normalcy. Paper presented at the 87th Annual Conference of the National Communication Association, Atlanta.

Rafaeli, Sheizaf and Fay Sudweeks. 1998. Interactivity on the Nets. In F. Sudweeks, M. McLaughlin and S. Rafaeli (eds.), *Network & Netplay: Virtual Groups on the Internet*, 173-190. Cambridge, MA: MIT Press.

Rössler, Patrick. 2002. Content Analysis in Online Communication: A Challenge for Traditional Methodology. In B. Batinic, U.-D. Reips, and M. Bosnjak (eds.), *Online Social Sciences*, 291-308. Seattle, WA: Hogrefe & Huber Publishers.

Sack, Warren. 2000. Conversation Map. Paper presented at the International Conference on Intelligent User Interfaces, New Orleans.

Schneider, Steve. 1997. Expanding the Public Sphere through Computer-Mediated Communication: Political Discussion about Abortion in a Usenet Newsgroup. Unpublished PhD dissertation, MIT.

Smith, Marc A. 1999. Invisible Crowds in Cyberspace. In M. A. Smith and P. Kollock (eds.), *Communities in Cyberspace*, 195-219. London: Routledge.

Smith, Marc A. 2003. Measures and Maps of Usenet. In C. Lueg and F. Danyel (eds.), *From Usenet to CoWebs*, 47-78. London: Springer.

Smith, Marc A. and Peter Kollock, eds. 1999. *Communities in Cyberspace*. London: Routledge.

Stromer-Galley, Jennifer. 2002. Motives for Political Talk Online: Implications for Political Conversation and Deliberation. Paper presented at the Annual conference of the National Communication Association, New Orleans.

Sudweeks, Fay and Simeon J. Simoff. 1999. Complementary Explorative Data Analysis. In S. Jones (ed.), *Doing Internet Research*, 29-56. Thousand Oaks: Sage.

Tai, Zixue. 2006. *The Internet in China: Cyberspace and Civil Society*. London: Routledge.

Tang, Dayong and Zhe Shi. 2001. *Xuni shequn yihuo gonggong lingyu* [Virtual community or public sphere]. In X. Deng and X. Li (eds.), *Internet Communication and News Media*, 393-411. Beijing: Beijing Broadcasting Institute Press.

Tay, Elaine. 2000. Global Chinese Fraternity and the Indonesian Riots of May 1998: The Online Gathering of Dispersed Chinese. <http://www.murdoch.edu.au/intersections/issue4/tay.html>

Wang, Gan. 2003. Yaolan luntan: Wangshang de muqin shequ [Cradle forum: Mothers' community online]. Unpublished post doctoral research report, Sociology Institute, China Academy of Social Sciences.

Whittaker, Steve, Loren Terveen, Will Hill, and Lynn Cherny. 2003. The Dynamics of Mass Interaction. In C. Lueg and D. Fisher (eds.), *From Usenet to Cowebs*, 79-91. London: Springer.

Wu, Mei. 2004. *Zhongwen wanglu luntan de zixun liudong* [Information flow in the Chinese net forums]. *Twenty-First Century* 81, 1, 113-23.

Wu, Mei. 2005. Informal Public Sphere: Theorizing & Measuring the Dynamics of Political BBSs in China. Paper presented at The 14th AMIC Annual Conference Media and Society in Asia: Transformations and Transitions, Beijing.

Wu, Mei. 2006a. E-public: Political Participation in Chinese Internet Discussion. Paper presented at The International Conference on Media and Democracy in the Knowledge Society, Seoul.

Wu, Mei. 2006b. A Quasi-News Medium: How Information Travels in Chinese Discussion Forums. Paper presented at The Digital Communication and Social Transformation: Chinese Societies and Beyond, Hong Kong.

Xiao, Li and Judy Polumbaum. 2006. News and Ideological Exegesis in Chinese Online Media: A Case Study of Crime Coverage and Reader Discussion on Two Commercial Portals. *Asian Journal of Communication* 16, 1, 40-58.