Framing Social Movement Identity with Cyber-Artifacts: A Case Study of the International Falun Gong Movement

Yi-Da Chen, Ahmed Abbasi, and Hsinchun Chen

Abstract Framing a collective identity is an essential process in a social movement. The identity defines the orientation of public actions to take and establishes an informal interaction network for circulating important information and material resources. While domestic social movements emphasize the coherence of identity in alliance, global or cyber-activism is now flexible in its collective identity given the rise of the Internet. A campaign may include diverse social movement organizations (SMOs) with different social agendas. This flexible identity framing encourages personal involvement in direct action. On the other hand, it may damage solidarity within SMOs and make campaigns difficult to control. To assess the sustainability of an SMO, it is important to understand its collective identity and the social codes embedded within its associated cyber-societies and cyber-artifacts. In this study, we took a cyber-archeology approach and used the international Falun Gong (FLG) movement as a case study to investigate this identity-framing issue. We employed social network analysis and Writeprint to analyze FLG's cyber-artifacts from the perspectives of links, web content, and forum content. In the link analysis, FLG's websites linked closely to Chinese democracy and human rights SMOs, reflecting FLG's historical conflicts with the Chinese government after the official ban in 1999. In the web content analysis, we used Writeprint to analyze the writings of Hongzhi Li and of his editors, and found that Hongzhi Li's writings center around the ideological teaching of Falun Dafa while the editors post specific programs to realize Li's teaching. In the forum content analysis, FLG comprehensively organizes several different concepts on a continuum: from FLG ideology to life philosophy and mysterious phenomena, and from mysterious phenomena to anti-Chinese Communist

Yi-Da Chen, Hsinchun Chen*
Artificial Intelligence Lab, Department of Management Information Systems, University of Arizona, Tucson, AZ 85721, USA
e-mail: ydchenb@email.arizona.edu; *hchen@eller.arizona.edu

Ahmed Abbasi
Sheldon B. Lubar School Business, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
e-mail: abbasi@uwm.edu

1

Party and persecution by conceptualizing the Chinese government as the Evil. By deploying those cyber-artifacts, FLG seamlessly connects different ideologies and establishes its identity as a Qi-Gong, religious, and activist group.

1 Introduction

The Internet nowadays is not merely a digital platform for exchanging data or information but is also a communication channel for individuals to share and promulgate their beliefs and ideas. People discuss topics of interest in forums or create their own blogs for posting experiences and opinions. Through hyperlinks, individuals can link their own posts to other web pages for citing the comments with which they agree or oppose. Gradually, "ideological networks" are formed on the Internet in which websites with similar ideas are connected together via hyperlinks. One may find a number of relevant websites of her/his interests by just following the hyperlinks of a few seed websites.

The Internet has also changed how social movement organizations (SMOs) advocate their ideology and mobilize their resources. They are no longer restricted by time or space. Their websites hold permanent campaigns appealing to a global audience [6, 7, 17]. The low cost of communication makes them more easily align together for large public actions [6, 7, 25]. For example, a demonstration against the war in Iraq in Washington, D.C., 2003, gathered around 100,000 people with different protest positions [7]. Furthermore, activist groups can increase individual participation by hyperlinking an inclusive ideological network which provides multiple entry points for potential supporters to join in [6]. The anti-Microsoft network, for instance, involves a great diversity of interest groups, including corporations, consumer protection organizations, and labor alliances [6].

In the traditional social movement theory, framing a coherent identity is a crucial process for an SMO to establish itself in social movements [19]. The collective identity helps participants to develop a trust relationship with their members and creates an informal interaction network for circulating important information and material resources [29]. The failure to achieve common framing can create tension and fragmentation within coordinated SMOs [7]. Global or cyberactivism, conversely, is flexible in identity framing [7]. Its advocate network may contain diverse social justice agendas, rich in personal appeal but thin in ideology [6, 7]. The ease of individuals joining and leaving a given network damages solidarity in SMOs and makes campaigns difficult to control [6]. To assess the sustainability of an SMO in the Internet era, it is important to understand how the SMO constructs identity and social codes within its associated cyber-societies [7].

In this study, we took a cyber-archaeology approach to investigate the framing of collective identity on the Internet and used the International Falun Gong (FLG)

movement as a case study. The cyber-archaeology approach adapts methods from archaeology and anthropology to analyze a community's cultural cyber-artifacts [16, 27]. In applying the cyber-archaeology approach to the study of SMOs, the approach involves (1) the identification of their associated cyber-societies, (2) collection of cyber-artifacts with automated procedures, and (3) analysis of cyber-artifacts from the perspective of the social movement theory. The FLG was chosen as our case study because it involves various identities, including Qi-Gong exercises, new religion, and activism [28, 30, 34, 21], and heavily uses the Internet as a vehicle for information dissemination [5]. Our goal was to investigate how FLG comprehensively maintains these three identities simultaneously within its cyber-societies.

The remainder of the study is organized as follows. We first review social movement theory and cyber-archaeology, and introduce two analytical tools used for examining cyber-artifacts: social network analysis (SNA) and Writeprint. Then we describe our research design which was based on the cyber-archaeology research framework, covering link, web content, and forum content analyses. Finally, we present our results and conclusions.

2 Literature Review

2.1 Social Movement Theory

Social movements are often recognized as irrational and disorganized activities working against social injustices. Unlike official political participation or lobbying, activists organize a broad range of individuals and seek to build a radically new social order through public protest [11]. However, the study of social movements reveals that they have much deeper organizational and psychological foundations than would appear. Activists do not just blindly deploy protest actions: they calculate the costs and benefits of the actions with present resources before initiation [29]. People are not irrationally supporting a social movement: they seek an identity and share a sense of belonging through their participation [29].

Collective action and resource mobilization are two intellectual currents which dominated the early development of social movement theory [19, 29]. Collective action is based on the idea that social movements are triggered by an overly rapid social transformation. In this school of thought, a society consisted of several balanced sub-systems and a movement reflective of the failure of social institutions or control mechanisms to rebalance them after a dramatic change [29]. In such a moment, the existing norms no longer provide a sufficient rationale for social action, and a group of individuals sees the situation as injustice and reacts to it by developing shared beliefs and new norms for behavior. The American Civil War and Civil Rights Movement are significant examples illustrating this point of view [29]. The framing of collective identity is an essential process for collective action [19, 29, 20]. For participants, the identity helps establish a trust relationship with others in the same group and excludes those whom they oppose [19, 29]. For the movement,

it defines the orientation of public action and constrains where the actions will take place [19, 20].

Resource mobilization examines the strategic components in social movements. It is based on two main assertions: (1) movement activities are not spontaneous and disorganized and (2) their participants are not irrational [24]. In other words, social movements are meaningful actions with specific purposes [29]. In this school of thought, the movements involve so-called "social movement entrepreneurs" who call public attention to problems and who recruit and mobilize supporters for action [19]. Public protests are derived from a careful calculation of costs and benefits as measured against present resources [29]. To achieve their goals, activists need to control the costs of their actions, organize discontent to reach social consensus, and create solidarity networks for circulating information and resources. From this point of view, social movements are an extension of formal political actions to pursue social interests [29].

2.2 Social Movement Organizations and the Internet

SMOs are non-governmental organizations which seek to realize the ideology of a social movement with clear goals [29]. Their coordination in a movement is suggested following a SPIN model: (1) Segmented, (2) Polycephalous, (3) Integrated, and (4) Networked [6, 29, 14]. According to the model, a social movement is composed of many SMOs which continuously die or rise [29]. It has multiple and sometimes competing leaders or influential centers [6, 14]. Those SMOs form a loose but integrated network with overlapping membership and shared ideas [6, 29, 14].

The Internet has changed the ways SMOs operate [6, 7, 25]. One advantage that the Internet brought to SMOs is the reduction of communication costs [6, 25]. In the pre-Internet era, SMOs relied on informal interaction networks of their supporters to circulate information and mobilize participants for actions [29]. They now use Computer-Mediated Communication (CMC), including e-mails and forums, to promote their ideology, recruit new members, and coordinate campaigns. Another significant change is that activists can hold permanent campaigns via their websites, thereby appealing to a global audience and transforming domestic social movements into global or cyber-activism [6, 7, 23]. Compared to traditional activism, global activism depends on the availability of technology networks in order to expand, involves diverse social justice agendas and multiple issues, and promotes personal involvement in direct action [7]. To reflect this transformation, Gerlach revised his SPIN model from "polycephalous" to "polycentric," meaning that global activist networks have many centers or hubs for supporters to join or leave, and are less likely to be managed by permanent leaders [6]. Therefore, they are thin in ideology but rich in personal identity and lifestyle narratives [6]. To assess the sustainability and quality of an SMO in the global movement, it is crucial to identify the social codes and values embedded in its CMC cyber-artifacts [7].

2.3 Cyber-Society Archaeology

Researchers have recognized the potential for conducting research grounded in social movement theory through analysis of SMO cyber-societies [10, 33]. In a cyber-society, all cultural artifacts of the community are recorded and preserved, and can be recovered by the social scientist if desired. The discourse among community members is rich, robust, filled with social cues, and highly revealing of individual perspectives. Traditional studies in archaeology and anthropology face the difficult task of determining the time associated with collected artifacts, but in an online environment, time is often measured and recorded precisely for each cyber-artifact. The accessibility of time information allows for analysis of the exact evolution of interactions and the production of cultural artifacts, which can reveal a great deal about social behavior.

Social movement research on cyber-societies requires adaptation of the traditional methods of studying conventional societies to virtual communities. Cyber-society researchers advocate using methods adapted from archaeology and anthropology to study a community's cultural cyber-artifacts [16, 27]. The application of methods used in archaeology to cyber-societies has been termed cyber-archaeology, and focuses on the analysis of CMC cyber-artifacts [15]. In particular, researchers have proposed the use of webometric methodologies in cyber-archaeology studies for the collection of cyber-artifacts [8]. Forms of analysis associated with webometrics include web content, link, technology, and usage analysis [4, 9]. Web content analysis focuses on the CMC communications and textual information provided on a web page. Link analysis considers the network of websites linked to and from a particular website. The sophistication of the platform for CMC communication is studied through web technology analysis. Web usage analysis focuses on user behavior and activity logs.

A systematic cyber-archaeology framework is proposed here to guide social movement research on SMO cyber-societies. The framework shown in Fig. 1 leverages traditional archaeology and anthropology research approaches and perspectives, as suggested by cyber-society researchers [16, 27]. CMC cyber-artifacts are collected, categorized, and visualized using webometric approaches [8]. The proposed framework for cyber-archaeology has three phases. In phase one, social movement research design, social researchers identify SMOs and cyber-societies of interest, and target cyber-artifacts for collection and analysis. Phase two consists of the automated collection and classification of cyber-artifacts across one or many identified cyber-societies. The third phase supports the analysis of cyber-artifacts by social researchers from the perspectives of social movement theory [22].

2.4 Social Network Analysis

The central theme in social movement studies revolves around how activists "organize" themselves in campaigns to achieve impact on governments and societies,
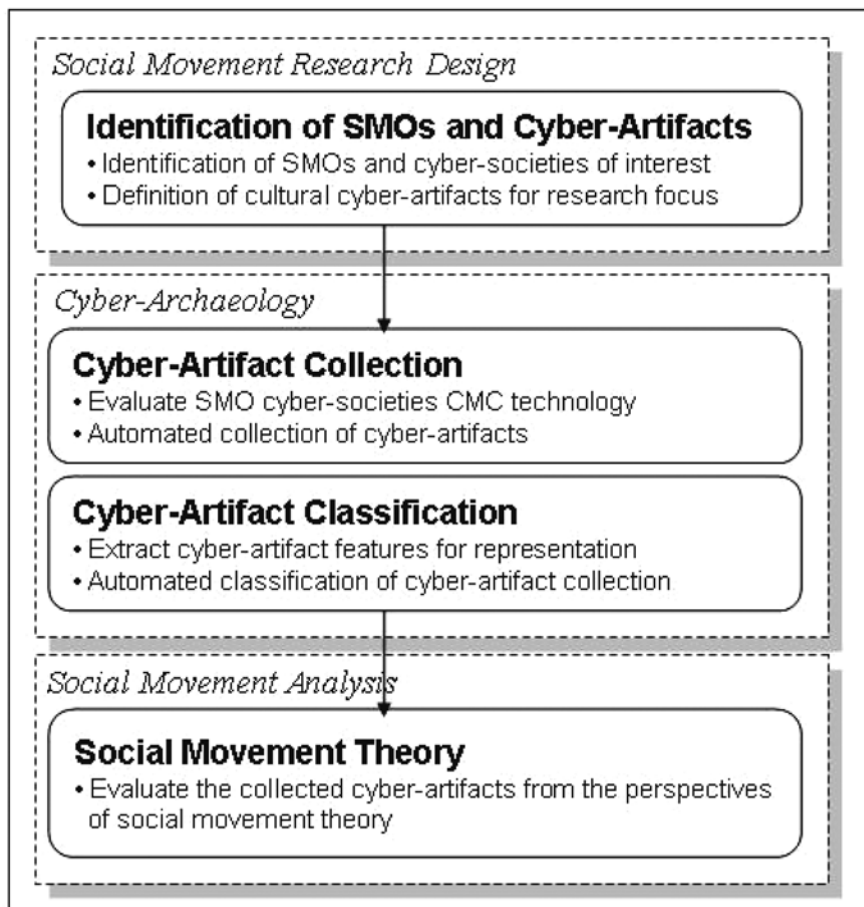
Social Movement Research Design

Identification of SMOs and Cyber-Artifacts
- Identification of SMOs and cyber-societies of interest
- Definition of cultural cyber-artifacts for research focus

Cyber-Archaeology

Cyber-Artifact Collection
- Evaluate SMO cyber-societies CMC technology
- Automated collection of cyber-artifacts

Cyber-Artifact Classification
- Extract cyber-artifact features for representation
- Automated classification of cyber-artifact collection

Social Movement Analysis

Social Movement Theory
- Evaluate the collected cyber-artifacts from the perspectives of social movement theory

Fig. 1 Cyber-archaeology framework

which makes SNA perfectly suitable for this kind of investigation. SNA is a graph-based methodology used to analyze social relations and their influence on individual behavior and organizational structure. It was developed by sociologists and has been applied in several academic fields, such as epidemiology and CMC [18]. In SNA, individuals or actors are represented as nodes in a simple network graph, called a social network, and tied with edges indicating relationships. The visualization of a social network can provide a basic sense of how actors affiliate with each other and what their roles are in the group. In a sample friendship network, as shown in Fig. 2, Persons A and G are considered the most active or "popular" persons since they are linked to the largest number of people. Person F is also important although he/she doesn't have as many connections as Persons A and G: Person F bridges two
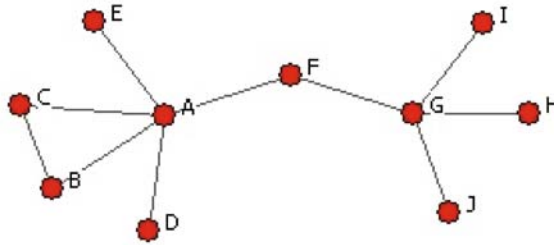
Fig. 2 Sample friendship network

different clusters of friends. In SNA, these three people are said to be central or prominent within the sample network.

Centrality measures are quantitative indicators for finding those "central" individuals from a network originally developed in communication scenarios. From a topological perspective, people who are able to receive or control the mainstream of message flow typically stand in a position similar to the central point of a star [12], such as the location of Person A in the sample network. Various centrality measures, such as degree and betweenness, can be employed to determine the importance of a node within a network. Degree is the number of edges that a node has. Since the central point of a star has the largest number of edges connecting it to the other nodes, a node with a higher degree is topologically considered to be more central to its network [12, 35]. Betweenness measures "the extent to which a particular node lies between the various other nodes" [32] because the central point also sits between pairs. The higher betweenness a node has, the more potential it has to be a gatekeeper controlling the connections (such as communications) between the others [32].

Through reconstructing the network of nineteenth-century women reform leaders in New York State, Rosenthal et al. [31] reported that weak ties played an important role in the women's movement: while strong ties linked major women's organizations in a movement, weak ties bridged several clusters and channeled the communication to diverse audiences. After the emergence of global activism, many researchers shifted their focus from SMO physical connections to SMO website hyperlinks. Ackland et al. [3] used their VOSON system to demonstrate the usefulness of network visualization in the analysis of linkage between environmental groups on the Internet. Garrido and Halavais [13] used hyperlink analysis to map the Zapatista online network and examine its affiliation with other SMOs. They found that the secondary tier of Zapatista related websites played a bridging role linking the Zapatista network to the global SMO network.

2.5 Writeprints

Because of its anonymous nature, the Internet has become a major medium for cybercrime ranging from illegal sales and phishing to terrorist communication. In

order to increase the awareness and accountability of users, many studies have been devoted to developing techniques to identify authors in the online environment [1]. Authorship identification is a process of matching unidentified writings to an author based on the similarity of writing styles between the known works of the author and unidentified pieces [1, 2]. Four major categories of style features have been extensively used to identify writing styles: lexical, syntactic, structural, and content-specific [1]. Lexical features include total number of words, words per sentence, and word length distribution. Syntax refers to the patterns used for the formation of sentences, such as punctuation and function/stop words. Structural features deal with the organization and layout of the text, such as the use of greetings and signatures, the number of paragraphs, and average paragraph length. Content-specific features are keywords that are important within a specific topic domain. Among these four categories, lexical and syntactic features are frequently used because of their high discriminatory ability and portability across domains [1, 2].

In 2006, Abbasi and Chen [2] proposed a visualization technique for authorship called Writeprints as shown in Fig. 3, which is analogous to the fingerprint biometric system. Unlike other studies of authorship visualization merely using *n*-gram features for discrimination, Writeprint was designed to apply to a large number of authors in an online setting. It uses all four major types of style features: lexical, syntactic, structural, and content-specific [2]. The generation of a "Writeprint" consists of two main steps: (1) reduce dimensionality and (2) create visualizations. After extracting features from a set of documents, the Writeprint adopts principal component analysis (PCA) to reduce dimensionality of feature usage vectors by using the two principal components or eigenvectors with the largest eigenvalues. Once the eigenvectors have been computed with PCA, a sliding window algorithm is used to extract the feature usage vector for the text region inside a window, which slides over the text. For each window instance, the sum of the product of the principal component (primary eigenvector) and the feature vector represents the *x*-coordinate of the pattern point while the sum of the product of the second component (secondary eigenvector) and the feature vector represents the *y*-coordinate of the data point. Each data point generated is then plotted onto a two-dimensional space to create the
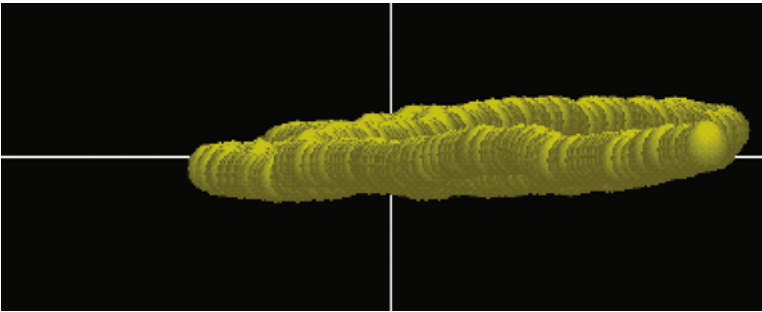

Fig. 3 Example of the Writeprint (each dot represents a text segment overlaying on the X–Y-coordinates)

Writeprint. Abbasi and Chen [2] reported that Writeprints outperformed the Support Vector Machine (SVM) in classification of online messages in their evaluation.

3 Research Design: A Case Study of the International FLG Movement

This study employed the proposed cyber-archaeology framework and used the Falun Gong (FLG, 法輪功) movement as a case study to investigate SMOs' collective identity on the Internet and their ideological ties with others.

3.1 The Falun Gong Movement

FLG was founded by Hongzhi Li (李洪志) and introduced to the public in Mainland China in 1992 [28]. It focuses on the concept of cultivation and has two main components: practice and principle. The practice component includes five sets of gentle exercise movements which are similar to Qi-Gong (氣功) exercises. The principal component emphasizes the importance of truthfulness, compassion, and forbearance (真善忍). FLG practitioners believe they can enhance their physical health and elevate their mind at the same time.

On July 20, 1999, the Chinese government designated FLG as an evil cult and banned all its public activities [34]. This suppression is widely believed to be related to the mass petition of Falun Gong practitioners in Zhongnanhai, Beijing on April 25, 1999 [28]. After the official ban, Hongzhi Li stayed in the United States and used FLG websites, such as Clearwisdom.net (法輪大法明慧網), to release his articles and continue his teaching. Currently, FLG has local websites of practitioners in over 38 countries and five continents. It holds several conferences annually in North America and Europe.

Before the suppression, there was no evidence that FLG or Hongzhi Li had any political agenda against the Chinese government. On January 23, 2001, a self-immolation incident by five FLG practitioners in Tiananmen Square in Beijing was reported by the international news media. In 2002, FLG websites began releasing accounts of persecution against practitioners in Mainland China, including pictures, stories, and persecution methods. In late 2004, the Epoch Times (大紀元), which is related to FLG, published "Nine Commentaries on the Communist Party" (九評共產黨) and held a "Quitting the CCP (Chinese Communist Party)" (退黨) campaign.

Several studies, each using a different approach, have investigated how FLG transformed from a simple Qi-Gong exercise group to a religious and social movement organization [28, 30, 34, 21]. Lu [21] applied the religious economy model to the formation of new religions to interpret the shift of FLG from a healing system to a religion. According to Lu's analysis [21], Hongzhi Li purposely introduced

his own theory of salvation, Falun Dafa, to differentiate FLG from other competing Qi-Gong groups. In addition, he used various organizational and doctrinal mechanisms to keep his practitioners and avoid schisms. For example, he claimed that he was the incarnation of the highest supernatural force and that the only master in Dafa [21]. Rahn [30] used the conflict between FLG and the Chinese government to explain FLG's role in social movements. He suggested that the persecution of FLG in China was the key to establishing FLG's identity as a human rights movement. By examining Hongzhi Li's messages in the Clearwisdom.net, he brought out a concern that the Fa-rectification teaching may induce violent behavior in FLG practitioners, because the frustration at achieving this ultimate goal may intensify "the battle between good and evil" [30]. But Rahn [30] asserted that FLG's positive public image as a human rights group can decrease the chances of a violent act.

The widespread use of the Internet for organizing practitioners is another research focus of those studying FLG. Bell and Boas [5] summarized three important functions of the Internet in the FLG movement: (1) disseminating Hongzhi Li's teachings, (2) strengthening the integrity of a globally dispersed community, and (3) bringing pressure on the Chinese government for lifting the ban. They also concluded that the use of the Internet might bring splinter sects challenging Li's authority [5]. This is consistent with Bennett's point of view that "the ease of joining and leaving polycentric issue networks means that it becomes difficult to control campaigns or to achieve coherent collective identity frames" [6].

3.2 Research Design

Figure 4 shows our research design. Our interest was in two types of cyber-artifacts in the FLG movement: websites and forums. SMO websites typically are official "entry points" of SMO networks advocating their ideologies and campaigns. Collecting those websites' hyperlinks and contents allowed us to investigate how FLG officially deploys itself on the Internet and connects to other SMOs. SMO forums, on the other hand, provide a relatively intimate view of how members interact with each other and discuss their SMO ideology. We used automatic programs to collect those cyber-artifacts and performed three analyses for our inquiries: link, web content, and forum content analyses.

3.2.1 Cyber-Artifacts Collection

Website Hyperlinks and Content: FLG has four core websites as listed in Table 1, distributing FLG news, Hongzhi Li's articles, and accounts of persecution of FLG practitioners in Mainland China. These four core websites are Clearwisdom.net (法輪大法明慧網), FalunInfo.net (法輪大法新聞社), FalunDafa. org (法輪大法), and EpochTimes.com (大紀元). Each website offers more than ten language versions and has multiple domain names, a reflection of the fact that the FLG movement is organized on a global scale.
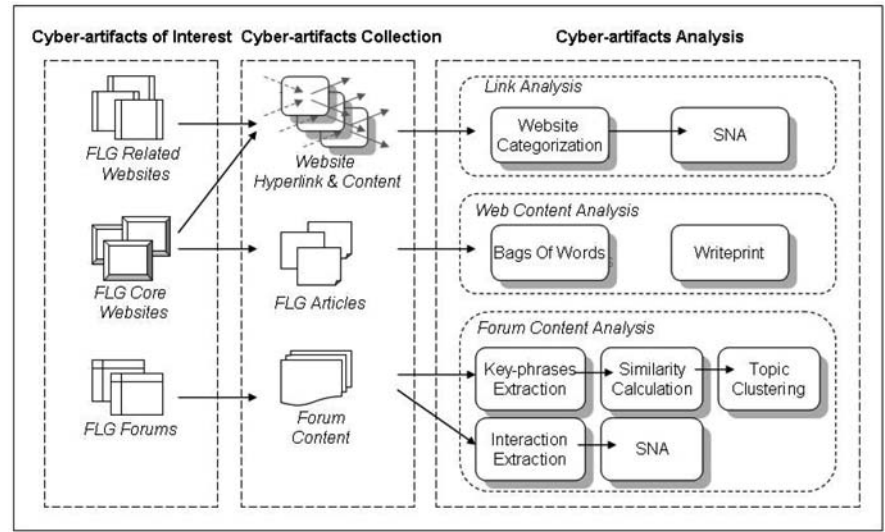
Fig. 4 Research design

Table 1 Four core FLG websites

Website	Content	Language	Domain name
Clearwisdom.net (法輪大法明慧網)	1.Falun Gong Information 2.Hongzi Li' Articles 3.Persecution Accounts 4.Practitioners' Sharing	10	27
FalunDafa.org (法輪大法)	1.Falun Gong Information 2.Local Contact Information and Websites	36	4
FalunInfo.net (法輪大法新聞社)	1.Persecution Accounts 2.Falun Gong News	12	14
EpochTimes.com (大紀元)	1.World News 2.Persecution Accounts 3.CCP Criticism	15	36

We automatically collected (spidered) FLG relevant websites, including those of other activist groups having hyperlinks to FLG websites, with two levels of inlinks and outlinks via 31 seed websites, which included the four core websites and another 27 FLG websites identified by Google search. A total of 425 relevant domain names were found during spidering and 172 were deemed relevant. Most of the relevant websites were found to be directly linked to the core websites as shown in Table 2.

FLG Articles: For the web content, we were particularly interested in studying the role of Hongzhi Li's articles in the FLG movement. From Clearwisdom.net, we collected 135 articles from Hongzhi Li and 74 articles from the editors for later

Table 2 Collecting FLG-related websites via seed websites

Seed Website	Outlink Level 1	Outlink Level 2	Inlink Level 1	Inlink Level 2	Total
FalunDafa.org (法輪大法)	85	3	3	0	91
EpochTimes.com (大紀元)	26	0	1	1	28
Clearwisdom.net (法輪大法明慧網)	15	4	1	1	21
GuangMing.org (澳洲光明網)	12	0	1	0	13
FalunInfo.net (法輪大法新聞社)	4	2	1	0	7
SoundOfHope.org (希望之聲電台)	2	2	0	0	4
GrandTrial.org (全球公審江澤民)	3	0	0	0	3
GlobalRescue.net (全球營救 FLG 學員)	2	0	0	0	2
ZhuiChaGouJi.org (追查迫害 FLG 組織)	0	1	0	0	1
NtdTV.com (新唐人電視台)	0	1	0	0	1
Minghui-School.org (明慧學校)	1	0	0	0	1
Total	150	13	7	2	172

comparison. Those articles concentrated on the discussion of three topics: teaching/principles of FLG, the position of FLG on political issues especially related to Mainland China, and summaries of various FLG conferences. The summary of these two sets of articles is shown in Table 3. Compared to the editors' articles, Hongzhi Li's are much longer in length.

Table 3 Summary of articles in Clearwisdom.net

Source	Number of articles	Words per article	Duration
Hongzhi Li	135	1,430	5/1999–2/2007
Editors	74	670	3/2000–12/2006
Total	209	1,161	5/1999–2/2007

Forum Content: We used Google search and website linkage to find FLG forums. Four forums were found, but only one forum, Falun Dafa Universal (世界法輪大法研究會), was, and is, still active and has more than 50 authors contributing to it. Therefore, we concentrated on the analysis of threads and messages in this forum. Falun Dafa Universal, located at city.udn.com (網路城邦), was established in 2005. It has 120 members and 28 discussion boards covering Hongzhi Li's articles, persecution accounts, and the FLG universal and science database. This forum circulates many articles from the four core websites. Thus, the average length of messages is long, 1,288 characters per message, but the average reply rate is low, 0.89 reply messages per thread. A total of 740 threads and 1,399 messages were collected for this forum.

3.2.2 Cyber-Artifact Analysis

Link Analysis: In order to understand the main ideas of these websites and how they linked together, we first classified their ideological types and performed SNA to analyze their network structure. Two measures of centrality in SNA are used to investigate which websites are prominent in this network: degree and betweenness. The degree of a node is the number of links it has, reflecting its activity level. Betweenness is a measure of the frequency with which a node lies on the shortest geodesic paths of pairs of other nodes. It can be used to detect the extent to which a node plays the role of a gatekeeper in controlling the communication of others [32].

Web Content Analysis: In order to highlight the characteristics of Hongzhi Li's writing, we used Bag of Words and Writeprints, developed by Abbasi and Chen [2], and compared his articles with other articles written by the editors of Clearwisdom.net.

Forum Content Analysis: At the forum level, we performed two types of analysis: thread topic and author interaction. In the thread topic analysis, we investigated how many topics are covered in this forum and how those topics relate to each other. Since Falun Dafa Universal is a Chinese forum, we first used MI, a Chinese phrase extraction tool developed by Ong and Chen [26], to extract key Chinese phrases from the threads and convert those threads into vectors of those key phrases. The top 20 key phrases based on frequency of appearance are shown in Table 4. We then used the cosine coefficient to calculate the similarity between threads and displayed those threads in a two-dimensional map. For author interaction analysis, we extracted the authors' responses to others' threads and performed SNA based on their interaction history.

4 Research Results

In this section, we present our research results.

4.1 Link Analysis

The 203 FLG relevant websites, including seed websites and those identified via the seeds, are classified into five main categories based on their web content: FLG

Table 4 Key Chinese phrases of forum messages

Rank	Phrase	Rank	Phrase
1	法輪 (Falun)	11	政府 (Goverment)
2	法輪功 (Falun Gong)	12	醫院 (Hospital)
3	中國 (China)	13	修煉 (Cultivation)
4	學員 (Practitioner)	14	問題 (Problem)
5	器官 (Organ)	15	國際 (International)
6	迫害 (Persecution)	16	人類 (Human)
7	美國 (the United States)	17	個人 (Individual)
8	蘇家屯 (Sujiatun Camp)	18	國家 (Country)
9	大法 (Dafa)	19	集中營 (Labor Camp)
10	社會 (Society)	20	人民 (People)
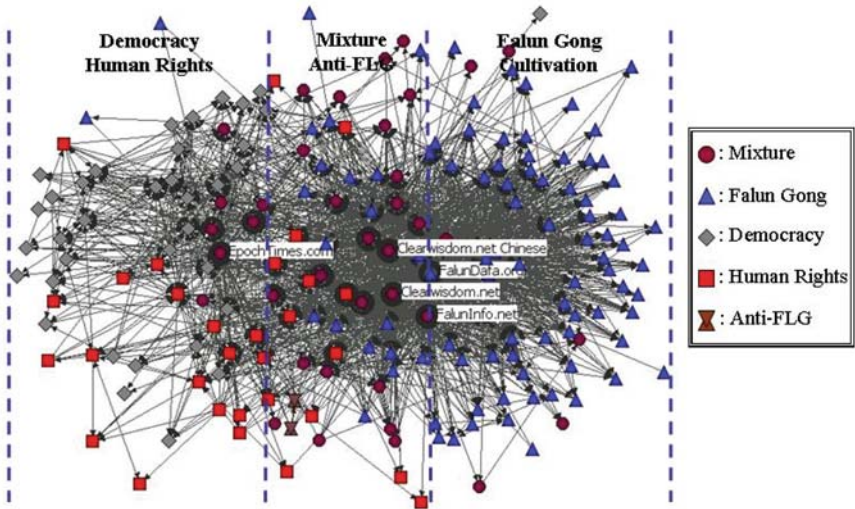
Fig. 5 Network of FLG relevant websites

cultivation, human rights, democracy, anti-FLG, and mixture (topics of more than one category). Two coders were hired for the website classification. The network of these websites, displayed with a spring embedded algorithm, is shown in Fig. 5. The network has three main components: human rights and democracy on the left-hand side, FLG cultivation on the right-hand side, and mixture (and anti-FLG) in the middle. The mixture websites, including Clearwisdom.net and EpochTimes.com, act as bridges connecting the other two main components. The human rights and democracy websites are somewhat mixed together.

We used two centrality measures, degree and betweenness, to identify the most prominent websites within this network. Here, the degree or in-degree is calculated by the number of inlinks and reflects the popularity of a website. The betweenness measures the potential that a website may be a gatekeeper controlling the interaction with other websites. The top ten prominent websites in this network are listed in Table 5. The four core FLG websites are at the top of the list.

Table 5   Top ten websites in FLG network based on centrality

Rank	Website	In Degree	Website	Betweenness
1	FalunDafa.org (法輪大法)	113	FalunDafa.org (法輪大法)	14657.33
2	FalunInfo.net (法輪大法新聞社)	99	EpochTimes.com (大紀元)	6166.15
3	Clearwisdom.net (法輪大法明慧網英文)	90	Clearwisdom.net Chinese version (法輪大法明慧網中文)	4318.54
4	Clearwisdom.net Chinese version (法輪大法明慧網中文)	88	GuangMing.org (澳洲光明網)	2533.59
5	EpochTimes.com (大紀元)	78	Clearwisdom.net (法輪大法明慧網英文)	2298.31
6	ZhengJian.org (正見)	65	FalunInfo.net (法輪大法新聞社)	2014.37
7	Fofg.org (法輪功之友)	54	ZhengJian.org (正見)	1335.54
8	ClearHarmony.net (歐洲圓明網)	50	SoundOfHope.org (希望之聲電台)	1276.49
9	SoundOfHope.org (希望之聲電台)	50	ClearHarmony.net (歐洲圓明網)	1077.36
10	NtdTV.com (新唐人電視台)	48	HriChina.org (中國人權)	792.51

We used the inlinks and outlinks of the connected websites to check the role of the FLG core websites in this network. The results are shown in Fig. 6. EpochTimes.com is mainly responsible for the linkage of human rights and democracy websites. Clearwisdom.net is located in the middle of the network and connects other major mixture websites. FalunDafa.org focuses on FLG cultivation and links local FLG practitioners' websites.
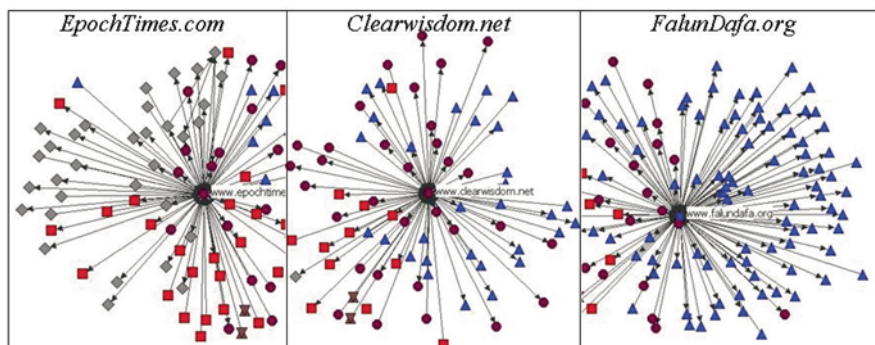
Fig. 6 Roles of FLG core websites in FLG network

4.2 Web Content Analysis

Writeprints illustrate the characteristics of words that authors frequently use to express their opinions or ideas. Figure 7 shows the Writeprint of Hongzhi Li. His discussion revolves around the teachings of Falun Dafa (法輪大法) and has neither significant temporal variation nor concentration on sub-topics as shown in Fig. 7(a). However, in his bag of words in Fig. 7(c), the word "evil" is used frequently.

The Writeprints of the editors of Clearwisdom.net, as shown in Fig. 8(a) and (b), had three significantly deviated areas which represented the topics of Dafa rectification (正法), righteous thoughts (發正念), and persecution of FLG practitioners (學員迫害真相). Their writings consistently revolved around these three major topics between 2000 and 2006. Comparing the Writeprints of Hongzhi Li and the editors allowed us to see their roles in the FLG movement. Hongzhi Li's articles focused on the central concepts of FLG cultivation but provide some hints of his political attitude (e.g., against evil). The editors' articles, on the other hand, provided their interpretations of Hongzhi Li's teaching.

4.3 Forum Content Analysis

4.3.1 Thread Topics

In the 740 threads collected from the forum "Falun Dafa Universal," ten main topics were identified based on the content of the threads: persecution accounts (學員迫害真相), FLG success story sharing (修煉心得分享), FLG ideology (法輪功哲學), FLG articles (法輪功書籍文獻), anti-Chinese Communist Party (anti-CCP, 反對中國共產黨), life philosophy (生活哲學), mysterious phenomena (宇宙科學與神秘現象), social issues (社會議題), health issues
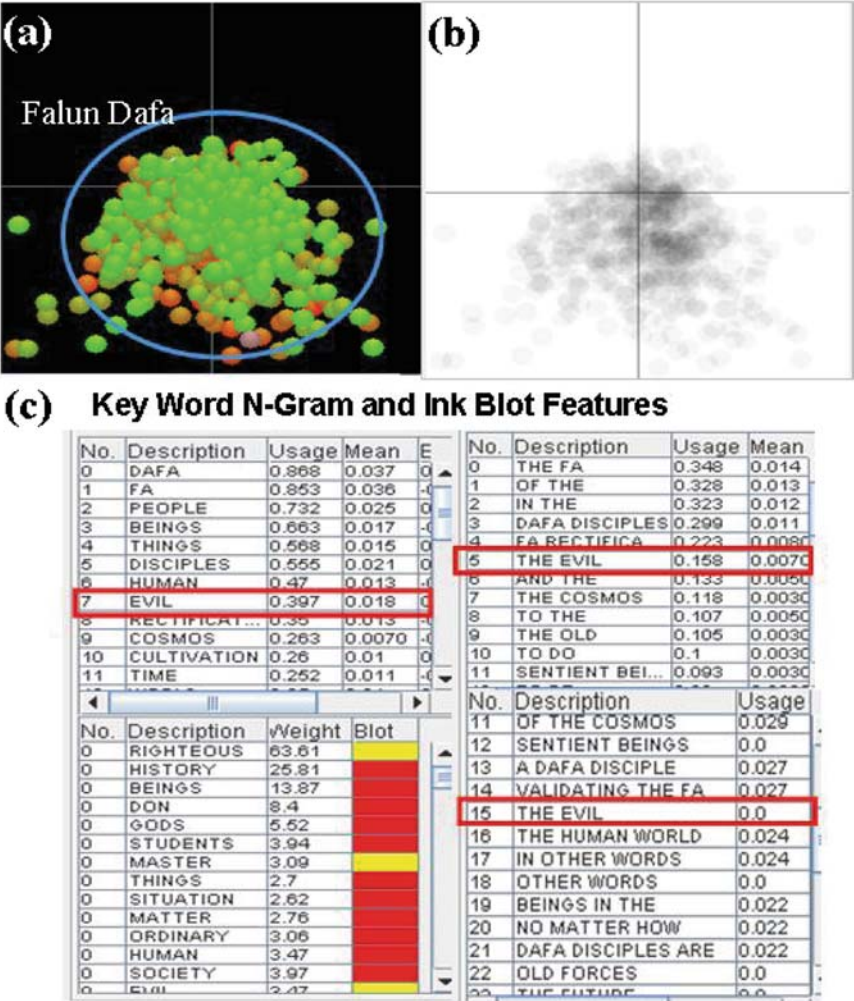
Fig. 7 Writeprint of Hongzhi Li

(健康議題), and general messages (網站管理訊息). The ten topics and their descriptions are listed in Table 6.

The distribution of threads over these ten topics is summarized in Table 7. Although life philosophy has the highest number of threads, anti-CCP has the highest average reply rate. Major discussions in this forum are often about anti-CCP topics.

Figure 9 displays the threads based on their similarity. We can see that the persecution accounts and anti-CCP are aligned and to some degree mixed together on the upper parts of the circle. Such a mixture
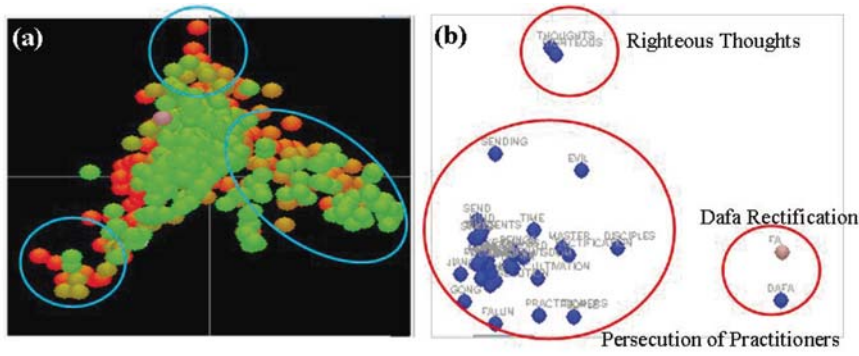
Fig. 8 Writeprint of the editors of Clearwisdom.net

Table 6 Ten main topics of forum threads

Main Topic	Description
Persecution accounts (學員迫害真相)	Detailed description of torture process and methods
FLG success story sharing (修煉心得分享)	Describe how a practitioner benefits from FLG
FLG ideology (法輪功哲學)	Share personal beliefs about Dafa rectification, cultivation, and righteous thoughts
FLG articles (法輪功書籍文獻)	Include articles and books from Hongzhi Li and FLG
Anti-CCP (反對中國共產黨)	Criticize CCP for organ harvest, human rights, and religious freedom
Life philosophy (生活哲學)	Share inspired life stories and words of wisdom
Mysterious phenomena (宇宙科學與神秘現象)	Distribute articles about the origin of the cosmos and unexplainable phenomena
Social issues (社會議題)	Discuss social issues, such as the role of news press and impact of violent video games
Health issues (健康議題)	Distribute health-related news and healthy recipes
General message (網站管理訊息)	Messages about the forum management and arguments

Table 7 Distribution of threads over ten topics

Main topic	Threads	Messages	Reply rate
Persecution accounts	100	214	1.14
FLG success story sharing	10	16	0.6
FLG ideologies	112	256	1.28
FLG articles	29	29	0
Anti-CCP	112	336	2
Life philosophy	166	255	0.54
Mysterious phenomena	87	107	0.23
Social issues	16	39	1.44
Health issues	86	105	0.22
General message	22	42	0.9
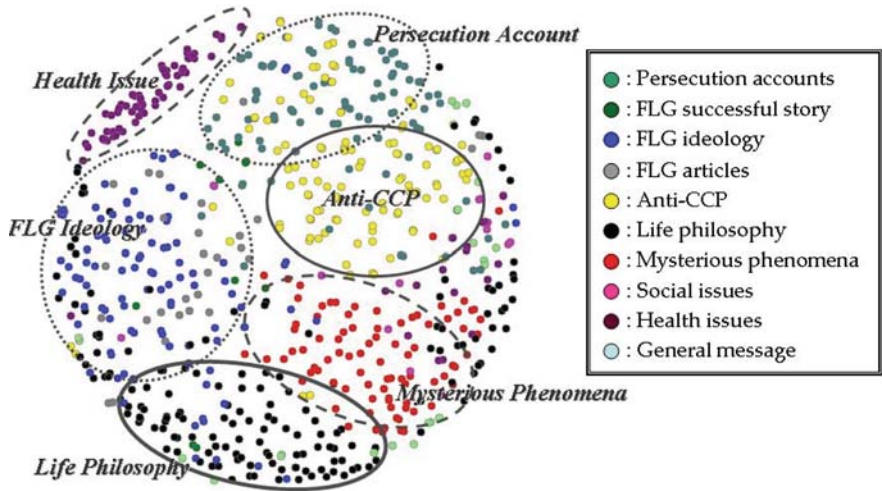
Fig. 9 Display of threads according to their similarity

is due to the high usage of the same key phrases, such as organ harvest (器官活摘). and labor camp (集中營). In the lower part, Falun Gong ideology is closely aligned with life philosophy and mysterious phenomena. From the relative positions of these three topics, we may infer that Falun Gong is similar to a religion, which not only teaches a certain life philosophy but also explains the origin of life.

From the display of threads, we can further see to which topics an author primarily contributes. Figure 10 shows the distribution of threads of the top two active authors in the forum. The author "Sujcs888" focused on Falun Gong ideology, life philosophy, and mysterious phenomena; while the author "LoveTender" targeted the topics of persecution and anti-CCP.
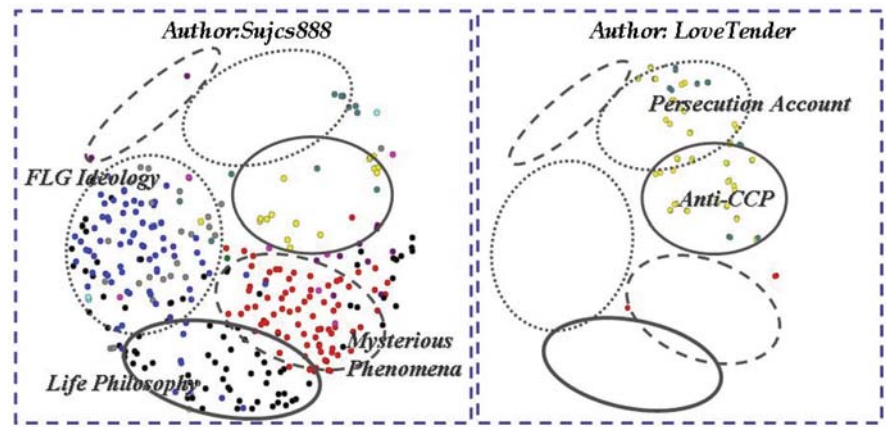
Fig. 10 Distribution of topics of the top two active authors in the forum

4.3.2 Author Interaction

In order to see which topics have shown more intense discussion among authors, we measured the interaction networks of the top five topics above with average degree and clustering coefficient. The average degree shows the overall activity or inter-action density of authors in a network. The clustering coefficient reflects clusters, which can indicate cliques or groups [35]. The results of these two measures are summarized in Table 8. The most intense interaction occurred in the discussion of FLG ideology and anti-CCP. However, compared to FLG ideology, anti-CCP had lower average degree but a much higher clustering coefficient. This implies that those authors were more likely to create discussion groups around this topic.

Table 8 Degree and clustering coefficient of five main topics

Main topic	Degree	Clustering coefficient
FLG ideology	2.400	0.086
Life philosophy	2.080	0.093
Mysterious phenomena	1.00	0.000
Persecution	1.733	0.000
Anti-CCP	2.261	0.128

Figure 11 shows the interaction networks relating to FLG ideology and the anti-CCP topic. Authors discussing FLG ideology seemed to follow the ideas and preaching of master practitioners (the node in the center of the network). In anti-CCP discussions, authors were more likely to share their opinions and interact with each other freely (as shown in small clusters of interactions).
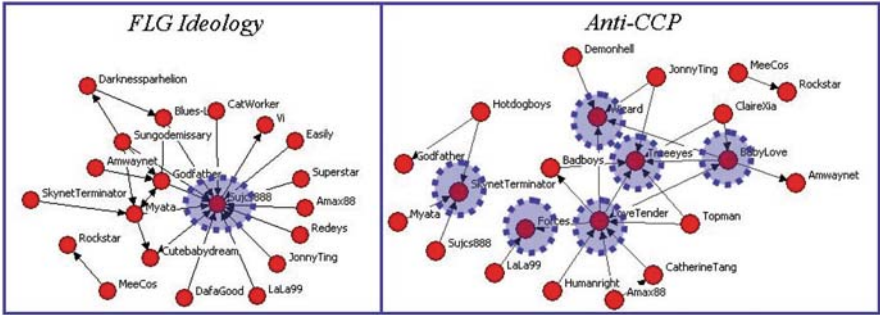
Fig. 11 Interaction network of authors in FLG ideology and anti-CCP

5 Conclusions

In this study, we took a cyber-archaeology approach and performed three separate analyses on the cyber-artifacts, including link, web content, and forum content analyses, to investigate the framing of FLG's identity in social movements. First, it is not surprising to see that FLG's websites closely link to two types of SMOs: Chinese democracy and human rights. This affiliation can be explained with the conflicts between FLG and the Chinese government: the official ban in 1999 and subsequent persecution of its practitioners. As Rahn [30] pointed out, those persecution accounts gave FLG a justified identity as a human rights movement. Second, by taking the cyber-archaeology perspective, we can see how FLG "strategically" deploys its cyber-artifacts forming its inclusive identity on the Internet. Not only are its four core websites located in the center of the FLG network, but they also bridge two seemingly unrelated groups of websites for different roles: activist and Qi-Gong groups. Each of the core websites plays a specific role in joining websites of different attributes. For example, most of human rights and democracies in the FLG network are connected via EpochTimes.com. In addition, we find a trace of its religious role in its forum: it tries to explain the origin and meanings of life. From topic clustering, we see how FLG includes and organizes several different concepts on a continuum: from FLG ideology to life philosophy and mysterious phenomena, and from mysterious phenomena to anti-CCP and persecution by conceptualizing the CCP as the "Evil." By deploying its cyber-artifacts, FLG smoothly connects different ideologies and establishes its inclusive role as a Qi-Gong, religious, and activist group.

As a religious group, Hongzhi Li without question is the spiritual leader of FLG. The Writeprint revealed a structural pattern resembling a religious hierarchy in the writings of Hongzhi Li and of the editors: Hongzhi Li's centered around the ideological teaching of Falun Dafa while the editors' posted specific programs outlining Li's teaching for the practitioners to follow. In the FLG forum, we also found that the authors exhibited different interaction patterns in the discussions of FLG ideology and of anti-CCP. They followed the monologue pattern of master-preaching

in the FLG ideology but had more interaction in the anti-CCP discussion. Does the structural difference between a religious and an activist organization cause tension in FLG's cooperation with other SMOs? While we don't find any evidence from analyzing its cyber-artifacts to support this, the human rights torch relay, mainly hosted by CIPFG.org (Coalition to Investigate the Persecution of Falun Gong in China) between 2007 and 2008, revealed potential coordination problems between FLG and its allies. The relay was widely conceived as an FLG campaign against the Beijing 2008 Olympic Games and used as its slogan, "The Olympics and crimes against humanity cannot coexist in China." However, Hongzhi Li posted a message on April 4, 2008, explaining that the human rights torch relay is for everyday people and "so it is not Dafa disciples that this event is for." It is interpreted as a call for FLG practitioners to focus on FLG persecution and truth-clarification work rather than involvement in human rights activities. How the FLG is going to reframe its identity in social movements will require careful monitoring and future studies.

References

1. Abbasi A, Chen H (2005) Applying Authorship Analysis to Extremist-Group Web Forum Messages. IEEE Intelligent Systems, 20: 1541–1672
2. Abbasi A, Chen H (2006) Visualizing Authorship for Identification. In: Intelligence and Security Informatics. San Diego, pp. 60–71
3. Ackland R, O'Neil M, Bimber B, Gibson RK, Ward S (2006) New Methods for Studying Online Environmental-Activist Networks. In: 26th International Sunbelt Social Network Conference, Vancouver
4. Almind TC, Ingwersen P (1997) Informetric Analyses on the World Wide Web: Methodological Approaches to Webometrics. Journal of Documentation 53: 404–426
5. Bell MR, Boas TC (2003) Falun Gong and the Internet: Evangelism, Community, and Struggle for Survival. Nova Religio 6: 277–293
6. Bennett W (2003) Communicating Global Activism. Information, Communication and Society 6: 143–168
7. Bennett WL (2005) Social Movements Beyond Borders: Understanding Two Eras of Transnational Activism. In: Porta Dd and Tarrow S (eds) Transnational Protest and Global Activism. Rowman & Littlefield Publishers, Inc., New York, pp. 203–226
8. Björneborn L, Ingwersen P (2001) Perspective of Webometrics. Scientometrics 50: 65–82
9. Björneborn L, Ingwersen P (2004) Toward a Basic Framework for Webometrics. Journal of the American Society for Information Science and Technology 55: 1216–1227
10. Clark JD, Themudo NS (2006) Linking the Web and the Street: Internet-based Dotcauses and the Anti-Globalization Movement. World Development 34: 50–74
11. Cohen R, Rai SM (2000) Global Social Movements. The Athlone Press, New Brunswick
12. Freeman LC (1978/79) Centrality in Social Networks: Conceptual Clarification. Social Networks 1: 215–239
13. Garrido M, Halavais A (2003) Mapping Networks of Support for the Zapatista Movement: Applying Social-Network Analysis to Study Contemporary Social Movements. In: McCaughey M, Ayers MD (eds) Cyberactivism: Online Activism in Theory and Practice. Routledge, New York

14. Gerlach LP (2001) The Structure of Social Movements: Environmental Activism and its Opponents. In: Arquilla J, Ronfeldt DF (eds) Networks and Netwars: The Future of Terror, Crime, and Militancy. Rand, Santa Monica, pp. 289–309
15. Jones Q (1997) Virtual-Communities, Virtual Settlements & Cyber-Archaeology: A Theoretical Outline. Journal of Computer-Mediated Communication 3
16. Jones Q, Rafaeli S (2000) What Do Virtual "Tells" Tell? Placing Cybersociety Research into a Hierarchy of Social Explanation. In: Hawaii International Conference on System Sciences, Hawaii
17. Keck ME, Sikkink K (1998) Activists Beyond Borders: Advocacy Networks in International Politics. Cornell University Press, New York
18. Klovdahl AS, Potterat JJ, Woodhouse DE, Muth JB, Muth SQ, Darrow WW (1994) Social networks and Infectious Disease: the Colorado Springs Study. Social Science & Medicine 38: 79–88
19. Langman L (2005) From Virtual Public Spheres to Global Justice: A Critical Theory of Internetworked Social Movements. Sociological Theory 23: 42–74
20. Larana E, Johnston H, Gusfield JR (1994) New Social Movements: From Ideology to Identity. Temple University Press, Philadelphia
21. Lu Y (2005) Entrepreneurial Logics and the Evolution of Falun Gong. Journal for the Scientific Study of Religion 44: 173–185
22. McAdam D, McCarthy J, Zald M (1999) Comparative Perspectives on Social Movements, Cambridge University Press, Cambridge
23. McCaughey M, Ayers MD (2003) Cyberactivism: Online Activism in Theory and Practice. Routledge, New York
24. Morris AD, Mueller CM (1992) Frontiers in Social Movement Theory. Yale University Press, New Haven
25. Myers DJ (1994) Communication Technology and Social Movements: Contributions of Computer Networks to Activism. Social Science Computer Review 12: 250–260
26. Ong TH, Chen H (1999) Updateable PAT-Tree Approach to Chinese Key Phrase Extraction using Mutual Information: A Linguistic Foundation for Knowledge Management. In: Proceedings of the Second Asian Digital Library Conference, Taipei
27. Paccagnella L (1997) Getting the Seats of Your Pants Dirty: Strategies for Ethnographic Research on Virtual Communities. Journal of Computer-Mediated Communication 3
28. Penny B (2003) The Life and Times of Li Hongzhi: Falun Gong and Religious Biography. The China Quarterly 175: 643–661
29. Porta Dd, Diani M (1999) Social Movements: An Introduction. Blackwell Publishers Ltd, Malden
30. Rahn P (2002) The Chemistry of a Conflict: The Chinese Government and the Falun Gong. Terrorism and Political Violence 14: 41–65
31. Rosenthal N, Fingrutd M, Ethier M, Karant R, McDonald D (1985) Social Movements and Network Analysis: A Case Study of Nineteenth-Century Women's Reform in New York State. The American Journal of Sociology 90: 1022–1054
32. Scott J (2000) Social Network Analysis: A Handbook, 2nd edn. Sage, London
33. Tesh SN (2002) The Internet and the Grass Roots. Organization & Environment 15: 336
34. Tong J (2002) An Organizational Analysis of the Falun Gong: Structure, Communications, Financing. The China Quarterly 171: 636–660
35. Wasserman S, Faust K (1994) Social Network Analysis: Methods and Applications. Cambridge University Press, New York