# Araña:

## Discovering and Characterizing Password Guessing Attacks in Practice

Mazharul Islam, Marina Sanusi Bohuk,

Paul Chung, Thomas Ristenpart, Rahul Chatterjee

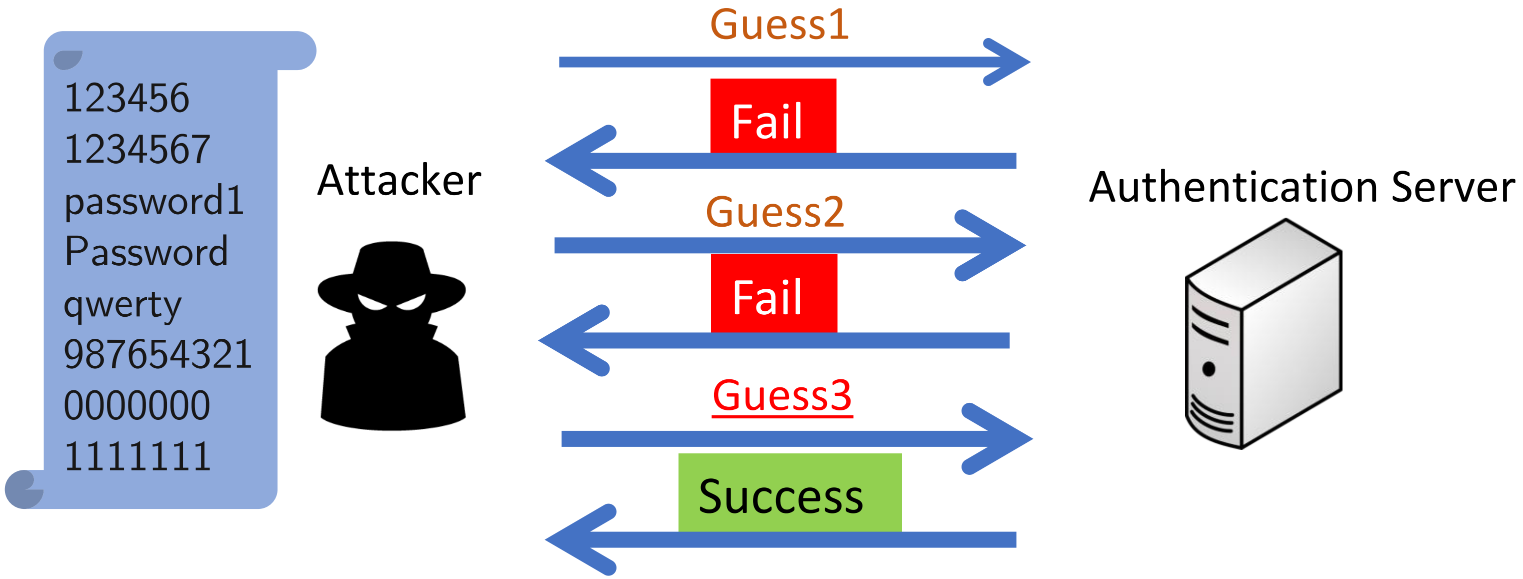
THE UNIVERSITY of WISCONSIN MADISON

MADS&P Security and Privacy Research Group at UW-Madison

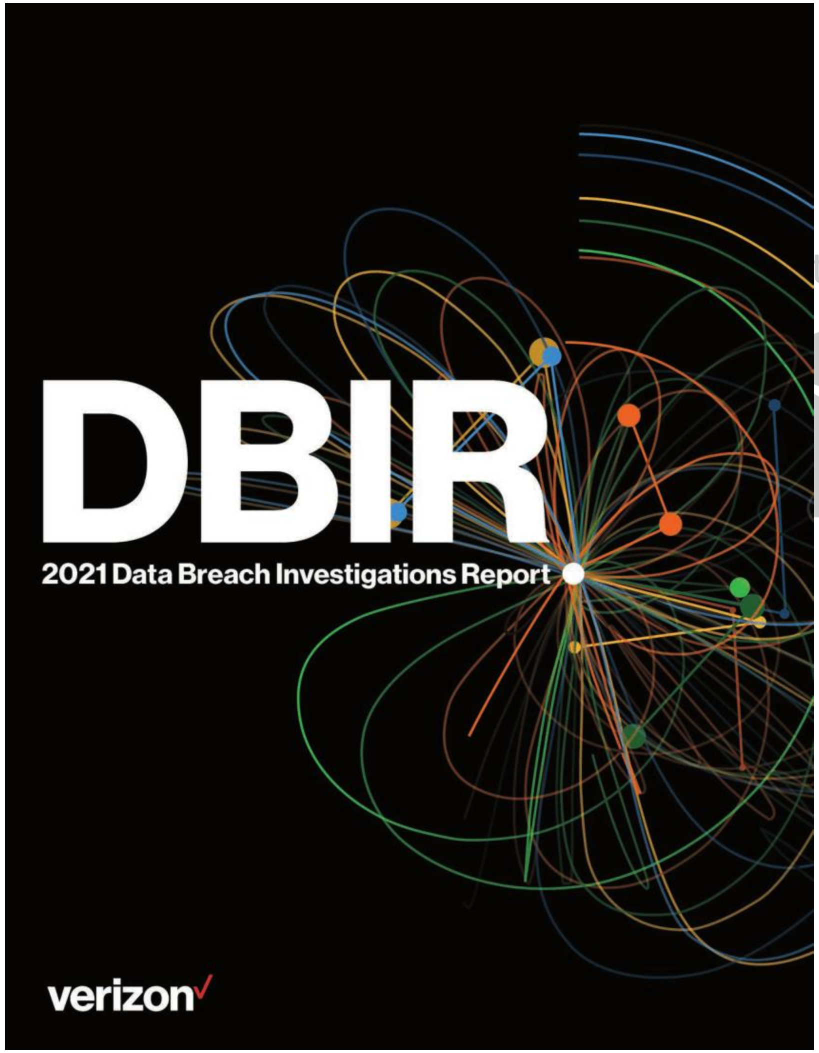
CORNELL TECH

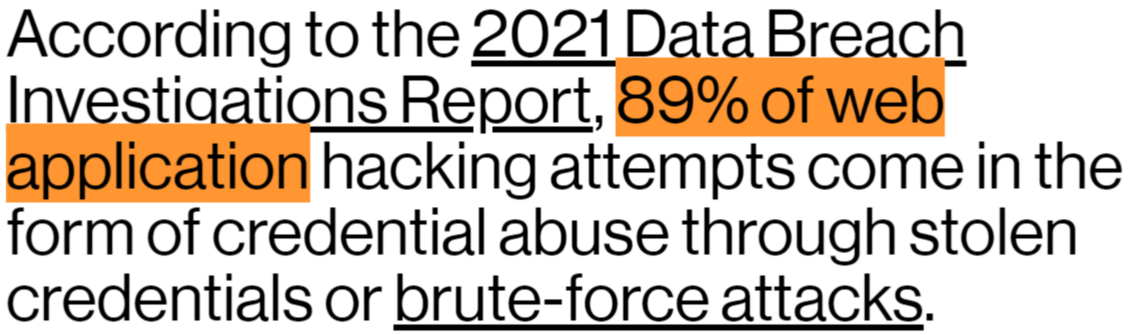
Image source: https://icon-icons.com/icon/Spider/109115

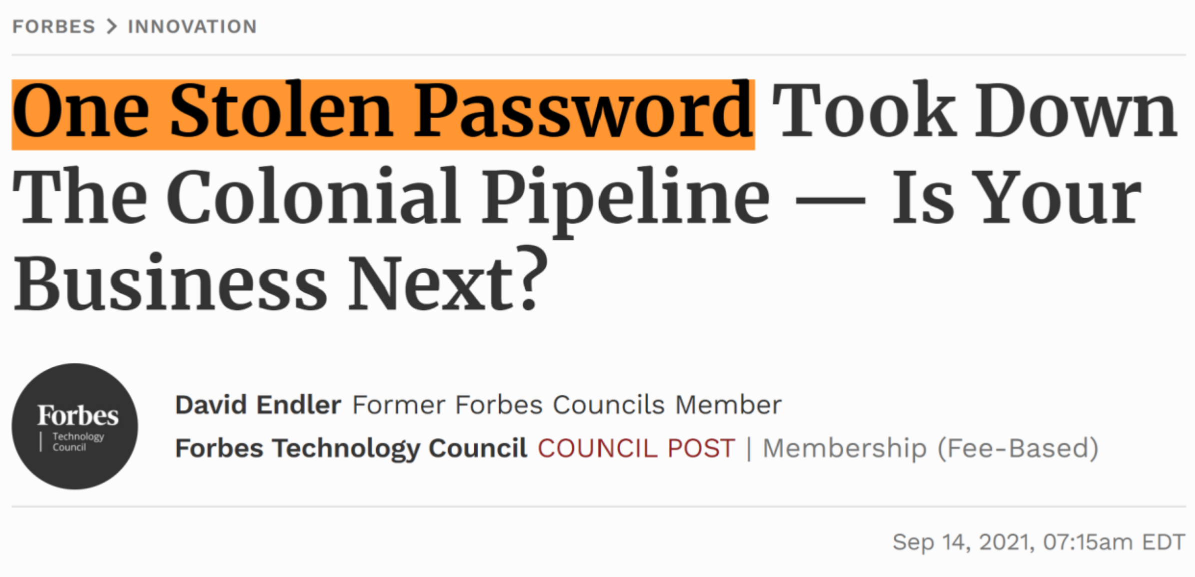
# Online password guessing attacks are damaging



# Online password guessing attacks are damaging




According to the 2021 Data Breach Investigations Report, 89% of web application hacking attempts come in the form of credential abuse through stolen credentials or brute-force attacks.


FORBES > INNOVATION

## One Stolen Password Took Down The Colonial Pipeline — Is Your Business Next?

David Endler Former Forbes Councils Member
Forbes Technology Council COUNCIL POST | Membership (Fee-Based)

Sep 14, 2021, 07:15am EDT

# Few prior works characterized guessing attacks

**Detection**

- Schechter et al. (Euro S&P ’16)
- Freeman et al. (NDSS ’17)
- Herley et al. (NDSS ’19)

- Bohuk et al. (USENIX ’22)

Our work

How do attackers execute password guessing attacks in practice?

How to detect password guessing attacks in practice?

# Araña: Detecting Password Guessing Attacks

- Framework for detecting password guessing attacks
- Discover 25 new attack clusters
- Characterized attack clusters

Attack Campaigns

Araña found **1,157** of new **compromised accounts**

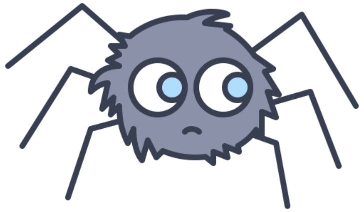


Image source: https://icon-icons.com/icon/Spider/109115

# Challenges to detect password guessing attacks

⇨ Large scale of real-world login dataset

Lack of ground truth on real word logins

Unknown attack strategies

Benign filters

Unsupervised clustering

Diverse feature types

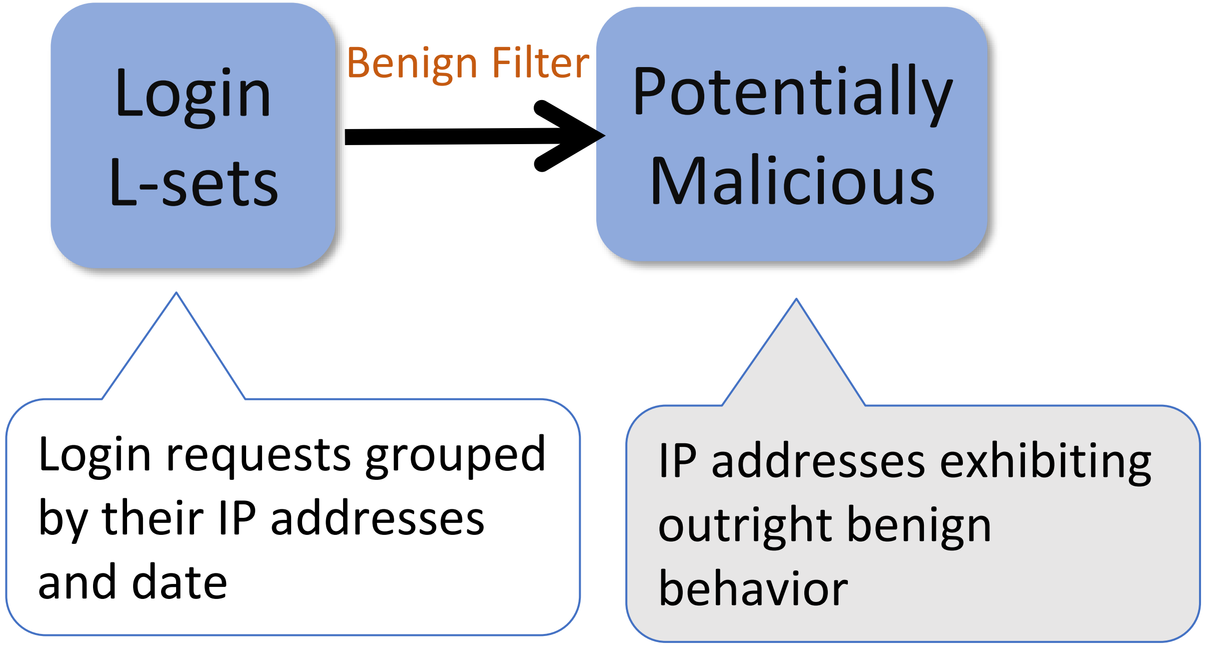
Araña  uses Filter, Cluster, Analysis

# Araña uses *Filter-Cluster-Analysis* Approach

Login
L-sets

Login requests grouped by their IP addresses and date

# Araña uses ***<u>Filter</u>***-Cluster-Analysis Approach



Login L-sets

Benign Filter

Potentially Malicious

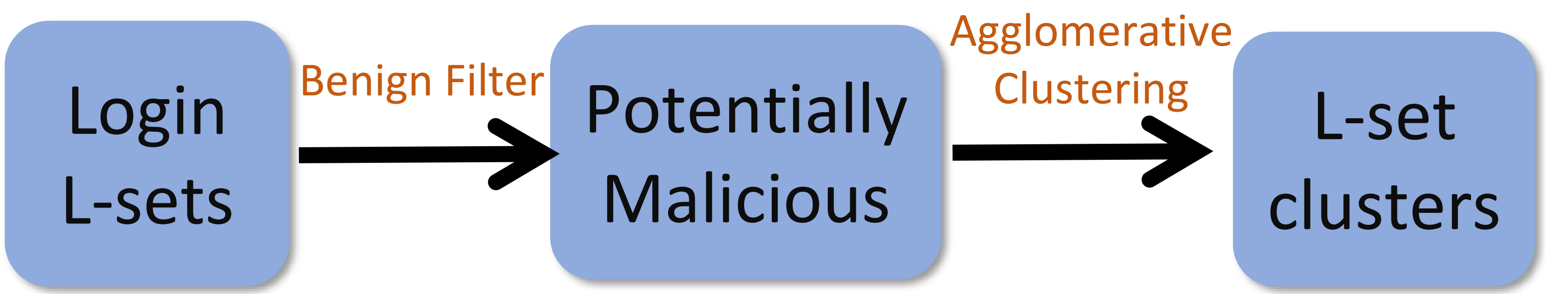
Login requests grouped by their IP addresses and date

IP addresses exhibiting outright benign behavior

## Benign filters

- High success rate
- Org's Private IP addresses
- Successful 2FA
- ….

# Araña uses Filter-***Cluster***-Analysis Approach



Login L-sets

Benign Filter

Potentially Malicious

Agglomerative Clustering

L-set clusters

Login requests grouped by their IP addresses and date

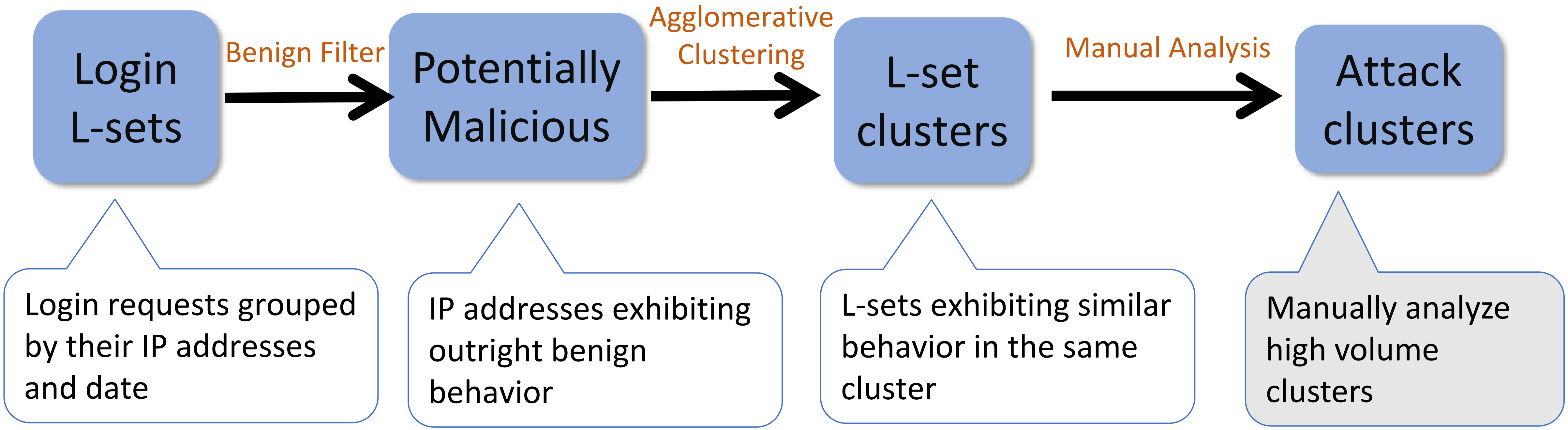
IP addresses exhibiting outright benign behavior

L-sets exhibiting similar behavior are clustered together

**Features**

- Client user agent, IP subnet, ISP
- Password features
- Volumetric features
- Timing, success rates etc.

# Araña uses Filter-Cluster-***Analysis*** Approach



Login L-sets

Benign Filter

Potentially Malicious

Agglomerative Clustering

L-set clusters

Manual Analysis

Attack clusters

Login requests grouped by their IP addresses and date

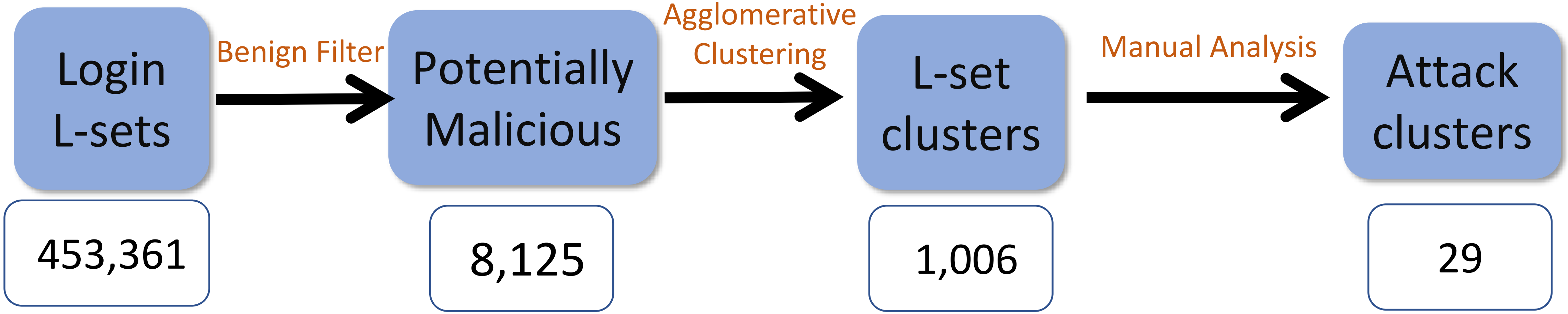
IP addresses exhibiting outright benign behavior

L-sets exhibiting similar behavior in the same cluster

Manually analyze high volume clusters

Sampling criteria
High precision detection of attack clusters

# Araña evaluated on real world login dataset



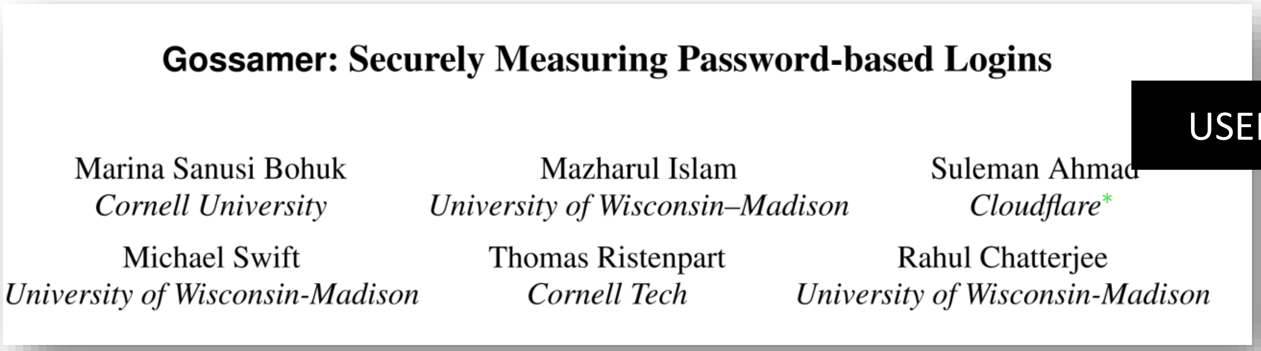
Login L-sets → Benign Filter → Potentially Malicious → Agglomerative Clustering → L-set clusters → Manual Analysis → Attack clusters

453,361

8,125

1,006

29



**Gossamer: Securely Measuring Password-based Logins**

Marina Sanusi Bohuk
*Cornell University*

Mazharul Islam
*University of Wisconsin–Madison*

Suleman Ahmad
*Cloudflare*[*]

Michael Swift
*University of Wisconsin-Madison*

Thomas Ristenpart
*Cornell Tech*

Rahul Chatterjee
*University of Wisconsin-Madison*

USENIX Security ’22

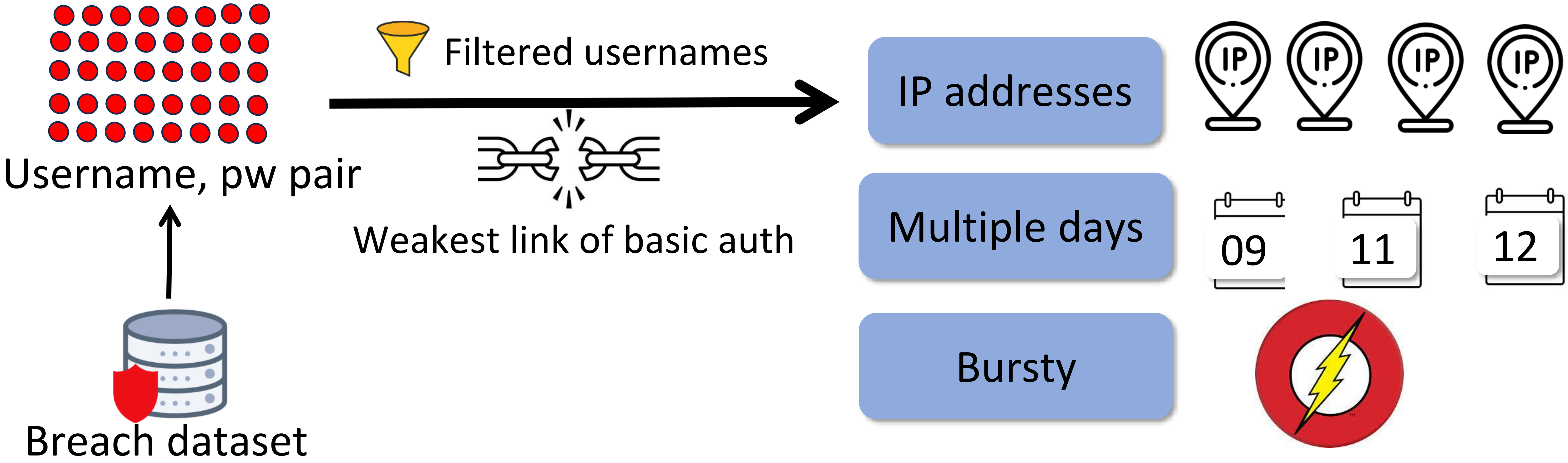
34 million logins for 7 months at 2 universities

# Araña detected thousands of compromised accounts

| Number of | Univ 1 | Univ  2 | Total |
|---|---|---|---|
| attack clusters | 9 | 20 | **29** |
| IP addresses | 756 | 1,668 | **2,424** |
| logins | 75,884 | 287,16 | **363,051** |
| users attacked | 11646 | 152278 | **163924** |
| compromised users | 41 | 1,116 | **1,157** |

Identified multiple attack strategies

# Attack cluster behaviors and strategies



Username, pw pair

Breach dataset

Filtered usernames

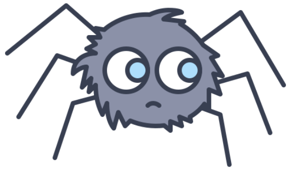
Weakest link of basic auth

IP addresses

Multiple days

Bursty

Araña can improve existing defenses



# Araña: A new framework for detecting password guessing attacks in practice

- Filter-Cluster-Analysis approach
- Identified several attack clusters (potentially attack campaigns)
- Help learn attack strategies to improve existing techniques

`https://github.com/islamazhar/Arana-Public`

Thanks!

@Mazharul_13 mazharul@cs.wisc.edu

Image source: https://icon-icons.com/icon/Spider/109115