

Stegobot: a covert social-network botnet

Shishir Nagaraja
Network and Distributed Systems Security Group
IIIT Delhi, India

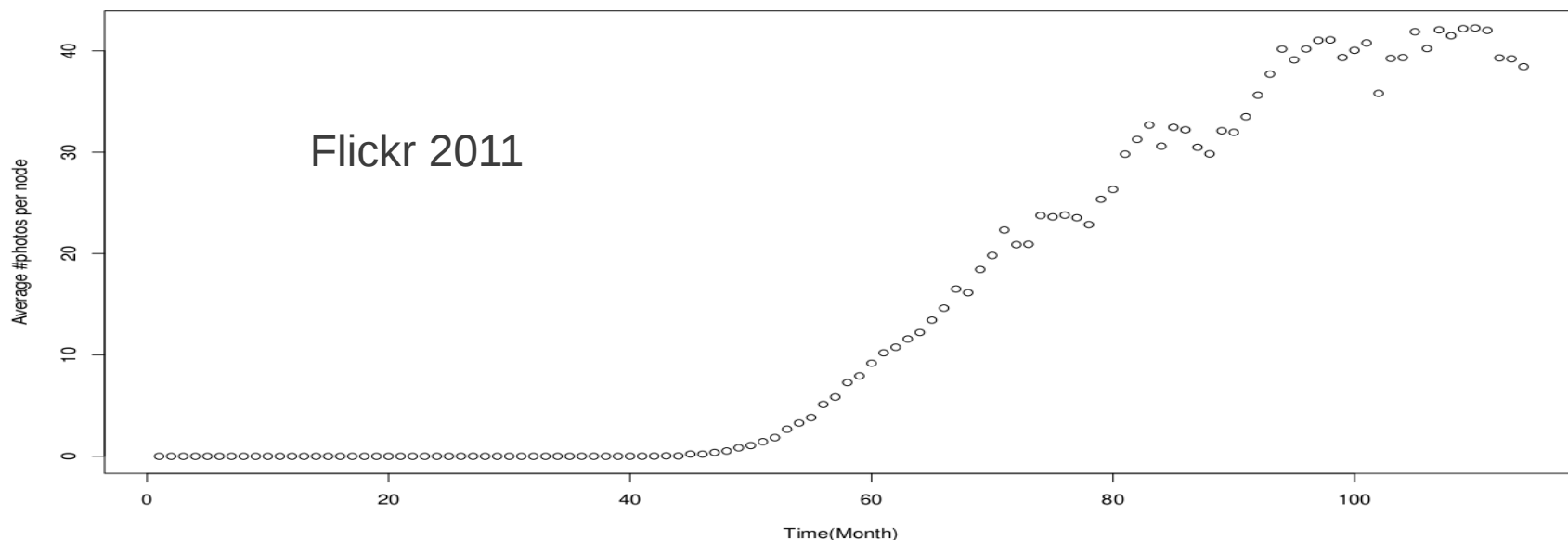
<http://www.hatswitch.org/~sn275>

Botnets

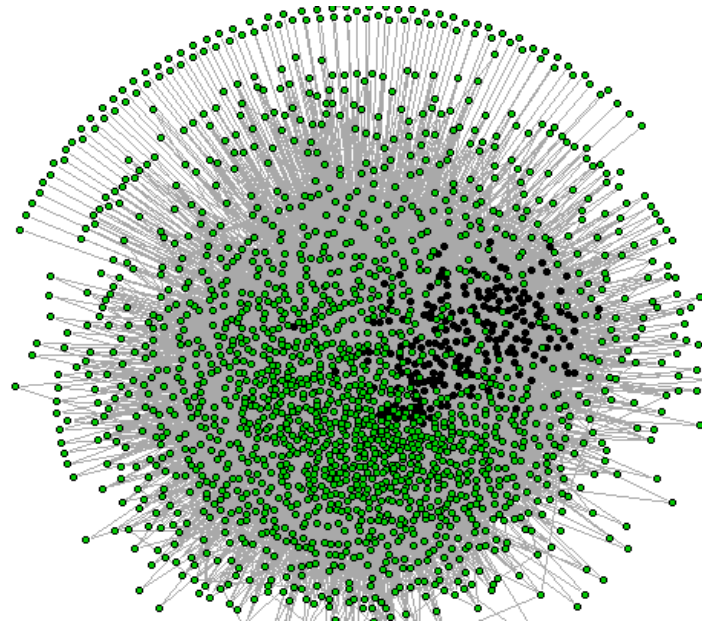
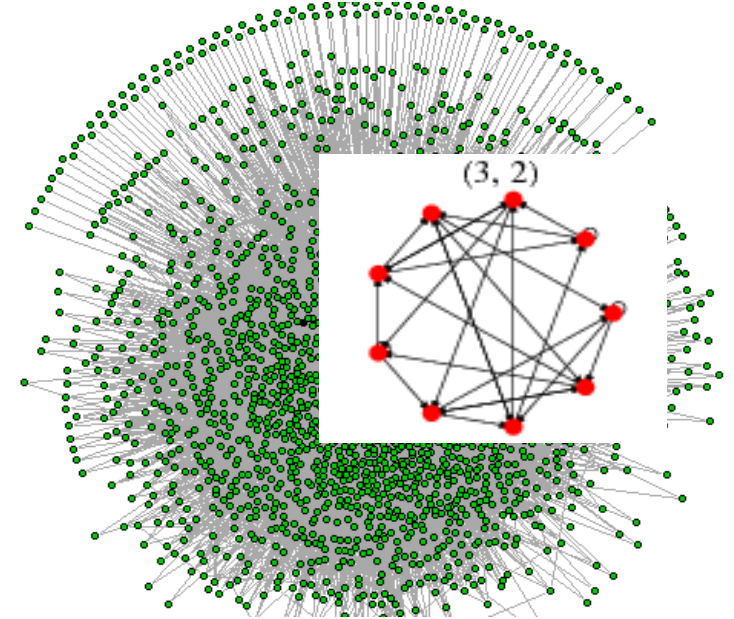
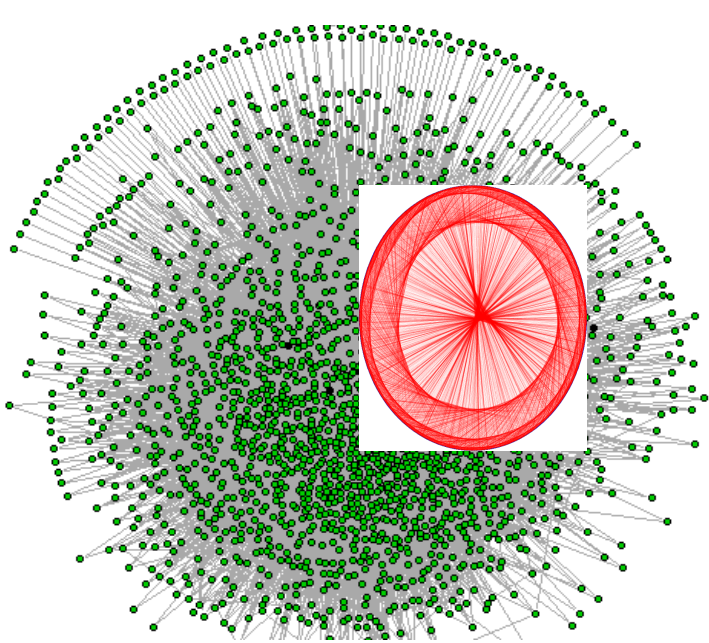
- Primary vehicle in online crime, DDOS attacks and information theft
- Social malware attacks is an emerging trend: Dalai Lama got attacked in 2008, Google in 2009 and 800 or so others were targets in 2010
- Botnets and anonymous communication networks have similar network properties: availability, resilience and undetectable C&C traffic.
- Standard threat model – global passive adversary

Designing a covert botnet

- Can we design a botnet using stego channels?
- New traffic links lower traffic analysis resistance
- New traffic patterns lower traffic analysis resistance
- Core idea: infect machines using social malware + use social image exchange behavior on OSN to create unobservable communication channels between infected machines



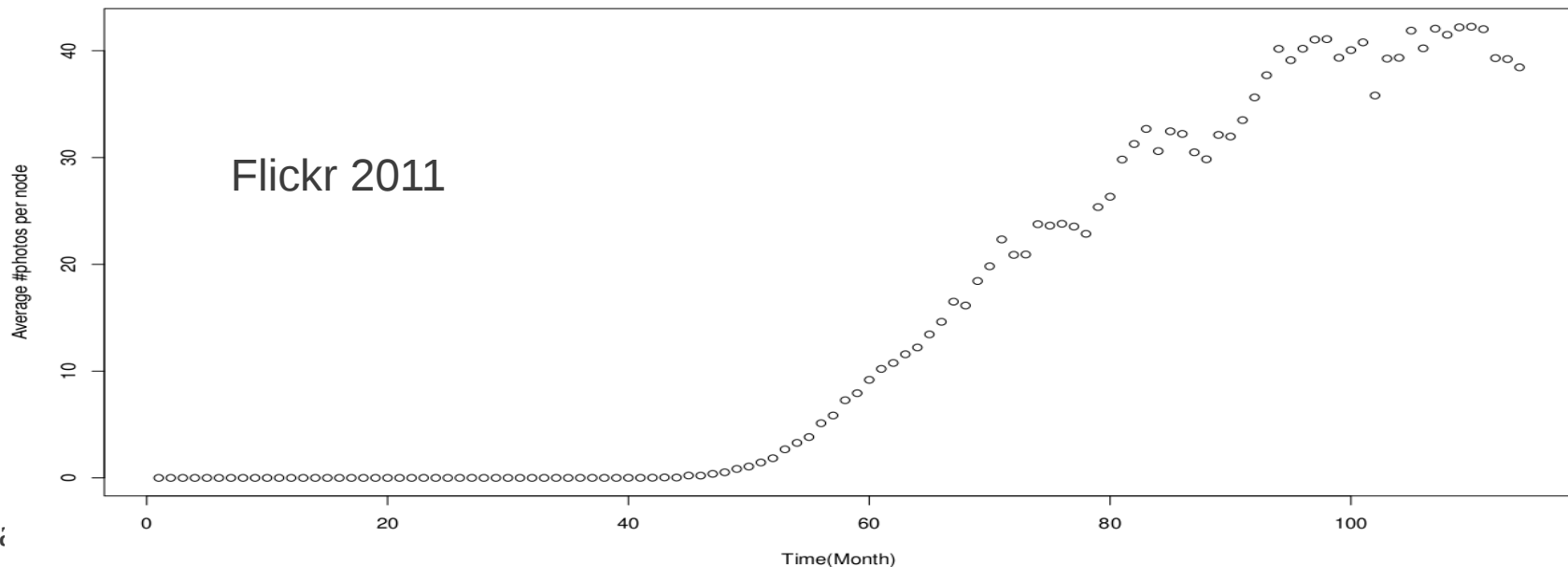
Botnet topologies



- C&C traffic
- Attack traffic

Designing a covert botnet

- Can we design a botnet using stego channels?
- New traffic links lower traffic analysis resistance
- Core idea: infect machines using social malware + use social image exchange behavior on OSN to create unobservable communication channels between infected machines



Attack vector (targeted malware)

- Hijack social trust
 - steal an email with an attachment
 - embed malware in the attachment
 - send/resend the email to the target
- Initial break
 - Social phish constructed with public information
 - Once the attacker gains a foothold, neighbors within the social network of the victim are compromised

Sample subverted email designed to achieve a foothold

Subject: Kalon Tripa Succession
From: "Pema Rinzin" <prinzintibet@yahoo.com>
Date: Thu, September 18, 2008 8:14 am
To: choejor@dalailama.com

Dear Sir,

Attached please find the final Tibetan translation of my English announcement for the Kalon Tripa succession initiative. Response to my press release on September 2nd has been very positive and I have been receiving lots of email and phone messages from Tibetans everywhere.

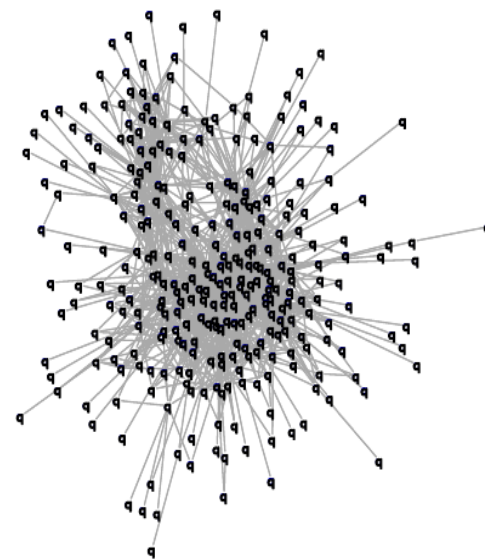
I am trying to get someone to translate the Kalon Tripa Hochoe into English, but if you already have it translated, please send it to me.

Any advice from you in this initiative of mine would be greatly appreciated.

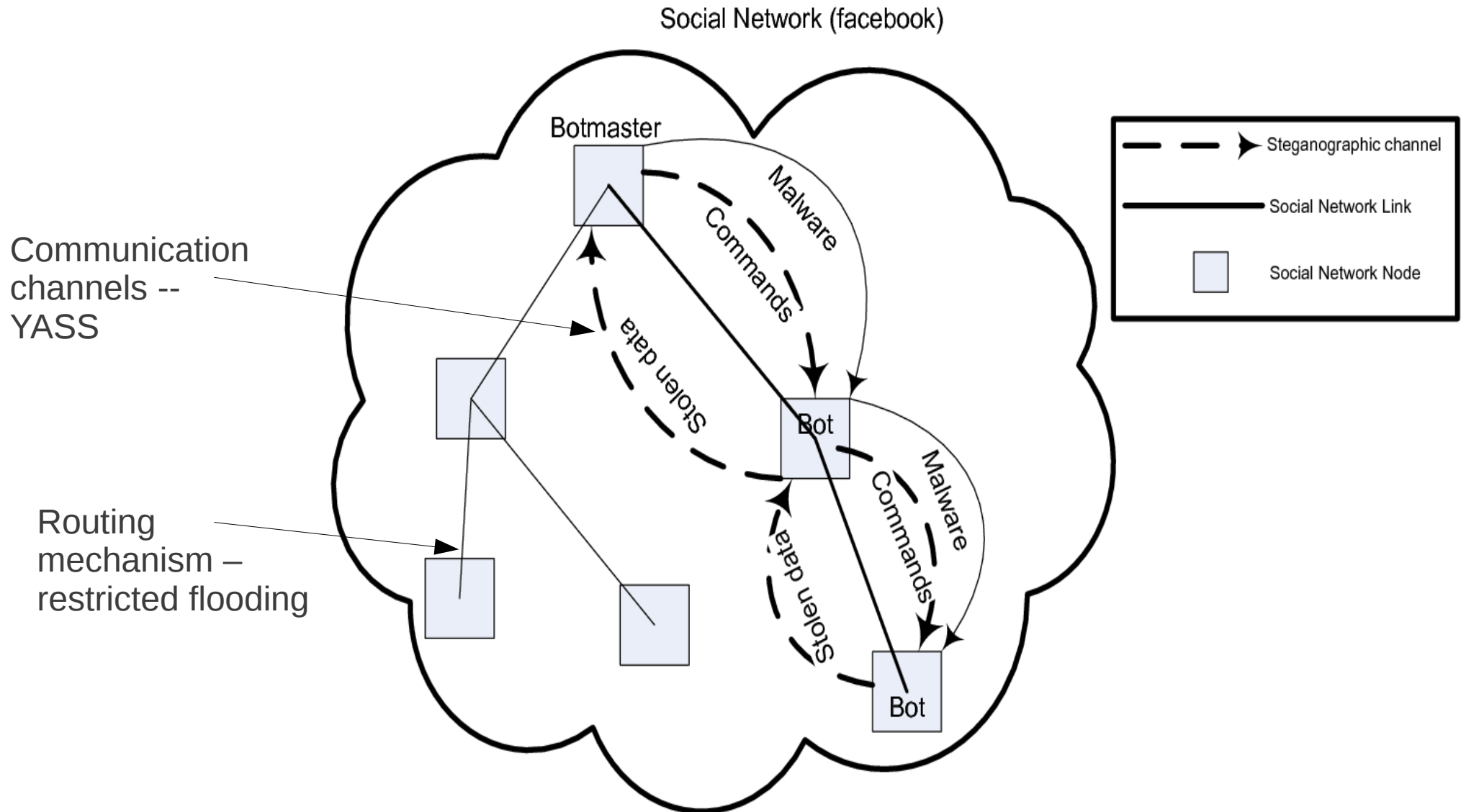
Yours sincerely,

Pema Rinzin
President
TAC

Official Photographer/webmaster
Office of His Holiness the Dalai Lama
Thekchen Choeling
P/O Mcleod ganj 176219
Dharamsala (H.P.)
India

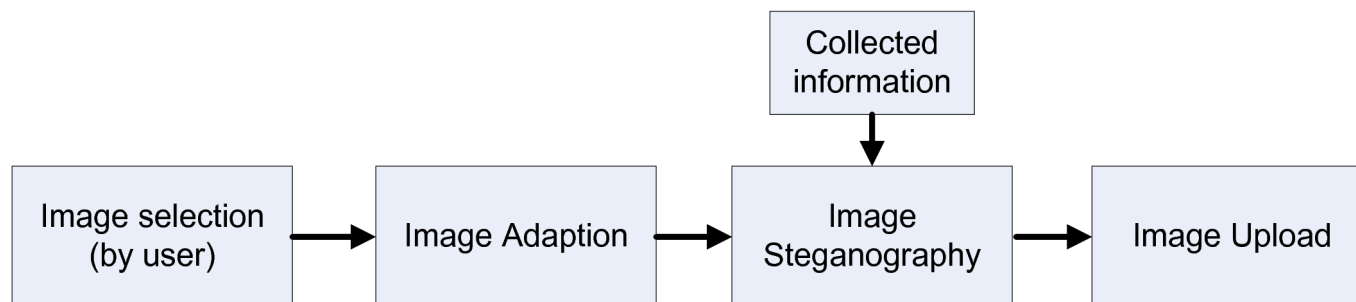


Stegobot architecture



Channel design

- Malware intercepts facebook image upload and embeds credit card information into it. FB sends notification to all neighbours.
- Image processing engine interference
- Facebook predictively caches images when neighbour visits victim page
- Channel efficiency is evaluated using the BER metric: $\frac{\#error\ bits}{\#total\ bits}$
- No interference: Stegobot doesn't upload or download the pictures



YASS parameters

Q – compression; q -- redundancy

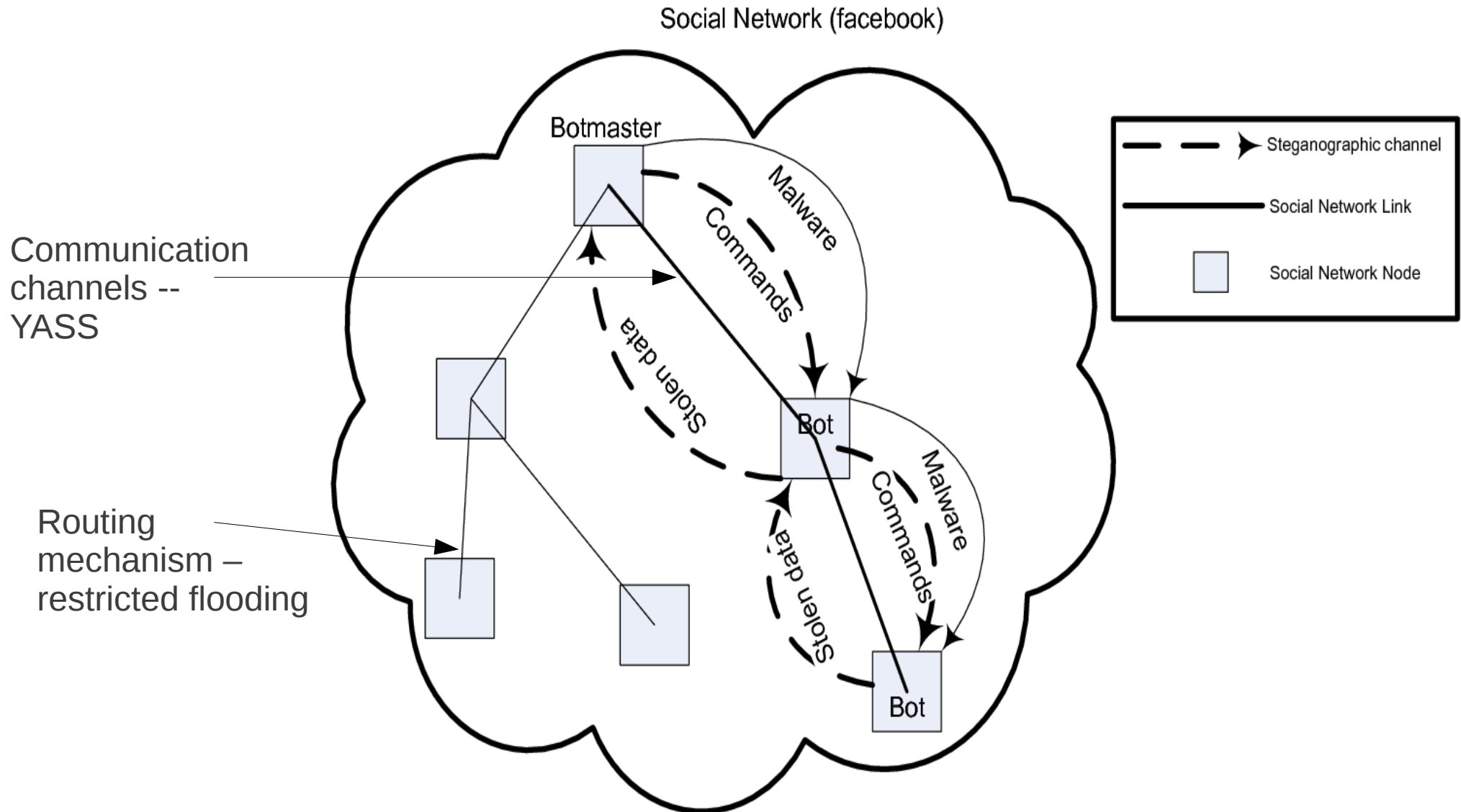
Table 1: Average BER (over 116 images) without removing 'bad images'

q	2	4	6	8	10	12	14	16	18	20
Q=65	0.3073	0.1320	0.0520	0.0227	0.0097	0.0047	0.0022	0.0010	0.0006	0.0003
Q=70	0.2966	0.1318	0.0529	0.0219	0.0096	0.0049	0.0025	0.0010	0.0005	0.0002
Q=75	0.3015	0.1557	0.0680	0.0283	0.0101	0.0067	0.0027	0.0010	0.0004	0.0000
Q=80	0.3086	0.1839	0.0846	0.0347	0.0143	0.0089	0.0034	0.0015	0.0008	0.0000
Q=85	0.3512	0.2618	0.1777	0.0854	0.0372	0.0183	0.0127	0.0053	0.0024	0.0013
Q=90	0.4287	0.3917	0.3639	0.3390	0.3146	0.2906	0.2567	0.2122	0.1591	0.1262

Table 2: Number of bits inserted in each image for different values of q

q	2	4	6	8	10	12	14	16	18	20
Data bits	40280	20140	13426	10070	8056	6173	5754	5035	4475	4028

Stegobot architecture

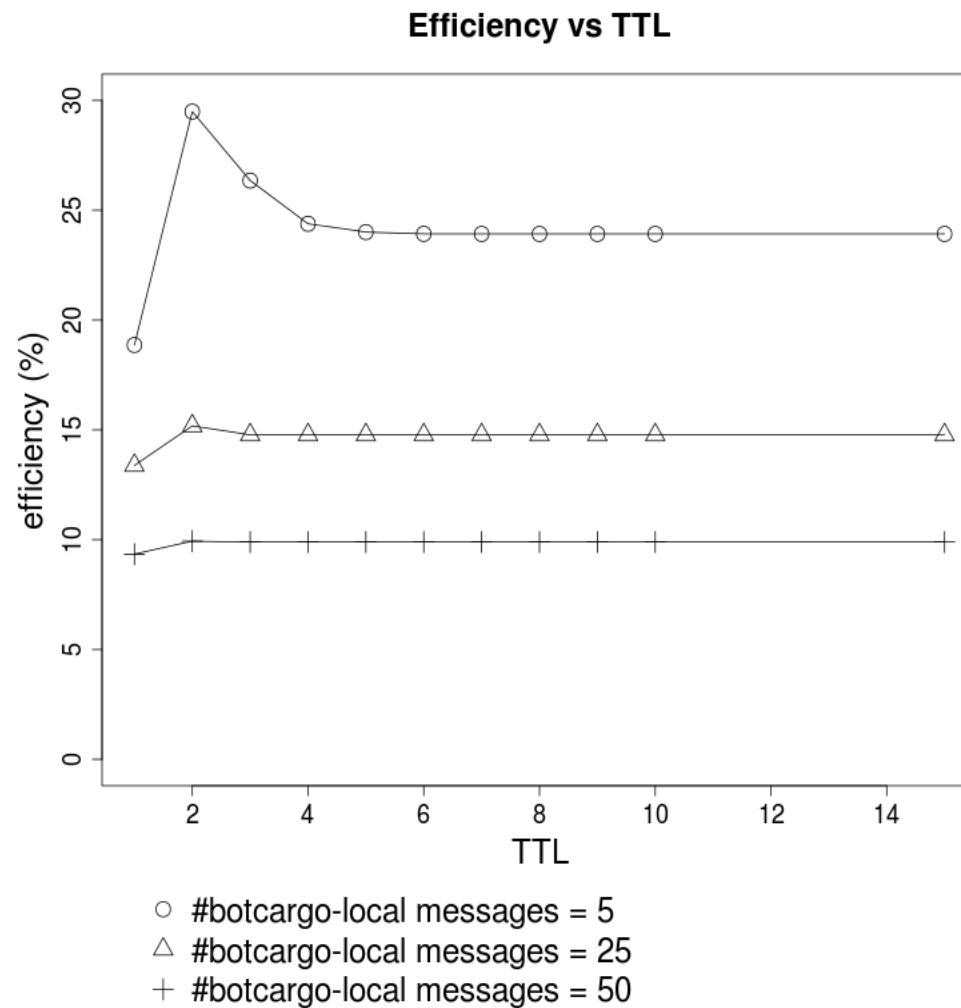


Routing mechanism

- Dataset: Flickr social network; monthly image posting behavior of ~15000 nodes over 40 months
- Assumed 50% infection, sub-graph of 7200 extracted.
- Now we had to find out if you can build a routing network over this.
- Really simple and robust but non-optimal routing algorithm: restricted flooding with $ttl = \log N$
- message queue: *local message*, *fwd_message*
- Routing efficiency averaged over randomly chosen botmaster nodes; each bot collects **k** image payload units of stolen information **per month**

Routing results

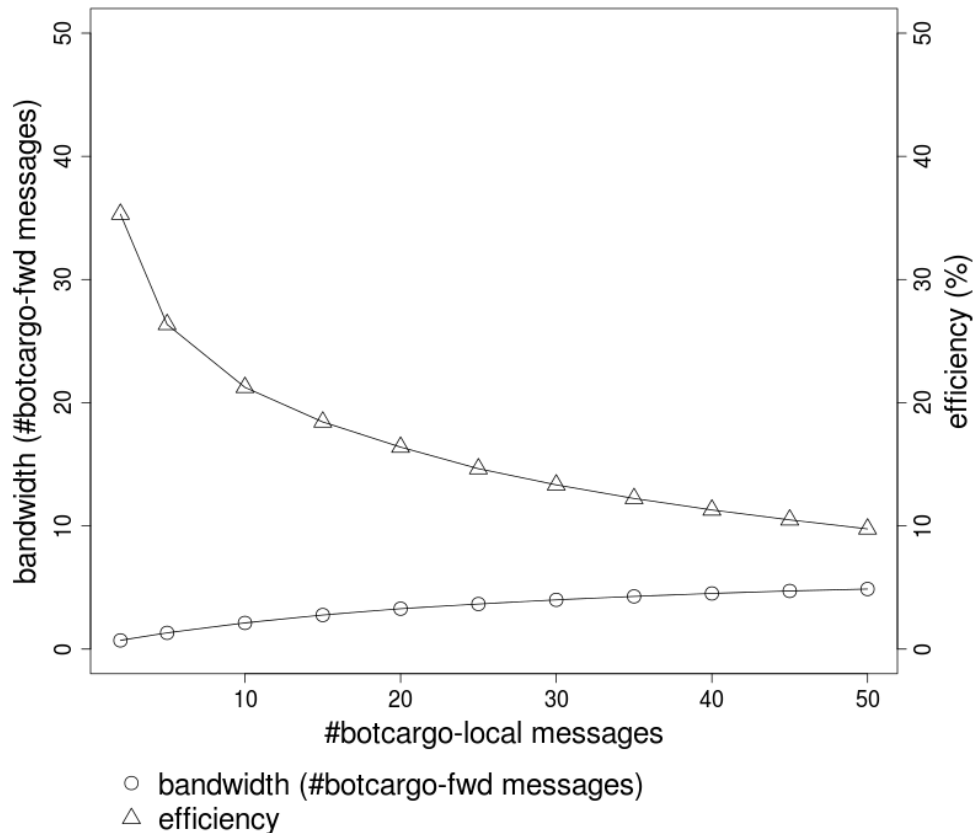
At the bots (efficiency of clearing the local queue)



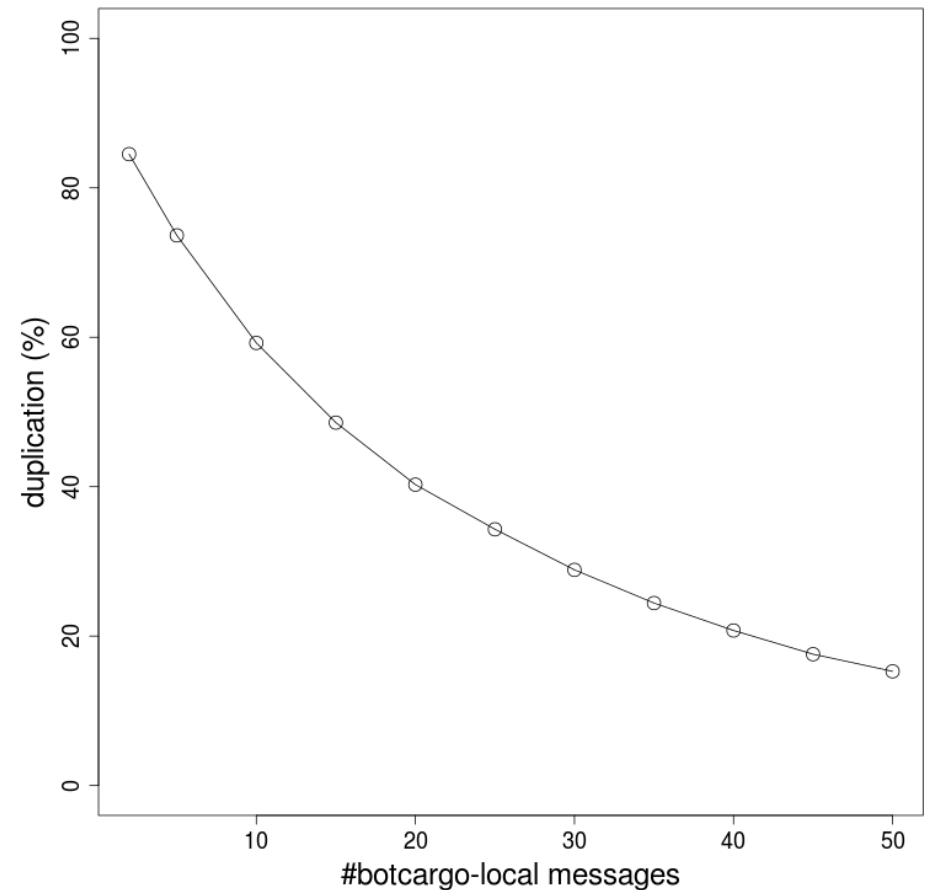
Routing b/w, efficiency, duplication

Bandwidth -- #unique messages reaching the botmaster

Bandwidth and Efficiency vs #botcargo-local messages

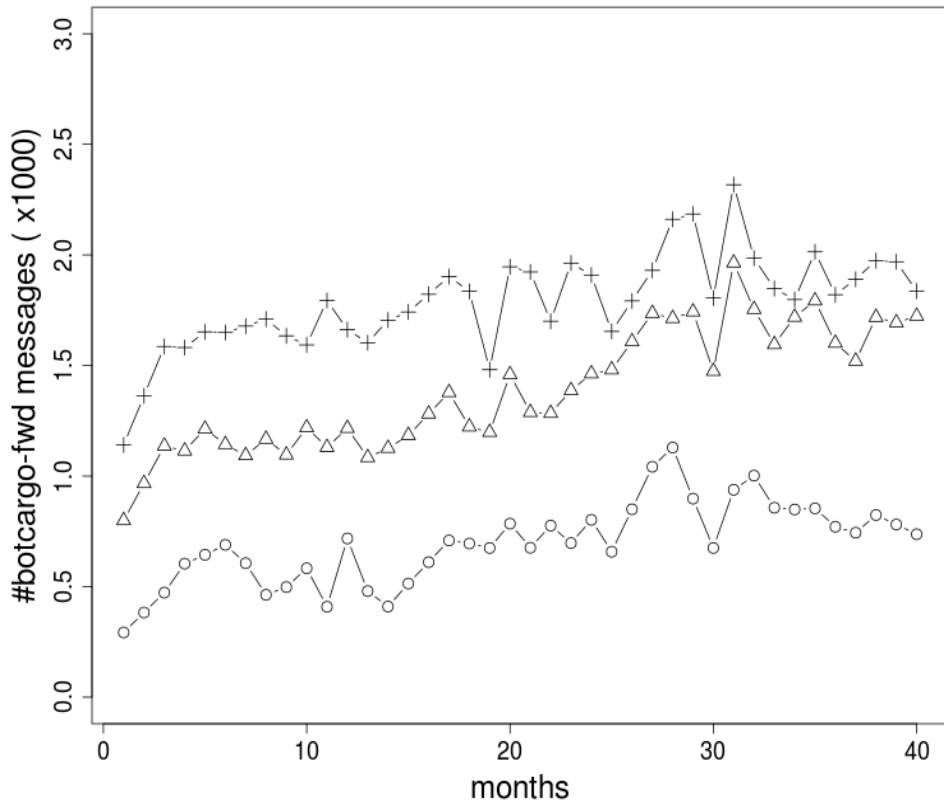


Duplication vs #botcargo-local messages



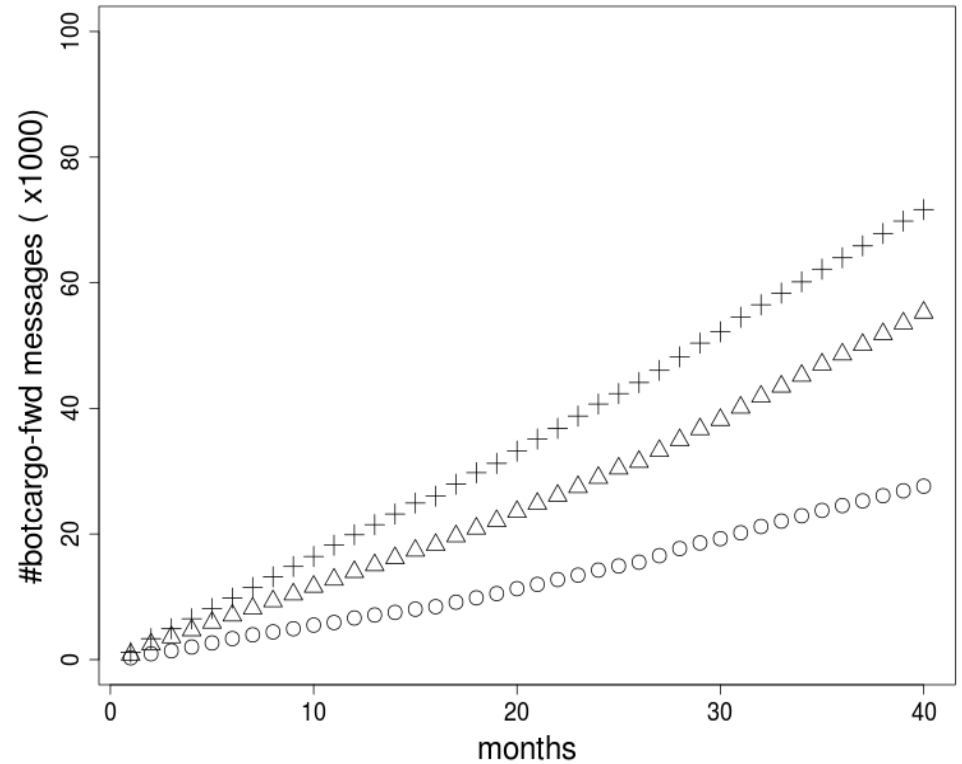
Network bandwidth

#botcargo-fwd messages received at botmaster across months



- #botcargo-local messages = 5
- △ #botcargo-local messages = 25
- + #botcargo-local messages = 50

#botcargo-fwd messages received at botmaster across months (Cumulative Distribution)



- #botcargo-local messages = 5
- △ #botcargo-local messages = 25
- + #botcargo-local messages = 50

Conclusions

- Building distributed systems over steganographic communication channels is fun!
- We have evaluated our proposed wicked system using real-world social behavior data.
- Even with a routing algorithm the botmaster can siphon off 82Mb per month ($q=2$) at the rate of 10kb per 700x700pixel image or 21.6Mb per month ($q=8$).
- Duplication rate of 50-80% indicates that with better routing algorithms much botnet bandwidth could at least be doubled or at best quadrupled.