

What Balkanizes the Internet? Access Denied or Access Unwanted

Harsh Taneja (University of Missouri), Angela Xiao Wu (The Chinese University of Hong Kong)[†] & Sushant Tripathy (Northwestern University)

Research Program

We challenge the assumption that state imposed access blockage isolates its citizens on the World Wide Web. We argue that, instead, users prefer culturally proximate content and access blockage only enhances this existing predisposition.

Our **preliminary study** of shared audience traffic to the world's top 1000 websites indicated the limited impact of China's Great Firewall on Chinese web user behavior (see right).

In order to observe sustained effects of access blockage, our **ongoing study** **historically analyzes patterns of global web usage from 2009 to 2013 with a focus on multiple countries** with varying levels of Internet penetration and access blockage. The initial findings confirm the limited effects of access blockage on user behavior; furthermore, they reveal the rapid rise of the global South on the WWW in the recent years, which may have resulted from various government interventions in these developing states.

Goal: Our research aims to establish the need to link Internet governance research with empirical analyses of user behavior.

What Balkanizes the WWW?

Dominant view: Access denied – state sponsored online filtering

Our view: Access unwanted – user preference based on cultural proximity

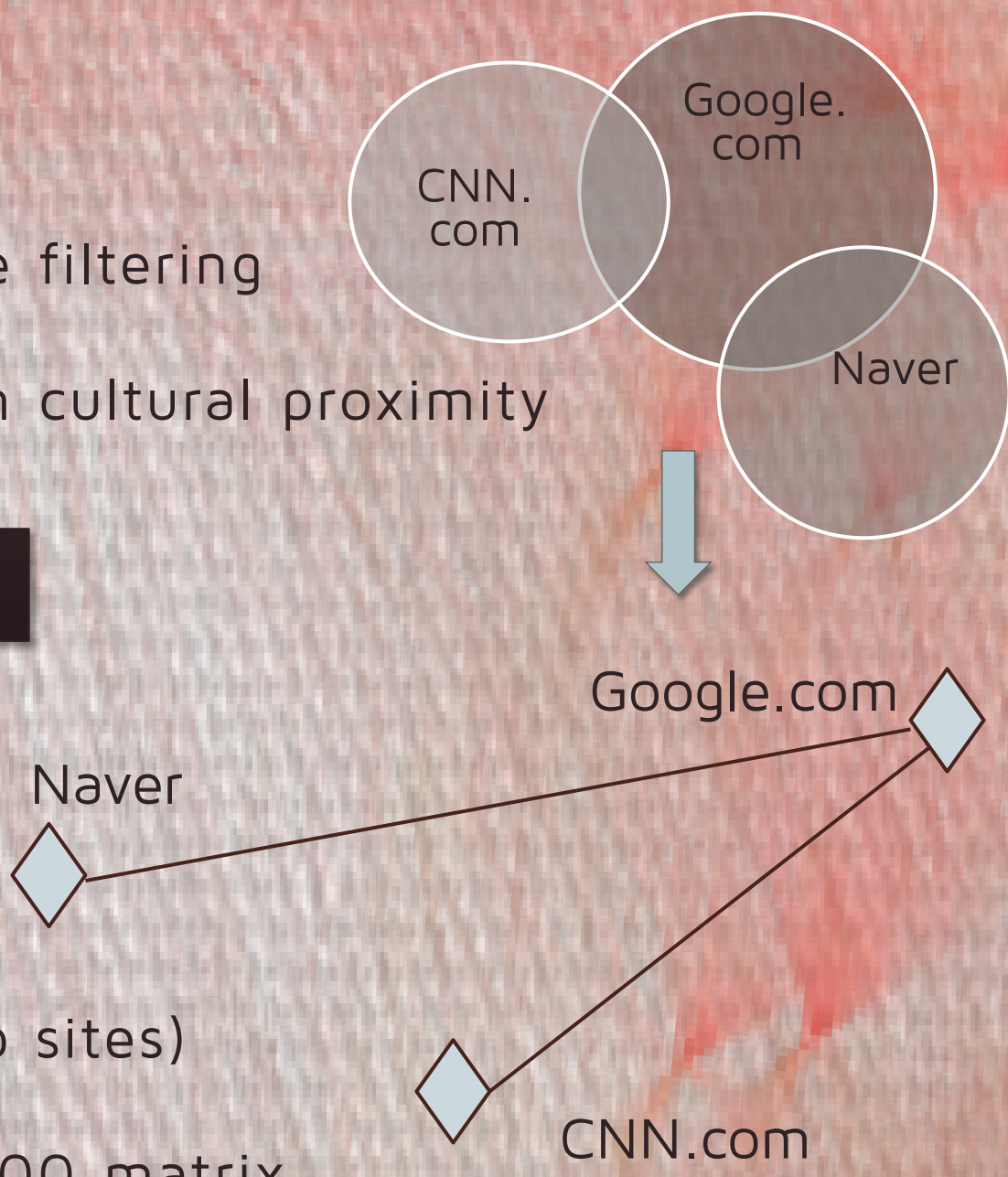
Analysis of WWW as an "Audience Network"

Nodes: Media outlets (websites)

Ties: Audience overlap

(greater than expected audience traffic common to two sites)

E.g., a sample of 1000 websites results in a 1000 x 1000 matrix



Ongoing Study

Analysis of multiple time periods and multiple markets

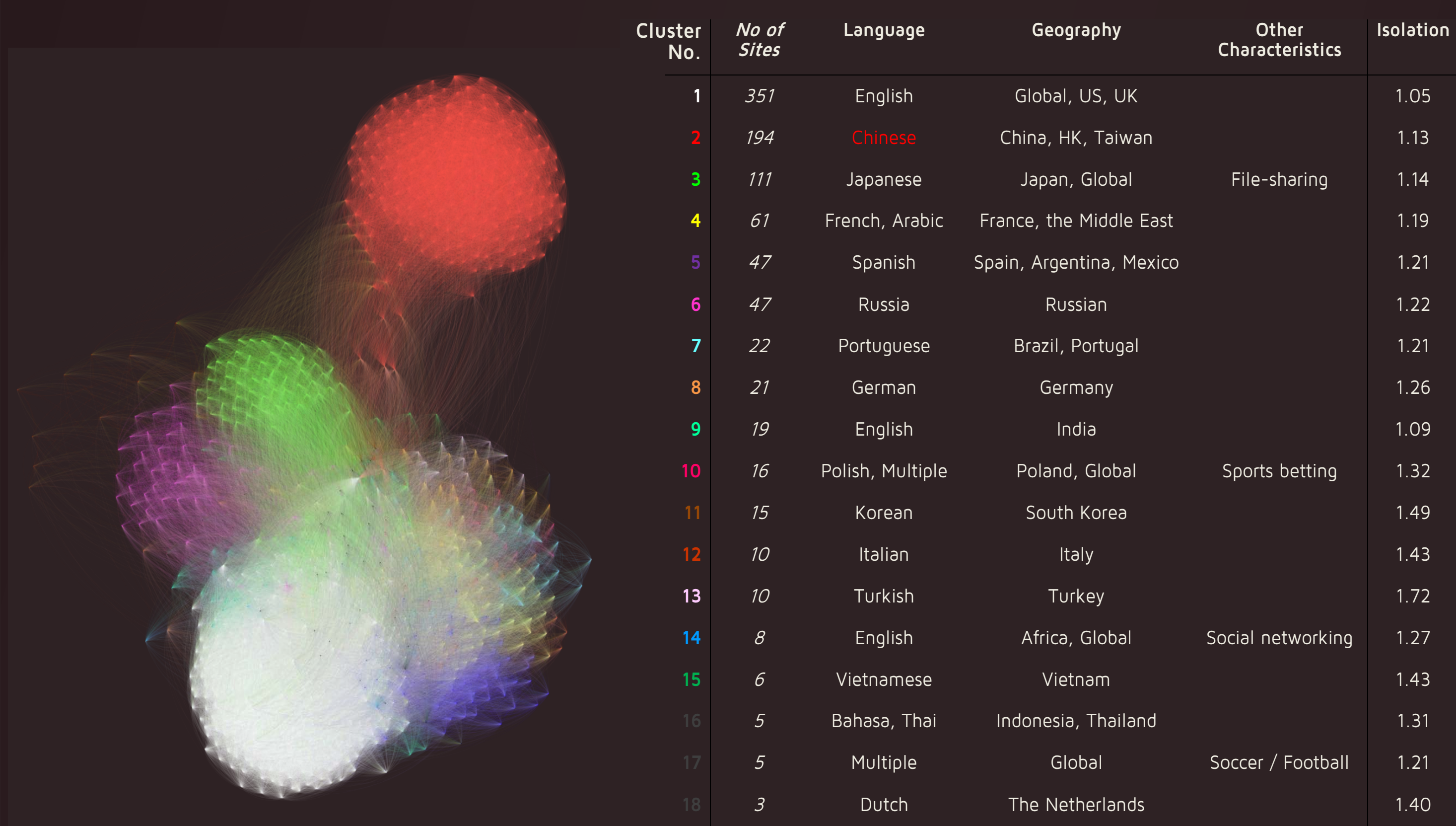
- Incorporates the historical dimension to study sustained effects of access blockage
- Compares four countries chosen based on variations of access blockage and culture

	User Base	Internet Adoption	Access Blockage
China	Big	Expanding	Pervasive
Japan	Big	Stable	No evidence
India	Small	Expanding	No evidence/Selective
South Korea	Small	Stable	Substantial

Preliminary Study (Taneja & Wu, in press, *The Information Society*, <http://arxiv.org/abs/1305.3311>)

Does the Great Firewall Isolate China's Internet Users?

Analysis of traffic between top 1000 websites globally (comScore, June 2012)



Minimal Impact of Great Firewall

- WWW clusters on geo-linguistic lines

- China as isolated as Japan
 - Korea, Poland and Turkey more isolated

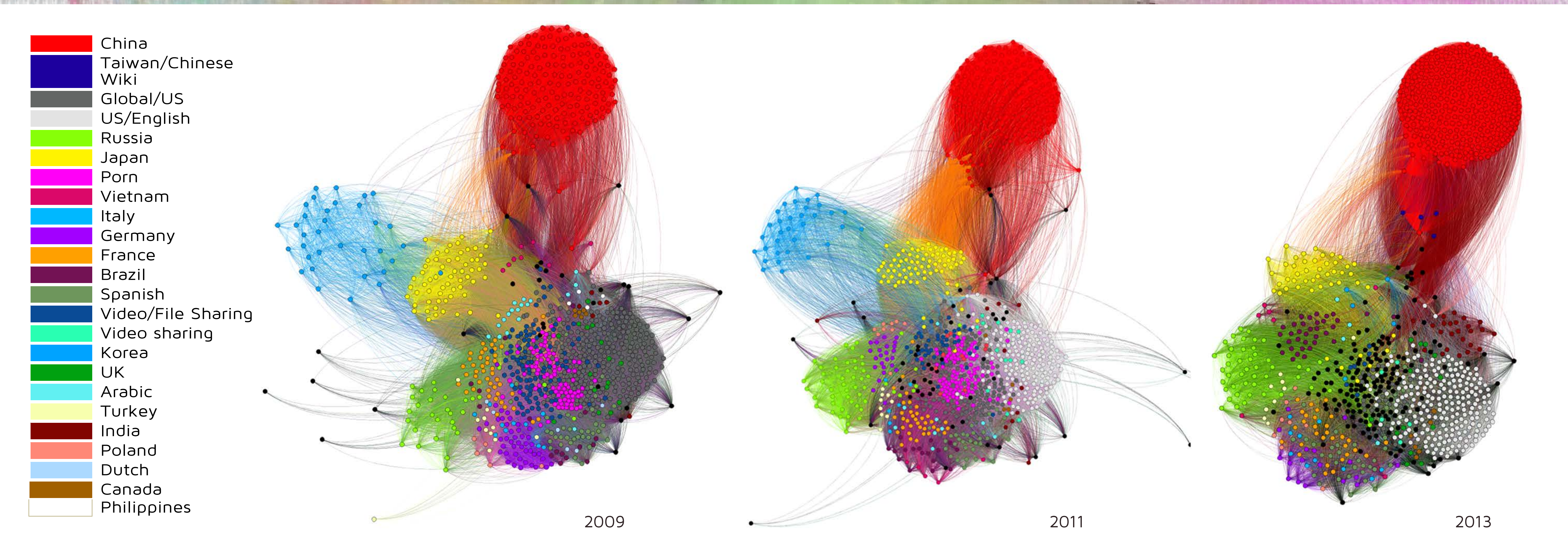
- Regardless of being blocked, none of the culturally distant websites made it to the Chinese cluster

- Does cultural proximity drive Internet balkanization more than access blockage?

Classification of Website	Cultural Proximity	Chinese Cluster	
		Non-Member	Member
In Chinese and China Focus	Highest	0 (0%)	189 (2%)
In Chinese and Taiwan & HK Focused	High	0	4 (75%)
In Chinese but Global Focus	Low	105 (24%)	1 (0%)
Neither in Chinese nor Greater China Focused	Lowest	674 (10%)	0
Total Number of Websites		779 (12%)	194 (4%)

Sample: Top 1000 sites in Septembers 2009, 2011 and 2013

Year	N	Network Centralization	Degrees Mean (SD)	Clustering Coefficient	Density	Distance Mean (SD)
2009	1018	0.448	432.29 (221.8)	0.855	0.425	1.578 (0.499)
2011	1022	0.515	407.55 (209.7)	0.851	0.399	1.609 (0.504)
2013	1030	0.572	359.7 (144.1)	0.857	0.35	1.651 (0.478)



Inferences

- WWW usage manifests as regional clusters based on language and geography
- Regional clusters form across regions with varying levels of access blockage
- Increase in regional clusters due to inclusion of users from and websites focusing on the global South

Next Steps

- Objective 1: Uncover various assumptions and agendas underlying the creation of various cyber-censorship measures for cross-country comparison
- Objective 2: Explore associations between the evolution of online regional cultures manifested through user access and preferences, and the role of policies not limited to access blockage
- Objective 3: Design and create a dataset containing relevant contextual variables to facilitate more research endeavors to bridge Internet governance and web use behavior