Every Rose Has Its Thorn
Censorship and Surveillance on Social Video Platforms in China

Jeffrey Knockel, Masashi Crete-Nishihata, Jason Q. Ng
Adam Senft, Jedidiah R. Crandall

Citizen Lab, Munk School of Global Affairs, University of Toronto
Dept. of Computer Science, University of New Mexico
USENIX FOCI 2015, August 10, Washington DC
4 most popular social video platforms in China
Over 1 billion registered users
17,547 sensitive keywords
Research Problem

• What are the underlying mechanisms and motivations behind censorship / surveillance on Chinese social media?
• Previous work shows inconsistencies in censorship / surveillance of Chinese social media (Crandall et al. 2013, MacKinnon 2008, Villeneuve 2008)

• Suggest general directives may be given by government, but companies have flexibility in implementation
Overview

• We reverse engineer four of the most popular Social Video Platforms in China: YY, 9158, Sina Show, GuaGua

• Found **client-side keyword censorship** in each SVP, and **keyword surveillance** in YY

• We reveal dataset of **17,547 unique keywords**, largest currently available to researchers!

• **Previously largest dataset 4,256 unique keywords** TOM-Skype + Sina UC (Knockel *et al.* 2012, Crandall *et al.* 2013)
登录能获得免费道具哦~
Real-time beautification camera

Real-time video rooms

Users sing karaoke or chat with each other in the video rooms

Users send virtual gifts to hosts and other users on air

Offer Android and iOS versions for all our major live social communities

Source: TianGe Corporate Presentation
<table>
<thead>
<tr>
<th>Company</th>
<th>Public / Private</th>
<th>Product</th>
<th>Registered Users</th>
<th>Monthly Active Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>YY Inc.</td>
<td>Nasdaq</td>
<td>YY</td>
<td>861.4 mn</td>
<td>117.4 mn</td>
</tr>
<tr>
<td>Tian Ge</td>
<td>SEHK</td>
<td>9158, Sina Show</td>
<td>245 mn</td>
<td>14.4 mn</td>
</tr>
<tr>
<td>Jinhua Changfeng</td>
<td>Private</td>
<td>GuaGua</td>
<td>70 mn</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Prohibited Topics on the Chinese Internet

- Endangering state security
- Divulging state secrets
- Subverting state power
- Jeopardizing state religious policy
- Spreading rumors
- Disrupting social order and stability pornography
- Gambling
- Violence
- Humiliating or slandering others

source:  http://www.china.org.cn/government/whitepaper/node_7093508.htm
We would like to… express our deep regret to the public. Sina is willing to work with other website operators to make vigorous efforts in the country’s anti-porn campaign and create a clean Internet environment”
- Sina Corporation

YY: “Content posted or displayed on our platform may be found objectionable by PRC regulatory authorities and may subject us to penalties and other severe consequences”

Content Filtering and Management

• Strict Terms of Service for users and performers
• Violations lead to fines, account suspensions
• Content Monitoring Teams
• Automated Content Monitoring
Keyword Censorship
Technical Analysis

• YY, 9158, Sina Show, GuaGua perform **client-side censorship**

• VV, Six Room, BoBo perform **server-side censorship**
YY Keyword Censorship

- YY v7.1 downloads 3 different keyword lists: *Finance, Normal, High*
YY Keyword Censorship

- Finance Keyword List (keywords related to phishing scams)
  
  http://do.yy.duowan.com/financekwordlist

- Plain text in UTF8-encoded XML

- When received, YY prints the following warning in the chat window:

  YY安全提示：聊天中若有涉及财产的操作，请一定要先核实好友身份，谨防受骗！

  [YY Security Tip: This chat seems to involve managing assets; please be sure to verify the identity of a friend to avoid being cheated!]
YY Keyword Censorship

• *Normal* Keyword List (sensitive keywords)

• Source: http://do.yy.duowan.com/NormalKWordlist.txt

• Base64-encoded list of UTF16-encoded keywords

• Keywords are asterisked out in the message before the message is sent
YY Keyword Censorship

- *High* Keyword List (sensitive keywords)
- Source: http://do.yy.duowan.com/HighKWordlist.txt
- Base64-encoded list of UTF16-encoded keywords
- Keywords are never sent. If a message containing one of these keywords is somehow received, then that message will show in the chat window as a blank message.
YY Keyword Surveillance

http://sere.hiido.com/do.action?
id=<id>&content=<content>

=id is a hex encoding of a hash computed as

\( \text{md5}(\lfloor \text{seconds since unix epoch} \rfloor/1000 \rfloor + ";username=report" + ";password=pswd@1234") \).
YY Keyword Surveillance

<content> is a base64 encoding of the following string:

```
type=2
;uid=<sending user id#>
;touid=<receiving user id#>
;keyword=<triggering keyword>
;txt=<entire triggering message>
```
Sina Show Keyword Censorship

• SinaShow v3.4 includes:
  • Built-in and pre-installed keyword lists
  • Downloaded lists
Sina Show Keyword Censorship

Outgoing censored message warning:

“系统过滤, 你发送的信息含有非法字符, 请重新输入!” [System filter: the message you sent contains illegal words; please re-enter!]

falungong
Incoming censored message warning:

“发送的消息有非法词汇,已经被自动屏蔽”
[The message sent contains illegal vocabulary; it has been automatically blocked]
Sina Show Keyword Censorship

- Installed with binary keyword database in Word_410.ucw
- Downloads updates from http://www.51uc.com/uc_interface/down_policy/Word_410.ucw
- Custom binary container stores GBK-encoded keywords
- Encryption: Blowfish in ECB mode with 8-byte key Dey,1b1E
Sina Show Keyword Censorship

- Sina Show has keywords built into many of its binary files, stored in the plain GBK-encoded
  - SinaShow.exe, ChatRoom.dll, Props.dll
  - UCClient.dll
9158 Keyword Censorship

- Analyzed 9158 v6.9
- Installed with two keyword lists: filnick.xml filter.xml
- filnick.xml replaces keywords with asterisks in usernames
- If chat message contains 6+ English letters all of the letters are replaced with asterisks
GuaGua Keyword Censorship

- Analyzed GuaGua v6.2.38
- Keywords built into RuleCenterPlug.dll
- Plain text, GBK-encoded
- Outgoing messages containing keywords are not sent, displaying warning message:

  “消息发送失败,含有违法或不文明字符!”
  [Failed to send message; it may contain illegal or uncivilized words!]
Our Results
## Keyword List Size (May 17 2015)

<table>
<thead>
<tr>
<th>List</th>
<th>Keywords</th>
<th>Unique</th>
</tr>
</thead>
<tbody>
<tr>
<td>YY Finance</td>
<td>48</td>
<td>18</td>
</tr>
<tr>
<td>YY Normal</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>YY High</td>
<td>13,482</td>
<td>13,242</td>
</tr>
<tr>
<td>9158 Nick</td>
<td>65</td>
<td>59</td>
</tr>
<tr>
<td>9158 Chat</td>
<td>318</td>
<td>318</td>
</tr>
<tr>
<td>Sina Show SinaShow.exe</td>
<td>1,224</td>
<td>910</td>
</tr>
<tr>
<td>Sina Show UCClient.dll</td>
<td>910</td>
<td>910</td>
</tr>
<tr>
<td>Sina Show Downloaded</td>
<td>3,711</td>
<td>3,206</td>
</tr>
<tr>
<td>GuaGua</td>
<td>58</td>
<td>58</td>
</tr>
</tbody>
</table>
Keyword lists clustered by Jaccard Similarity

- YY Normal
- YY High
- Sina UC DL 5
- Sina UC DL 2
- Sina UC Builtin 2
- Sina UC DL 1
- Sina UC Builtin 1
- Sina Show DL
- GuaGua
- 9158 Nicknames
- Skype 5.0 Builtin
- Skype 3.6 Builtin
- LINE
- Sina UC Builtin 5
- Sina UC Builtin 4
- Sina UC Builtin 3
- Skype 5.1 Surveil
- Skype 5.0 DL
- Skype 5.5 Surveil
- Skype 4.0 DL
- Skype 3.6 DL
- Sina Show UC
- Sina Show Builtin
- Sina UC DL 4
- Sina UC DL 3
Keyword lists clustered by similarity(x, y) = max(% of x in y, % of y in x)
Content Analysis

• Keywords human / machine translated, contextual analysis

• Grouped into 80 content categories under six themes based on code book developed by Crandall *et al.* (2013)
• Gambling. e.g., 多伦多娱乐城 “Toronto casino”

• Drugs, firearms, other illicit goods, e.g., 冰毒制作 “Methamphetamine production”

• Prostitution, sexual interests. e.g, 妓女 “Prostitute” 性伴侣 “Sexual partner”
### Political Theme

- **37 content categories**
  - e.g., CPC, religious movements, ethnic minorities, corruption, democracy groups, and terrorism
• CPC General: 共产党 “Communist Party”

• CPC Critical: 共匪 “Communist gangsters”
• Includes references to Islam, Uyghur groups and culture, terrorism, and general terms

• English: uyghurcanadian

• Chinese: 突厥斯坦 “Turkestan”

• Uyghur: türkistan “Turkestan”

• Arabic script: ปาร์ติ์คุ์จ ปั่นเข้า ที่เรียก: “instructions to make explosives”
<table>
<thead>
<tr>
<th>Theme</th>
<th>Categories</th>
<th>Client</th>
<th>Number of Records</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>People</strong></td>
<td>People -- CPC member / Gov official</td>
<td>YY</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9158</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>guagua</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>People -- Dissident/Activist/ Rights defender</td>
<td>YY</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9158</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>People -- Relatives / associates of government officials</td>
<td>YY</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9158</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>People -- context unclear</td>
<td>YY</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>People -- Victims of Crime</td>
<td>YY</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
<td>10</td>
</tr>
</tbody>
</table>
• CPC Leaders: 习近平 “xí jìnpíng”
• Homonyms: 习尽平 “xí jìn píng”
• Nicknames: 习包子 “Steamed Bun Xi”
习包子 “Steamed Bun Xi”

source: http://chinadigitaltimes.net/2014/01/sensitive-words-steamed-buns-rumor-dictatorship/
<table>
<thead>
<tr>
<th>Theme</th>
<th>Categories</th>
<th>Client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>16th CPC National C.</td>
<td>Sina Show</td>
</tr>
<tr>
<td>Event</td>
<td>18th CPC National C.</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>2008 Olympics</td>
<td>Sina Show</td>
</tr>
<tr>
<td>Event</td>
<td>2012 Bloomberg Cen.</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>2012 Kashgar Riots</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>2012 Scarborough Sh.</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>2014 Kunming attack</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>APEC 2014</td>
<td>Sina Show</td>
</tr>
<tr>
<td>Event</td>
<td>Arab Spring</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Bo Xilai Scandal</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Carrefour Boycott</td>
<td>Sina Show</td>
</tr>
<tr>
<td>Event</td>
<td>Diaoyu / Senkaku islands protests</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Flight MH370</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Forced demolitions/e.</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Google leaving China</td>
<td>Sina Show</td>
</tr>
<tr>
<td>Event</td>
<td>Jasmine Revolution</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Jiang Zemin death ru.</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Occupy Central HK</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Occupy movement</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>South Korea Yellow S.</td>
<td>YY</td>
</tr>
<tr>
<td>Event</td>
<td>Tibetan Self-immolati..</td>
<td>Sina Show</td>
</tr>
<tr>
<td>Event</td>
<td>Urumqi Riots (July 2009)</td>
<td>YY</td>
</tr>
</tbody>
</table>

Number of Records

0  20  40  60
YY: references largest number of events (18) 
90% of YY event keywords related to June 4 1989
Recent Events in Keyword Lists

2014 - Occupy Central Protests

2014 - APEC Summit

2014 - Kunming Terrorist Attack
Discussion

• Inconsistencies in the content and implementation in keyword lists across companies and platforms

• Very limited overlap within SVPs and in comparison to chat clients

• Supports theories that companies are under general pressure from the government to perform censorship and surveillance, but have flexibility in implementation
Discussion

- Targeted content relates to government criticism and collective action
  - King *et al.* analyze post censorship in 1,382 Chinese social media sites.
  - Contend content related to collective action censored but content critical of government permitted
  - Our results reveal sensitive keywords in SVPs and chat clients related to collective action and government criticism
  - Calls into question generalizability of King *et al.* theory
### Collective Action Keywords

<table>
<thead>
<tr>
<th>Theme</th>
<th>Categories</th>
<th>Client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>Event -- 2012 Kashgar Riots</td>
<td>YY</td>
</tr>
<tr>
<td></td>
<td>Event -- Arab Spring</td>
<td>YY</td>
</tr>
<tr>
<td></td>
<td>Event -- Diaoyu / Senkaku islands protests</td>
<td>Sina Show</td>
</tr>
<tr>
<td></td>
<td>Event -- Jasmine Revolution</td>
<td>YY</td>
</tr>
<tr>
<td></td>
<td>Event -- Occupy Central HK</td>
<td>YY</td>
</tr>
<tr>
<td></td>
<td>Event -- Tibetan Self-immolati.</td>
<td>Sina Show</td>
</tr>
<tr>
<td></td>
<td>Event -- Urumqi Riots (July 2009)</td>
<td>9158</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
</tr>
<tr>
<td></td>
<td></td>
<td>YY</td>
</tr>
<tr>
<td>Political</td>
<td>Protest</td>
<td>9158</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sina Show</td>
</tr>
<tr>
<td></td>
<td></td>
<td>YY</td>
</tr>
<tr>
<td></td>
<td>Tuidang</td>
<td>Sina Show</td>
</tr>
<tr>
<td></td>
<td></td>
<td>YY</td>
</tr>
</tbody>
</table>

Number of Records
Keyword List Updates

- Hourly data collection: YY (Feb. 7, 2015), 9158 (Feb. 24 2015), Sina Show (March 11 2015)
- YY Normal (3 updates), YY High (21 updates)
- 9158 Chat (8 updates)
- Sina Show Downloaded (0 updates)
- GuaGua does not download updates
YY Updates

• 32 keywords referencing names of Christian songs added April 23, removed April 30 (YY Normal)

• e.g, 主耶稣永远陪伴我 “Jesus is always with me”

• May have been in response to controversy over recent church demolitions
9158 Updates

- Keywords added on May 28 2015
- 戴建勇 “Dai Jianyong” (name of Shanghai artist)
- 麻痹哥 “Paralysis brother”
Data: https://github.com/citizenlab/chat-censorship

Special Thanks:
Jakub Dalek, Mamatjan Juma, Greg Fay, Zubayra Shamseden, and Greg Wiseman

Sponsors:
Open Technology Fund Information Controls Fellowship Program
John D. and Catherine T. MacArthur Foundation