Catching Social Media Advertisers with Strategy Analysis

Meng Jiang

University of Illinois at Urbana-Champaign

www.meng-jiang.com

*The First International Workshop on Computational Methods for **CyberSafety***

Advertising Strategies

❑Before we talk about social media …

❑Different strategies of advertising
  ❑Bug your customer every week/month/season
  ❑Dive into your customers' community …

❑Analyzing the strategies is a marketing issue
  ❑$$$

❑Analyzing the strategies is also a safety issue!

My Experience: It is a safety issue!

Meng Jiang
October 13 at 4:16pm · 🌐 ▼

This is an awful company! Everyone in Urbana should be careful. They sent sales people annoying our neighbors and us saying that every apartment in Urbana has to change their energy supply to theirs. They ask for our previous bill and use the account number on the bill to change our plan. They are lying - Ameren Illinois is using Homefield Energy and the service is fine. They speak fast and don't give you a chance to challenge. They keep knocking our neighbors' door now. We've no idea if they will do things more evil - they've collected lots of residents' information. After reported, the property holder has announced "they are not allowed to sell their services on our property and they need to leave". God bless America.

The No. 1 Provider of Green Energy in North America.
Clearview Energy
Commercial & Industrial

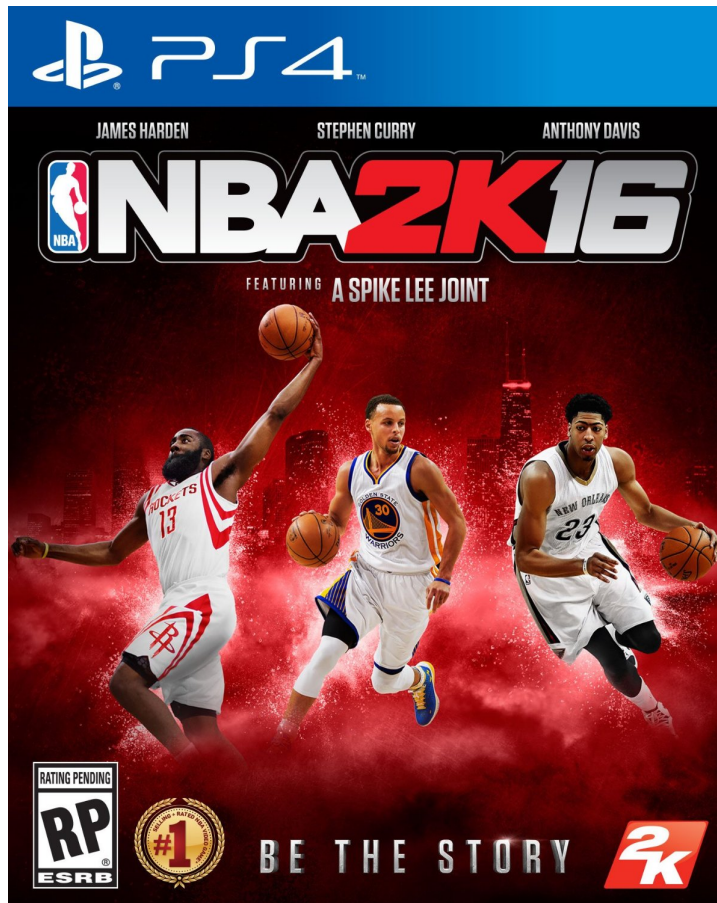clearview ENERGY™

8,150 Likes
42 talking about this

👍 Like Page

Some Strategy Examples
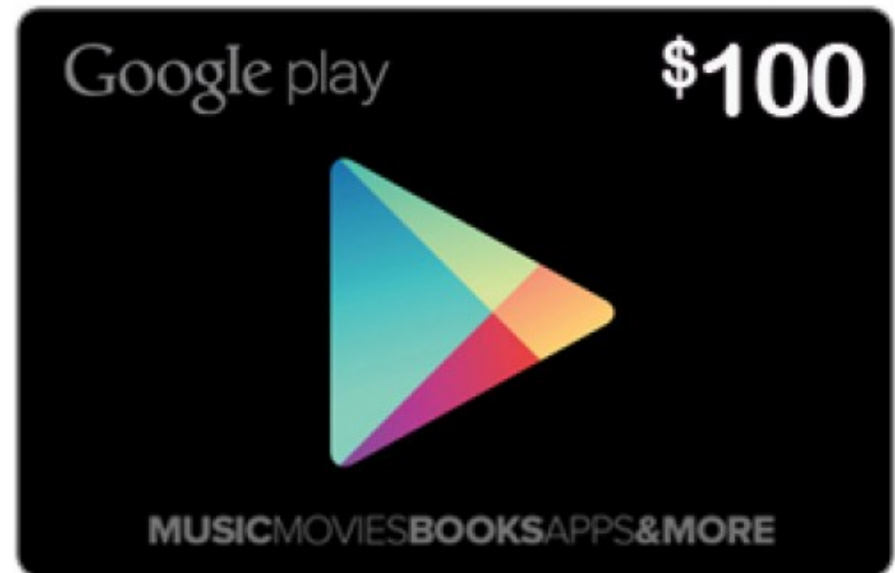
☐S1. Celebrity branding

Some Strategy Examples

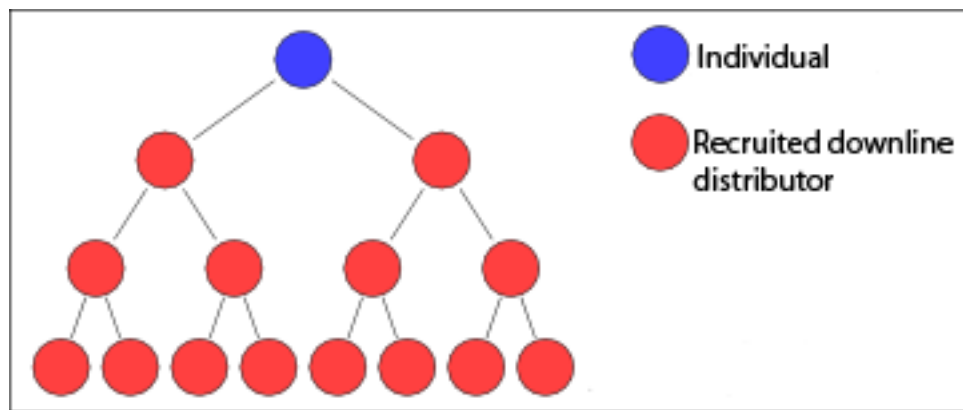❑S2. Collaborative advertising

Some Strategy Examples

❑S3. Gift advertising

Some Strategy Examples
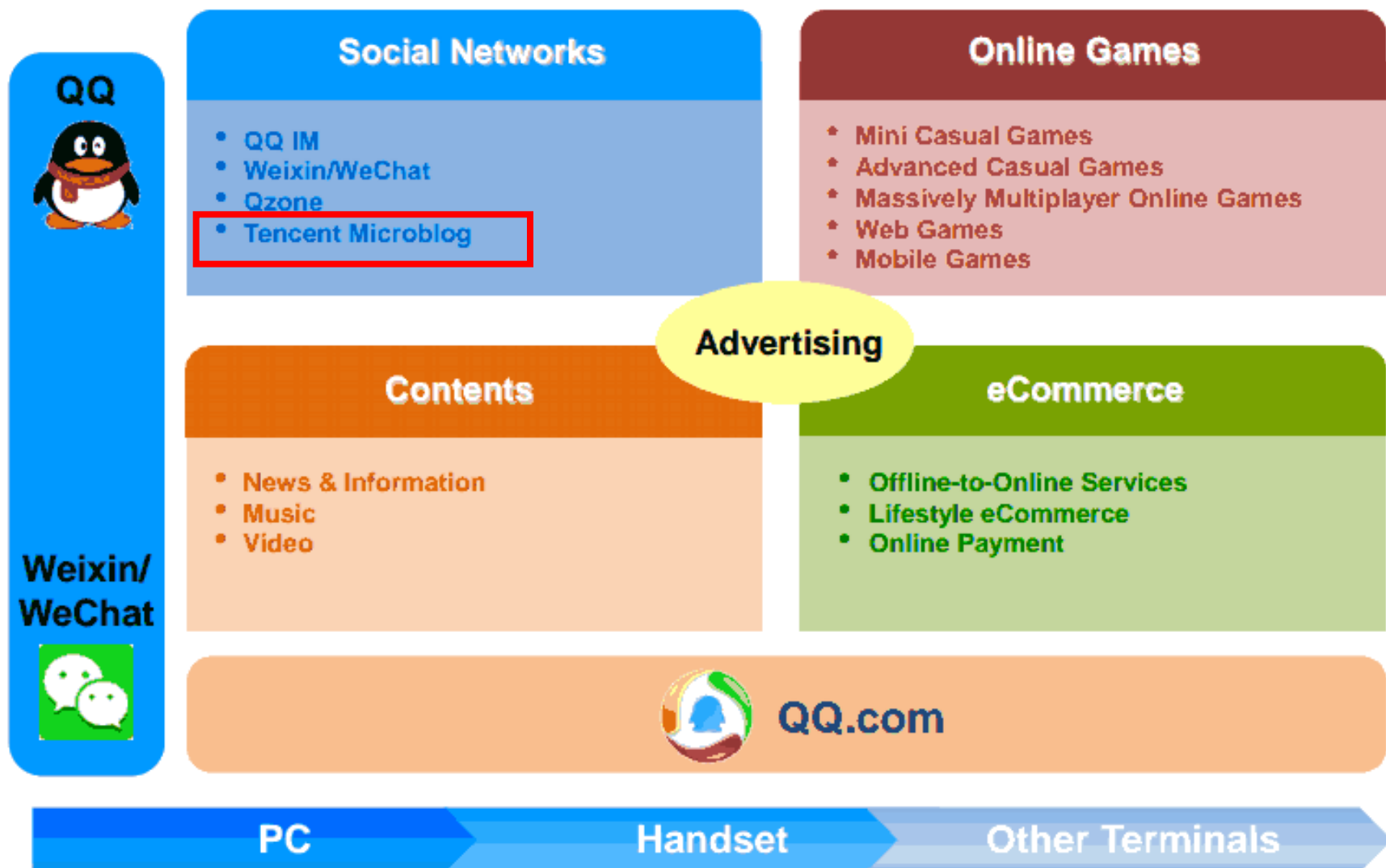
❑S4. Multi-level marketing

Social Media, say, Twitter-like Network?

- ❑ Q: Do we spot these in social media advertising?
- ❑ S1. Celebrity branding
- ❑ S2. Collaborative advertising
- ❑ S3. Gift advertising
- ❑ S4. Multi-level marketing

Diversified Product Portfolio

Fulfilling Online Lifestyle Needs

QQ

Weixin/WeChat

Social Networks

- QQ IM
- Weixin/WeChat
- Qzone
- Tencent Microblog

Online Games

- Mini Casual Games
- Advanced Casual Games
- Massively Multiplayer Online Games
- Web Games
- Mobile Games

Advertising

Contents

- News & Information
- Music
- Video

eCommerce

- Offline-to-Online Services
- Lifestyle eCommerce
- Online Payment
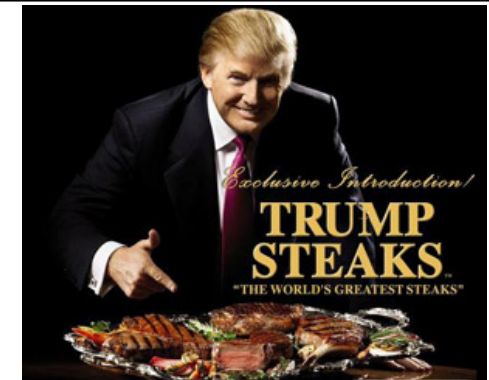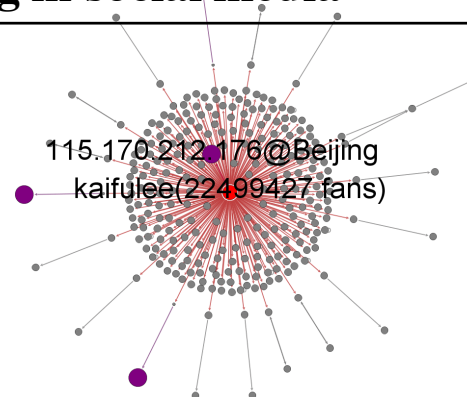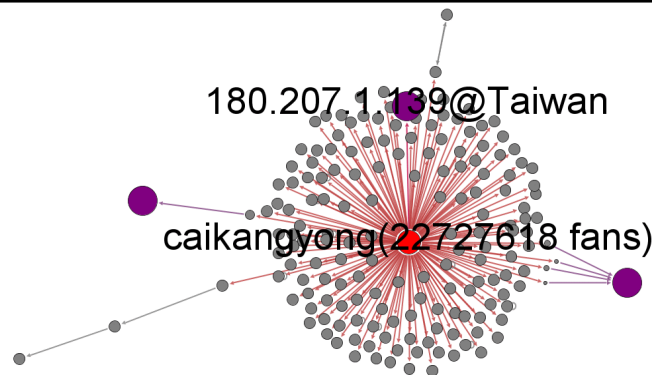
QQ.com

PC → **Handset** → **Other Terminals**

This Talk

❑Tech difference: for Twitter/Weibo, marketers can set up botnets and scripts

❑Q1: Do they take those 4 strategies? Do they have different strategies using botnets and scripts?

❑Q2: Can we detect the botnet advertisers? How accurate? Are botnets more smarter than advertisers that use the 4 strategies?

Comparing Social Media Advertising and Traditional Advertising: S1

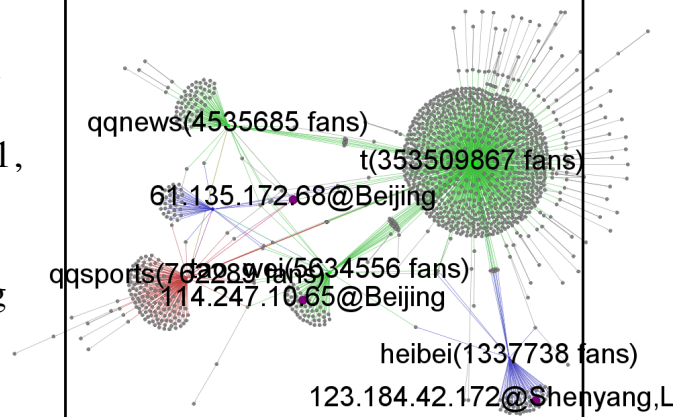| Celebrity branding/advertising in social media		Old-school celebrity branding
180.207.1.139@Taiwan caikangyong(22727618 fans)	115.170.212.176@Beijing kaifulee(22499427 fans)	
Statistics: The root user *@caikangyong* has 22.7M followers. The thread infected 5,138 users (**97.6% from followers**) and 3,246 devices (**97.6% from followers**). The number of retweets is 5,214 (**97.4% from followers**).	**Statistics:** The root user *@kaifulee* has 22.5M followers. The thread infected 6,346 users (**95.7% from followers**) and 5,357 devices (**97.3% from followers**). The number of retweets is 6,425 (**95.3% from followers**).	**Strategy:** celebrity branding is a type of advertising in which a celebrity becomes a brand ambassador and uses his or her status in society to promote or endorse a product, service or charity. (Wikipedia)

Comparing Social Media Advertising and Traditional Advertising: S2

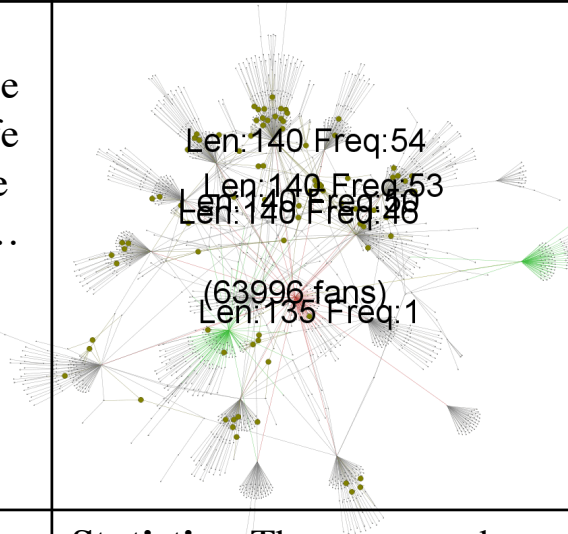| Collaborative advertising in social media		Old-school collaborative advertising
Retweet: by (1) @t with 353.5M followers, (2) @qqnews with 4.5M followers, (3) @heibei with 1.3M followers, (4) @taowei with 5.6M followers **Root tweet:** #China 0-1 Iraq# November 11, 2014 Brazil World Cup Asian Zone 20 the fourth round, the Chinese team 0-1 Iraq, Younis at stoppage time scored the winning goal. **http://url.cn/1qkiWX** Root user @qqsports has 762K followers.	qqnews(4535685 fans) t(353509867 fans) 61.135.172.68@Beijing qqsports(762289 fans) taowei(5634556 fans) 114.247.10.65@Beijing heibei(1337738 fans) 123.184.42.172@Shenyang,L	
Strategy: in social media, no matter how popular an account is, only his followers can directly receive his message. The more big nodes in the network retweet/share the content and URL, the larger number of the infected followers the message has.	**Statistics:** The root user has 762,289 followers. The thread infected 21,807 users (**9.0% from followers**) and 15,446 devices (**7.8% from followers**). The number of retweets is 23,625 (**9.1% from followers**).	**Strategy:** collaborative advertising is the process of sharing the same goal to increase brand and influence. For example, recruiting several famous basketball players can promote the product in different channels (i.e., the fan groups of the stars).

Comparing Social Media Advertising and Traditional Advertising: S3

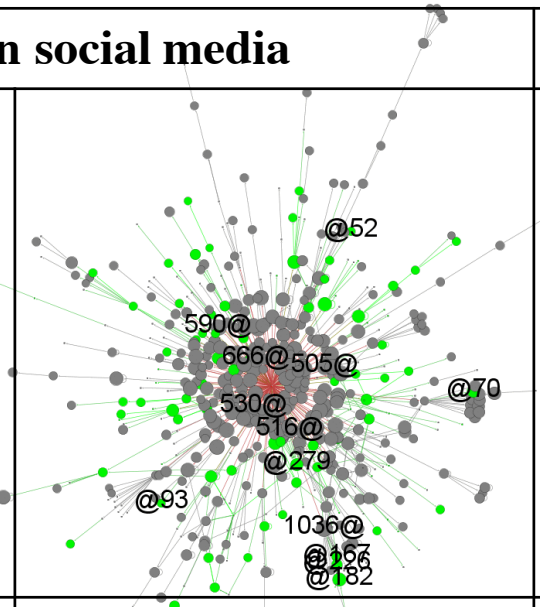| Gift advertising in social media		Old-school gift advertising
Retweet: Our failures, one to the wrong time, the second is lost to the distance, the third is lost to himself. We always feel life is very long, very far away tomorrow, the dream will never be realized in waiting … **Root tweet:** [6 tips to have thin legs] … (2) jogging should be uniform, not when the fast slowly; … (5) … at a 90 degree angle standing; … ❤ Want to know how to become a beauty Follow @zsyjkbk	Len:140 Freq:54 Len:140 Freq:53 Len:140 Freq:48 (63996 fans) Len:135 Freq:1	$100 Amazon.com Gift Card amazon.com gift card $100
Strategy: when the root user's followers retweet the message, they add **irrelevant but attractive** content (e.g., about "failure", "life", "dream") as a "**gift card**" to replace the original text. Thus, the message can be widely diffused over the online network.	**Statistics:** The root user has 63,996 followers. The thread infected 12,977 users (**4.5% from followers**) and 7,911 devices (**6.2% from followers**). The number of retweets is 14,904 (**7.6% from followers**).	**Strategy:** if you register an account or purchase a product, you will get a $100 gift card. Such a strategic behavior has been used to attract customers to a new business since the Mobil Oil Company introduced the first retail gift card in 1995.

Comparing Social Media Advertising and Traditional Advertising: S4

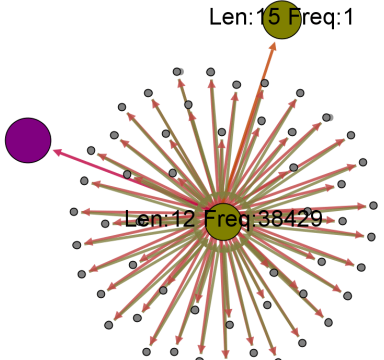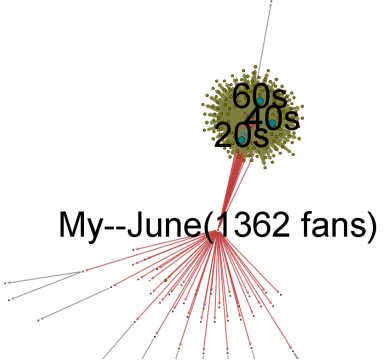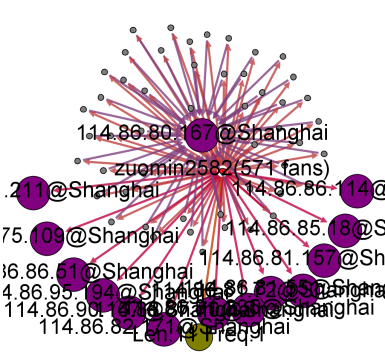| Multi-level marketing in social media		Old-school multi-level marketing
Retweet: #Qunaer# fair! good luck, Come together!@ lixianglin131407 @ snes986 @ cornett20096 **Root tweet:** #Qunaer# activity Awards: #ipad2 # Participation: 1) become a fan @qunaer 2) forwarding this entry microblogging AT five friends; --- activities after the end of a winner drawn, presented an award ipad2! fair and just! Publication time: at 16:00 on Dec 30th		
Strategy: the company manipulates their accounts to mention legitimate users with "@XXX". The mentioned users become the "downline" to diffuse the message in multiple levels of the network. The message often refers to an activity of continue mentioning friends.	**Statistics:** The root user has 113,026 followers. The thread infected 1,060 users (**69.2% from followers**) and 1,013 devices (**45.2% from followers**). The number of retweets is 14,282 (**91.5% from followers**).	**Strategy:** Multi-level marketing is a marketing strategy in which the sales force is compensated not only for sales they generate, but also for the sales of people that they recruit. This recruited sales force (referred to "downline") provide multiple levels of compensation.

Solution 5: Synchrony Strategies in Social Media

| Comment synchrony	URL synchrony	Time synchrony	Device synchrony
similar comments	Len:15 Freq:1 Len:12 Freq:38429	60s 40s 20s My--June(1362 fans)	114.86.80.167@Shanghai zuomin2582(571 fans) .211@Shanghai 114.86.86.114@ 75.109@Shanghai 114.86.85.18@S 86.86.61@Shanghai 114.86.81.157@Sha 4.86.95.194@Shanghai 114.86.90.185@Shanghai 114.86.82.27@Shanghai Len:14 Freq:9
Strategy: (Galaxy Note ads) The botnets frequently retweet with similar phrases.	**Strategy:** (Porn URL) The botnets frequently retweet the same URL.	**Strategy:** (Galaxy Note ads) The botnets operate in lockstep with several fixed time intervals.	**Strategy:** (Galaxy Note ads) The botnets operate on the same group of devices in Shanghai.
Statistics: The root user has 8,573 followers. The thread infected 3,059 users (**98.4% from followers**) and 18 devices (**61.1% from followers**). The number of retweets is 9,777 (**99.5% from followers**).	**Statistics:** The thread infected 38,122 users (**100% from followers**) and 18,008 devices (**99.989% from followers**). The number of retweets is 38,432 (**99.995% from followers**).	**Statistics:** The root user has 1,362 followers. The thread infected 598 users (**98.0% from followers**) and 12 devices (**100% from followers**). The number of retweets is 4,202 (**99.6% from followers**).	**Statistics:** The root user has 571 followers. The thread infected 676 users (**100% from followers**) and 42 devices (**100% from followers**). The number of retweets is 9,504 (**100% from followers**).

Observation: Features

❑ Comment features and mentioning features
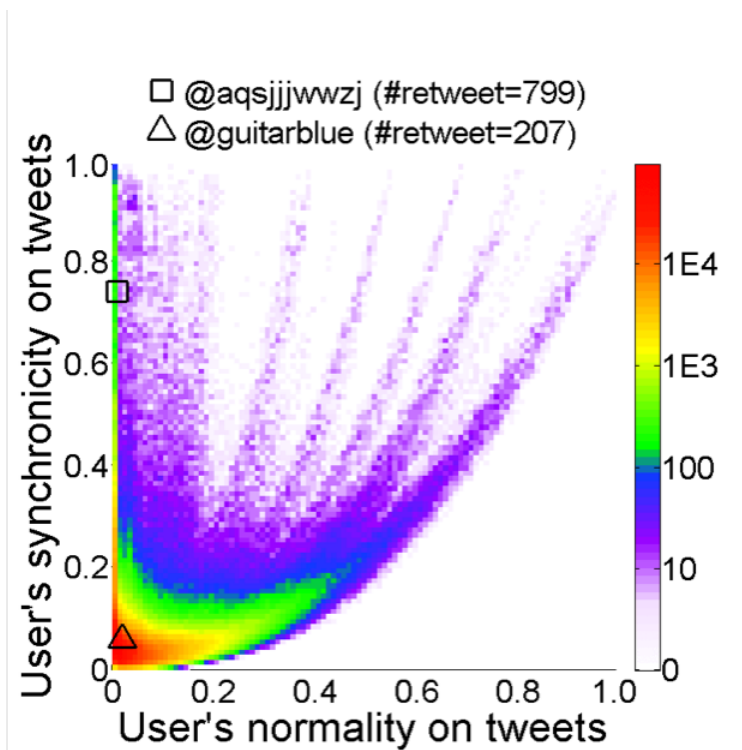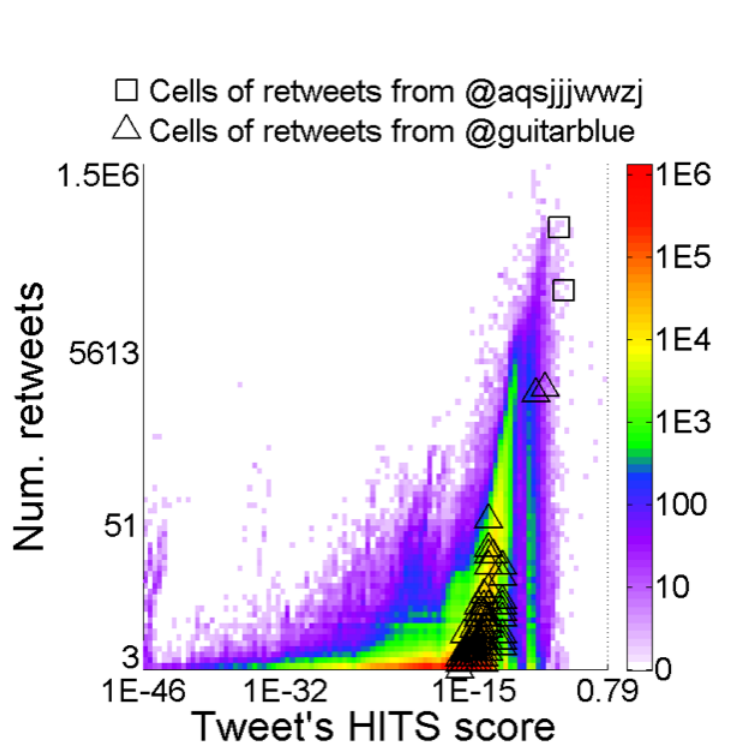   ❑ Every dot is a "tweet cascade"
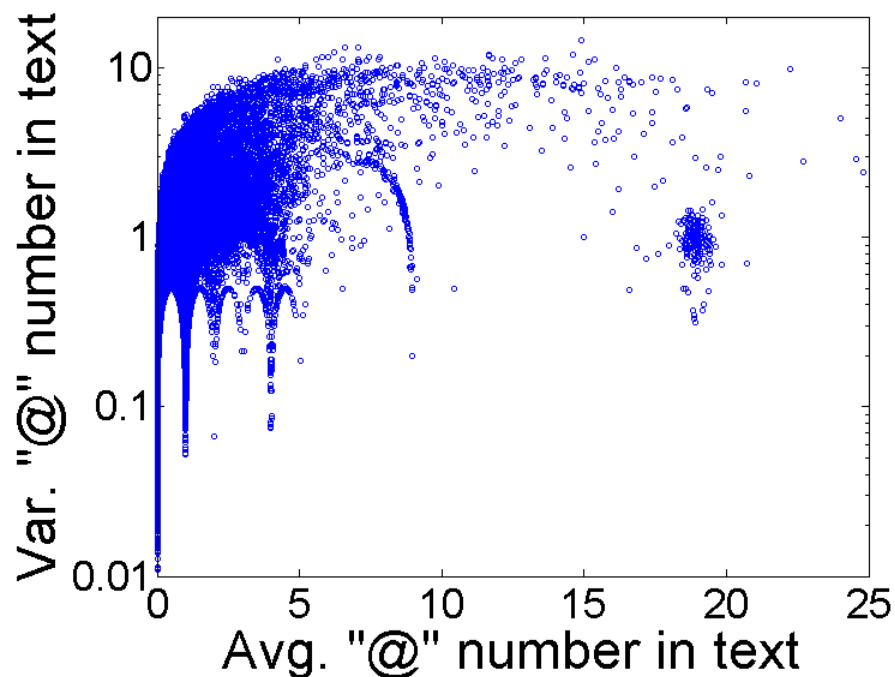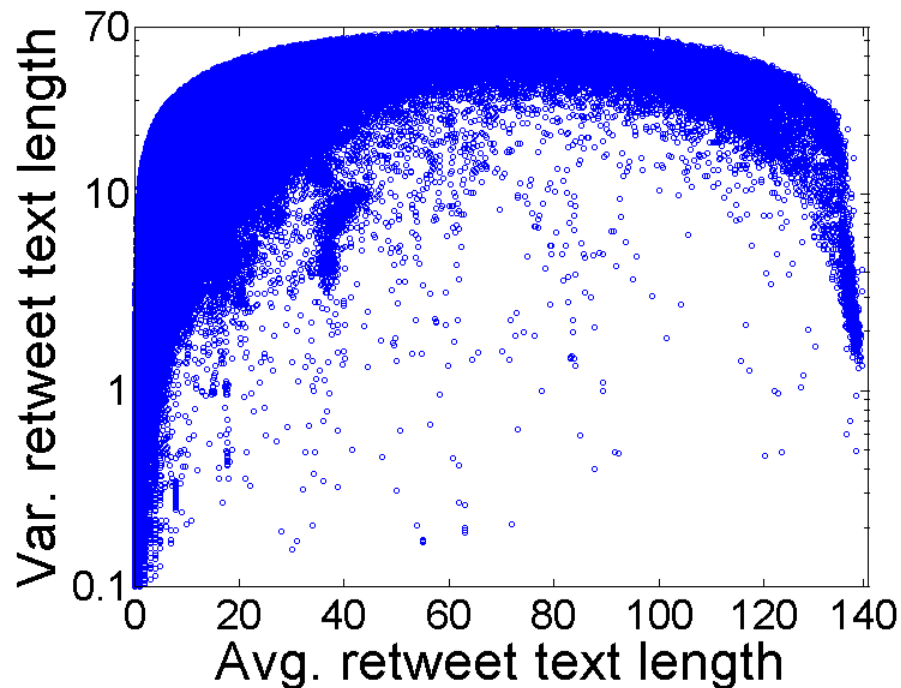
Observation: Features

❑ Lockstep behavioral features
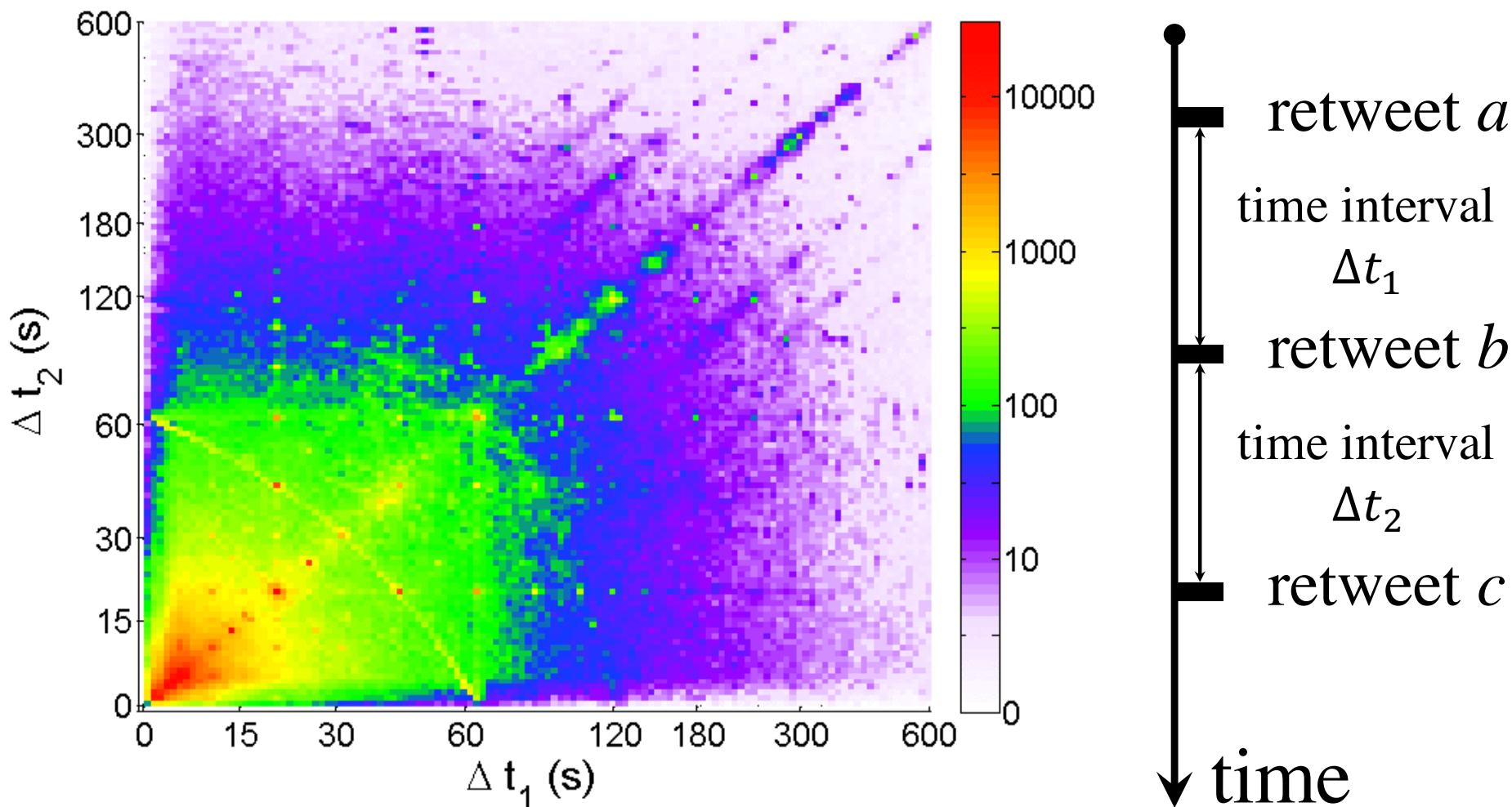
Observation: Features

❑ Message synchrony [CatchSync, KDD'14]

Modeling: Features

Advertising strategy	Feature definition
S1: Celebrity branding	F1: the number of followers of the root users
S2: Collaborative advertising	$F2(k)$: the number of infected users who have more than k followers
	F3: the number of infected users who share the same device of the root user
S3: Gift advertising	F4: the largest frequency of the length of the comments
	F5: the most frequent length of the comments
S4: Multi-level marketing	$F6(k)$: the number of users who were mentioned more than k times
	$F7(k)$: the number of users who mentioned more than k users in total
	F8-F9: the average value (variance) of the number of mentions in the comments
S5-1: Comment synchrony	F10-F11 : the message synchronicity (normality) of the user [2]
	F12-F13: the average value (variance) of the length of the retweet comments by the user
S5-2: URL synchrony	F14: the percentage of comments that have at least one URL by the user
	F15-F16: the average value (variance) of the number of URLs in the comments by the user
S5-3: Time synchrony	F17: the most frequent time interval Δ_t between two retweets in a thread by the user
	F18-F19: the average value of the number of retweets (the time period) in a thread by the user
S5-4: Device synchrony	F20: if the user operates on the most frequent device in a thread

Experimental Results

❑**T1:** Which strategy the tweet used? 5-class classification.

Method	Parameters	Accuracy
Random	-	0.200
S1 (F1)	-	0.447
S2 (F2-F3)	F2(1000)	0.326
S3 (F4-F5)	-	0.352
S4 (F6-F9)	F6(5), F7(10)	0.257
S5 (F10-F20)	-	0.525
SocAdDet (F1-F20)	F2(100), F6(5), F7(10)	0.852
	F2(1000), F6(2), F7(10)	0.855
	F2(1000), F6(5), F7(5)	0.867
	F2(10000), F6(10), F7(20)	0.776
	F2(1000), F6(5), F7(10)	**0.889**

S1: Celebrity branding	F1: the number of followers of the root users
S2: Collaborative advertising	F2(k): the number of infected users who have more than k followers
	F3: the number of infected users who share the same device of the root user
S3: Gift advertising	F4: the largest frequency of the length of the comments
	F5: the most frequent length of the comments
S4: Multi-level marketing	F6(k): the number of users who were mentioned more than k times
	F7(k): the number of users who mentioned more than k users in total
	F8-F9: the average value (variance) of the number of mentions in the comments

Experimental Results

❑**T2:** Detecting botnet advertisers (spammers in synchrony): binary classification.

Method	Accuracy
CatchSync [2] (F10-F11)	0.725
Comment synchrony (F10-F13)	0.796
URL synchrony (F14-F16)	0.725
Time synchrony (F17-F19)	0.831
Device synchrony (F20)	0.645
SocAdDet (F10-F20)	**0.923**

S5-1: Comment synchrony	F10-F11 : the message synchronicity (normality) of the user [2] F12-F13: the average value (variance) of the length of the retweet comments by the user
S5-2: URL synchrony	F14: the percentage of comments that have at least one URL by the user F15-F16: the average value (variance) of the number of URLs in the comments by the user
S5-3: Time synchrony	F17: the most frequent time interval Δ_t between two retweets in a thread by the user F18-F19: the average value of the number of retweets (the time period) in a thread by the user
S5-4: Device synchrony	F20: if the user operates on the most frequent device in a thread

Summary

- Comparing social media advertising strategies and traditional strategies
    - Celebrity branding
    - Collaborative advertising
    - Gift advertising
    - Multi-level marketing
    - Synchrony (social media botnet)
- Strategy classification: 0.889
- Spammer detection: 0.923
- Faster but not smarter!
- *Future work*: Can we early predict?!

Thank you!

Catching Social Media Advertisers with Strategy Analysis