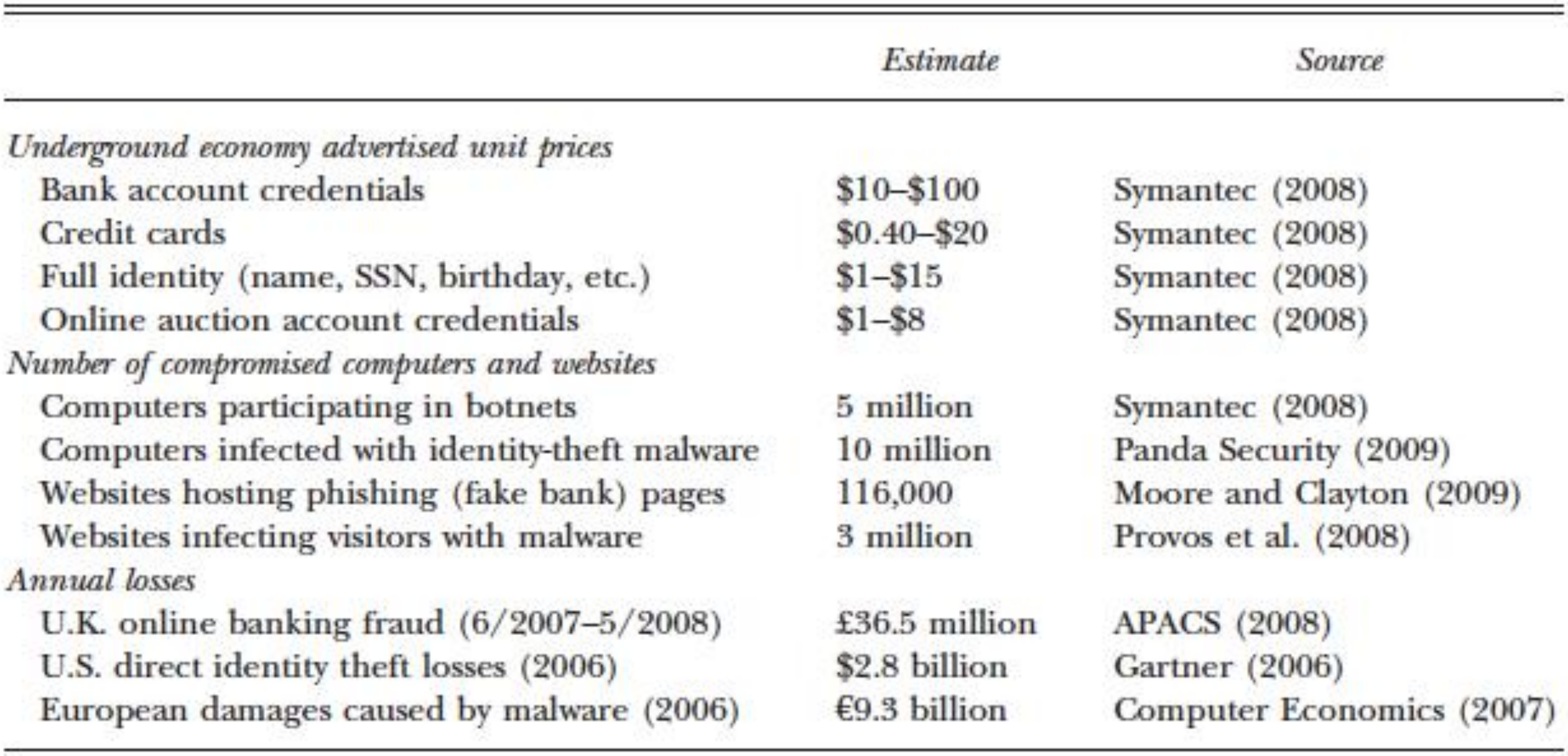
Lecture 27

Markets, Mechanisms and Machines

David Evans and Denis Nekipelov

Information and crime

- Technological progress increases productivity across activities

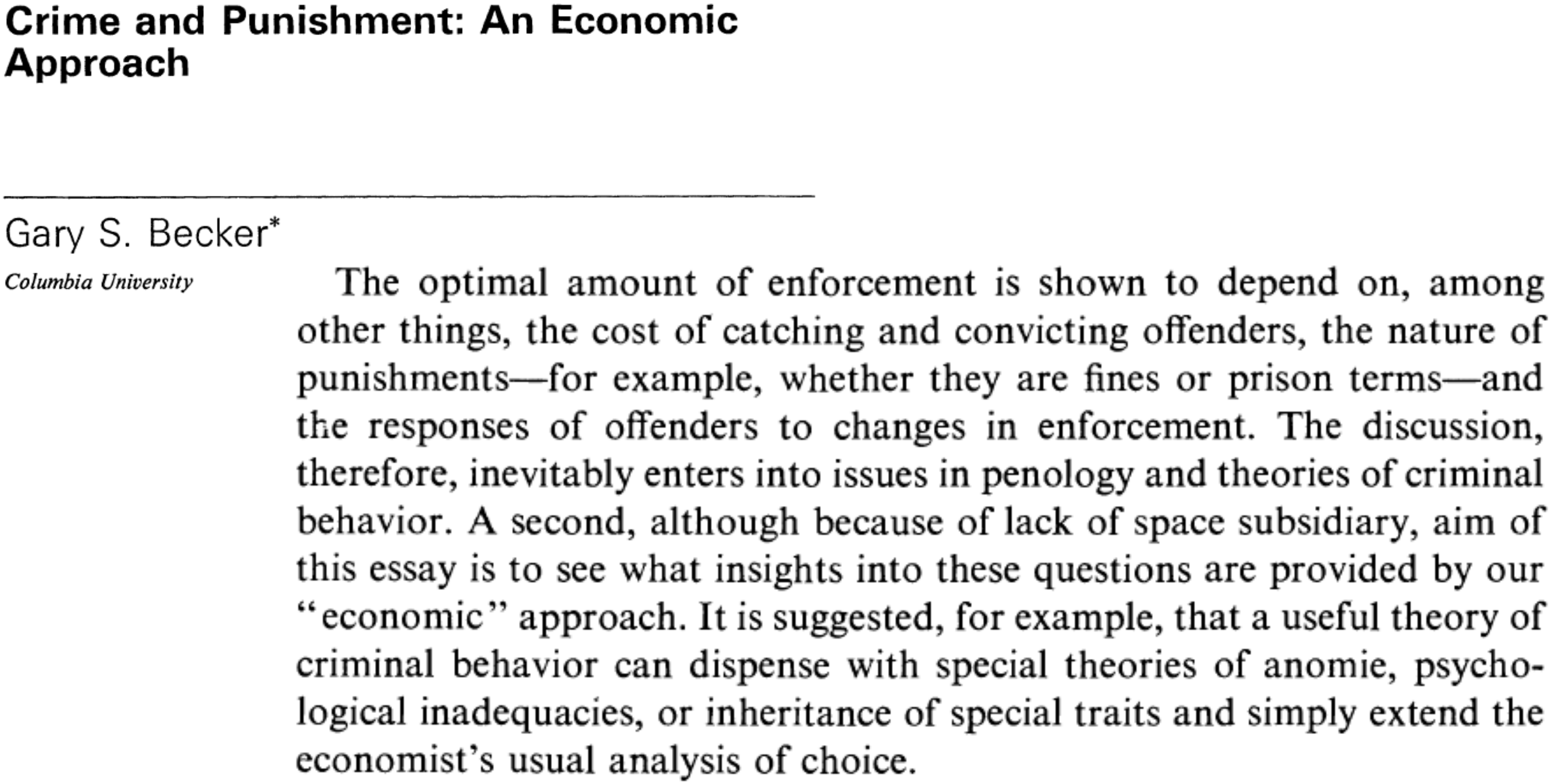


	Estimate	*Source*
Underground economy advertised unit prices		
Bank account credentials	\$10–\$100	Symantec (2008)
Credit cards	\$0.40–\$20	Symantec (2008)
Full identity (name, SSN, birthday, etc.)	\$1–\$15	Symantec (2008)
Online auction account credentials	\$1–\$8	Symantec (2008)
Number of compromised computers and websites		
Computers participating in botnets	5 million	Symantec (2008)
Computers infected with identity-theft malware	10 million	Panda Security (2009)
Websites hosting phishing (fake bank) pages	116,000	Moore and Clayton (2009)
Websites infecting visitors with malware	3 million	Provos et al. (2008)
Annual losses		
U.K. online banking fraud (6/2007–5/2008)	£36.5 million	APACS (2008)
U.S. direct identity theft losses (2006)	\$2.8 billion	Gartner (2006)
European damages caused by malware (2006)	€9.3 billion	Computer Economics (2007)

(Moore, Clayton , Anderson, 2009)

Information and crime

- (Becker, 1968): Crime as a commodity



Crime and Punishment: An Economic Approach

Gary S. Becker*

Columbia University

The optimal amount of enforcement is shown to depend on, among other things, the cost of catching and convicting offenders, the nature of punishments—for example, whether they are fines or prison terms—and the responses of offenders to changes in enforcement. The discussion, therefore, inevitably enters into issues in penology and theories of criminal behavior. A second, although because of lack of space subsidiary, aim of this essay is to see what insights into these questions are provided by our "economic" approach. It is suggested, for example, that a useful theory of criminal behavior can dispense with special theories of anomie, psychological inadequacies, or inheritance of special traits and simply extend the economist's usual analysis of choice.

Information and crime

- Unlike “conventional” crimes, crimes involving information rely on technology
- That makes them potentially more productive
- It also involves different demographic than regular crime also impacting composition of labor force
- With some technologies being digital goods, they can be costlessly copied and shared
- There are also significant network effects
- The impact may affect large communities or companies
- Can be “weponized” to affect entire countries

Information and crime

- In 2007 government of Estonia decided to move monument to Soviet soldier to outskirts of Tallinn
- Russian-language press started spreading information claiming the monument is getting destroyed
- On April 26 riots in Tallinn have started; 156 people were injured, one person died and 1,000 people were detained
- On April 27 April Estonia hit by major cyber-attacks for next several weeks
- Shut down all online services and ATMs of Estonian banks, media and government
- Massive waves of spam were sent by botnets that overwhelmed targeted servers

Information and crime

- Similar attacks of different magnitude occur on daily basis

Map of occurring attacks

- How are they organized?

Spam

- Spam historically refers to unsolicited commercial email and related undesirable online communication

Spam

- Typical advertising is method for “monetizing” content or service that is valued by consumer
- Spam imposes an externality on consumers without benefit and ability to opt out
- Like traffic congestion in distracts attention of consumers, takes over storage and bandwidth
- Historically the market has evolved from few companies that offered spamming services to organized network of botnets and spammers
- Imposes large costs on email services; more recently also threatens physical properties of networks

Spam

- Exploiting structure of SMTP (Simple Mail Transfer Protocol)
- SMTP protocol is mechanism of communication between two message transfer agents across TCP connection
- Sender-SMTP establishes two-way connection with receiver-SMTP
- Email sent via request–response transactions between client and server
- Email header contains the receivers email address that contains information of target server who’s IP address is resolved through DNS system
- Once destination is established, message gets delivered

Spam

- 1994 attorneys Canter and Siegel hire programmer to automatically post to every USENET newsgroup with ads for “green card” application services
- In 1995 first commercial spamware Floodgate offered for $100
- Able to harvest email addresses from classified ads, AOL Member directory, etc.
- Companion software Goldrush allowed sending out 1000’s of emails per hour (~$0.0001 per email)

Spam

- (Stone-Gross, Holz, Stringhini, and Vigna, 2011) infiltrated Spamdot.biz forum
- Strict vetting process to join (3 referrals from existing members)
- 91.3 % users are Russian speakers, remaining 8.7% use English
- 1,929 users with 35,423 public and 11,638 private messages.
- Forum is divided into: spam community and vendor services
- Forum operates based on a system of trust, members often review each other's products and services

Spam

- Lists of emails are hot commodities
- Determinants of value
  - Validity
  - If emails have recently been targeted by another spam group
  - Localization of email addresses (.us, .uk, .ru) or regionalized by IP-based geolocation
  - Whether email belongs to free email service (Gmail, Hotmail, Yahoo)
- Email addresses from free email services are ½ the price of standard emails (due to more sophisticated spam filters)
- Typical price per million emails $25 to $50, with discounted prices for bulk purchases

Spam

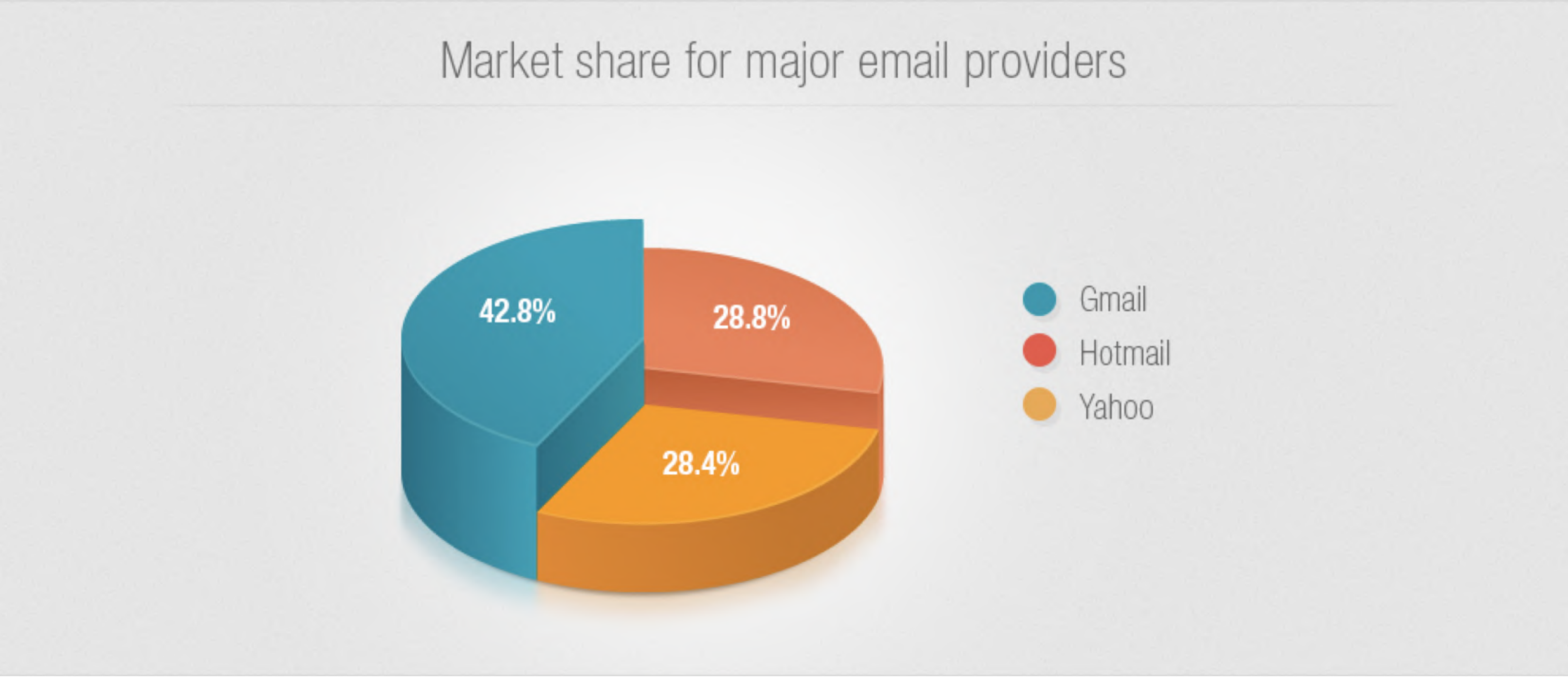
- SMTP protocol has basic built-in spam protection: address is validated through DNS (can automatically delete messages from specific IP addresses)
- This allows IP blacklisting after it was authenticated (Yahoo!: this method rejects 80% of arriving emails)
- Email is post-processed using ML techniques to do supervised learning
- Ground truth given by human labels of spam versus non-spam, algorithm learns to classify emails correctly
- Featurize email contents to find words or phrases in email body that improve error rate of classifier
- Risk of “false positives”!

Spam

- Spam classifiers rely on crowdsourcing
- Webmail providers encourage users to press “mark as spam” button to generate more labels for spam classifier
- (Rao and Riley, 2012) report that out of random sample of 1.3 million active Yahoo! Mail 6% ever marked messages as “spam,” others simply delete them
- At the same time, spammers take advantage of “not spam” button
- Yahoo! reports that 63% of “not spam” votes were cast by users who never cast “spam” vote and come from users few specific IP addresses

Spam

- Webmail market is highly concentrated and has resources to invest in sophisticated anti-spam technologies



Spam

- Blacklisting made it impossible for spammers with fixed IP addresses to operate
- Using new IP addresses became cheaper with botnets
- Use malware to form network of computers
- Infected computers form hierarchy transmitting messages from central servers
- IP blacklisting is useless with botnets: spam emails originate from 1000’s of changing IP addresses assigned by DHCP (used by most internet providers for residential clients)
- ISP’s started to prohibit client’s computers work as mail servers

Spam

- In 2009, 6 botnets generated > 90% of botnet spam
- Rustock was largest botnet on record capable of sending 30 billion of emails per day
- In March 2011 software was used to reverse-engineer location of command servers and network was shut down

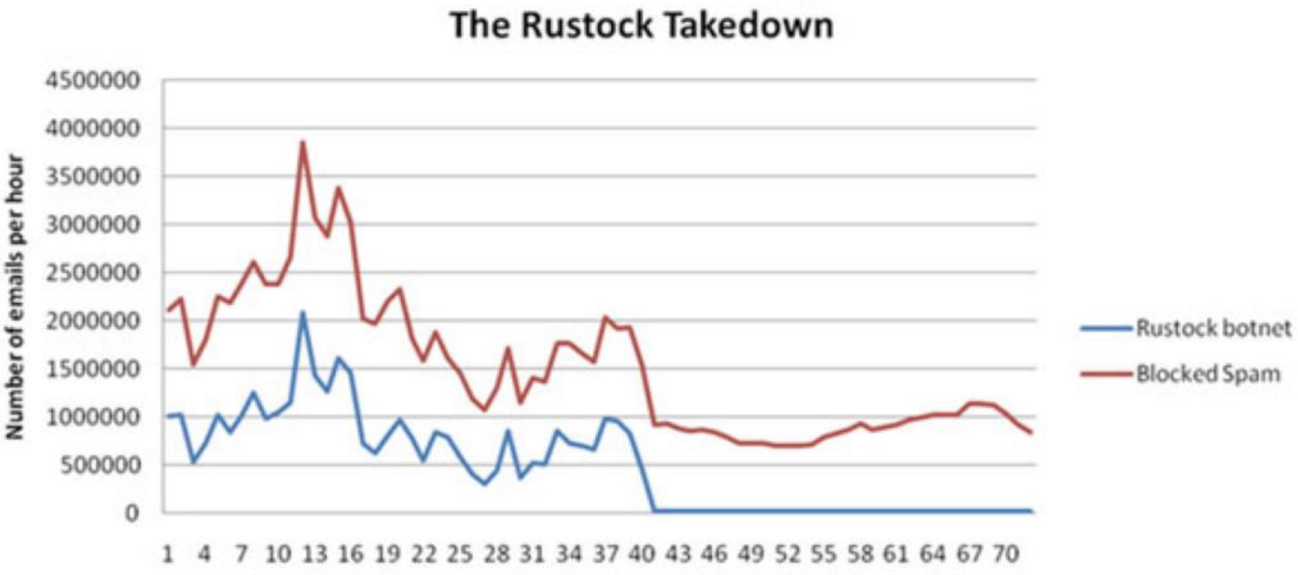
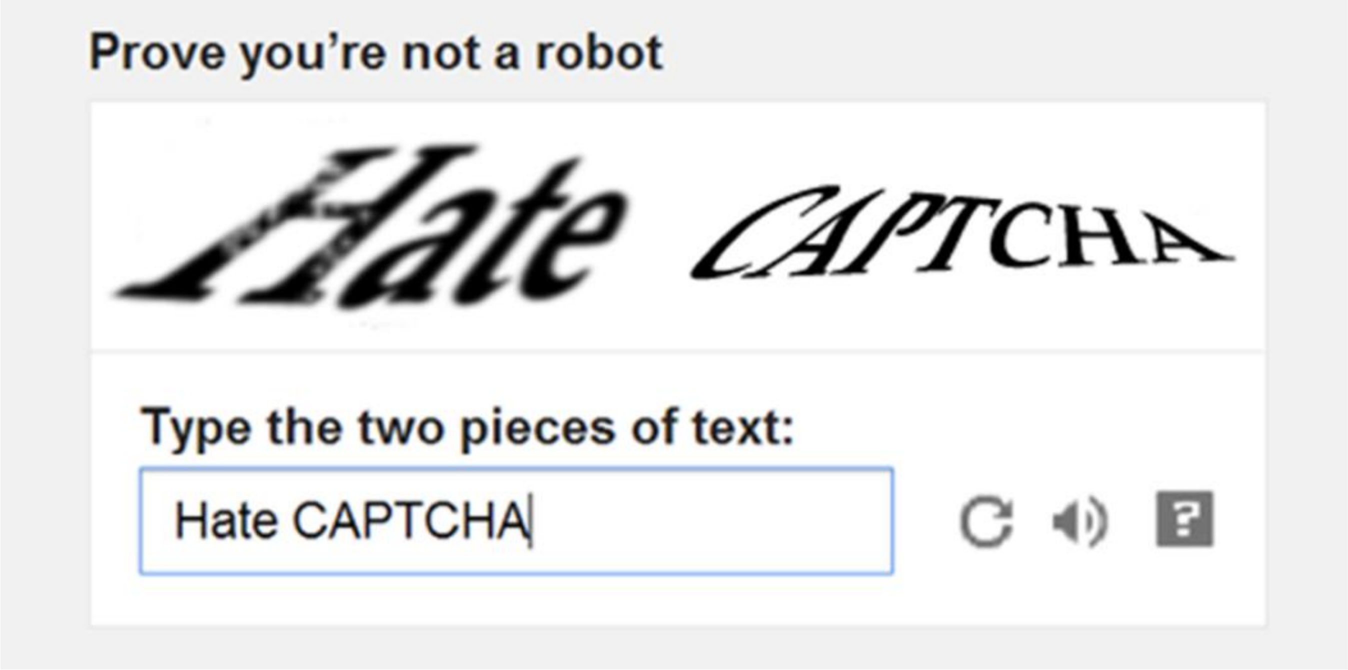


Figure 1. Rustock spam volume for hourly intervals from March 15 to 17

Spam

- Rustock used email accounts on Microsoft’s webmail
- To prevent possibility of automated sign-up for email accounts commercial providers adopted CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)



Prove you’re not a robot

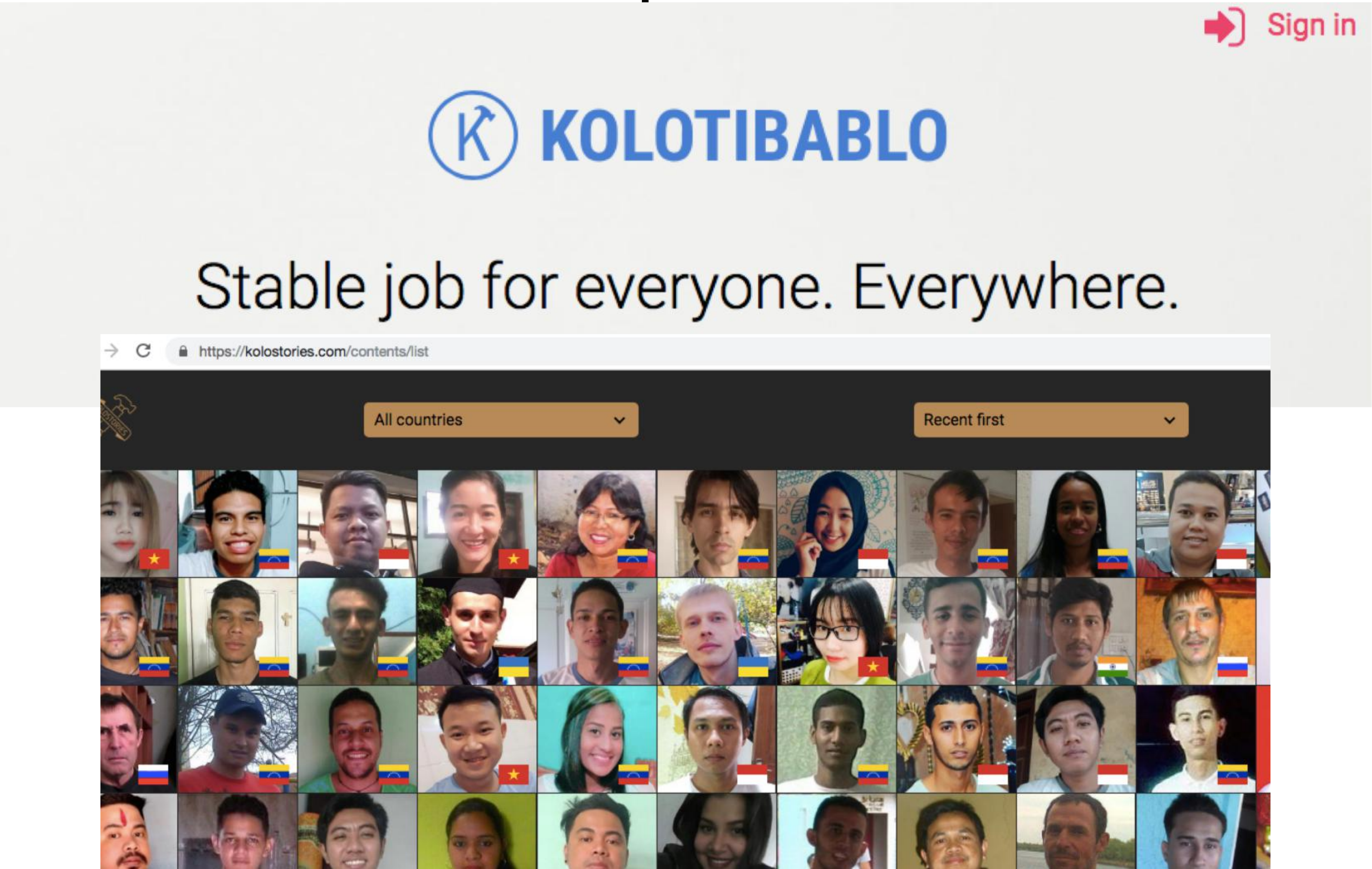
Type the two pieces of text:

Hate CAPTCHA

Spam

- Formal market arose for resolving CAPTCHA
- Market maker solicits offers from buyers of CAPTCHA-breaking and interacts with workers offering labor
- CAPTCHA-breaking services offered by DeCaptcher
- Each CAPTCHA transmitted to worker at PixProfit
- Resolved CAPTCHA sent to DeCaptcher
- Use separate piece of software to transmit CAPTCHA and its solution
- Process takes ~30 seconds
- Market wage rapidly decreasing with laborers recruited from poorer developing countries

Spam



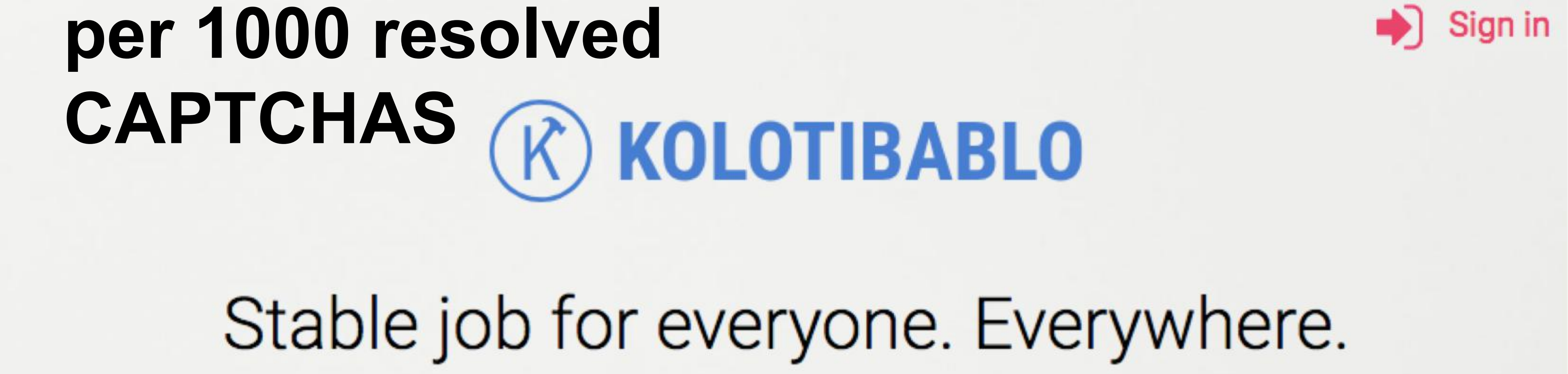
KOLOTIBABLO

Stable job for everyone. Everywhere.

https://kolostories.com/contents/list

Rates $0.3 to $1 per 1000 resolved CAPTCHAS

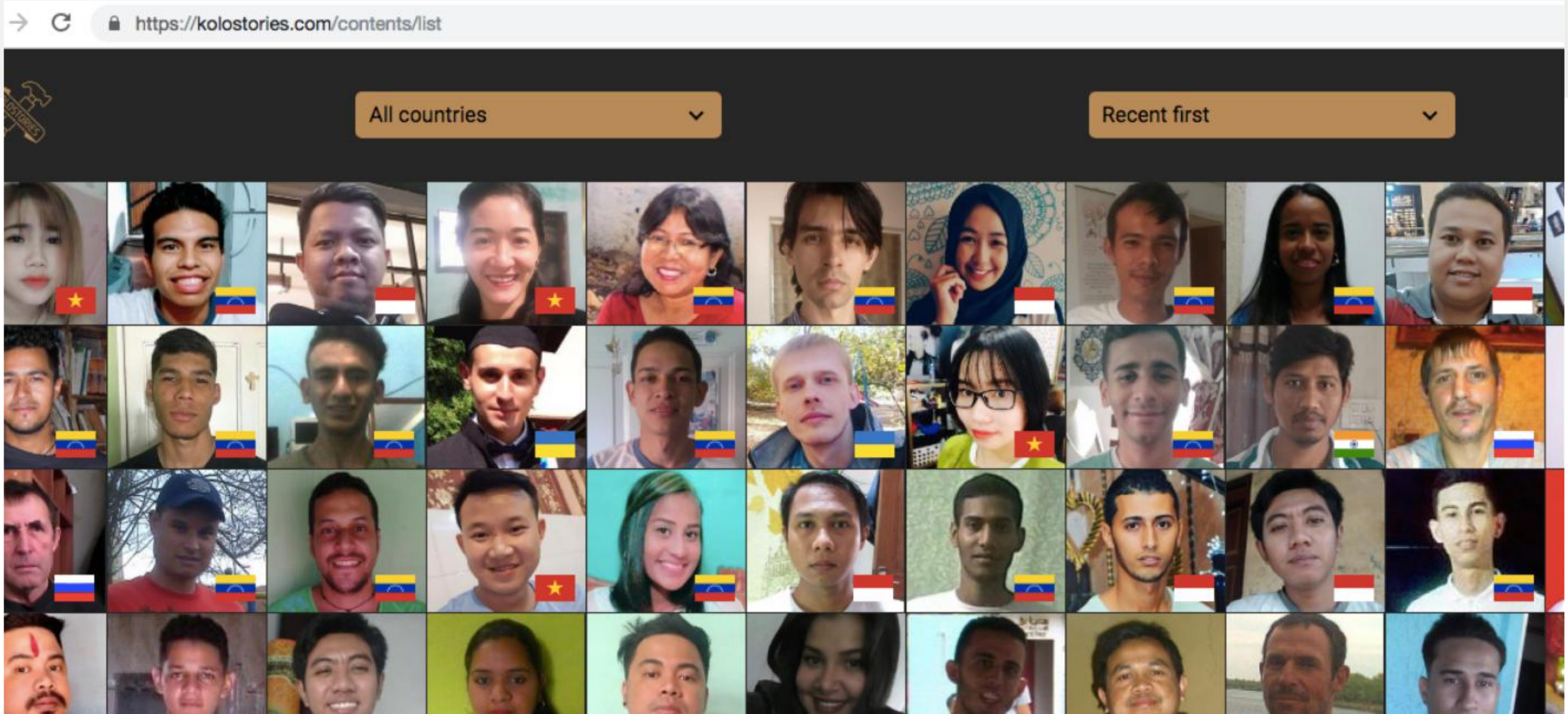
Spam



Sign in

KOLOTIBABLO

Stable job for everyone. Everywhere.



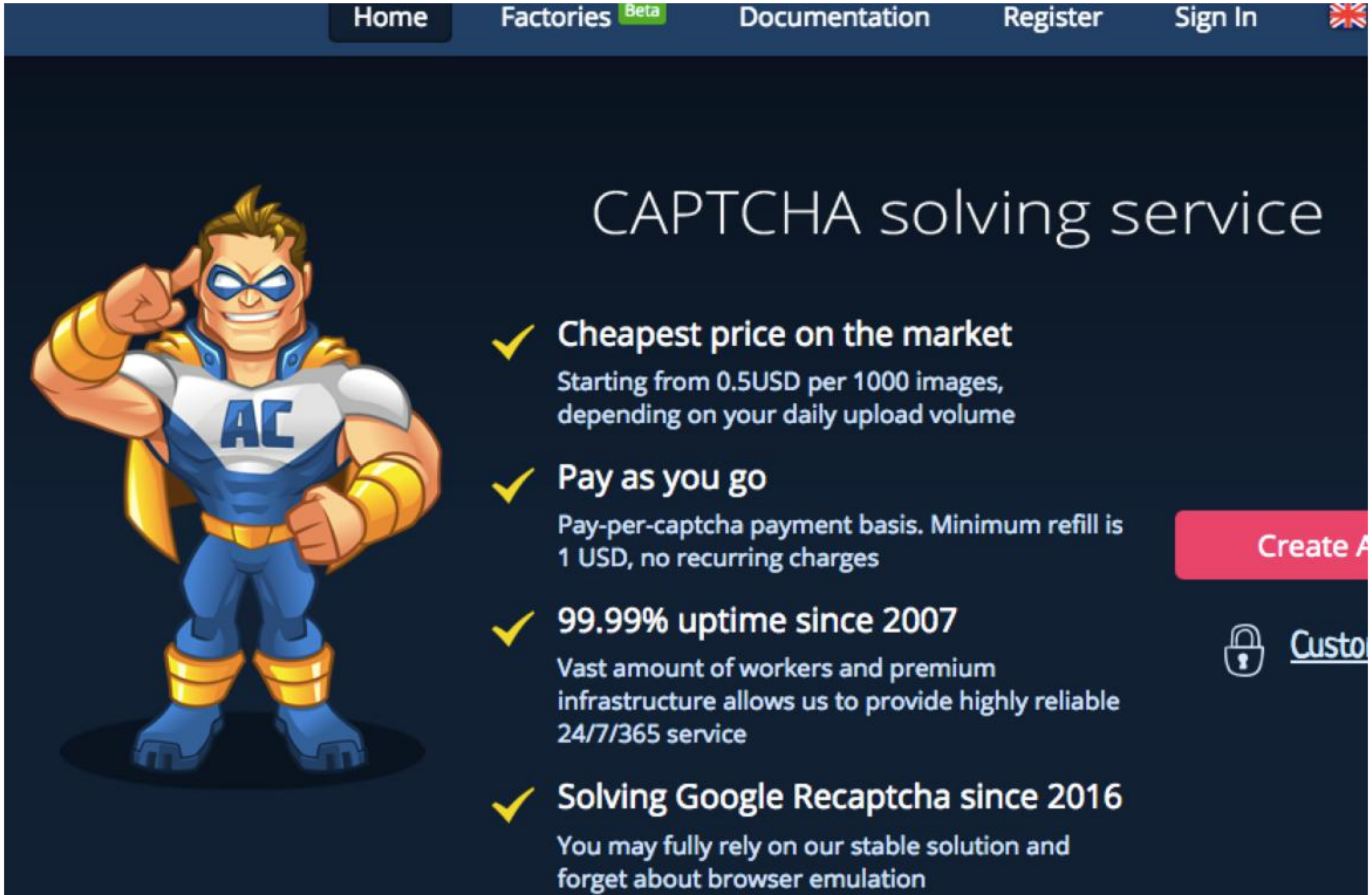
https://kolostories.com/contents/list

All countries

Recent first

Spam

- CAPTCHA-breaker services operate as “general contractors”
- Motoyama, Levchenko, Kanich, McCoy, Voelker, and Savage (2010) report 10-15 second response times with 90% accuracy



CAPTCHA solving service

✓ **Cheapest price on the market**
Starting from 0.5USD per 1000 images, depending on your daily upload volume

✓ **Pay as you go**
Pay-per-captcha payment basis. Minimum refill is 1 USD, no recurring charges

✓ **99.99% uptime since 2007**
Vast amount of workers and premium infrastructure allows us to provide highly reliable 24/7/365 service

✓ **Solving Google Recaptcha since 2016**
You may fully rely on our stable solution and forget about browser emulation

Spam

- Most spam illegal in the US (CAN-SPAM Act of 2003)
    - Unsolicited emails must have valid return addresses and opt-out options
- Technically spam only includes unsolicited email from companies that had no economic relation with given user
- Email offer from BestBuy after you made purchase is not spam
- Spam market is market for “unsolicited online advertising”
    - Merchant recruits intermediaries (a.k.a. spammers) to advertise products paying share of final purchase amount
    - Merchant conceals its identity to attract more potential customers

Spam

- Just like in online advertising, user experience related to viewing of ad (email)
- Spam industry contains multiple actors: consumers, advertisers and publishers (spammers)
- Publishers and advertisers are typically not the same entity (spammers do not operate storefronts)
- Transactions between parties are based on negotiated rates related to cost per acquisition
- Stages of spam activity
    - Advertising
    - Click support
    - Realization

Spam

- Advertising
  - All activities on reaching potential customers and making them click on URL
  - Delivery of spam uses botnets, Webmail spam, and IP prefix hijacking
  - Growing market specialization where botnets can be rented on contract basis
  - Most commercial antispam offerings focus on delivery aspect of spam, but spam is still delivered
  - That means that effort invested in spam delivery yields sufficient return

Spam

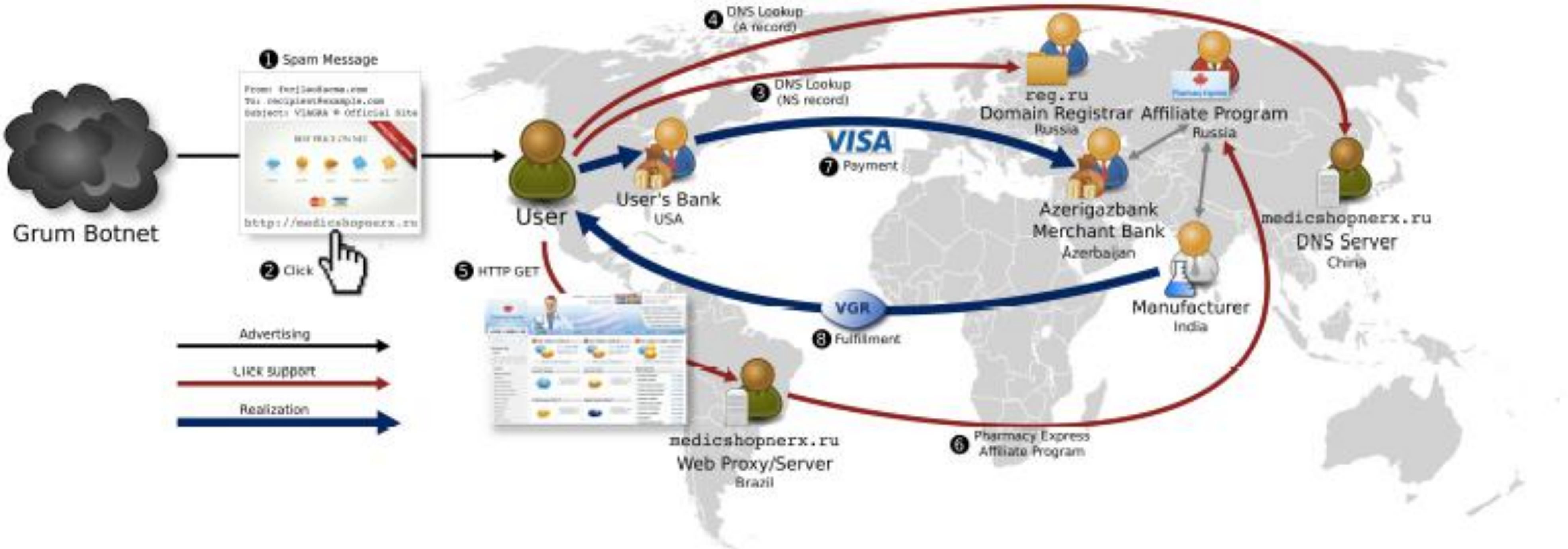
- Click support
  - After advertising is delivered spammer needs some recipients to respond (click on embedded URL)
  - Can directly advertise URL, but this is risky due to blacklisting and site takedowns
  - Typically advertised URL redirects to additional URLs
    - Can use legitimate third party that controls DNS name resource for edirection site (free hosting, URL shorteners, compromised Web sites)
    - Spammers or their affiliates manage the DNS name resources (“throwaway” domain such as minesweet.ru redirects to more persistent domain such as greatjoywatches.com)
  - Domaine name typically purchased from domain reseller who purchases domains in bulk via multiple sources and sells to underground trade
  - Sometimes offered as “package” from spam “affiliate program”

Spam

- Realization
  - Seller receives payment through conventional payment networks
  - Stores try to support standard credit card payments
  - Store arranges to fulfill and ship order
  - Products are acquired through and delivered through B2B websites (Alibaba, ECPlaza, ECTrade) that offer brand and off-brand drugs, replicas of luxury products
  - Suppliers ship directly and stores do not pay for storing, warehousing and shipping goods themselves

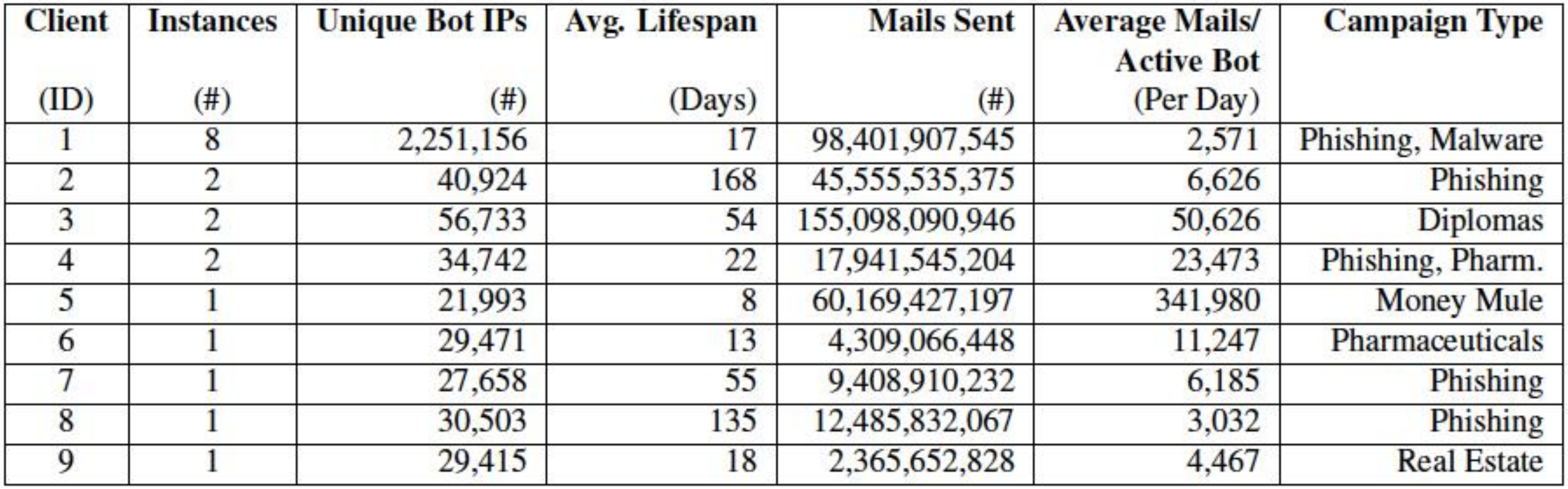
Spam

- (Levchenko et al., 2011) developed spam feeds to identify examples of spam, followed advertised URLs, and botnet infiltration algorithms



Spam

- Supply of spam is provided by botnets
- Major merchants are advertised by multiple botnets, and botnets compete with each other for clients
- Botnet may either rent out its services to independent spammers, or send its own spam on behalf of merchant



Client (ID)	Instances (#)	Unique Bot IPs (#)	Avg. Lifespan (Days)	Mails Sent (#)	Average Mails/ Active Bot (Per Day)	Campaign Type
1	8	2,251,156	17	98,401,907,545	2,571	Phishing, Malware
2	2	40,924	168	45,555,535,375	6,626	Phishing
3	2	56,733	54	155,098,090,946	50,626	Diplomas
4	2	34,742	22	17,941,545,204	23,473	Phishing, Pharm.
5	1	21,993	8	60,169,427,197	341,980	Money Mule
6	1	29,471	13	4,309,066,448	11,247	Pharmaceuticals
7	1	27,658	55	9,408,910,232	6,185	Phishing
8	1	30,503	135	12,485,832,067	3,032	Phishing
9	1	29,415	18	2,365,652,828	4,467	Real Estate

Table 1: Statistics for individual spam operations run by Cutwail.

(Stone-Gross, Holz, Stringhini, and Vigna, 2011)

Spam

- (Stone-Gross, Holz, Stringhini, and Vigna, 2011) get access to 13 Cutwail C&C servers
- Content of spam varies by client
- Client-1 coordinated phishing campaigns (e.g., Google
- Mail, Friendster, etc)
- Client-9 was advertising only Russian real estate.
- Phishing is most popular campaign type
- Malware campaigns (mail included malicious link or attachment) are second popular
- Next was “pharmacies” and “education”

Spam

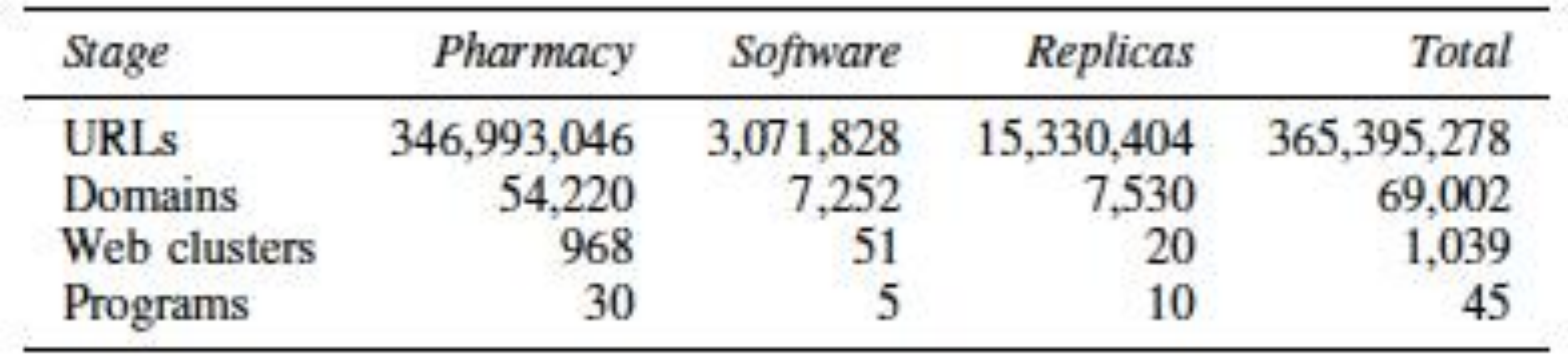
- Market for botnet installation services (“loads”)
- Sources of loads: drive-by-download attacks using HTML iframes and other malware
- 10,000 malware installations offered at $300– $800.
- Price varied by location
  - US-based computed are more valuable than those in Asia
  - Latency and quality of infected computersimportant for operation
- Loads per 1000 cost $13 in Asia, $35 in Europe, $125 US
- Bots that have not been blacklisted (“clean”) sell at higher prices, they are valuable for spam campaigns

Spam

- Controllers of spam botnets put significant effort into maintaining sufficient number of bots
- Bot populations can drop by 50% per day
- With sufficiently large sustained botnet, groups launch their own spam campaigns or rent out parts of botnet to other parties
- Contracts based on sharing a percentage of sales (SpamIt paid 40% commission) or purchase of spam-as-a-service for ~$100–$500 per million emails
- Botnets can be rented out for larger campaigns (> 100 million emails per day) for ~$10,000 per month
- Renter of botnet is offered “free trial”

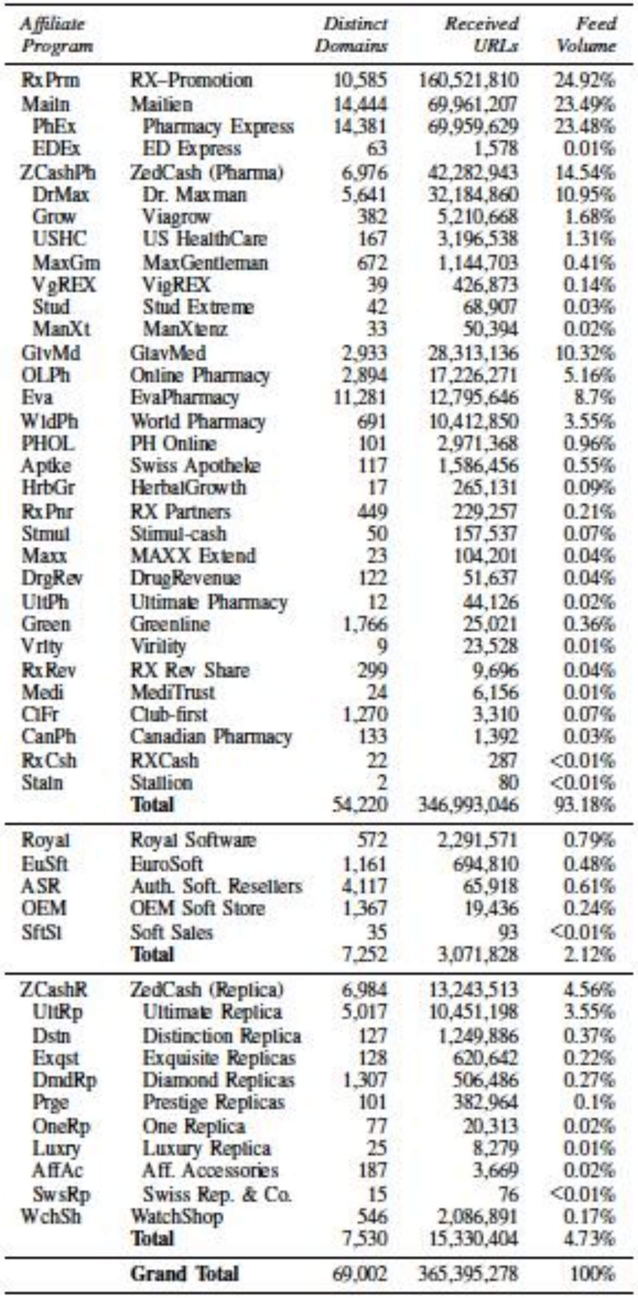
Spam

- (Levchenko et al., 2011) studied the structure of “global spam Economy” by following the URL delivered in spam emails



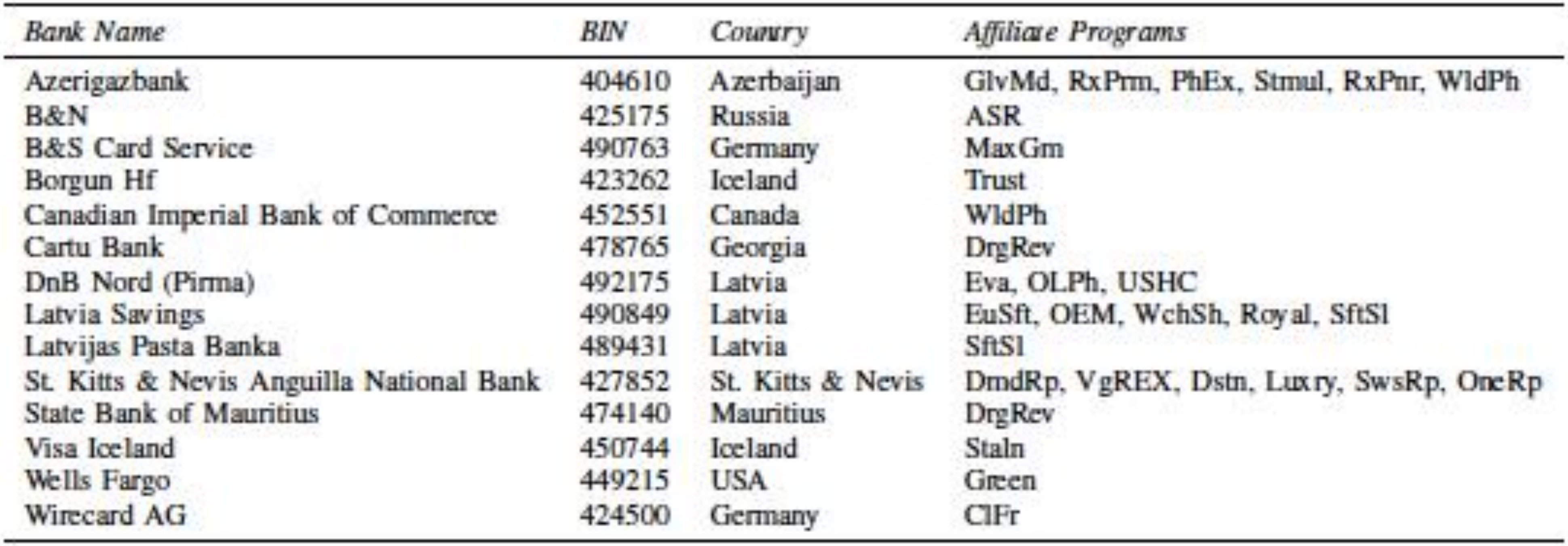
Stage	*Pharmacy*	*Software*	*Replicas*	*Total*
URLs	346,993,046	3,071,828	15,330,404	365,395,278
Domains	54,220	7,252	7,530	69,002
Web clusters	968	51	20	1,039
Programs	30	5	10	45

Spam



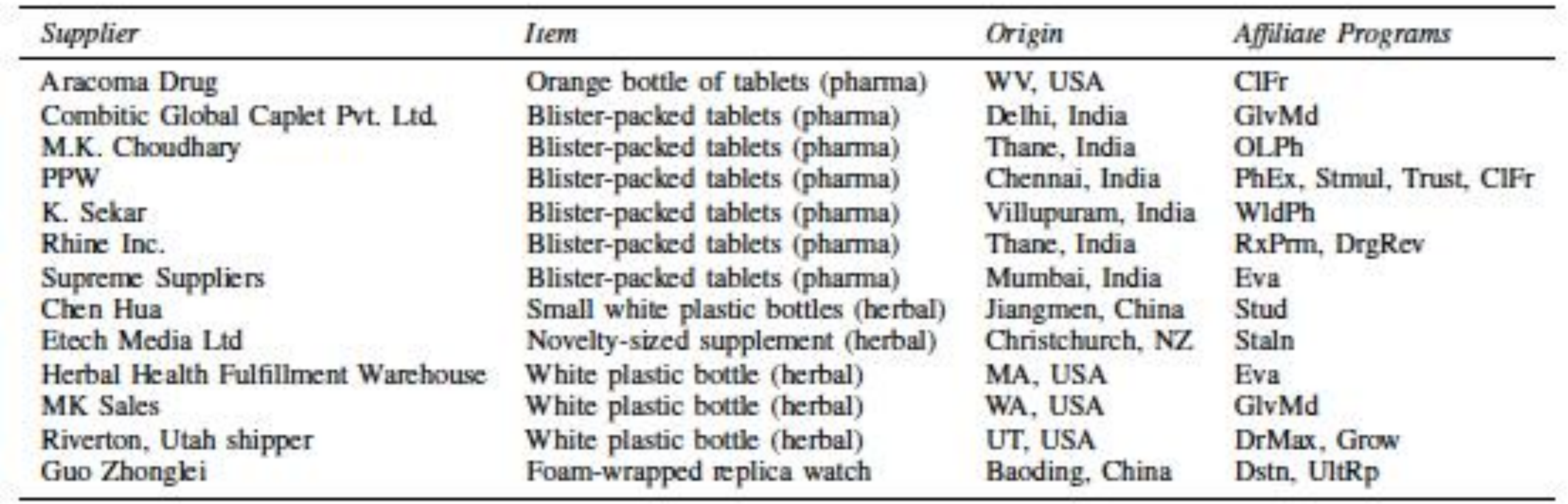
Affiliate Program		*Distinct Domains*	*Received URLs*	*Feed Volume*
RxPrm	RX–Promotion	10,585	160,521,810	24.92%
Mailn	Mailien	14,444	69,961,207	23.49%
PhEx	Pharmacy Express	14,381	69,959,629	23.48%
EDEx	ED Express	63	1,578	0.01%
ZCashPh	ZedCash (Pharma)	6,976	42,282,943	14.54%
DrMax	Dr. Maxman	5,641	32,184,860	10.95%
Grow	Viagrow	382	5,210,668	1.68%
USHC	US HealthCare	167	3,196,538	1.31%
MaxGm	MaxGentleman	672	1,144,703	0.41%
VgREX	VigREX	39	426,873	0.14%
Stud	Stud Extreme	42	68,907	0.03%
ManXt	ManXtenz	33	50,394	0.02%
GlvMd	GlavMed	2,933	28,313,136	10.32%
OLPh	Online Pharmacy	2,894	17,226,271	5.16%
Eva	EvaPharmacy	11,281	12,795,646	8.7%
WldPh	World Pharmacy	691	10,412,850	3.55%
PHOL	PH Online	101	2,971,368	0.96%
Aptke	Swiss Apotheke	117	1,586,456	0.55%
HrbGr	HerbalGrowth	17	265,131	0.09%
RxPnr	RX Partners	449	229,257	0.21%
Stmul	Stimul-cash	50	157,537	0.07%
Maxx	MAXX Extend	23	104,201	0.04%
DrgRev	DrugRevenue	122	51,637	0.04%
UltPh	Ultimate Pharmacy	12	44,126	0.02%
Green	Greenline	1,766	25,021	0.36%
Vrlty	Virility	9	23,528	0.01%
RxRev	RX Rev Share	299	9,696	0.04%
Medi	MediTrust	24	6,156	0.01%
ClFr	Club-first	1,270	3,310	0.07%
CanPh	Canadian Pharmacy	133	1,392	0.03%
RxCsh	RXCash	22	287	<0.01%
Staln	Stallion	2	80	<0.01%
	Total	54,220	346,993,046	93.18%
Royal	Royal Software	572	2,291,571	0.79%
EuSft	EuroSoft	1,161	694,810	0.48%
ASR	Auth. Soft. Resellers	4,117	65,918	0.61%
OEM	OEM Soft Store	1,367	19,436	0.24%
SftSl	Soft Sales	35	93	<0.01%
	Total	7,252	3,071,828	2.12%
ZCashR	ZedCash (Replica)	6,984	13,243,513	4.56%
UltRp	Ultimate Replica	5,017	10,451,198	3.55%
Dstn	Distinction Replica	127	1,249,886	0.37%
Exqst	Exquisite Replicas	128	620,642	0.22%
DmdRp	Diamond Replicas	1,307	506,486	0.27%
Prge	Prestige Replicas	101	382,964	0.1%
OneRp	One Replica	77	20,313	0.02%
Luxry	Luxury Replica	25	8,279	0.01%
AffAc	Aff. Accessories	187	3,669	0.02%
SwsRp	Swiss Rep. & Co.	15	76	<0.01%
WchSh	WatchShop	546	2,086,891	0.17%
	Total	7,530	15,330,404	4.73%
	Grand Total	69,002	365,395,278	100%

Spam



Bank Name	*BIN*	*Country*	*Affiliate Programs*
Azerigazbank	404610	Azerbaijan	GlvMd, RxPrm, PhEx, Stmul, RxPnr, WldPh
B&N	425175	Russia	ASR
B&S Card Service	490763	Germany	MaxGm
Borgun Hf	423262	Iceland	Trust
Canadian Imperial Bank of Commerce	452551	Canada	WldPh
Cartu Bank	478765	Georgia	DrgRev
DnB Nord (Pirma)	492175	Latvia	Eva, OLPh, USHC
Latvia Savings	490849	Latvia	EuSft, OEM, WchSh, Royal, SftSl
Latvijas Pasta Banka	489431	Latvia	SftSl
St. Kitts & Nevis Anguilla National Bank	427852	St. Kitts & Nevis	DmdRp, VgREX, Dstn, Luxry, SwsRp, OneRp
State Bank of Mauritius	474140	Mauritius	DrgRev
Visa Iceland	450744	Iceland	Staln
Wells Fargo	449215	USA	Green
Wirecard AG	424500	Germany	ClFr

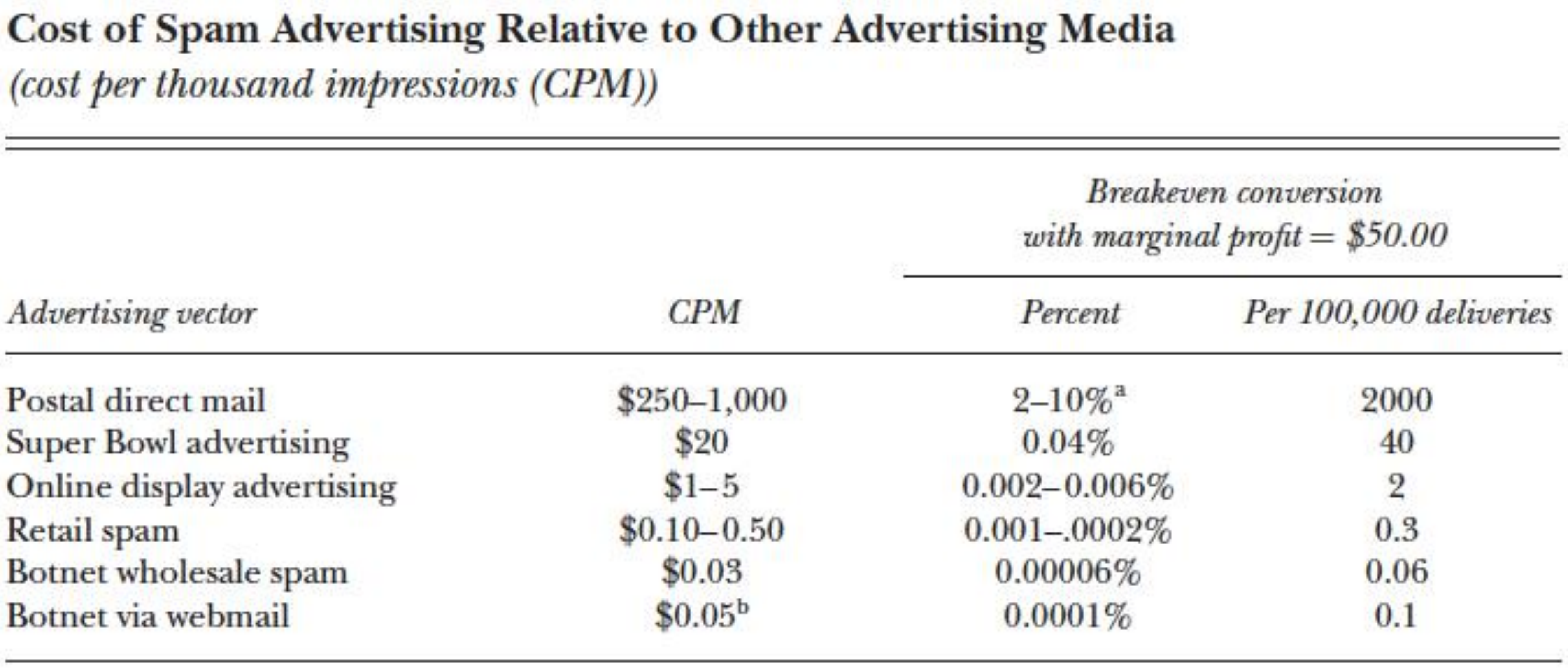
Spam



Supplier	*Item*	*Origin*	*Affiliate Programs*
Aracoma Drug	Orange bottle of tablets (pharma)	WV, USA	ClFr
Combitic Global Caplet Pvt. Ltd.	Blister-packed tablets (pharma)	Delhi, India	GlvMd
M.K. Choudhary	Blister-packed tablets (pharma)	Thane, India	OLPh
PPW	Blister-packed tablets (pharma)	Chennai, India	PhEx, Stmul, Trust, ClFr
K. Sekar	Blister-packed tablets (pharma)	Villupuram, India	WldPh
Rhine Inc.	Blister-packed tablets (pharma)	Thane, India	RxPrm, DrgRev
Supreme Suppliers	Blister-packed tablets (pharma)	Mumbai, India	Eva
Chen Hua	Small white plastic bottles (herbal)	Jiangmen, China	Stud
Etech Media Ltd	Novelty-sized supplement (herbal)	Christchurch, NZ	Staln
Herbal Health Fulfillment Warehouse	White plastic bottle (herbal)	MA, USA	Eva
MK Sales	White plastic bottle (herbal)	WA, USA	GlvMd
Riverton, Utah shipper	White plastic bottle (herbal)	UT, USA	DrMax, Grow
Guo Zhonglei	Foam-wrapped replica watch	Baoding, China	Dstn, UltRp

Spam

- Overall returns from spam (in perspective)



Cost of Spam Advertising Relative to Other Advertising Media
(cost per thousand impressions (CPM))

		Breakeven conversion with marginal profit = $50.00	
Advertising vector	*CPM*	*Percent*	*Per 100,000 deliveries*
Postal direct mail	$250–1,000	2–10%[a]	2000
Super Bowl advertising	$20	0.04%	40
Online display advertising	$1–5	0.002–0.006%	2
Retail spam	$0.10–0.50	0.001–.0002%	0.3
Botnet wholesale spam	$0.03	0.00006%	0.06
Botnet via webmail	$0.05[b]	0.0001%	0.1

(Rao and Riley, 2012)