

A Domain-Agnostic Approach to Spam-URL Detection via Redirects

Heeyoung Kwon



Mirza Basim Baig



Leman Akoglu



Carnegie Mellon University

Era of Spam

 **toby** @Leisahmdlw · 7h
 @agomezromero USA Government trying to shutdown Bitcoin network
 more here: bit.ly/1lzmyiC
 Expand Reply Retweet Favorite

 **A. M. S.** @AgungMahardilan · 7h
 Yah terus.. :| @Maisieojdqy: @AgungMahardilan USA Government trying t
 shutdown Bitcoin network read more here: bit.ly/1gDKsOh
 Expand Reply Retweet Favorite

 **Buzzy** @buzzardbob · 7h
 Hum ? I don't care RT @Wackerly_59899: @buzzardbob USA Government
 trying to shutdown Bitcoin network read more here: bit.ly/1lzmkbo
 Expand Reply Retweet Favorite

 **Sondra Underkoffler** @Dorlakezjm · 7h
 @marlamax13 USA Government trying to shutdown Bitcoin network read
 more here: bit.ly/1gC
 Expand



ha ha check this out..she is soo busted



CLICK HERE to see the status update that got a girl expelled from school!! you got to see this she's in such trouble

January 18 at 4:50pm via ef3 · Like · Comment

LinkedIn

Chandi Nagaraj has sent you a message.

Date: 8/22/2012

Subject: RE: "It Works!" (I think you can get 37 checks in 29 days with this too)

FREE! "How To Get Up To 37 Checks Per Month, Earn Upwards Of \$4,954.55 While You Sleep At Night, And Recruit WITHOUT Ever Having To Pick UPhone"

====> www.maximum-leverage.com/free.php?aff=knagaraj

See You At The TOP!

Chandi & Kavya

Market For The Future
9455 Lanham Severn Road
Lanham, Maryland 20706
United States

or contact details, Email us @

[message](#)

Era of Spams

The image displays a collage of social media spam examples and a line graph. On the left, there are four tweets from various users, including @Leisahmdlw, @agomezromero, @AgungMahard, and @marlamax13, all containing suspicious links and promotional text. In the center, a blue line graph titled "Social media spam is on the rise. It grew 658% since mid - 2013" shows a sharp increase in spam volume from 2013 to 2014. On the right, a screenshot of a LinkedIn message from Chandi Nagaraj is shown, containing a link and a reference to a "message". Below the graph, a small image of a woman is shown with a caption that reads "expelled from school!! you got to see this she's in such trouble".

LinkedIn
Chandi Nagaraj has sent you a message.

[1]
t 37 checks in 29 days with this too)
Month, Earn
Night, And
hone"

[hp?aff=knagaraj](#)

details, Email us @

[message](#)

expelled from school!!
you got to see this
she's in such trouble

January 18 at 4:50pm via ef3 · Like · Comment

Year	Spam Volume (Relative)
2010	Low
2011	Low
2012	Low
2013	Low
2014	High (658% increase)

[1] Social Media Spamming Grew By 658% Between 2013 And 2014: Entertainment, Financial And News Categories Main Target, <https://dazeinfo.com/2014/12/15/social-media-spamming-growth-2014-facebook-twitter-entertainment/>

Popular Solutions

- IP blacklisting
 - Popular for social media and URL shortening services
 - False negative rates between 40.2 to 98.1%
 - Slow and unscalable
- Account based approach
 - Limited ability to detect compromised accounts
 - Require a history of malicious behavior
 - Not generalizable to different services

Popular Solutions

- IP blacklisting

URL-level decisions are required

- able to filter individual post
- more generalizable

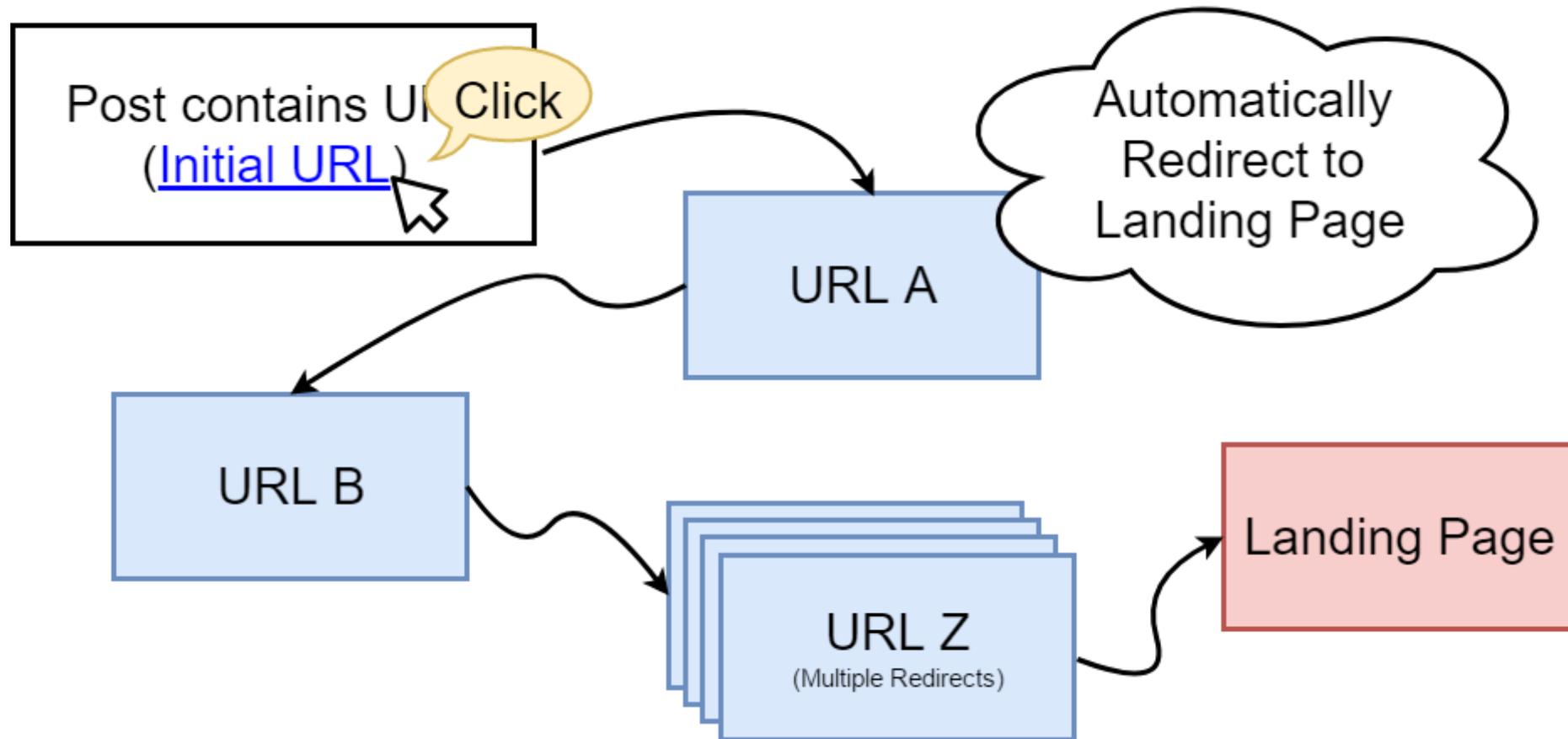
- A

- Require a history of malicious behavior

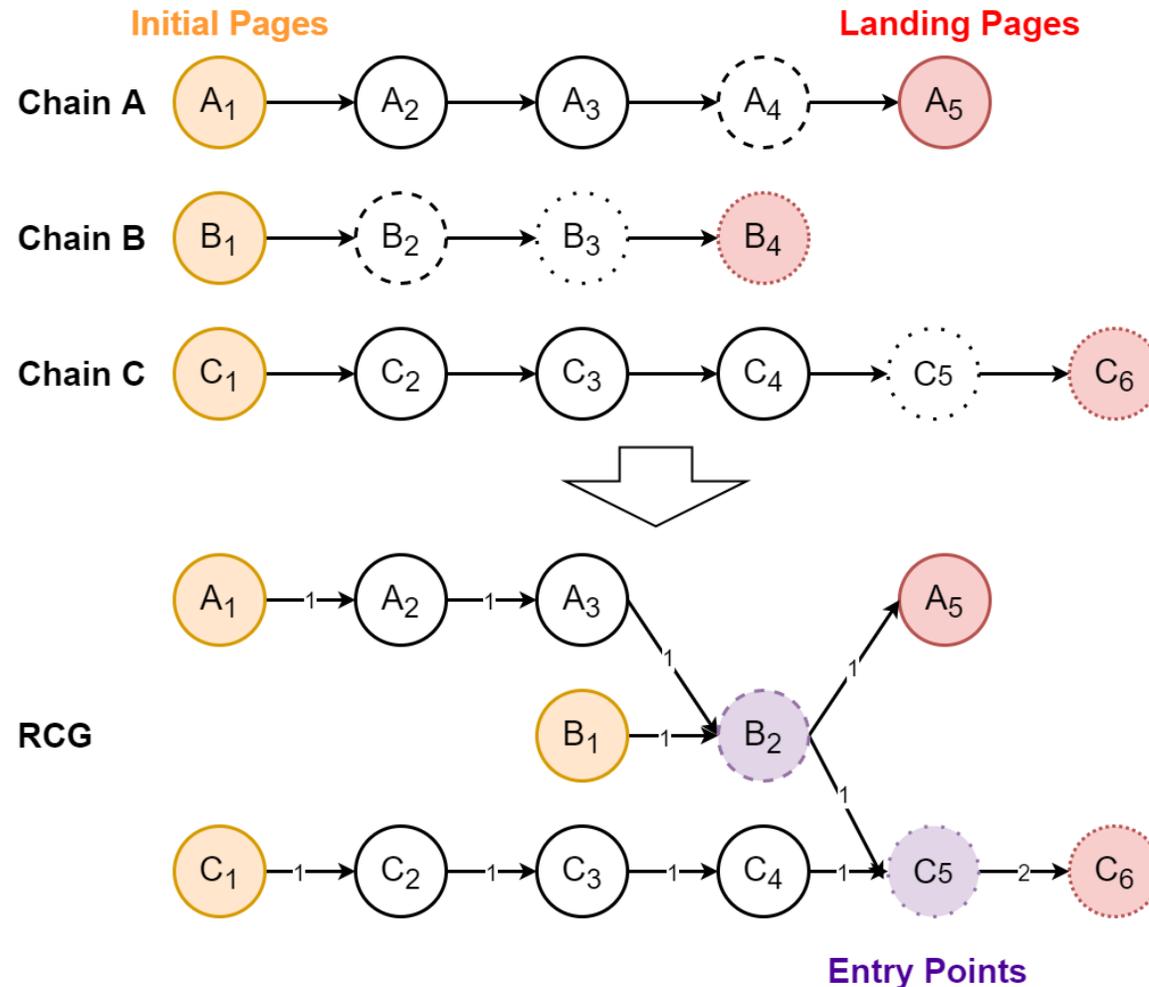
Domain-Agnostic Approach

- Leverages widespread of redirect chains by spammers
- Extracts robust features to capture the nature of spammers' behavior
- Can be applied into different domains

Redirect Chain



Redirect Chain Graph



- Identify same URLs
- Aggregate chains
- Find Entry points
 - Largest in-weight node in each chain

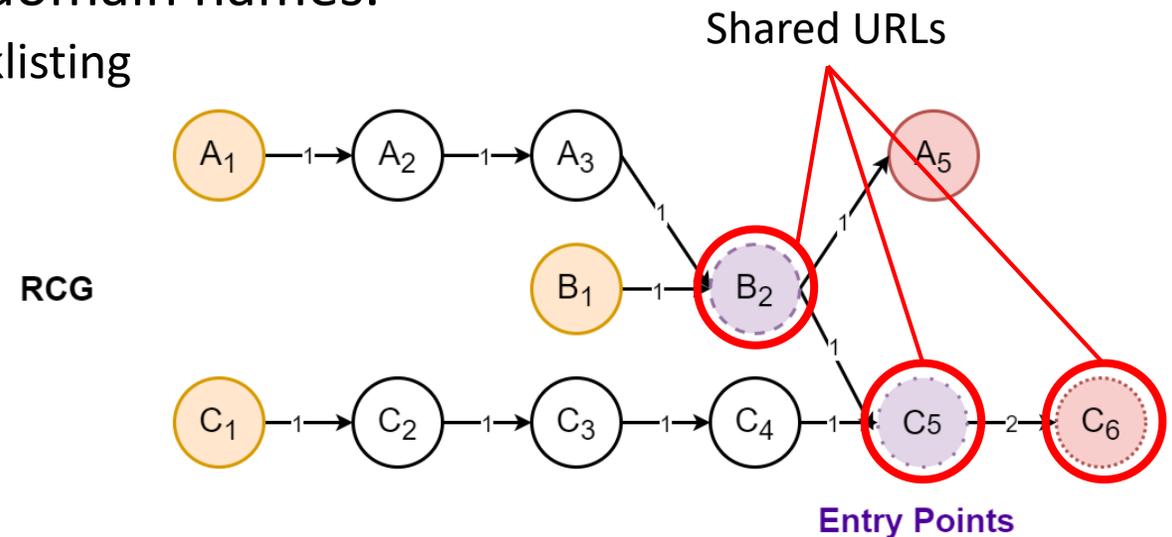
Feature Design

- Three groups of Features that characterize spammers' behavior
 - Shared resources
 - Heterogeneity
 - Flexibility

Features – Shared Resources

- To reduce costs, sharing resources is inevitable
 - Reuse of URLs
 - Same servers hosting many different domain names.
 - To evade and stay ahead of domain blacklisting

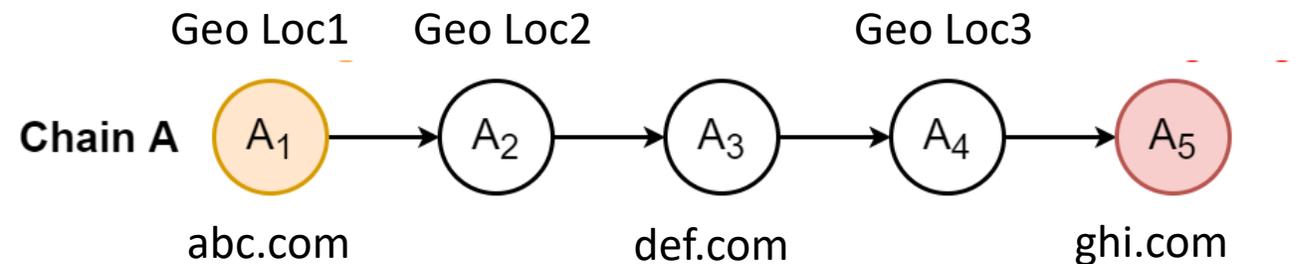
- Total 17 features



Features – Heterogeneity

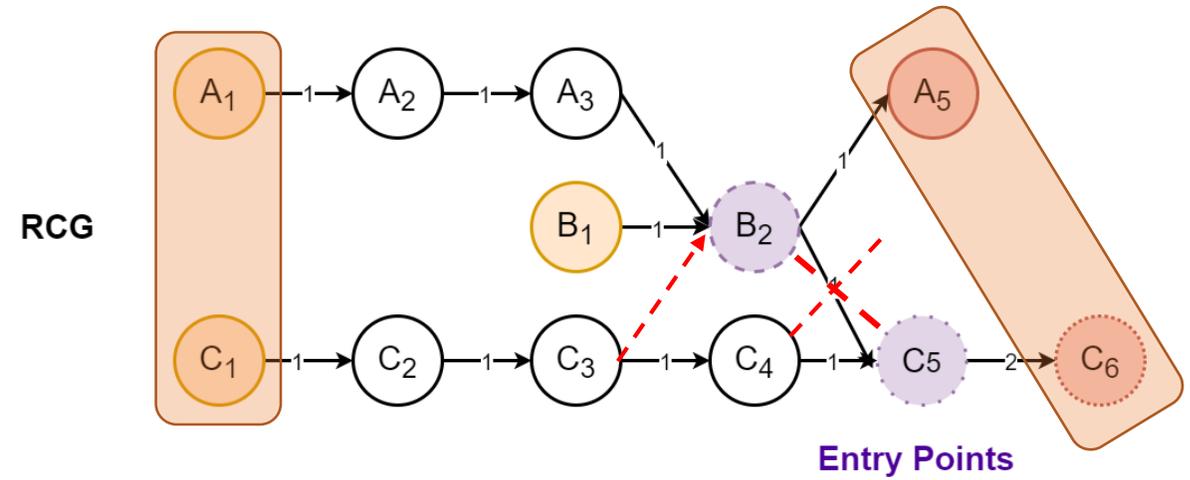
- “Don't put all your eggs in one basket”
 - Place servers to different geo-locations
 - Use of compromised servers and bot machines

- Total 12 features



Features – Flexibility

- Two types of flexibility:
 - For luring more users
 - Multiple different initial URLs
 - For evading detection
 - Using multiple landing URLs with redundant content
 - Same URLs with different IPs
 - Dynamicity and selectivity using long redirect chains
- Total 10 features



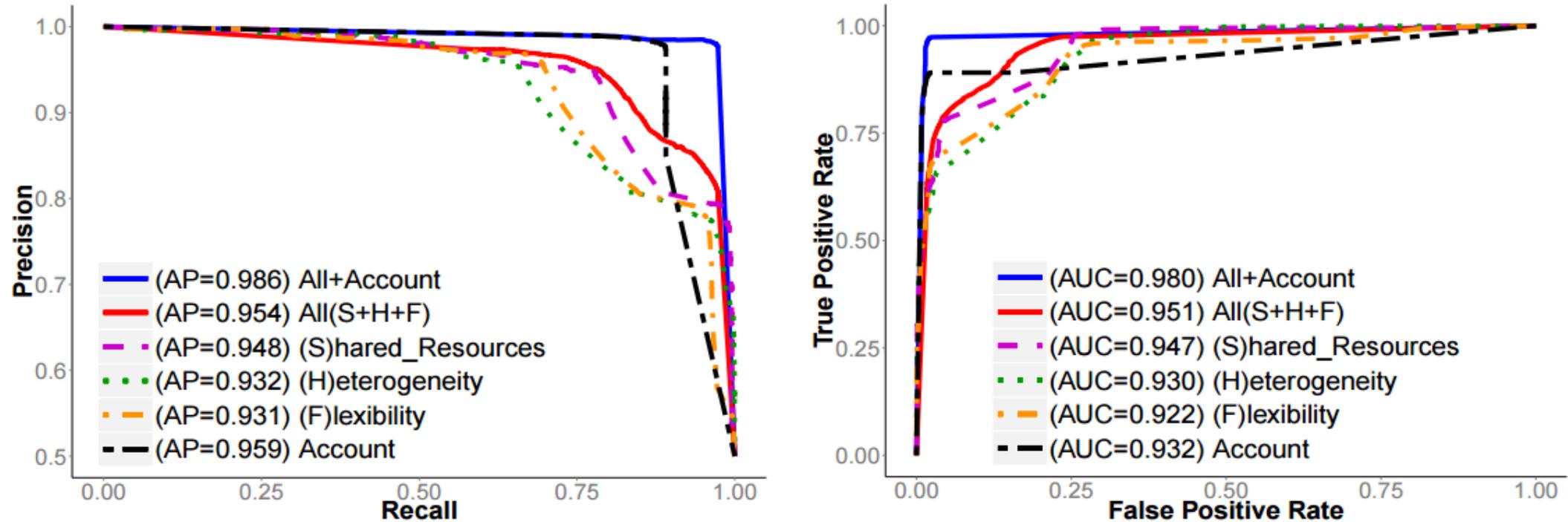
Dataset

- Tweets
 - 3,764,395 tweets have URLs
 - 3,871,911 initial URLs are identified
- Redirect Chain
 - Chain lengths are vary from 1 to 46
 - 99% of chains are less than length 6
- Redirect Chain Graph
 - 4,874,256 nodes
 - 3,839,633 edges

Experiment

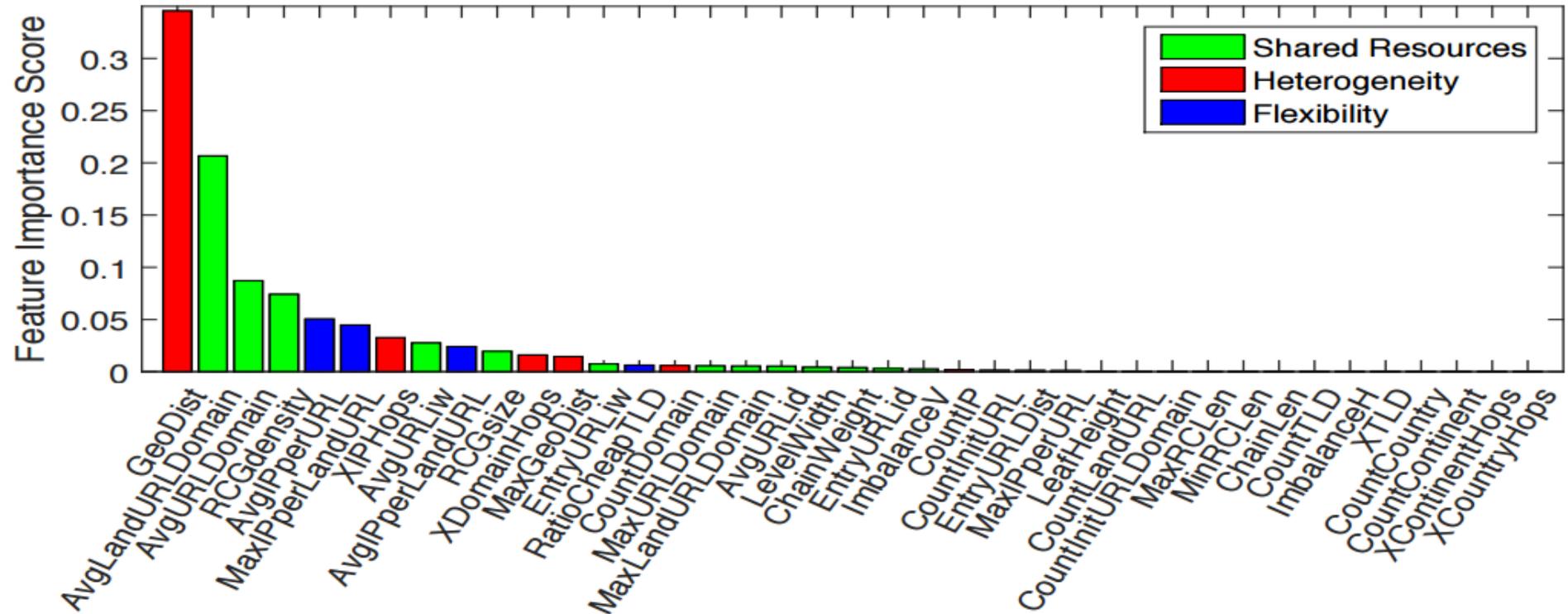
- Supervised Detection
 - Compare between context-free and context-aware detection
- Semi-supervised Detection
 - Small fraction of labels are revealed (1% or 5%)
 - Loopy belief propagation (LBP) through user-URL bipartite graph

Result – Supervised methods



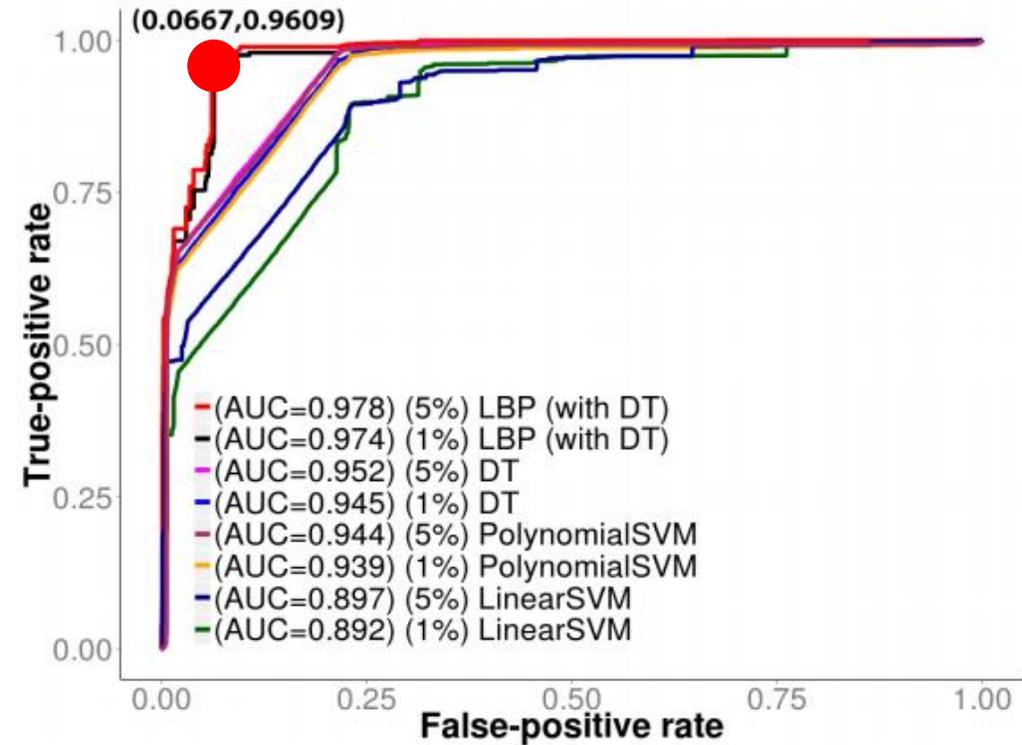
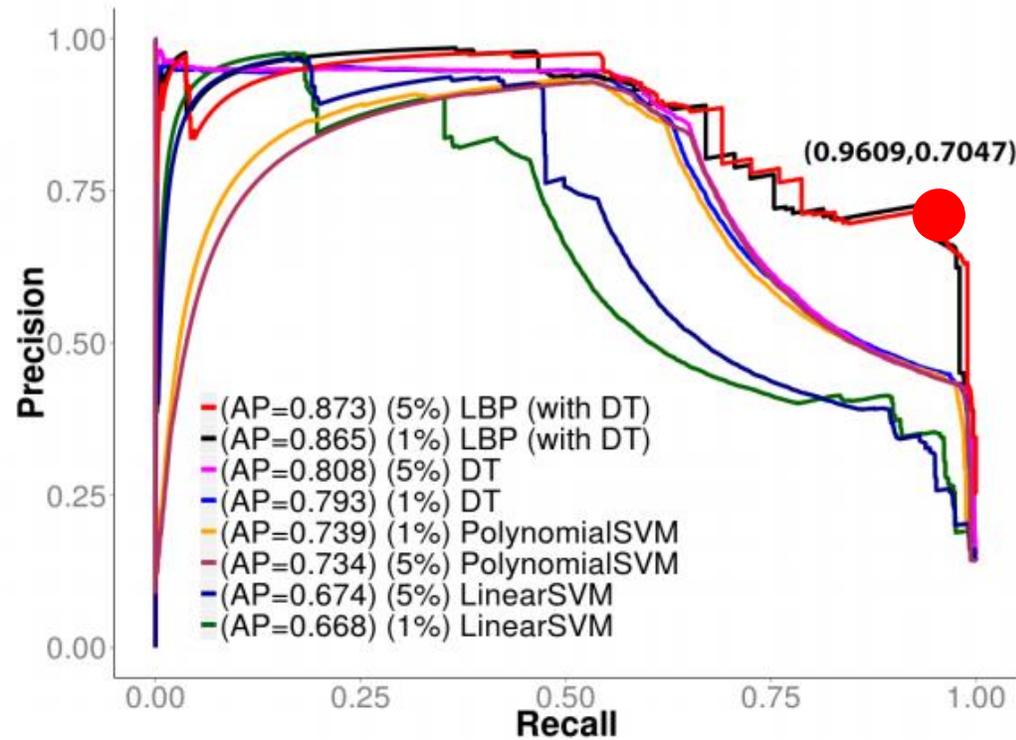
- Context-free features achieve competitive performance

Result – Feature importance score



- Top features evenly come from all three categories

Result – Semi-supervised methods



- Red dots show the performance at threshold 0.5

Conclusion

- Alternative approach to detect spam URL using Redirect Chain Graph
 - Context-free
 - Adversarially robust
 - Semi-supervised

data available at: <http://cs.stonybrook.edu/~heekwon>

Thank you!