Seeking a fix:
Measuring, analyzing and disrupting unlicensed online drug sales

Nicolas Christin
Carnegie Mellon University
nicolasc@cmu.edu

Joint work with: N. Leontiadis, T. Moore, K. Soska

"Traditional" view of attackers

- Bound by computational and mathematical limitations…
- … but by little else
- High expertise assumed

(from my intro. security class)

Active attackers

- Or, what can Mallory do?
  - Can eavesdrop on all protocol runs
  - Can **replay** messages at will
  - Can **inject** fabricated messages in the network
    - For instance fabricated from pieces of old messages
  - Can **modify** a principal's message
  - Can **initiate multiple parallel protocol sessions**
  - Can perform **dictionary attack** on passwords
  - Can perform **exhaustive attack** on non-random (or poorly random) nonce

- Sound security engineering shouldn't rest on assumptions about possible attacker's weaknesses

Attackers in practice

Attackers in practice

Attackers in practice

- Most security attacks carried out by entities that are
  - Financially interested
  - Economically rational
  - Not necessarily overly sophisticated
  - Heavily reliant on commoditization
    - Purchase "services" from others

- **Thesis: Security defenses (interventions) that disrupt profit likely very successful**

Overarching research agenda

- Economics can tell us which intervention strategies (security defenses) most likely to succeed
- … but we need sound economic models of attackers' behavior…
- … **and thus we need data about how attackers operate**

→ Network measurements

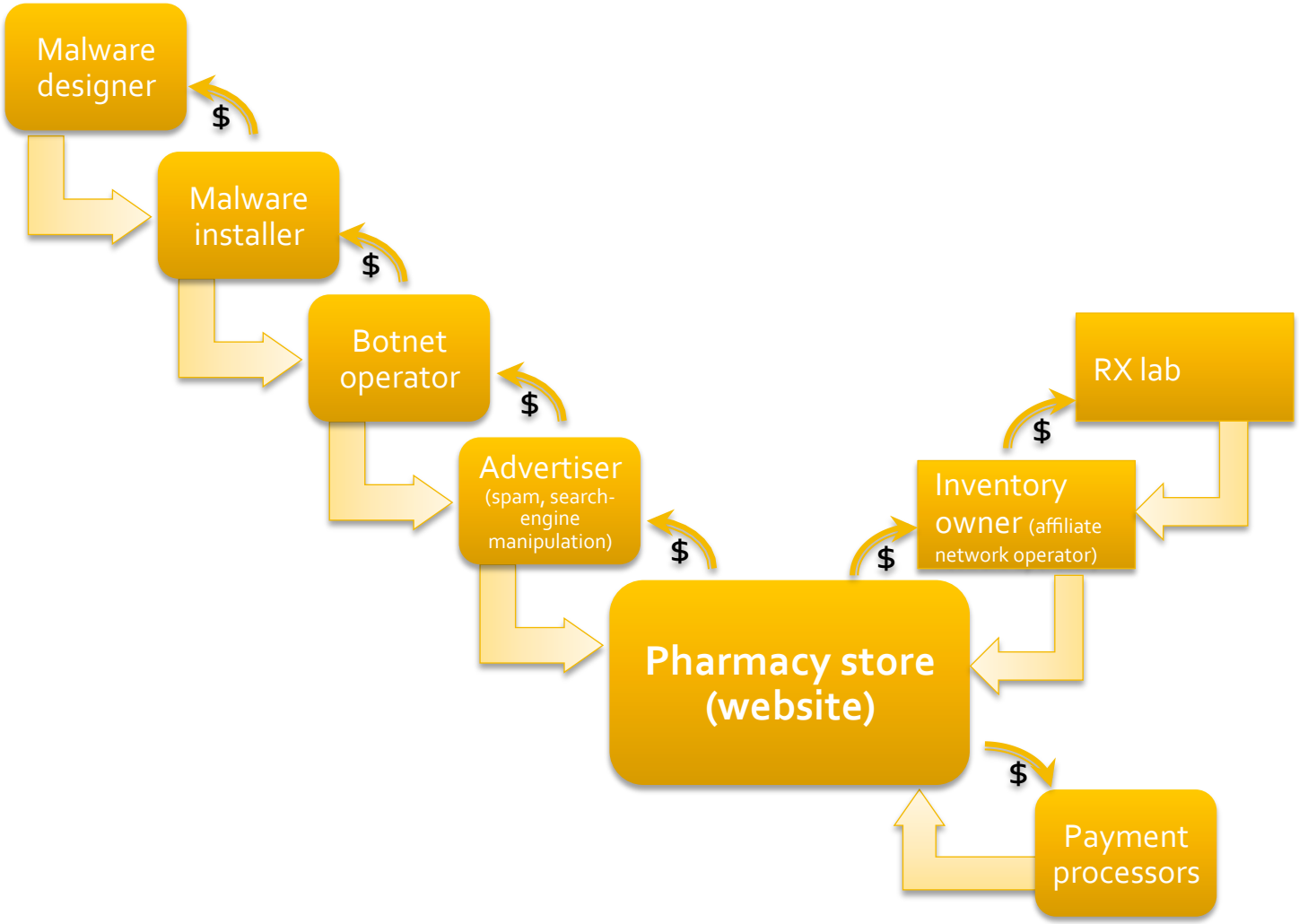
Note: Modeling defender (user) behavior is equally interesting and important… but that is a story for another day (combines game theory, and behavioral experiments).

Our case study: Online sale of drugs

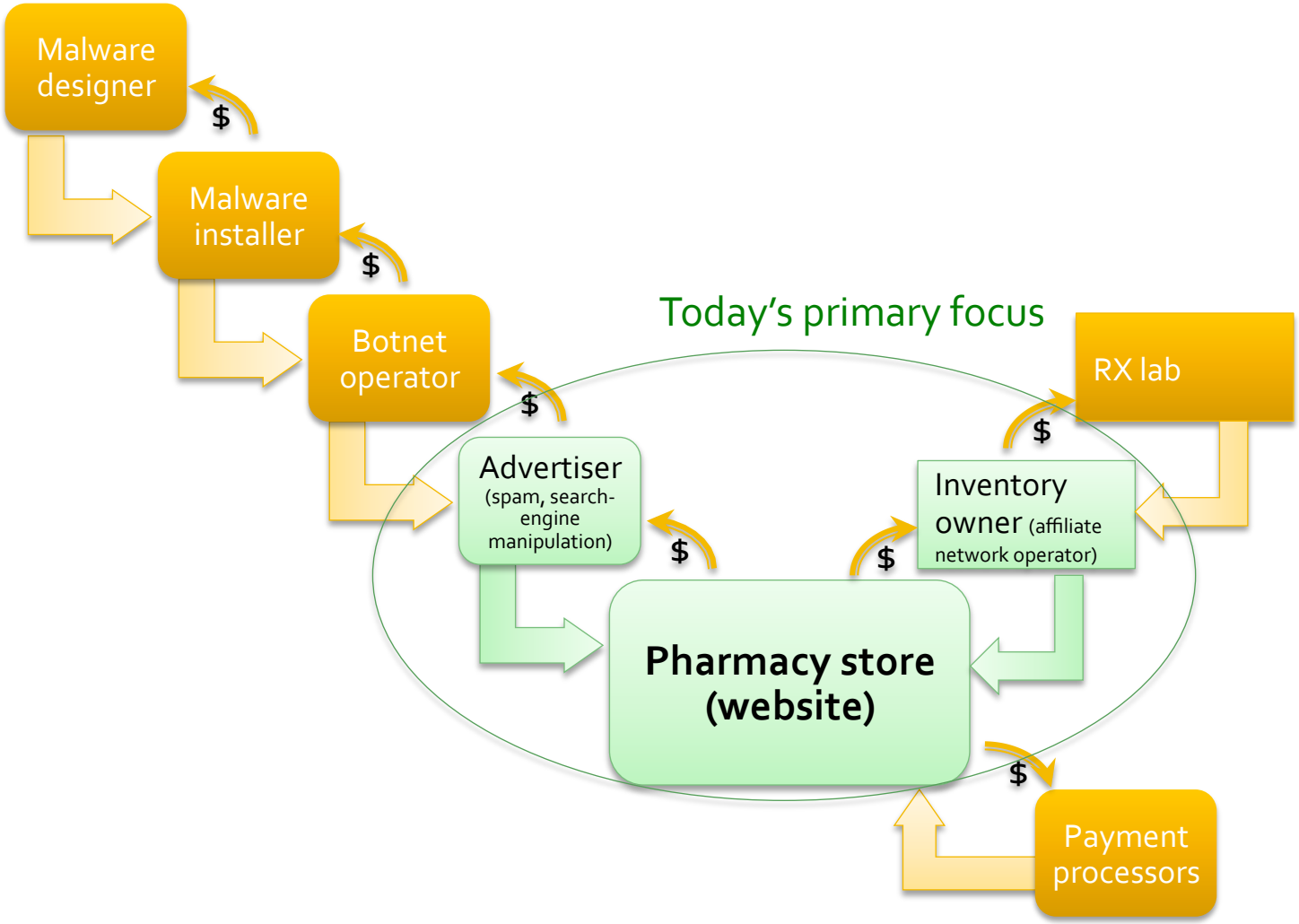
- One of the best known illicit online trades
  - Who hasn't received email spam for prescription drugs?

- Potentially most dangerous form of online crime
  - Wrong dosage can kill: cf. Ryan Haight
- Complex supply chain that can tell us a lot about the online criminal ecosystem

Supply chain: high-level overview



Supply chain: high-level overview



Bibliography (and talk outline)

Measuring and Analyzing Search-Redirection Attacks in the Illicit Online Prescription Drug Trade
Joint work with N. Leontiadis and T. Moore. In *Proceedings of the 20th USENIX Security Symposium (USENIX Sec'11)*. San Francisco, CA, USA. August 2011.

Pick Your Poison: Pricing and Inventories at Unlicensed Online Pharmacies
Joint work with N. Leontiadis and T. Moore. In submission, 2013. Available upon request.

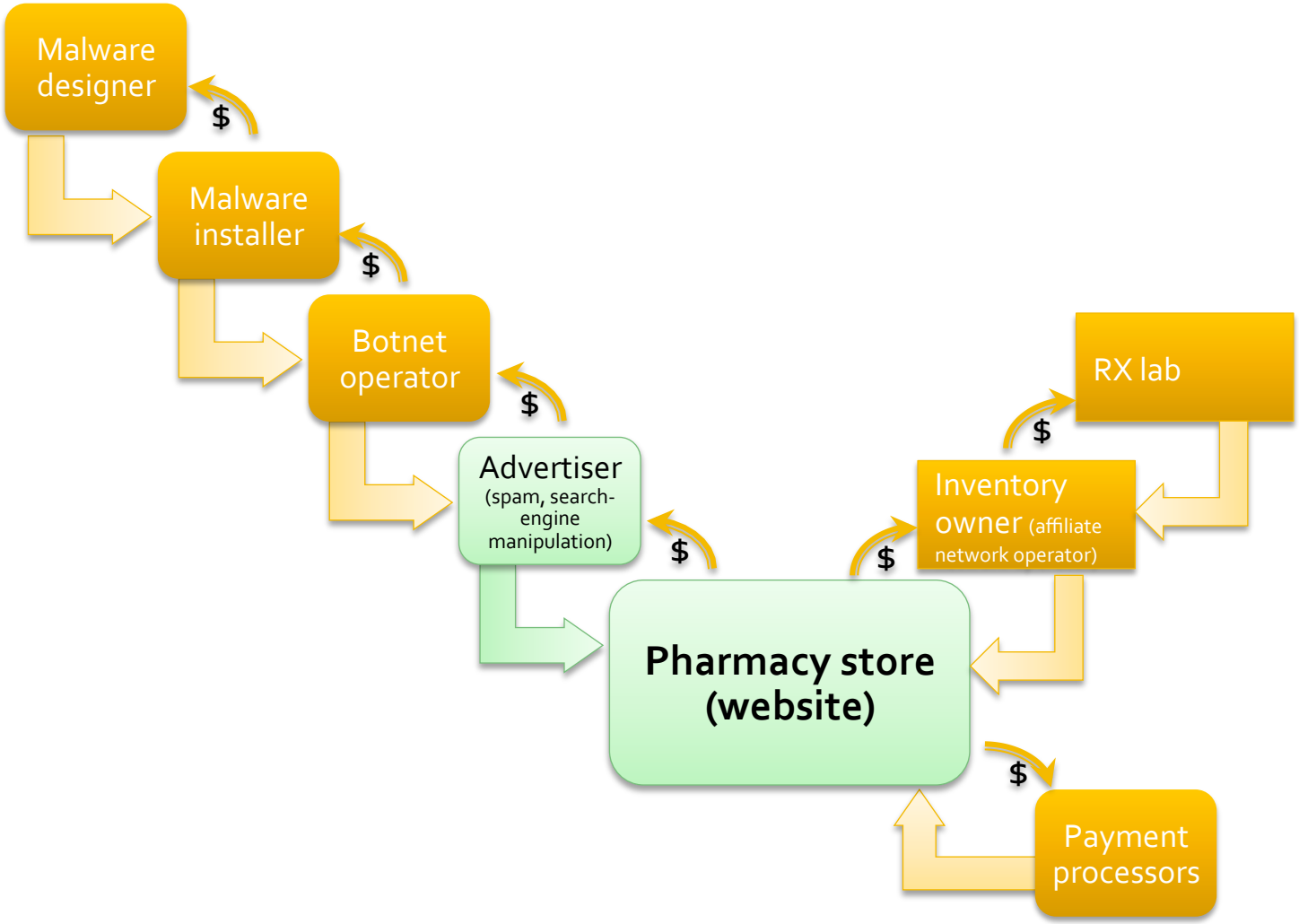
Traveling the Silk Road: An measurement analysis of a large online marketplace.
To appear in WWW 2013. Preliminary version also available as https://arxiv.org/abs/1207.7139

Dissecting One Click Frauds
Joint work with S. Yanagihara and K. Kamataki. In *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS 2010)*, pages 15-26. Chicago, IL, USA. October 2010.

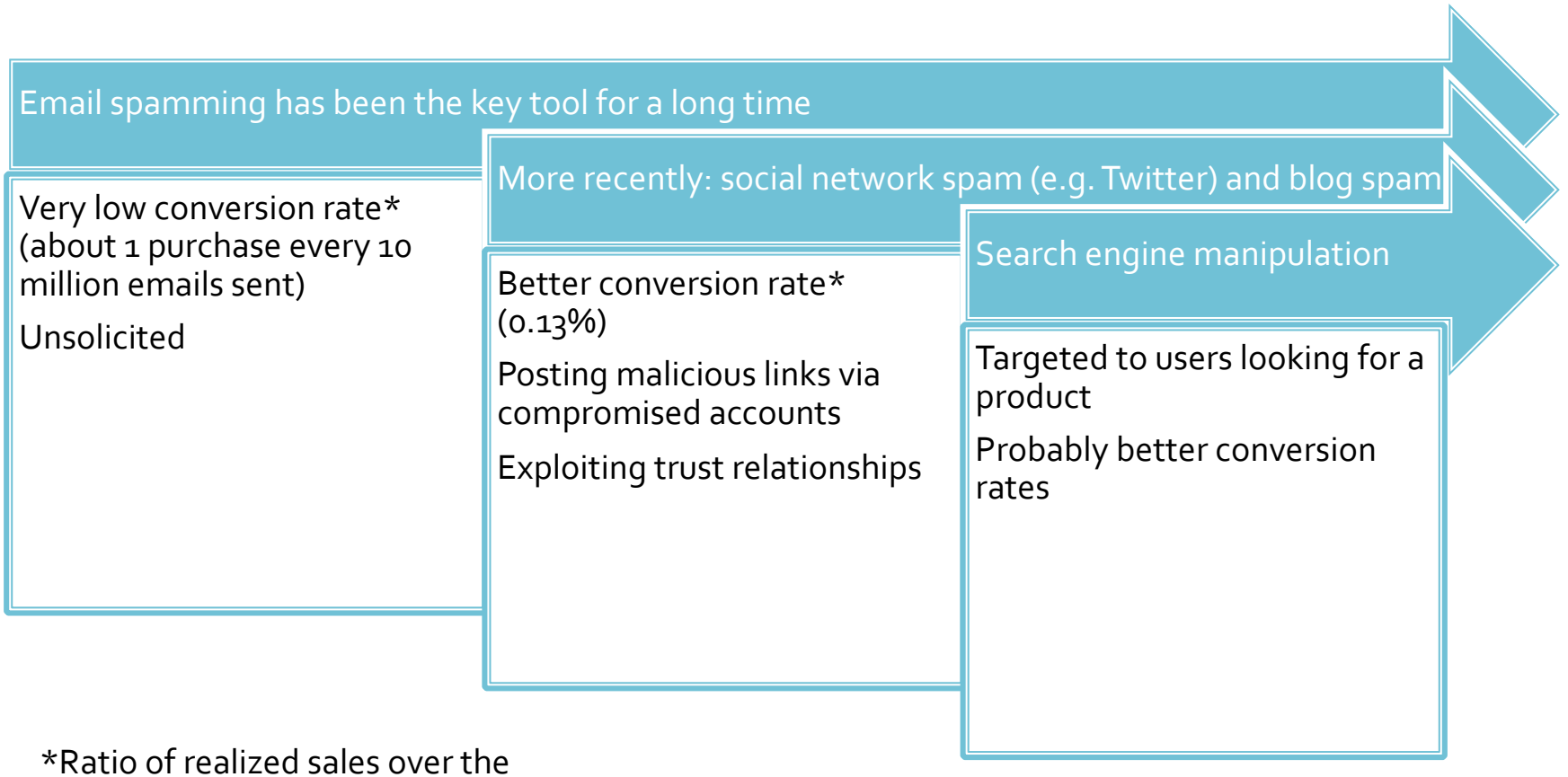
Fashion Crimes: Trending-Term Exploitation on the Web
Joint work with T. Moore and N. Leontiadis. In *Proceedings of the 18th ACM Conference on Computer and Communications Security (CCS 2011)*. Chicago, IL, USA. October 2011.

https://www.andrew.cmu.edu/user/nicolasc/papers.html

Part I: Advertising unlicensed drugs

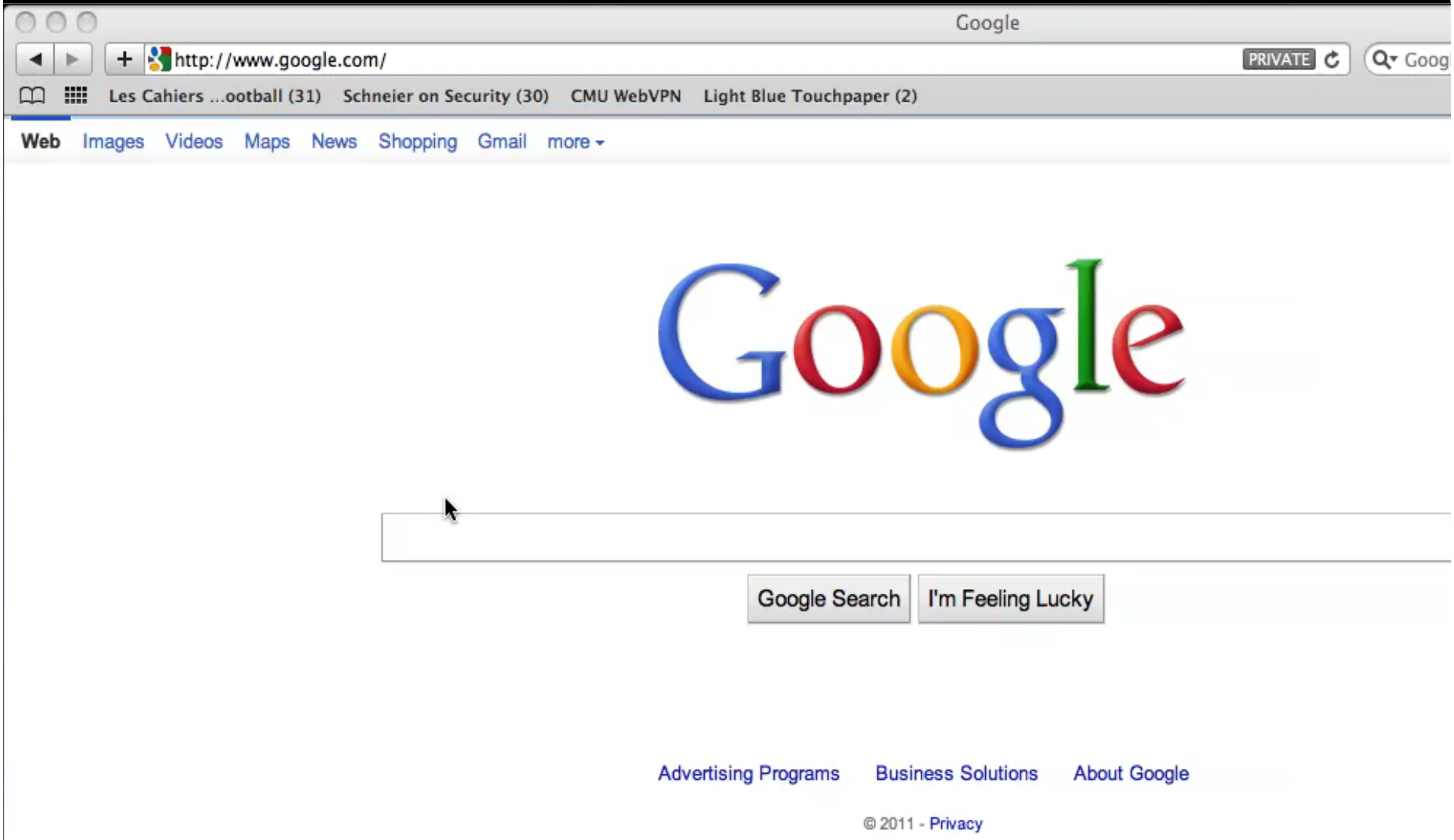


Illicit online advertising [LMC, USENIX Sec' 2011]

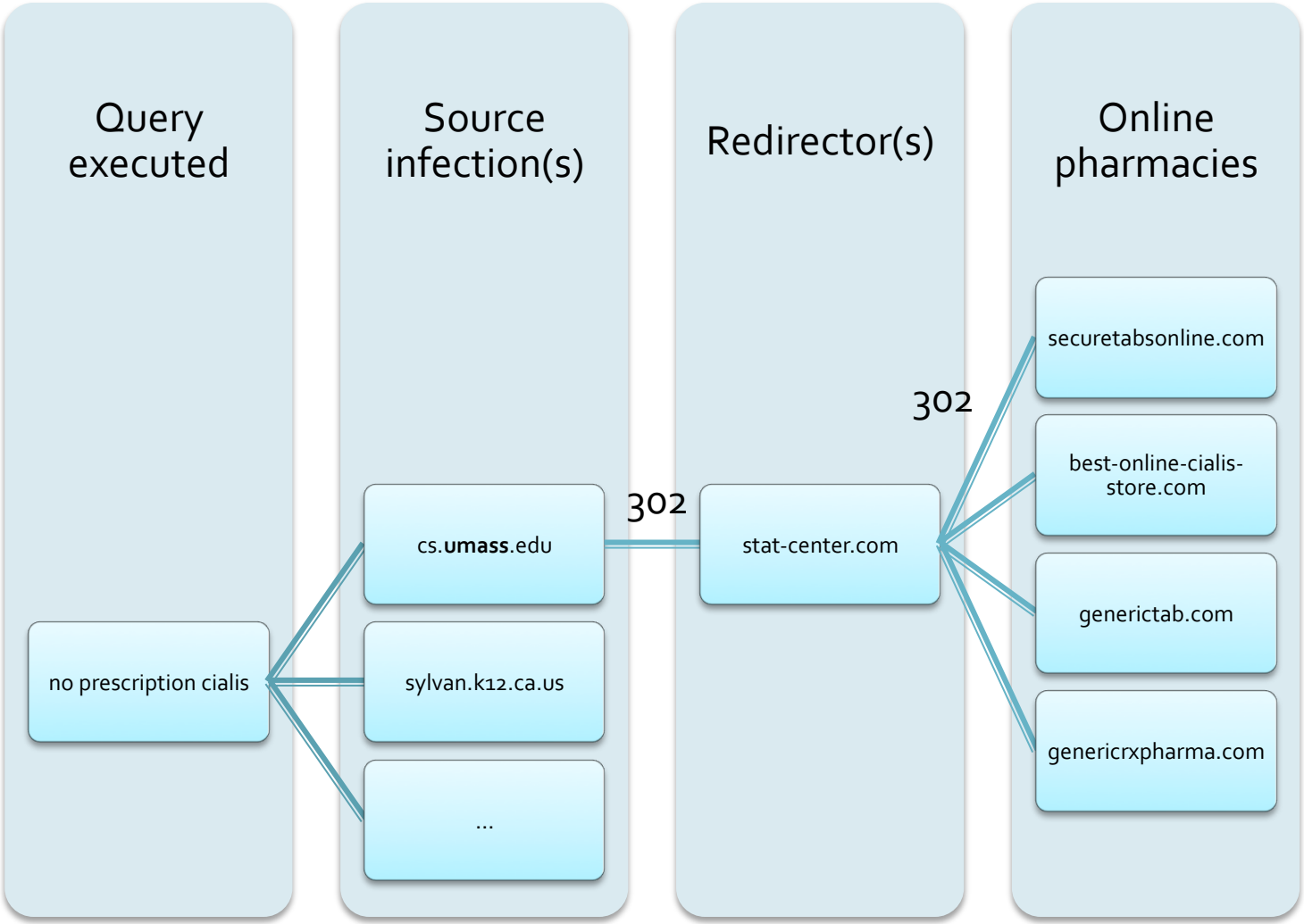


*Ratio of realized sales over the number of emails/clicks

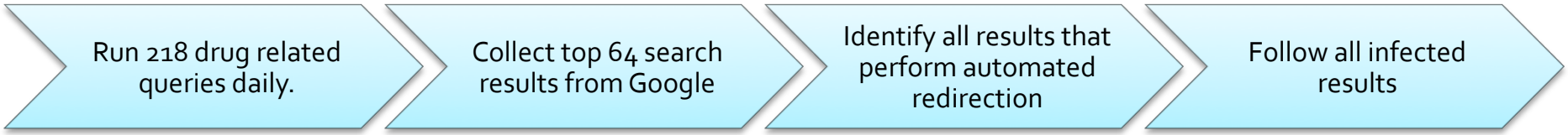
Search-redirection attack



Attack modus operandi: Redirection chains



Data collection process



- Daily collection from 4/12/2010.
- Using data until 10/21/2010.
- Complemented by a second 10-week dataset
- Collection process is still running.

- The limit is defined by Google Search API
- Storing all results for later processing
- Will also examine position information

- A search result defines the website that a user will be redirected to when clicking on the link
- If the browser is redirected instead to a different website (domain), the result is infected.

- Follow each result identified as infected from previous step
- Follow all redirections that might occur
- Record all the redirection information

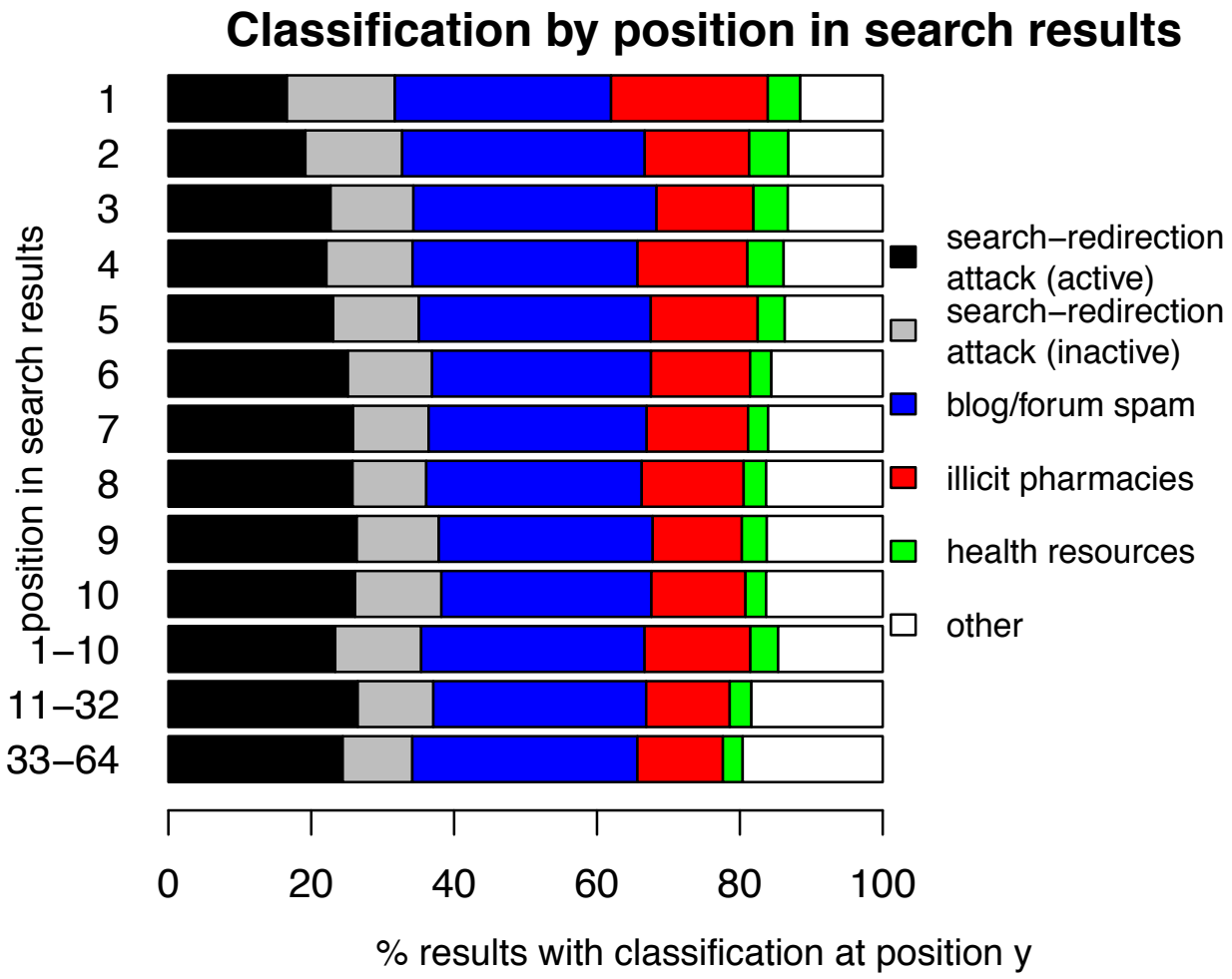
Some of the 218 queries used

vicodin no prescription
cheap valium non prescription
buy ativan online injecting pills
buy xanax valium online florida
order vicodin si levitra online
buy xanax valium online florida
color of adipex pills safest place to buy online
vicodin without prescription
generic cialis free sample
cheap tadalafil
20 mg ambien overdose
prozac side effects
ambien buy online
alprazolam online without prescription buy cheap

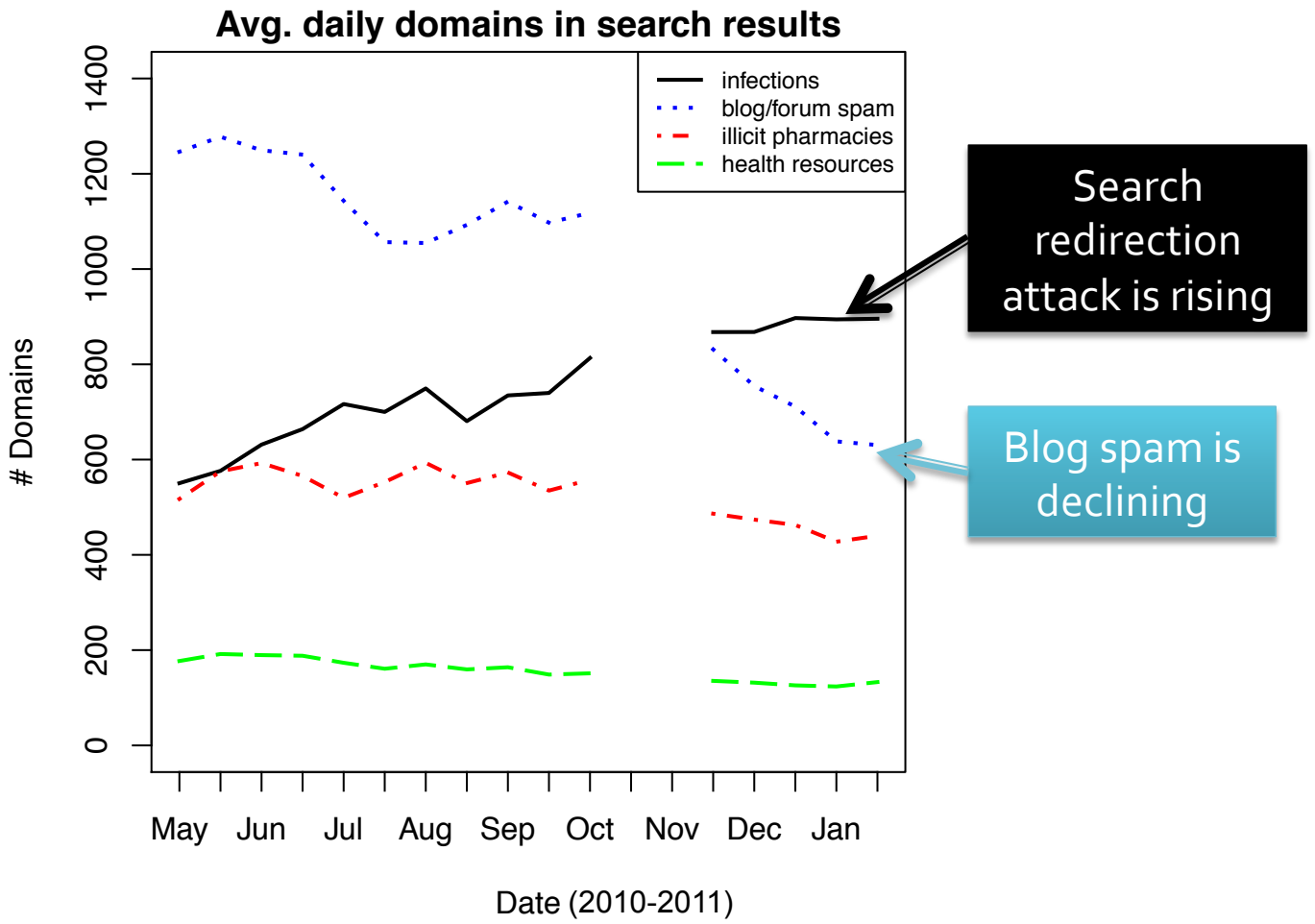
Search results classification

	URI (number)	URI (%)	Domains (#)	Domains (%)
Source infections	73909	53.8	4652	20.2
Active	*44503*	*32.4*	*2907*	*12.6*
Inactive	*29406*	*21.4*	*1745*	*7.6*
Health resources	1817	1.3	422	1.8
Pharmacies	4348	3.2	2138	9.3
Legitimate	*12*	*0.01*	*9*	*0.04*
Illicit	*4336*	*3.2*	*2129*	*9.2*
Blog/forum spam	41335	30.1	8064	34.9
Uncategorized	15945	11.6	7766	33.7
Total	**137354**	**100**	**23042**	**100**

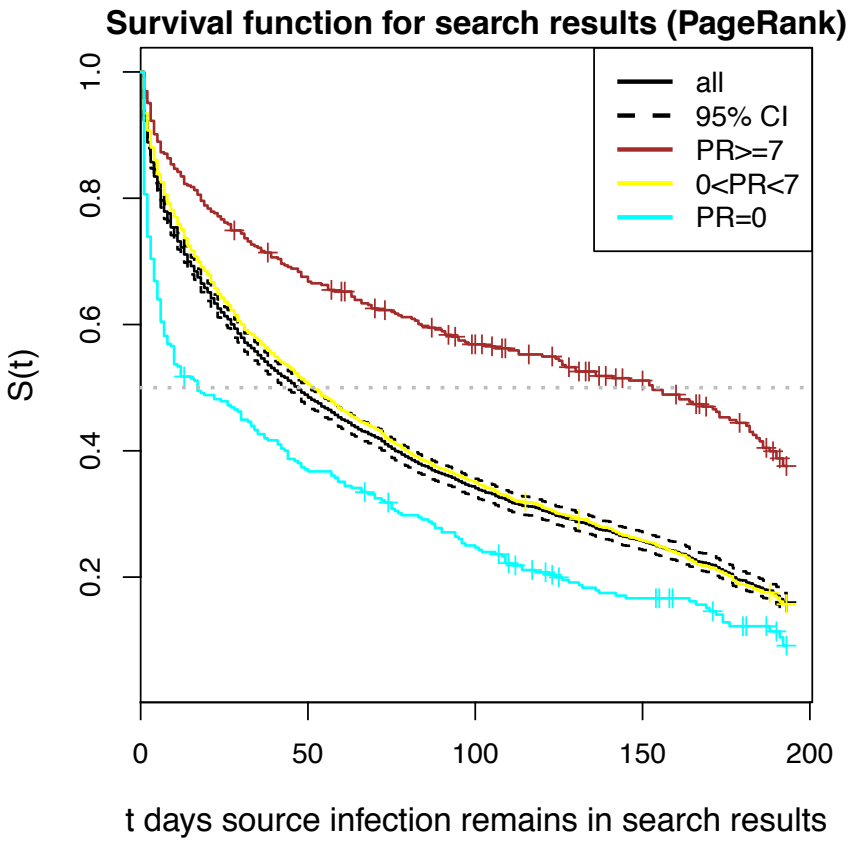
An equal opportunity attack...

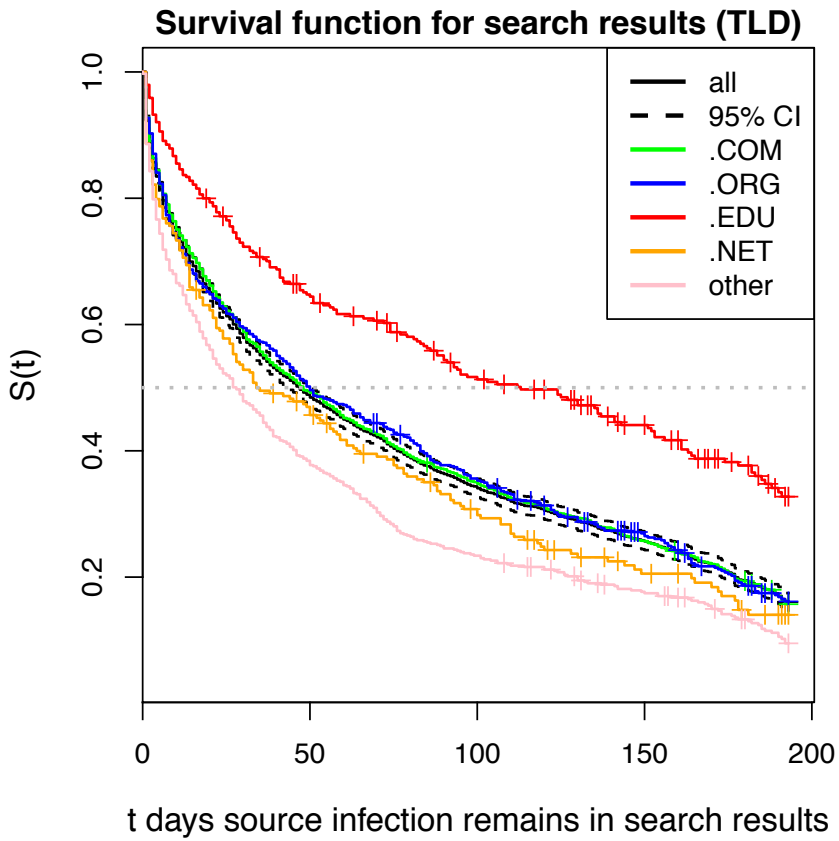


… which overtook web spam



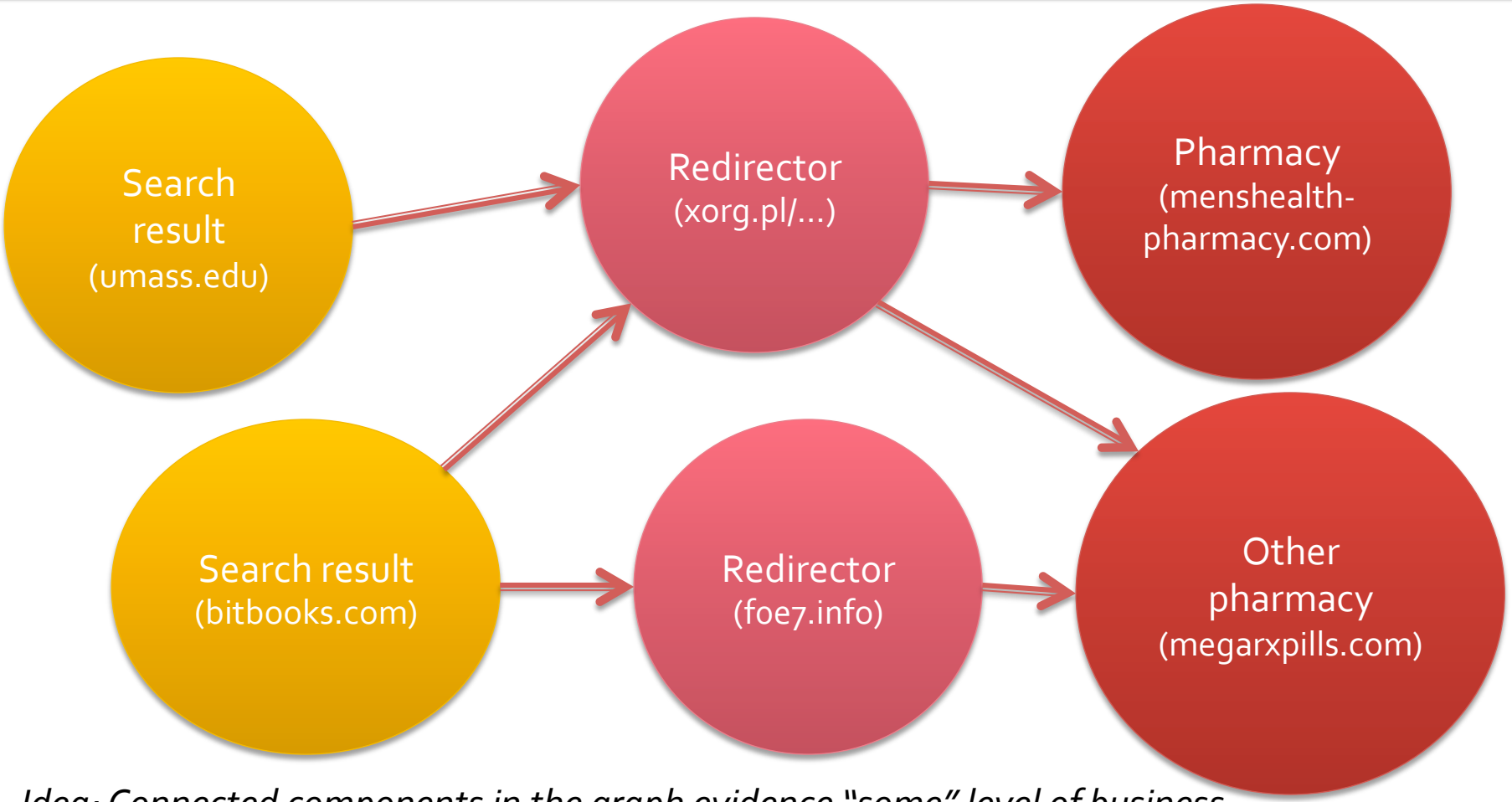
Persistent infections





.edu sites particularly attractive, as well as high PageRank sites (often sites fall in both categories)

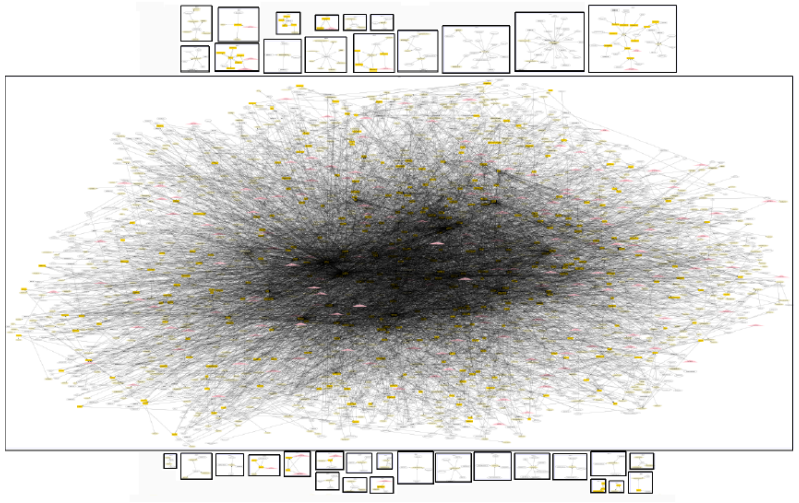
Uncovering relationships in search results



Idea: Connected components in the graph evidence "some" level of business relationships between the nodes they connect

Connected components

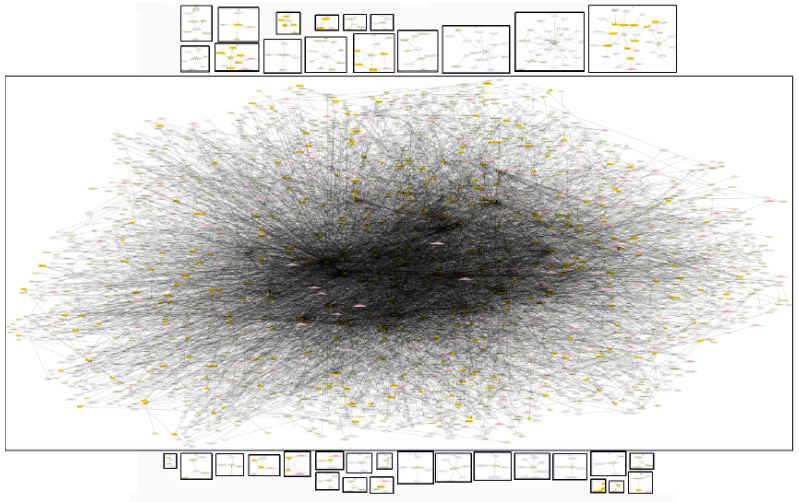
- 34 connected components
- One connected component contains
  - 96% of all infected domains
  - 90% of all redirection domains
  - 92% of all pharmacies



- Is one person responsible for all of this?!

Connected components

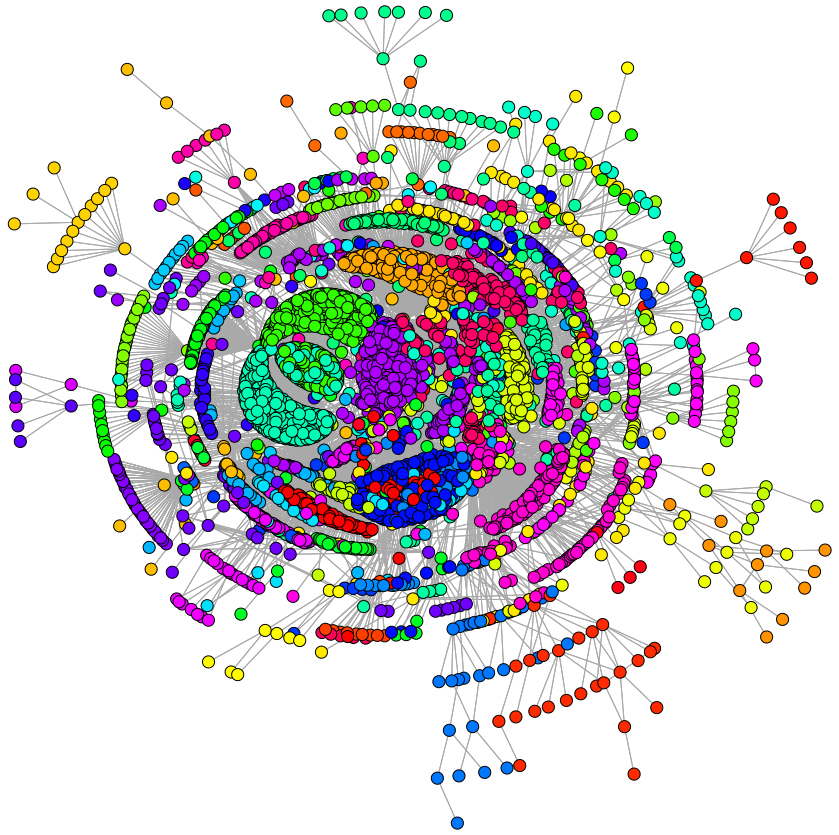
- 34 connected components
- One connected component contains
  - 96% of all infected domains
  - 90% of all redirection domains
  - 92% of all pharmacies



- Is one person responsible for all of this?!
  - **NO!**
  - Some advertisers work for several different affiliate networks
  - Certain domains are (ab)used by multiple advertisers

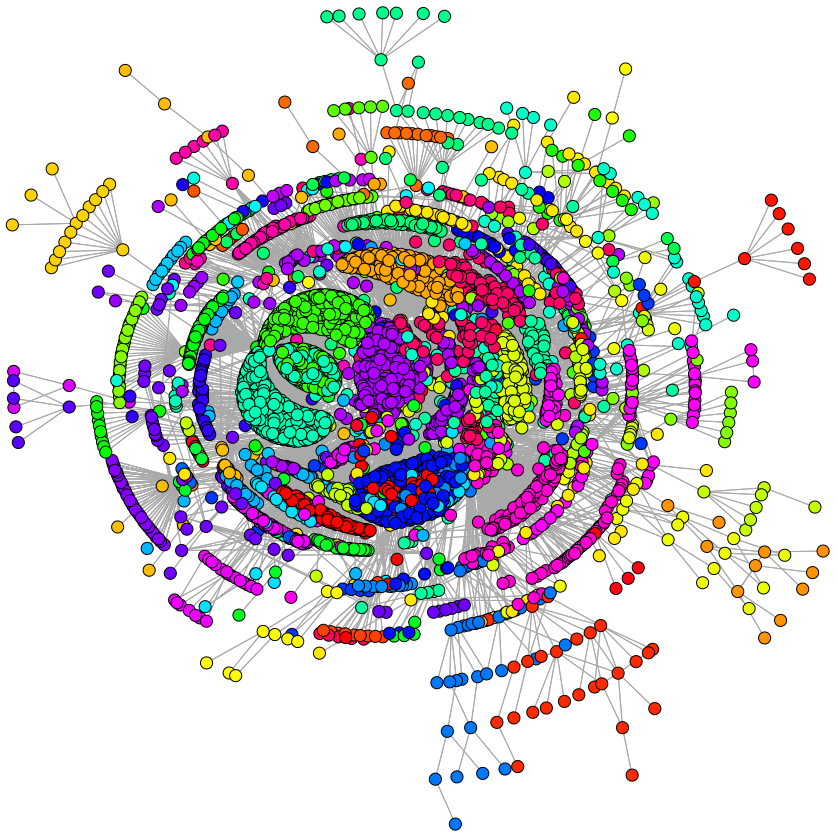
Identifying the main players

- Run (spinglass) clustering algorithm in big connected component
- Each cluster represented by different color
- Evidence of separate organized groups/ campaigns more loosely connected to each other
  - About 10-12 large groups



Identifying the main players

- Clusters may denote
  - A set of hosts compromised at the same time
  - A set of hosts all compromised by the same people
- Interesting network properties
  - All major redirectors hosted on 11 ASes



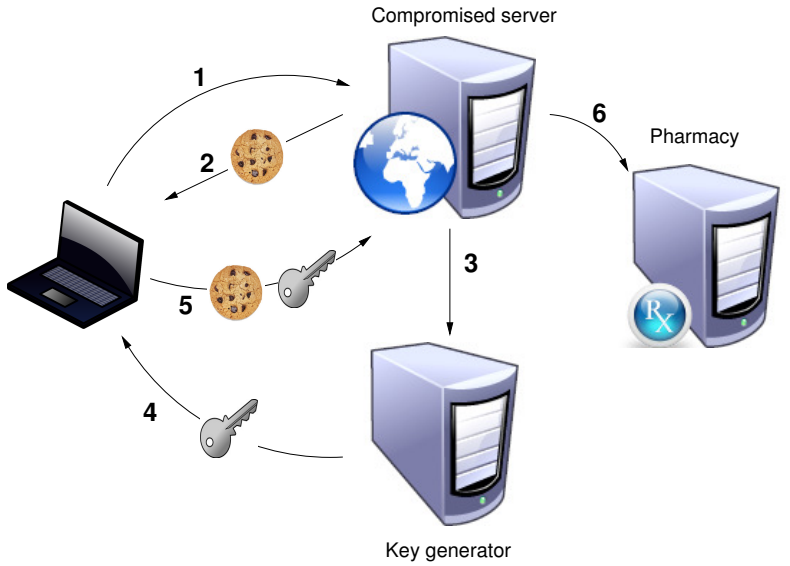
Possible interventions (client side)

- Blanking HTTP referrer field
  - Default behavior with HTTPS
    - Firefox 14+ (rel. 7/2012) automatically uses HTTPS connections to Google (and other search engines)

- **Not necessarily enough**
  - One may infer that a blank referrer indicates provenance from Google (as opposed to a page from the same university, for instance)

Novel attack variants (2012+)

- Inlining the store front in the compromised website
  - A click then leads you to the pharma store

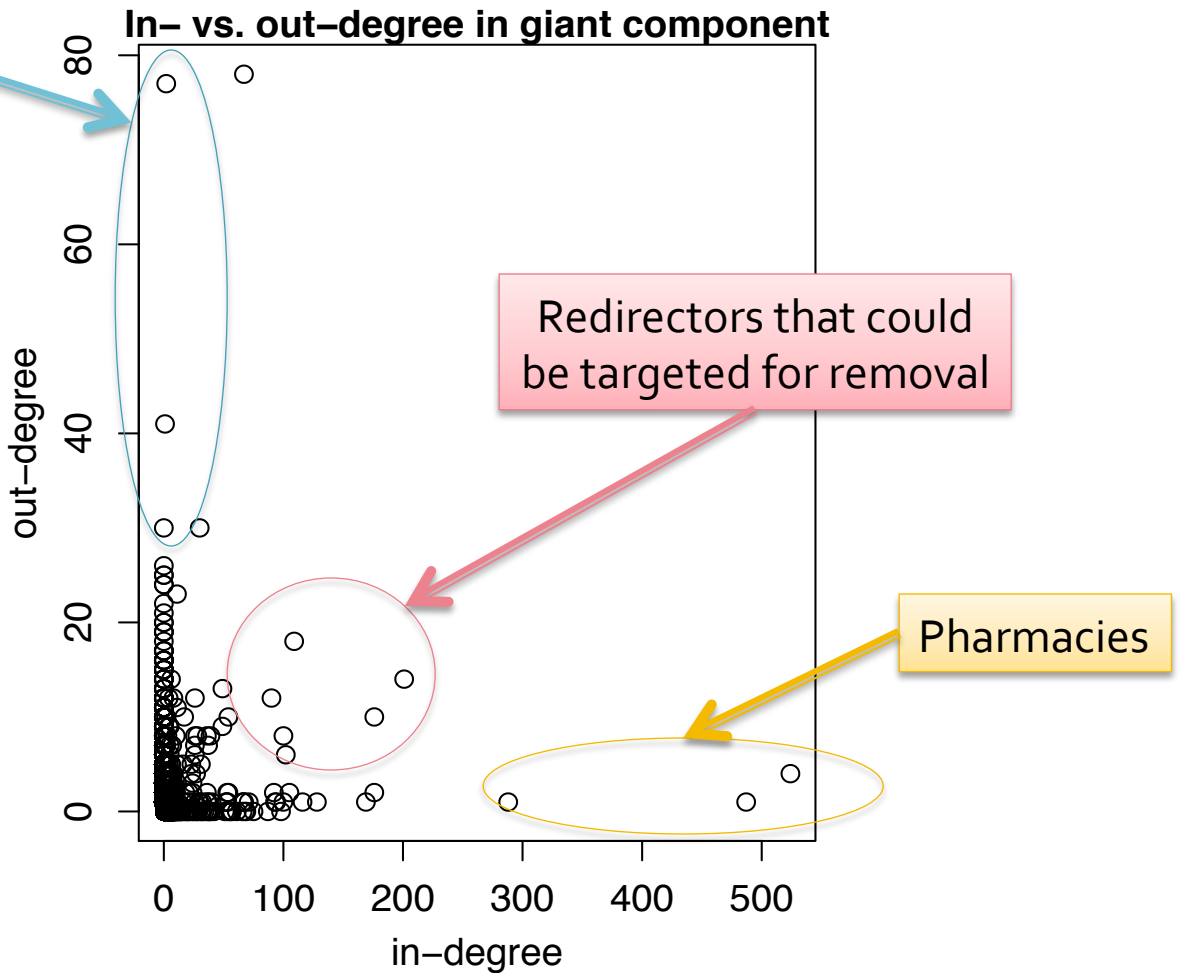


- Stateful server
- Goal: escape detection by automated crawlers

- Attackers are quickly adapting to monitoring and/or countermeasures

Possible interventions (network side)

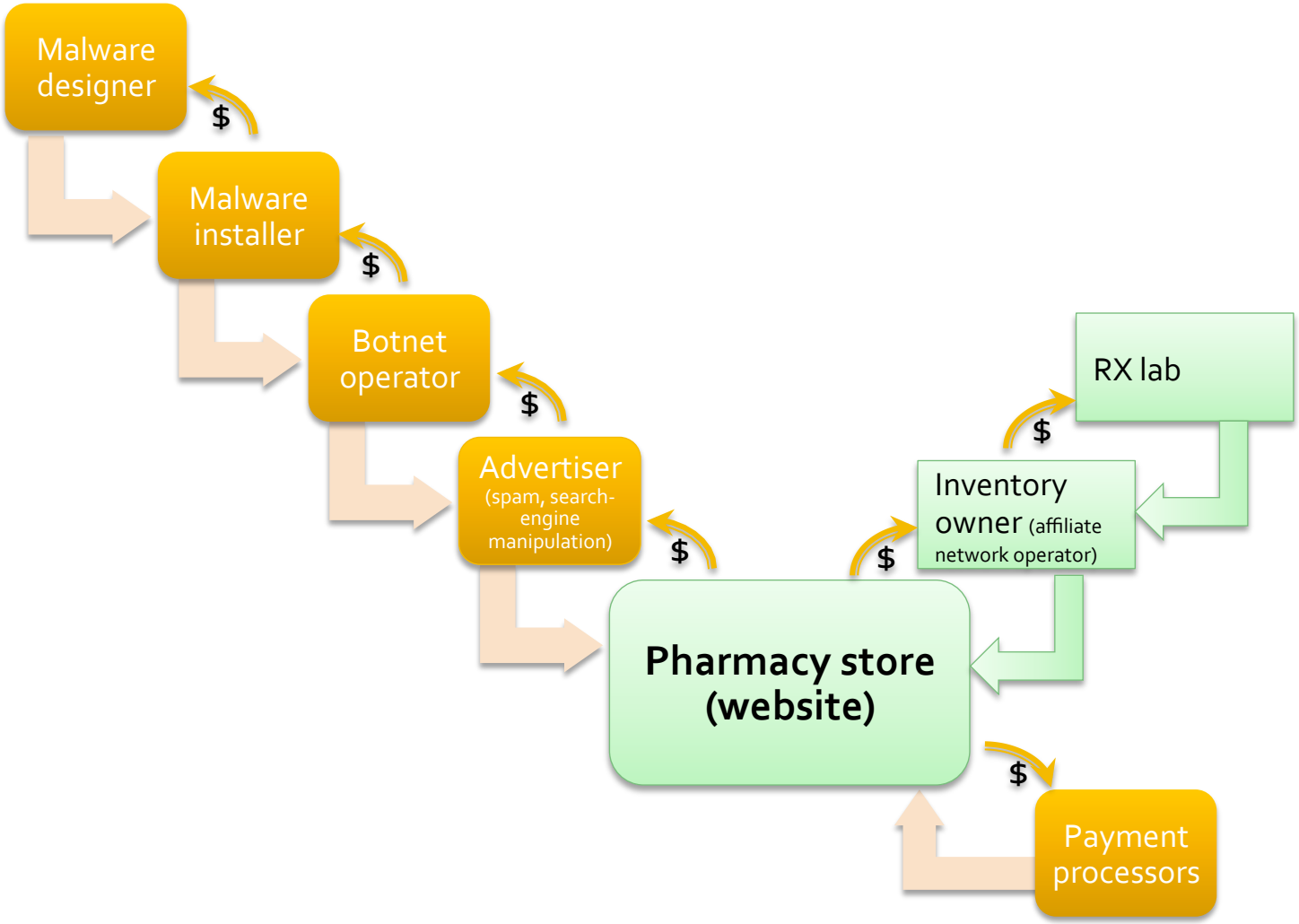
Valuable infected sites (e.g., umass.edu)



Redirectors that could be targeted for removal

Pharmacies

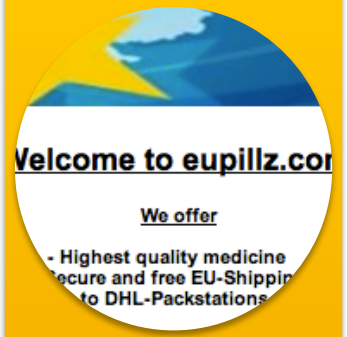
Part II: Procuring unlicensed drugs



Different types of pharmacies

Licensed pharmacies

- Endorsed by National Association Board of Pharmacies



Not licensed, not blacklisted

- No endorsement
- Usually means they do not meet requirements for licensing
- Perhaps haven't been checked – basically status unknown

"Not recommended"

- Blacklisted by the NABP
- Usually means they do clearly fail one requirements for licensing
- Set of 9679 shops

Unlicensed, abusive

- Present in our corpus of pharmacies using search-engine manipulation
- Not necessarily blacklisted by the NABP!

Black market

- No claims of legitimacy
- RX drugs only part of the business
- Can find narcotics as well
- Tor hidden service (anonymous)

Most legitimate ← → Least legitimate

Inventory analysis [LMC, EC' 13]

Inventory analysis [LMC, EC'13]

- **Research question 1:** How do unlicensed online pharmacies even survive?
  - What do unlicensed pharmacies offer that allows them to compete with licensed pharmacies?

Inventory analysis [LMC, EC'13]

- **Research question 1:** How do unlicensed online pharmacies even survive?
  - What do unlicensed pharmacies offer that allows them to compete with licensed pharmacies?

Cheaper prices / Volume discounting

Different inventories (e.g., counterfeits)

Inventory analysis [LMC, 2013 (in submission)]

- **Research question 1:** How do unlicensed online pharmacies even survive?
  - What do unlicensed pharmacies offer that allows them to compete with licensed pharmacies?
- **Research question 2:** Can we identify patterns in drug suppliers?

Inventory analysis [LMC, 2013 (in submission)]

- **Research question 1:** How do unlicensed online pharmacies even survive?
  - What do unlicensed pharmacies offer that allows them to compete with licensed pharmacies?

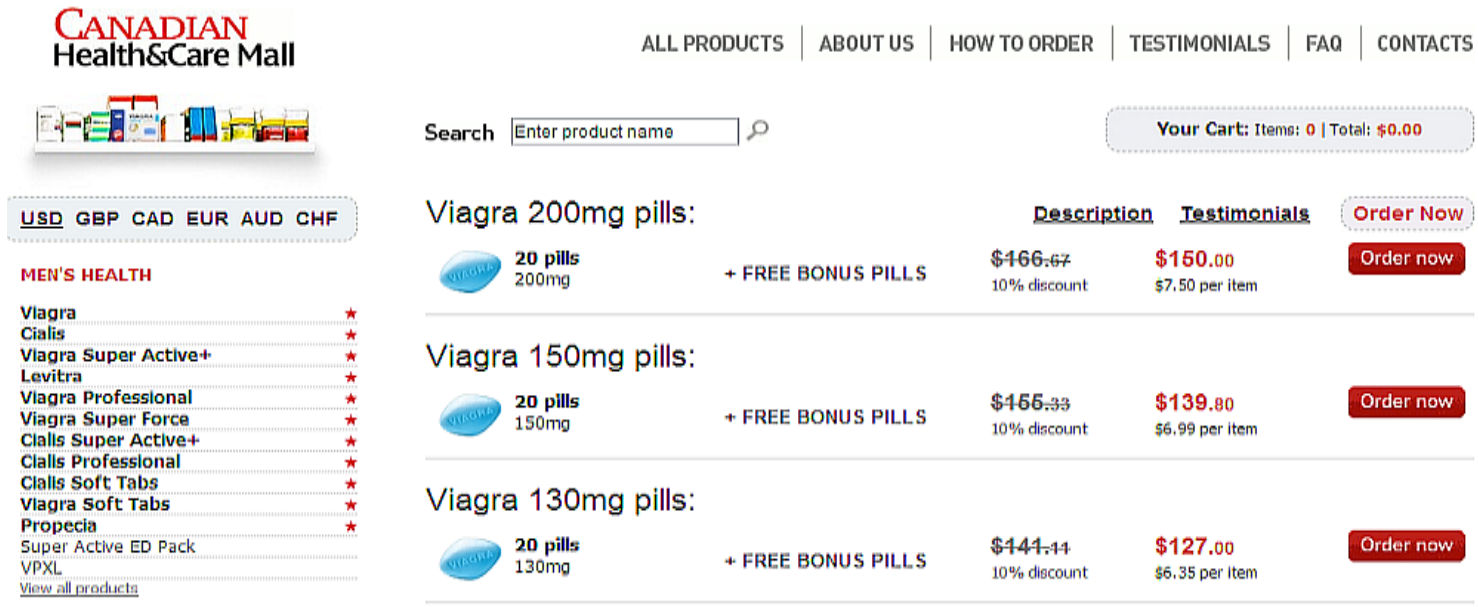
- **Research question 2:** Can we identify patterns in drug suppliers?

Data collected

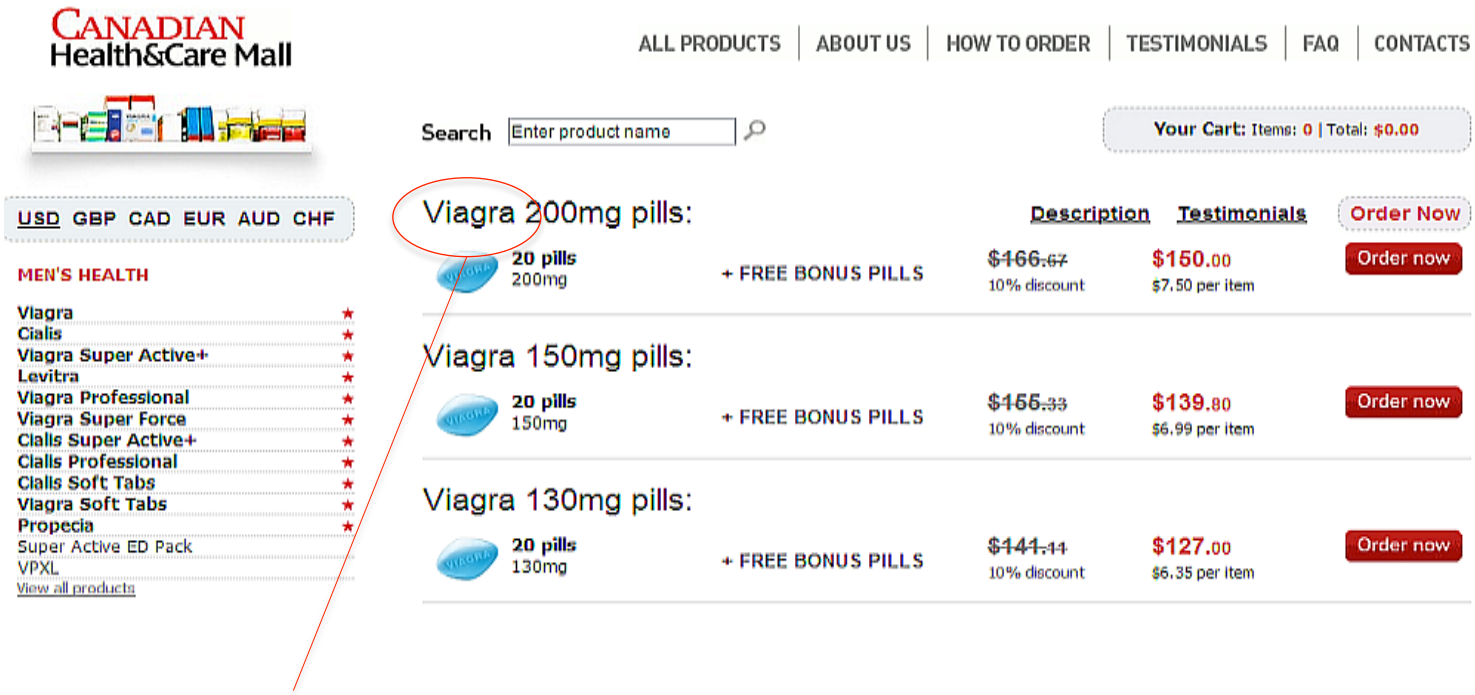
- Scraped for prices and inventories:
  - 265 unlicensed pharmacies (doing search-redirection attacks) collected between April 3, 2012 and October 16, 2012
  - 265 "blacklisted" pharmacies
    - Randomly sampled out of a corpus of 9000+ NABP "not recommended" pharmacies
    - No overlap with the unlicensed pharmacy corpus
- We also collected for comparison

  (but I won't use this data in this presentation):
  - The entire Silk Road black market between February and July 2012 (see [Christin, WWW 2013])
  - *familymeds.com*, a licensed online pharmacy

Scraping

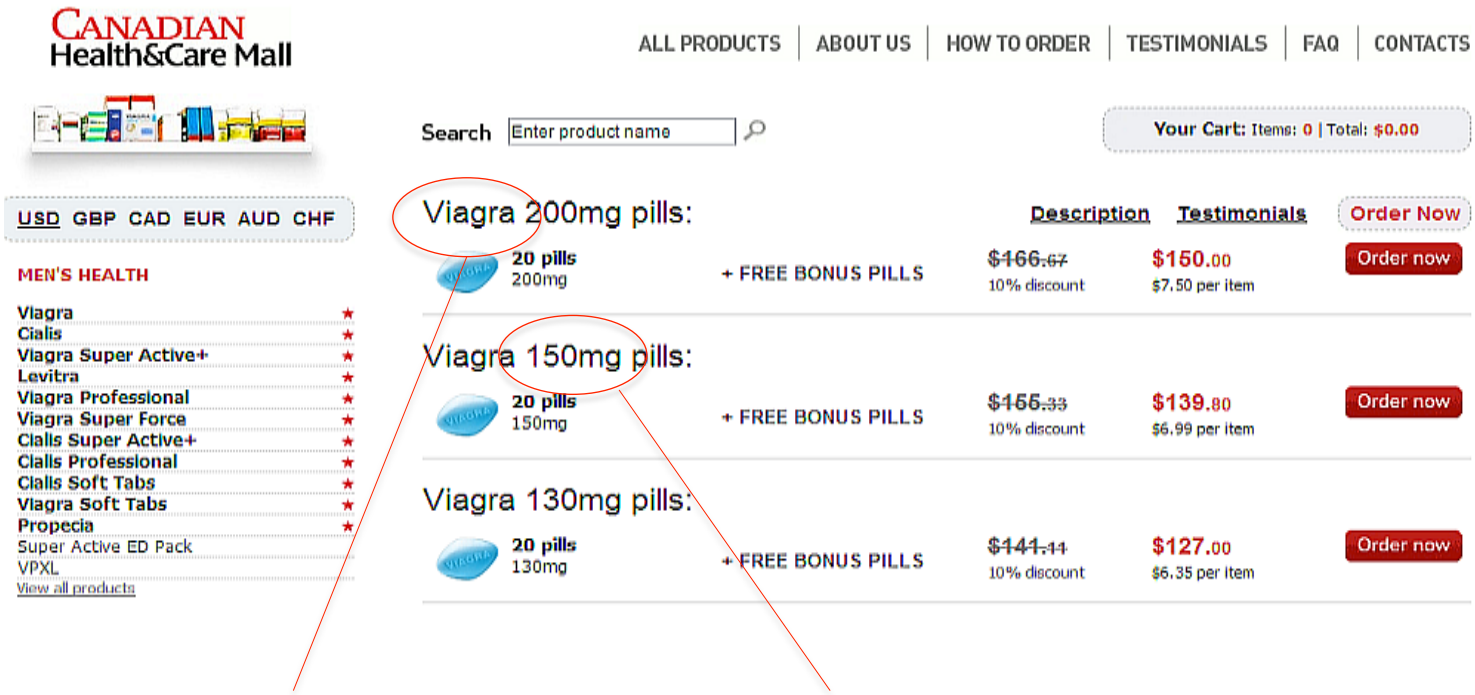

CANADIAN
Health&Care Mall
ALL PRODUCTS | ABOUT US | HOW TO ORDER | TESTIMONIALS | FAQ | CONTACTS
Search Enter product name
Your Cart: Items: 0 | Total: $0.00
USD GBP CAD EUR AUD CHF
MEN'S HEALTH
Viagra
Cialis
Viagra Super Active+
Levitra
Viagra Professional
Viagra Super Force
Cialis Super Active+
Cialis Professional
Cialis Soft Tabs
Viagra Soft Tabs
Propecia
Super Active ED Pack
VPXL
View all products
Viagra 200mg pills:
Description Testimonials Order Now
20 pills
200mg
+ FREE BONUS PILLS
$166.67
10% discount
$150.00
$7.50 per item
Order now
Viagra 150mg pills:
20 pills
150mg
+ FREE BONUS PILLS
$166.33
10% discount
$139.80
$6.99 per item
Order now
Viagra 130mg pills:
20 pills
130mg
+ FREE BONUS PILLS
$141.11
10% discount
$127.00
$6.35 per item
Order now

Scraping



Drug name/product (subsequently mapped to active ingredient/condition)

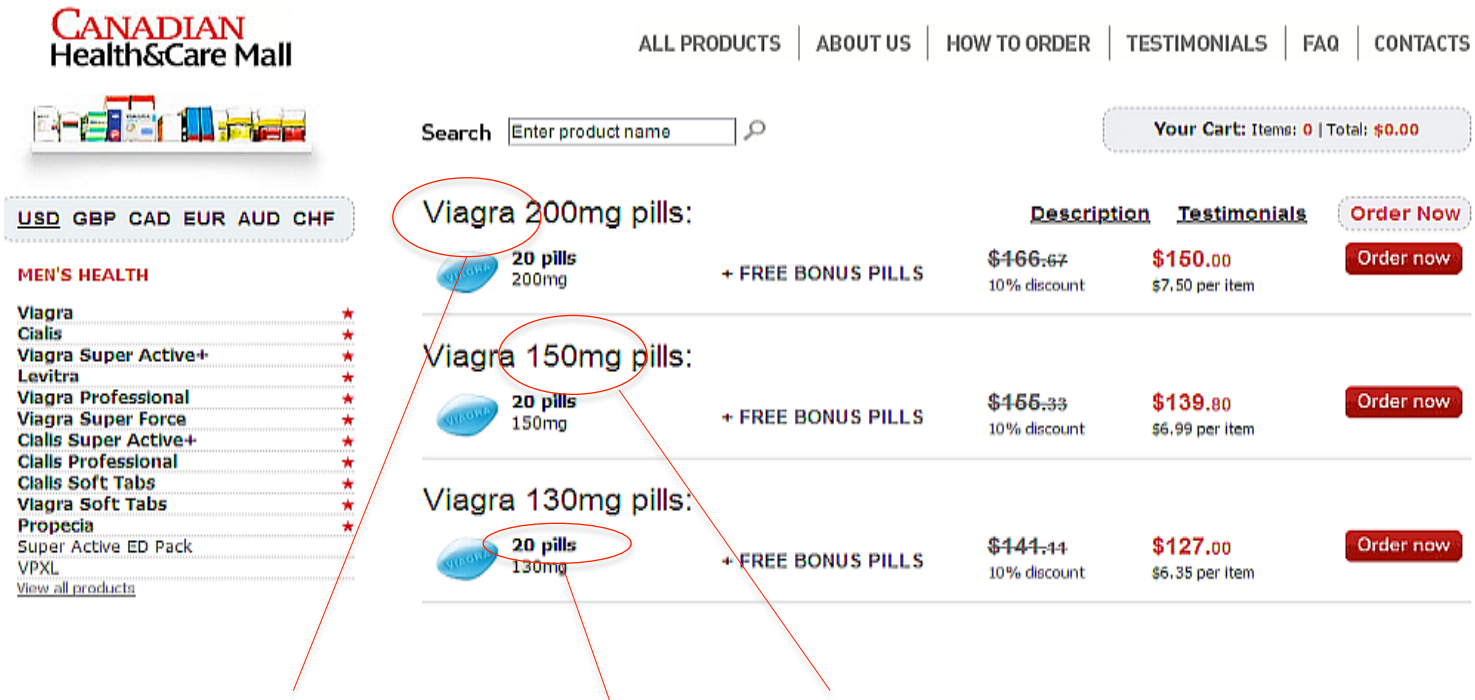
Scraping



Drug name/product (subsequently mapped to active ingredient/condition)

Dosage

Scraping

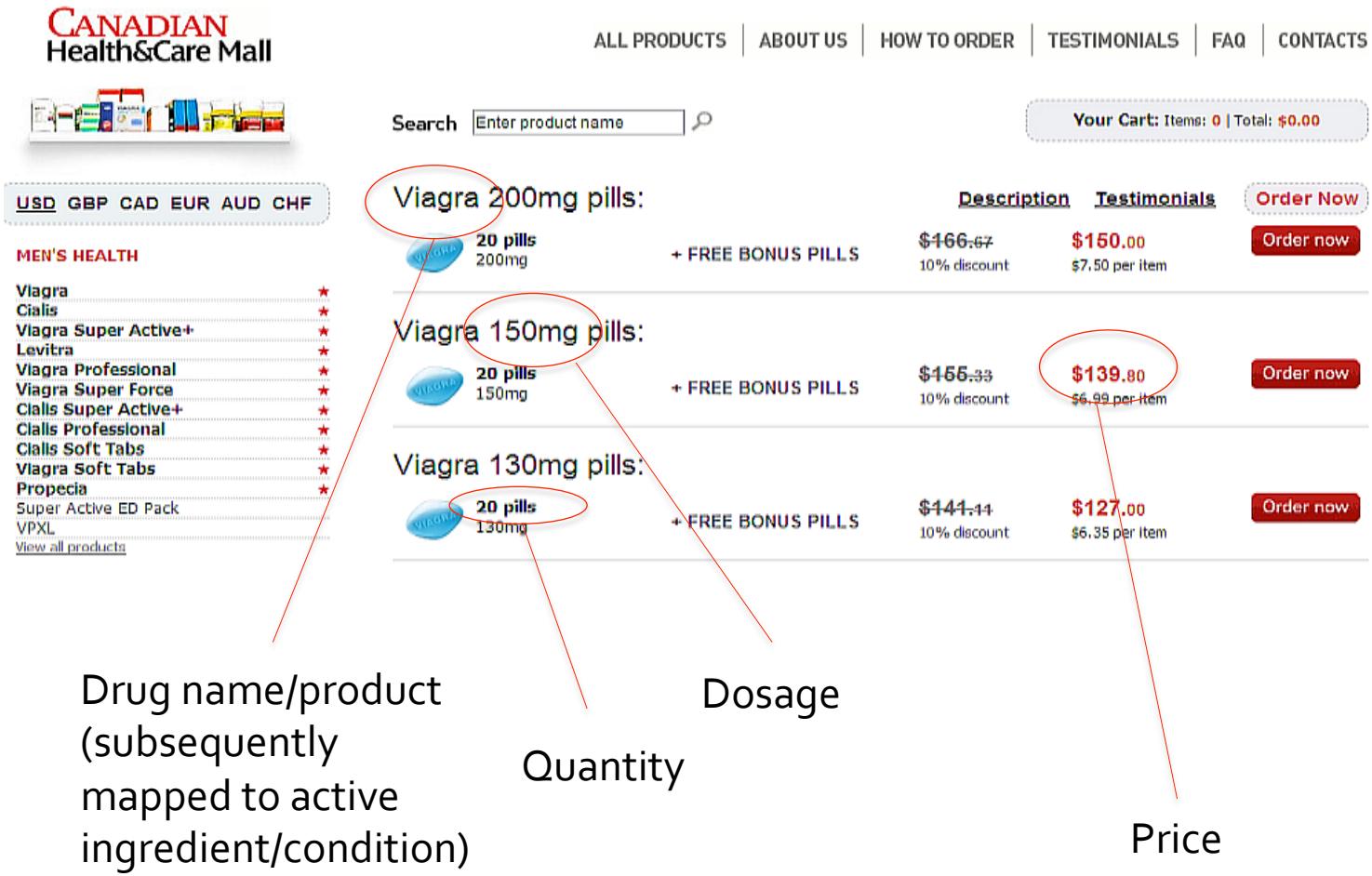


Drug name/product (subsequently mapped to active ingredient/condition)

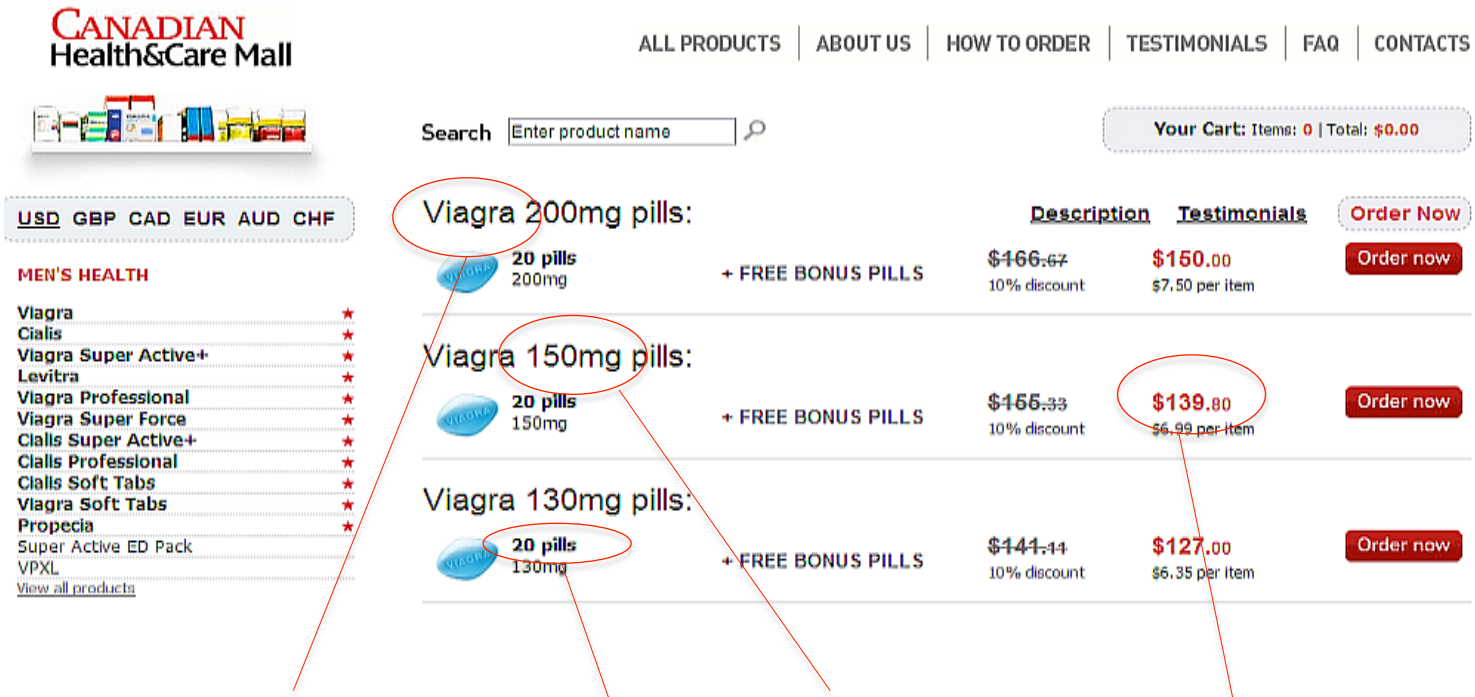
Quantity

Dosage

Scraping


CANADIAN
Health&Care Mall
ALL PRODUCTS | ABOUT US | HOW TO ORDER | TESTIMONIALS | FAQ | CONTACTS
Search Enter product name
Your Cart: Items: 0 | Total: $0.00
USD GBP CAD EUR AUD CHF
MEN'S HEALTH
Viagra
Cialis
Viagra Super Active+
Levitra
Viagra Professional
Viagra Super Force
Cialis Super Active+
Cialis Professional
Cialis Soft Tabs
Viagra Soft Tabs
Propecia
Super Active ED Pack
VPXL
View all products
Viagra 200mg pills:
Description
Testimonials
Order Now
20 pills
200mg
+ FREE BONUS PILLS
$166.67
10% discount
$150.00
$7.50 per item
Order now
Viagra 150mg pills:
20 pills
150mg
+ FREE BONUS PILLS
$166.33
10% discount
$139.80
$6.99 per item
Order now
Viagra 130mg pills:
20 pills
130mg
+ FREE BONUS PILLS
$141.11
10% discount
$127.00
$6.35 per item
Order now
Drug name/product (subsequently mapped to active ingredient/condition)
Quantity
Dosage
Price

Scraping



Drug name/product (subsequently mapped to active ingredient/condition)

Quantity

Dosage

Price

Total = 1,451,587 distinct (drug, active ingredient, dosage, unit) tuples collected
1,661 different drug names

Identifying common suppliers: Inventory overlap

- How much overlap is there between distinct unlicensed pharmacies' inventories?
- Jaccard distance:

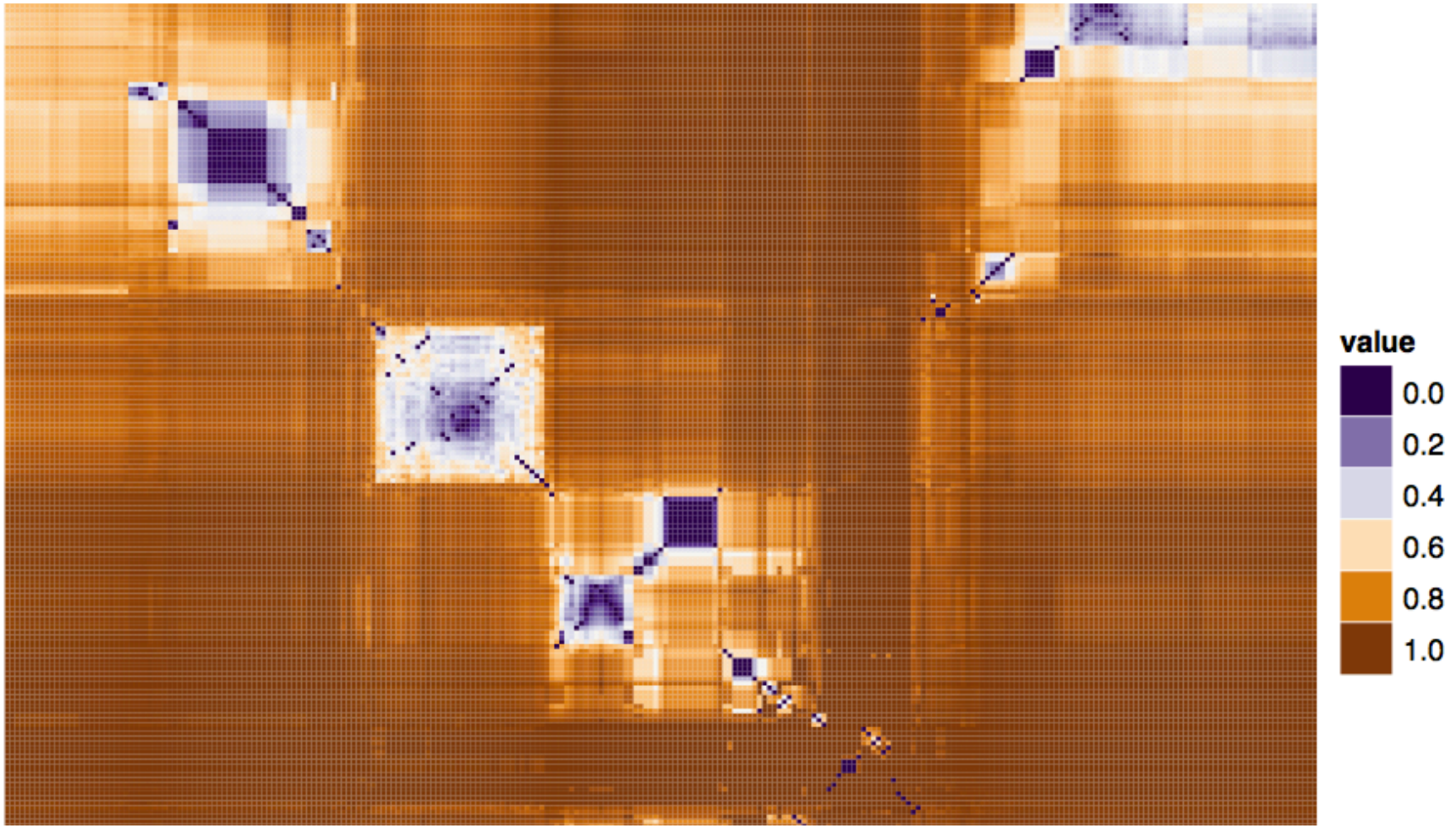
$$J_\delta(A,B) = 1 - \frac{|A \cap B|}{|A \cup B|}$$

Inventory of pharmacy $\mathscr{A}$

Inventory of pharmacy $\mathscr{B}$

  - Identical inventories $\Leftrightarrow J_\delta(A,B) = 0$
  - No overlap at all $\Leftrightarrow J_\delta(A,B) = 1$

Identifying common suppliers: Inventory overlap

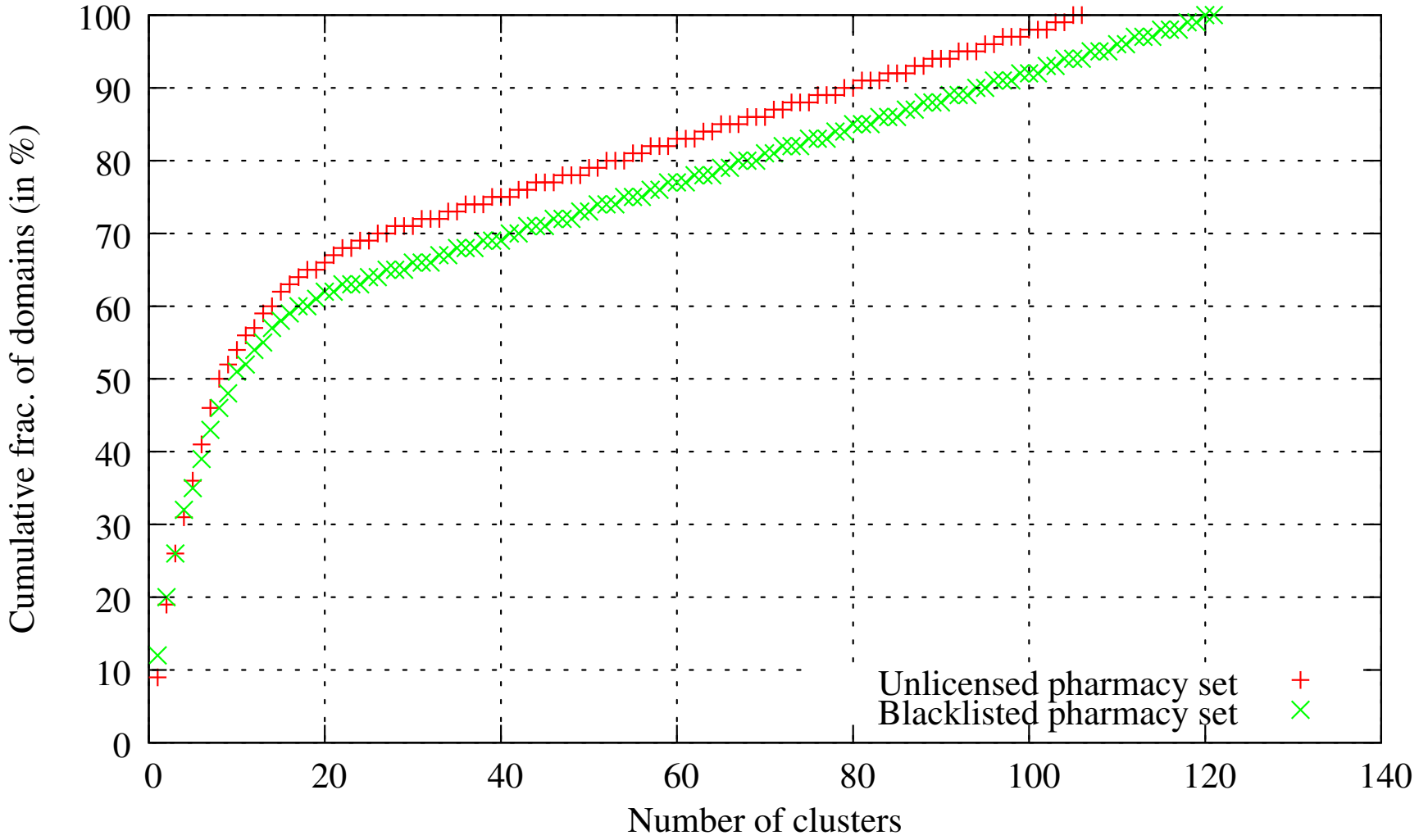


Each point *(i,j)* represents $J_\delta(i,j)$; columns and rows are then reordered to highlight clusters

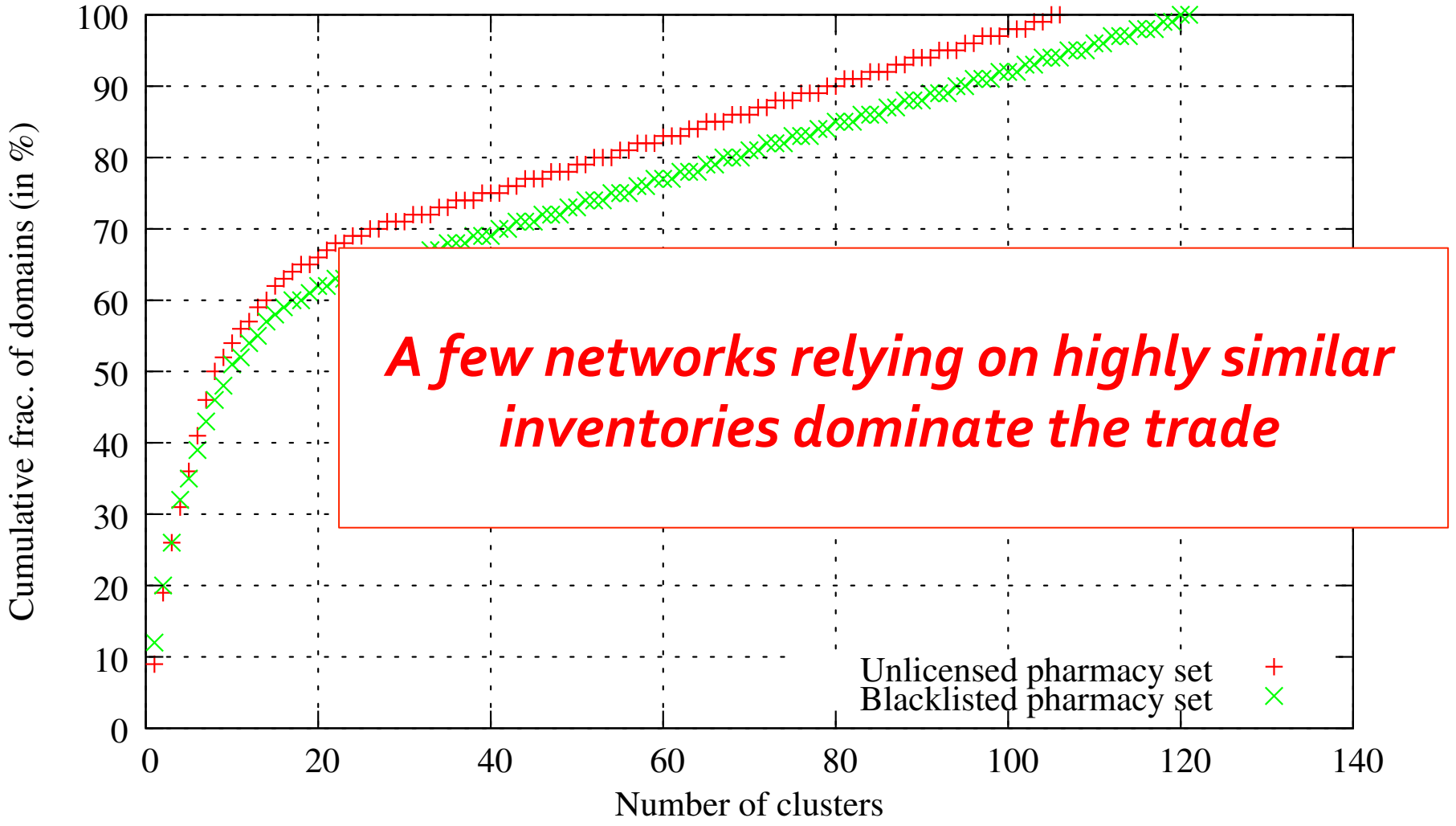
Clustering inventories

- Inventory A and Inventory B belong to the same cluster iff $J_\delta(A, B) < t$
  - t is an arbitrary threshold, $0<t<1$
- Distance between two groups of inventories X, Y:
  - Minimum linkage: $J_\delta(X, Y) = \min\{J_\delta(x, y) : x \in X, y \in Y\}$
  - Maximum linkage: $J_\delta(X, Y) = \max\{J_\delta(x, y) : x \in X, y \in Y\}$
  - Average linkage: $J_\delta(X, Y) = \frac{1}{|X| \cdot |Y|} \sum_{x \in X} \sum_{y \in Y} J_\delta(x, y)$

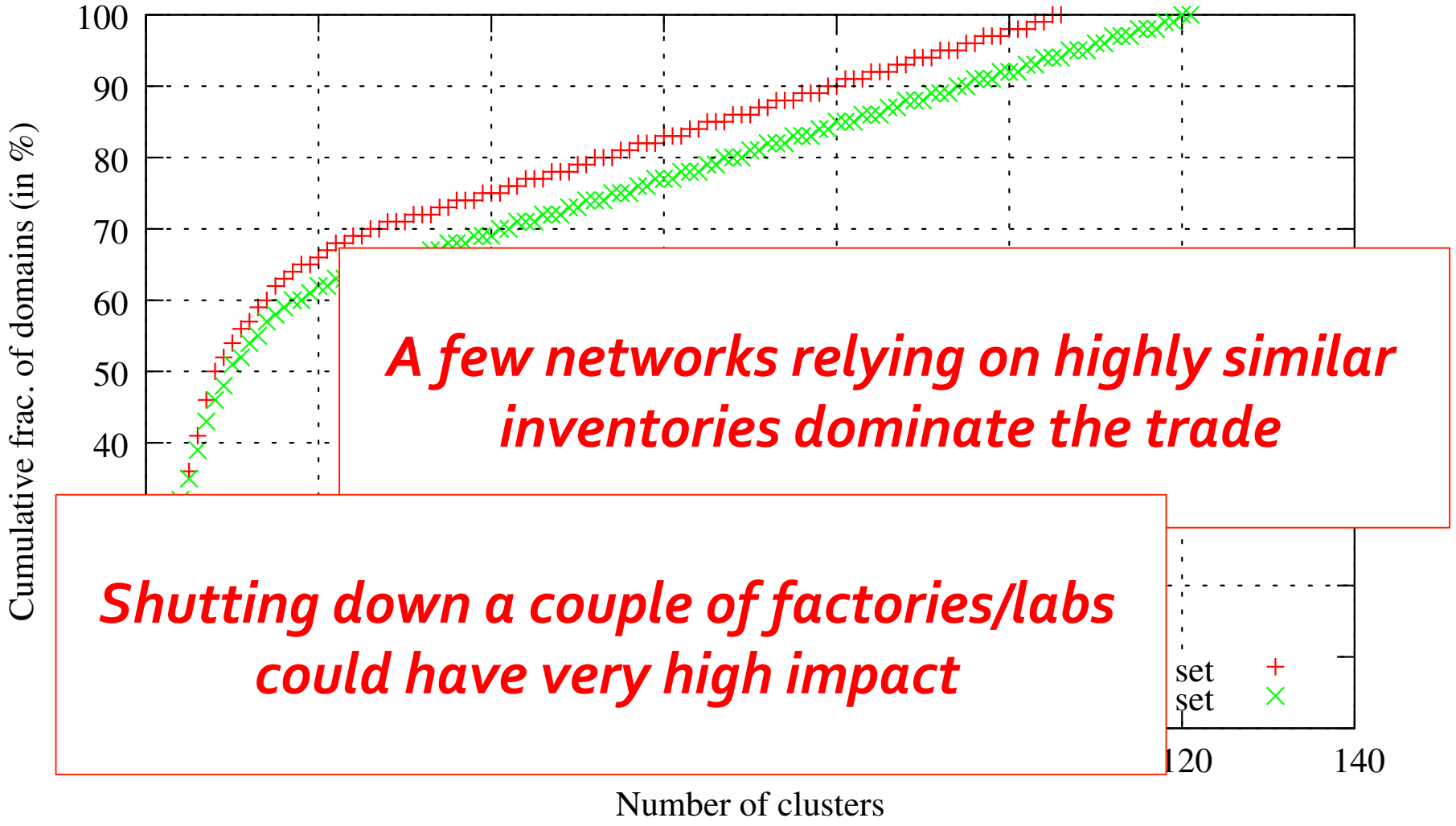
Clustering inventories: Average linkage, *t=0.31*



Clustering inventories: Average linkage, *t=0.31*



Clustering inventories: Average linkage, *t=0.31*



A few networks relying on highly similar inventories dominate the trade

Shutting down a couple of factories/labs could have very high impact

Summary of results

- Unlicensed online pharmacies use
  - Elaborate advertising mechanisms to circumvent boycotts
    - E.g., can't advertise w/ Google adwords
  - A limited number of suppliers
    - Eight groups of pharmacies with similar inventories represent more than half the pharmacy population
- Intervention strategies
  - Focus on **concentration points**
    - Limited number of advertisers
    - Limited number of labs providing inventories
  - Probably much more scalable than other approaches
    - E.g., shutting down pharmacies one-by-one

Summary of results

- Unlicensed online pharmacies use
  - Elaborate advertising mechanisms to circumvent boycotts
    - E.g., can't advertise w/ Google adwords
  - A limited number of suppliers
    - Eight groups of pharmacies with similar inventories represent more than half the pharmacy population
- Intervention strategies
  - Focus on **concentration points**
    - Limited number of advertisers
    - Limited number of labs providing inventories
  - Probably much more scalable than other approaches
    - E.g., shutting down pharmacies one-by-one

How do you evaluate success?

Some successful interventions

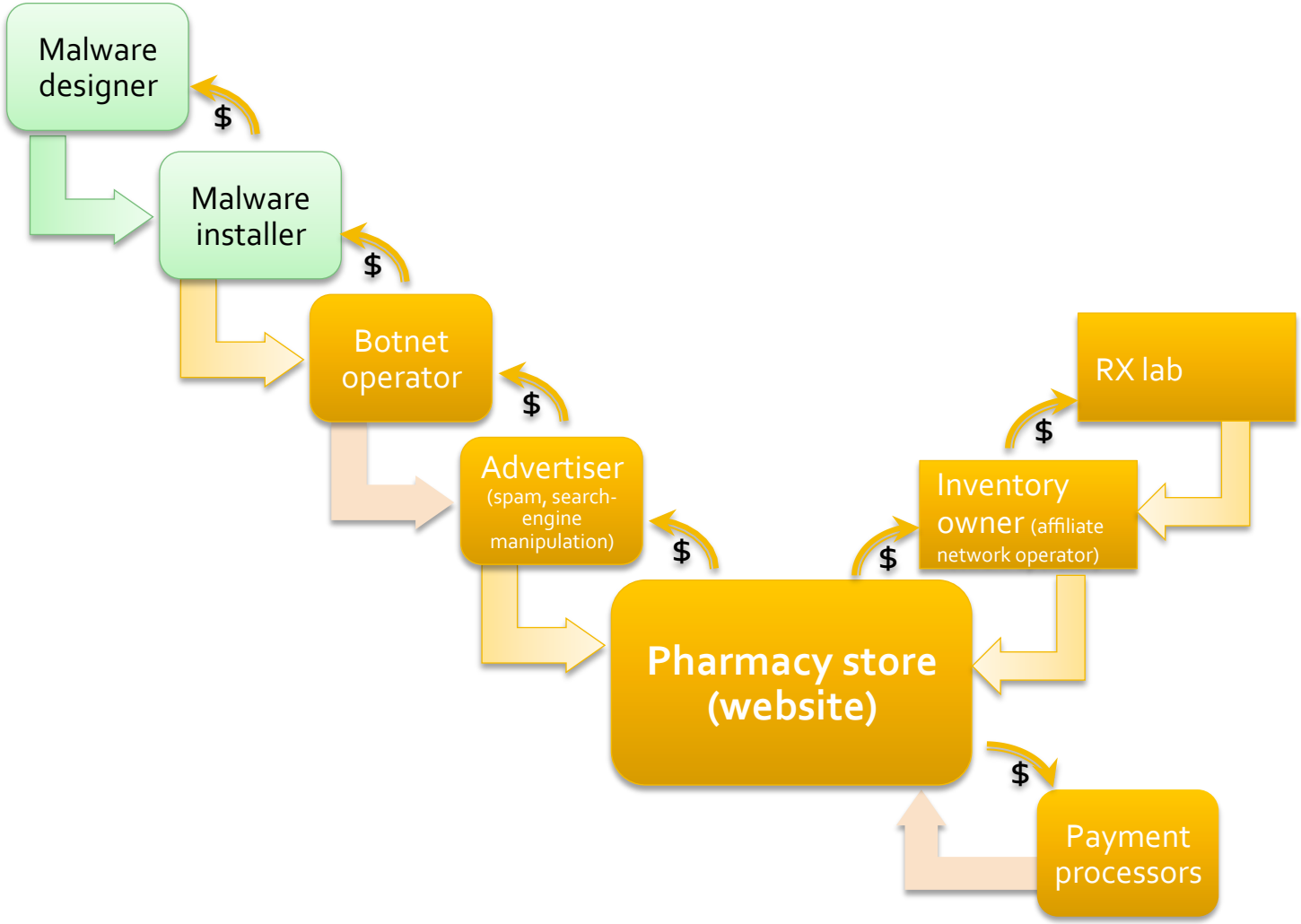
- Payment processors
  - UCSD/Berkeley work on pharmacies/email spam
  - Visa/banks started clamping down on processors
  - Significant disruptions reported by operators in forums

- Carders markets
  - Large exchanges of credit card and financial data
  - FBI identified, infiltrated and took down the largest market
  - Those forums haven't really recovered since

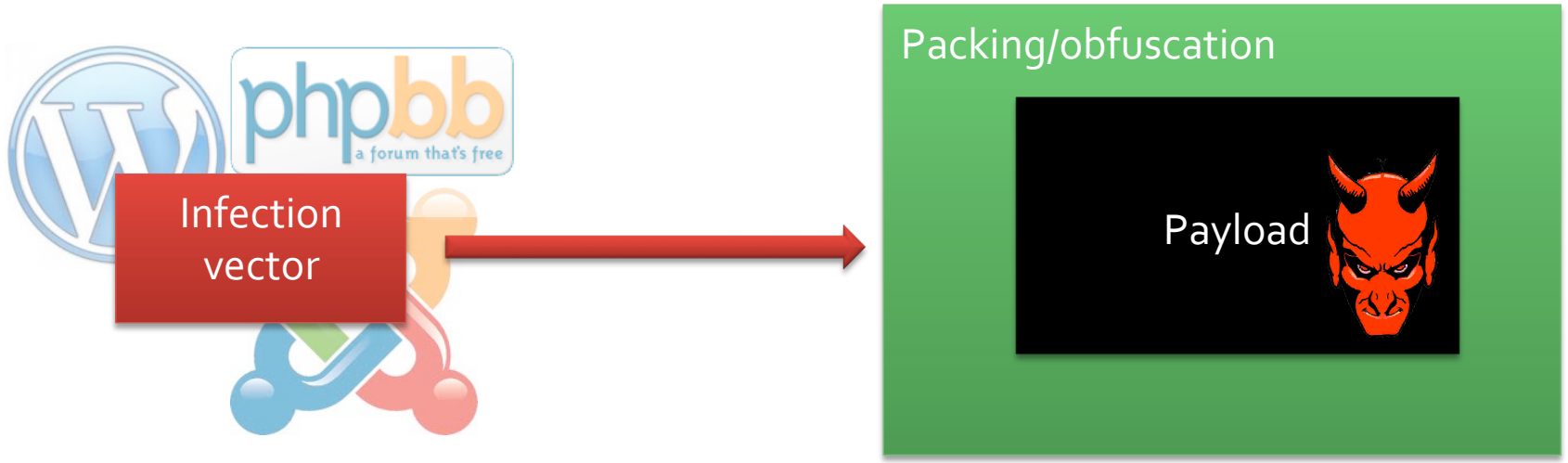
Ongoing research thesis

- All financially-motivated online criminal activities exhibit similar concentration or "choke" points
- Successful intervention mechanisms will have to primarily target these choke points

Next: Analysis of web malware



Web-server malware



- Gather web-server malware
- Cluster by infection vectors (how are vulnerable sites found and infected)
- Cluster by packing/obfuscation techniques
- Code similarity analysis (after reverse engineering) of payload

➔ could yield interesting information about malware designers

Web-based malware

- E.g., drive by downloads
- Completely different!
  - Doesn't attack a crummy WordPress install
  - Instead compromise a browser or even an entire O/S
- The game here is to bring potential victims to the server that delivers the malware
  - Again, traffic "optimization"
    - Trending term campaigns, etc …

Take-away points

- Financially-motivated online crime is an important part of all security attacks
- Attackers are rational: reducing profits will lead them to abandon certain attacks
  - E.g., trending-term abuse has been greatly curbed by Google's change in algorithms [CCS 2011]
- Complex eco-system, similar to industrial supply chains
- Modeling supply chains → discover "choke" points
- Targeting those choke points offers viable intervention strategies

Relevant papers

Measuring and Analyzing Search-Redirection Attacks in the Illicit Online Prescription Drug Trade

Joint work with N. Leontiadis and T. Moore. In *Proceedings of the 20th USENIX Security Symposium (USENIX Sec'11)*. San Francisco, CA, USA. August 2011.

Pick Your Poison: Pricing and Inventories at Unlicensed Online Pharmacies

Joint work with N. Leontiadis and T. Moore. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC'13).* Philadelphia, PA, USA. June 2013

Traveling the Silk Road: An measurement analysis of a large online marketplace

In *Proceedings of the 25th International World Wide Web Conference (WWW 2013)*. Rio de Janeiro, Brazil. May 2013.

Related/further reading:

Dissecting One Click Frauds

Joint work with S. Yanagihara and K. Kamataki. In *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS 2010)*, pages 15-26. Chicago, IL, USA. October 2010.

Fashion Crimes: Trending-Term Exploitation on the Web

Joint work with T. Moore and N. Leontiadis. In *Proceedings of the 18th ACM Conference on Computer and Communications Security (CCS 2011)*. Chicago, IL, USA. October 2011.

https://www.andrew.cmu.edu/user/nicolasc/papers.html

Questions?

Thank you!

Nicolas Christin nicolasc@cmu.edu
Carnegie Mellon University ECE/CyLab