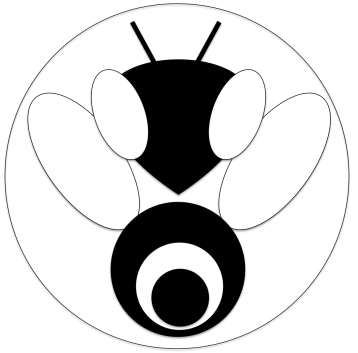
BEEWOLF

Catching Worms, Trojan Horses and PUPs: Unsupervised Detection of Silent Delivery Campaigns



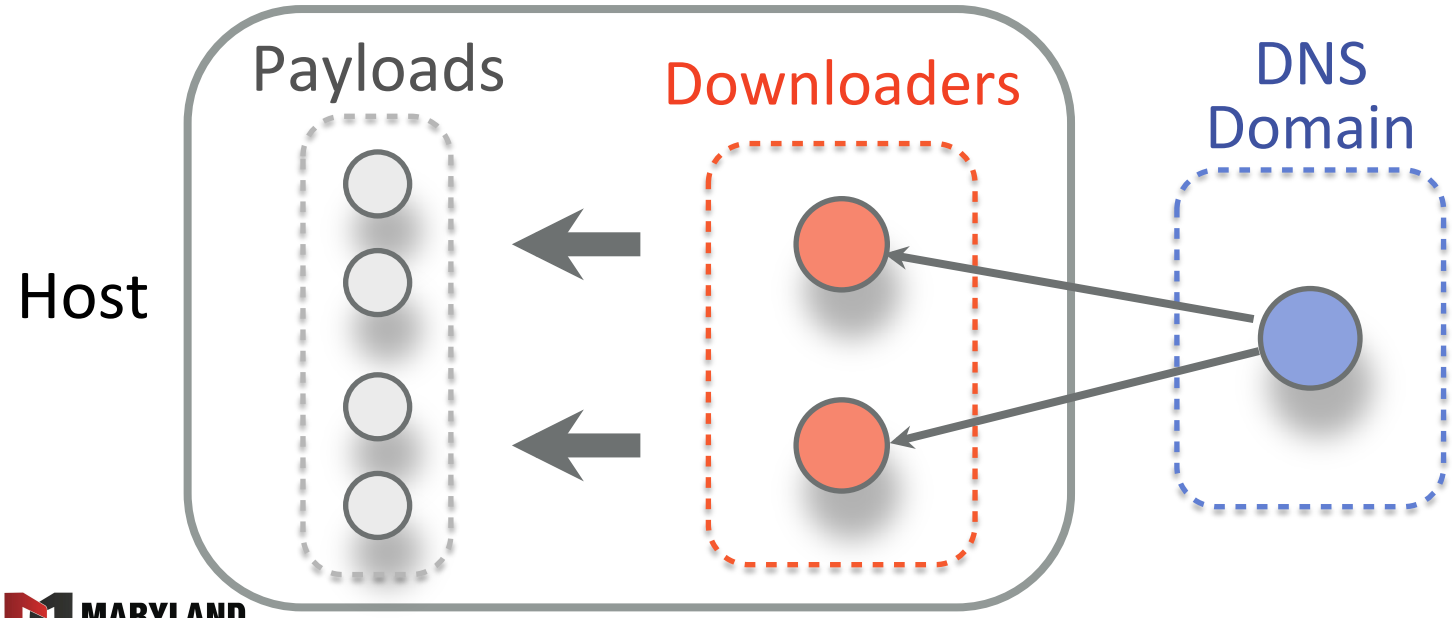
BumJun Kwon, Virinchi Srinivas,

Amol Deshpande, Tudor Dumitraș

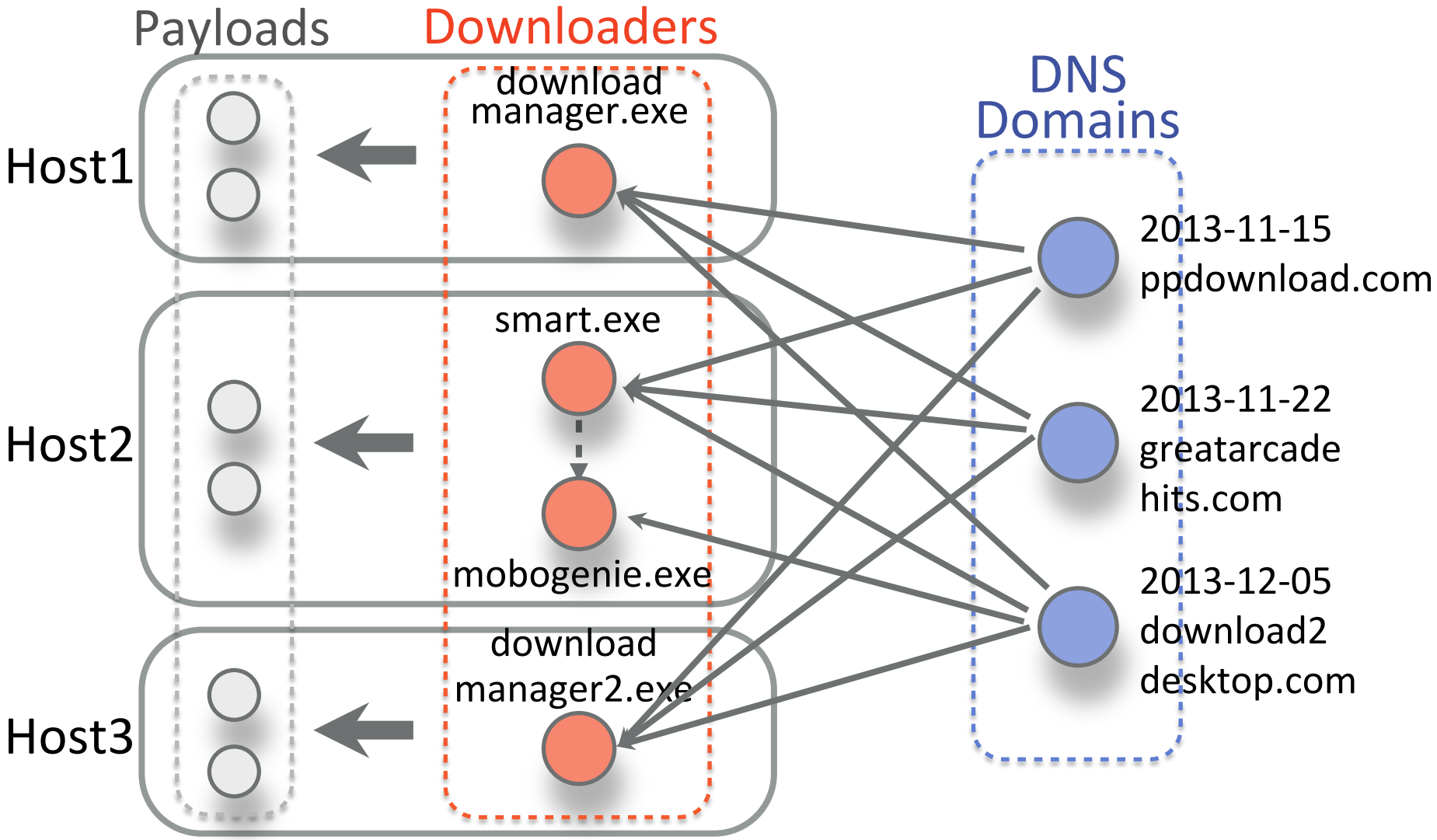
University of Maryland—College Park

Malware Delivery Campaigns

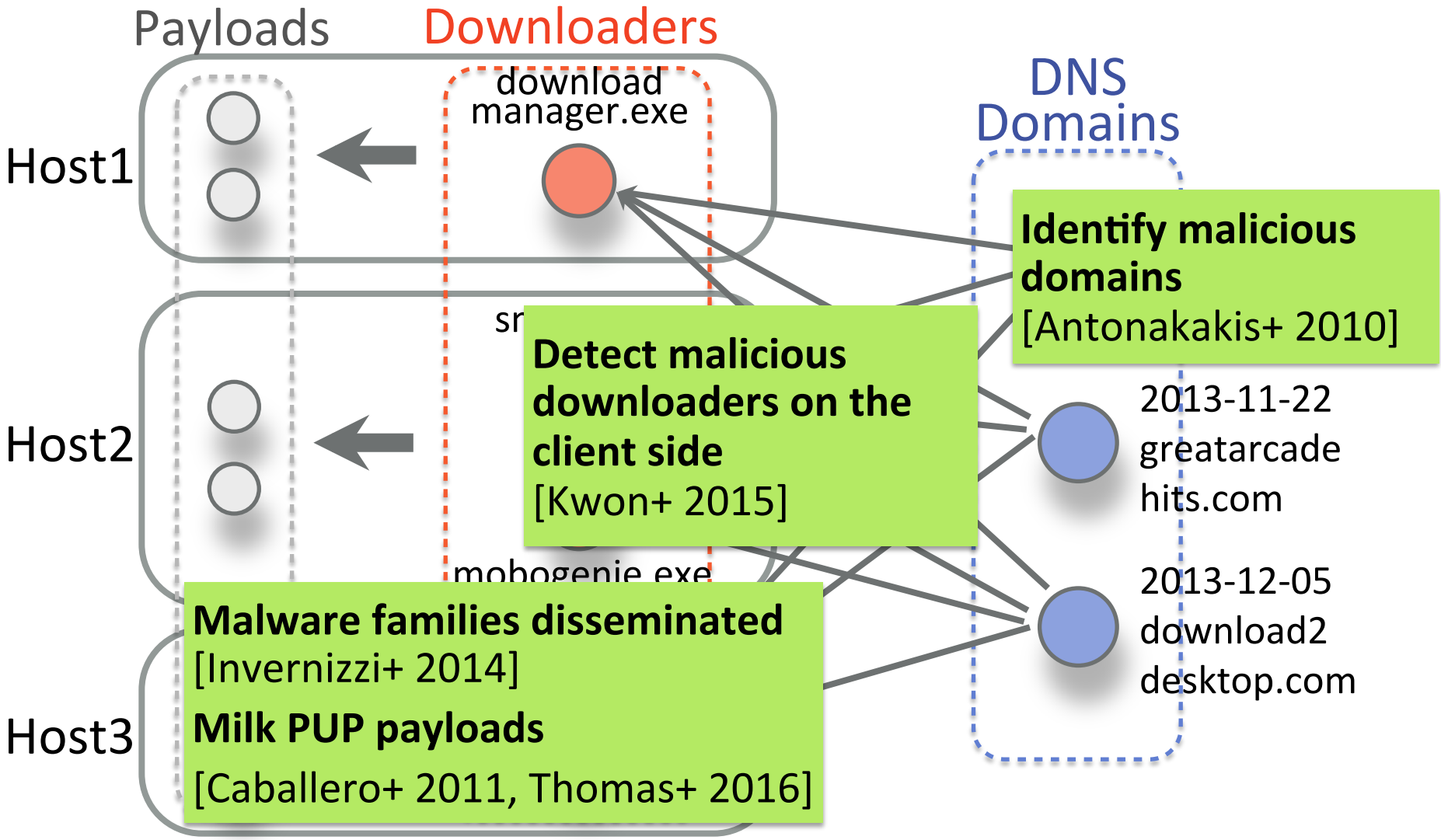
- Business model
  - Charge fees for delivering malware or PUPs
- Key method
  - **Orchestrate Silent delivery campaigns**



Silent Delivery Campaigns

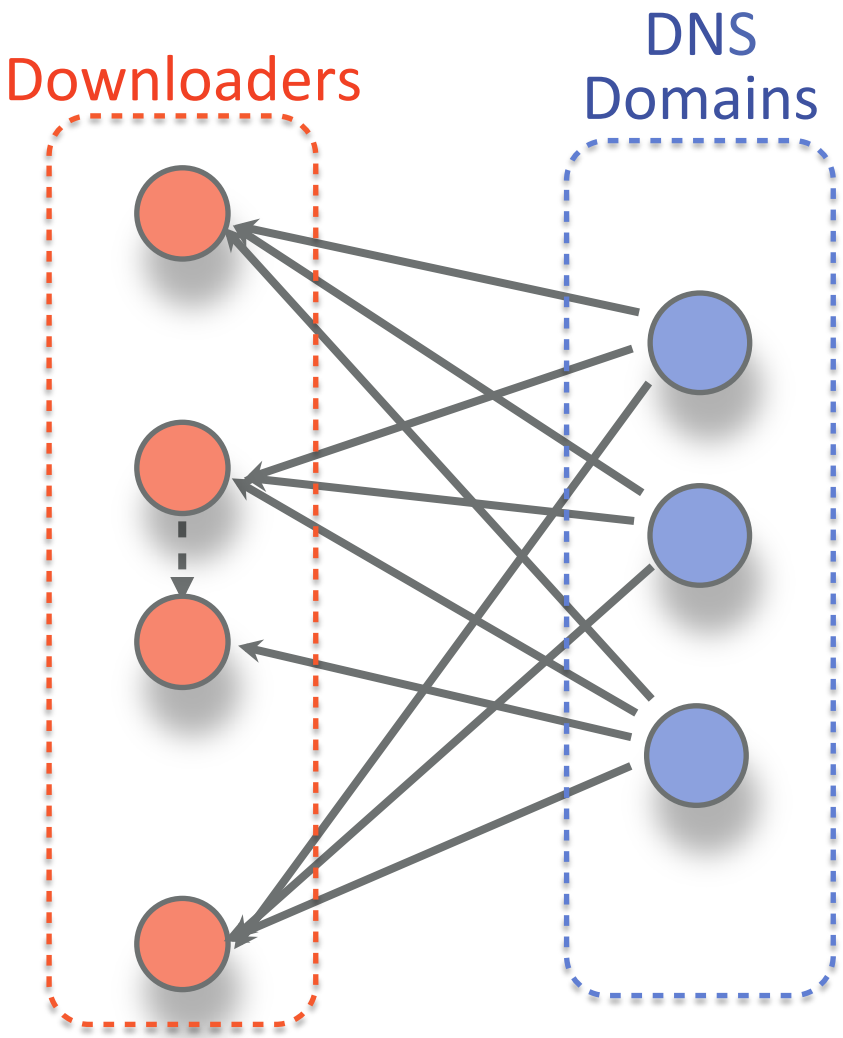


Silent Delivery Campaigns



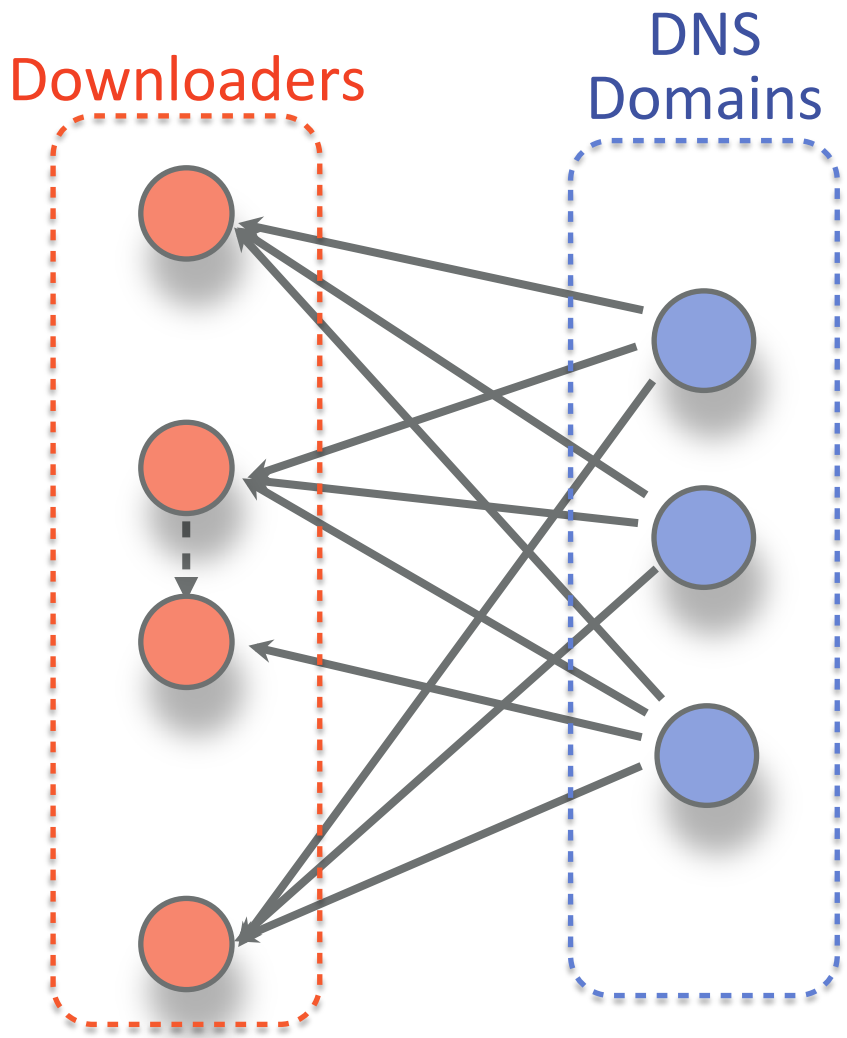
Lockstep Behavior

[Beutel+ 2013, Cao+ 2014, Jiang+ 2015]



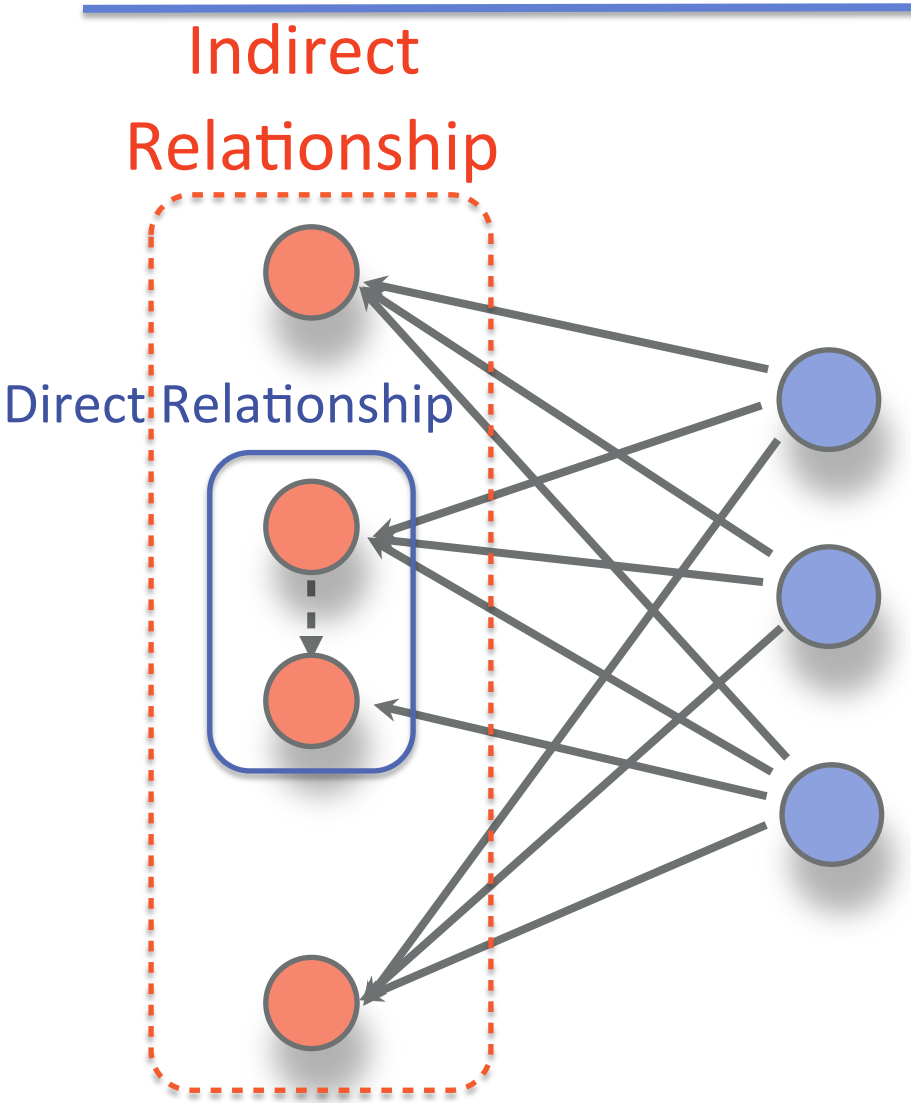
- Require seed nodes
- Require interpreting events defined by multiple features
- Not designed for streaming data

We Introduce Beewolf



- Propose an **unsupervised** and deterministic technique
- Orthogonal to the work that use machine learning
- Operate on a stream of download events
- Reveal the indirect relationships

Understanding Indirect Relationships

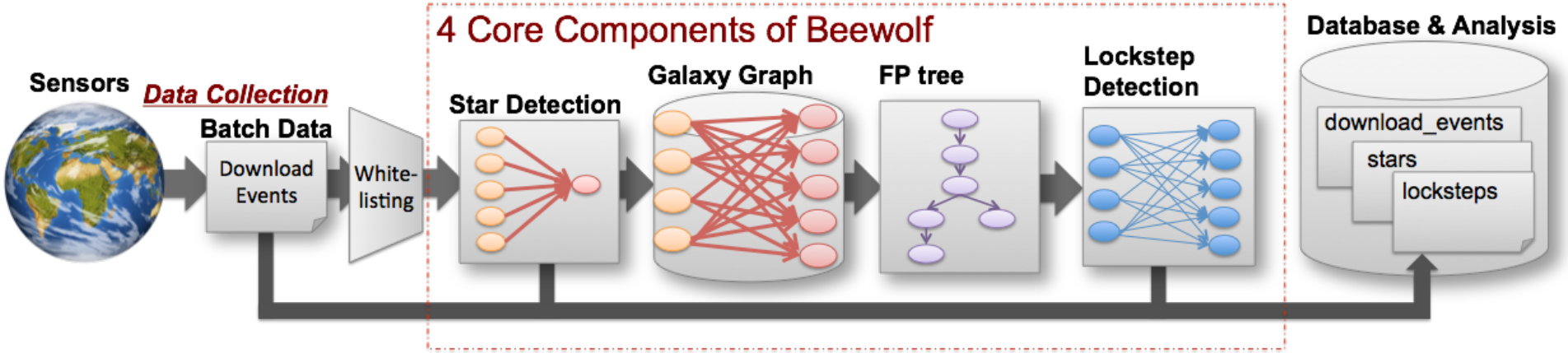


- Expose **hidden dependencies** in the underground economy
- Suggest suitable interventions for disrupting the malware delivery

Outline

- System overview
- Lockstep analysis
  - Attribution
  - Observations
- Evaluation
  - Streaming
- Conclusion

System Overview



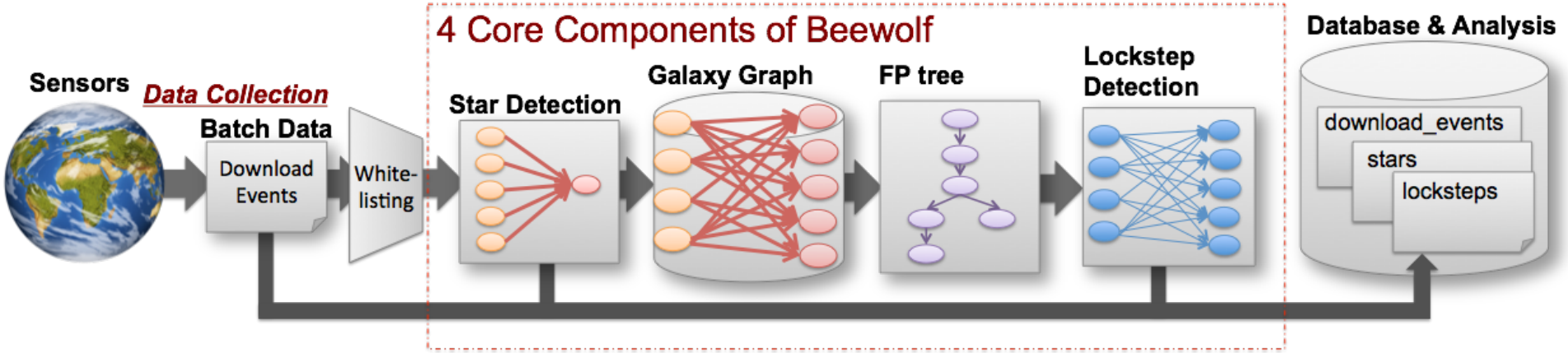
- Beewolf
  - Two modes: offline/streaming
  - Input: download event data
  - Whitelisting: download events from benign downloaders

Data Set: Download Activity in The Wild

- Download activity
  - Kwon et.al. The Dropper Effect paper (CCS'15)
  - Download event: downloader, second level domain name (domain), payload, sever timestamp
  - Year 2013

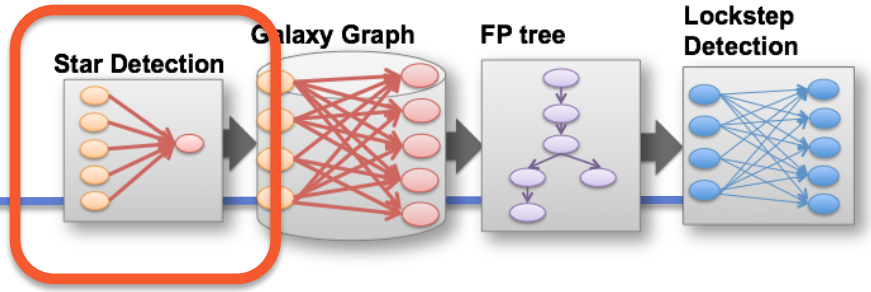
- Ground truth for labeling
  - VirusTotal
  - NSRL (National Software Reference Library)
  - Underground forums, Reason Labs knowledge base

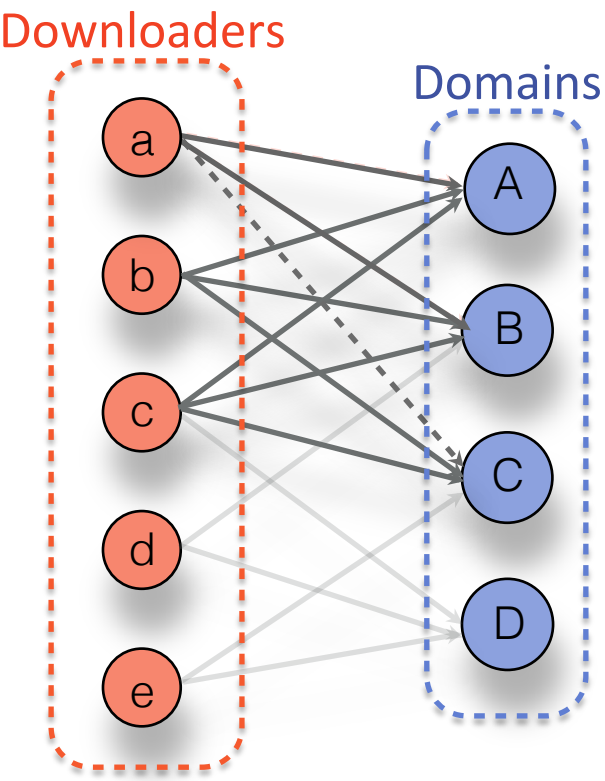
System Overview Cont'

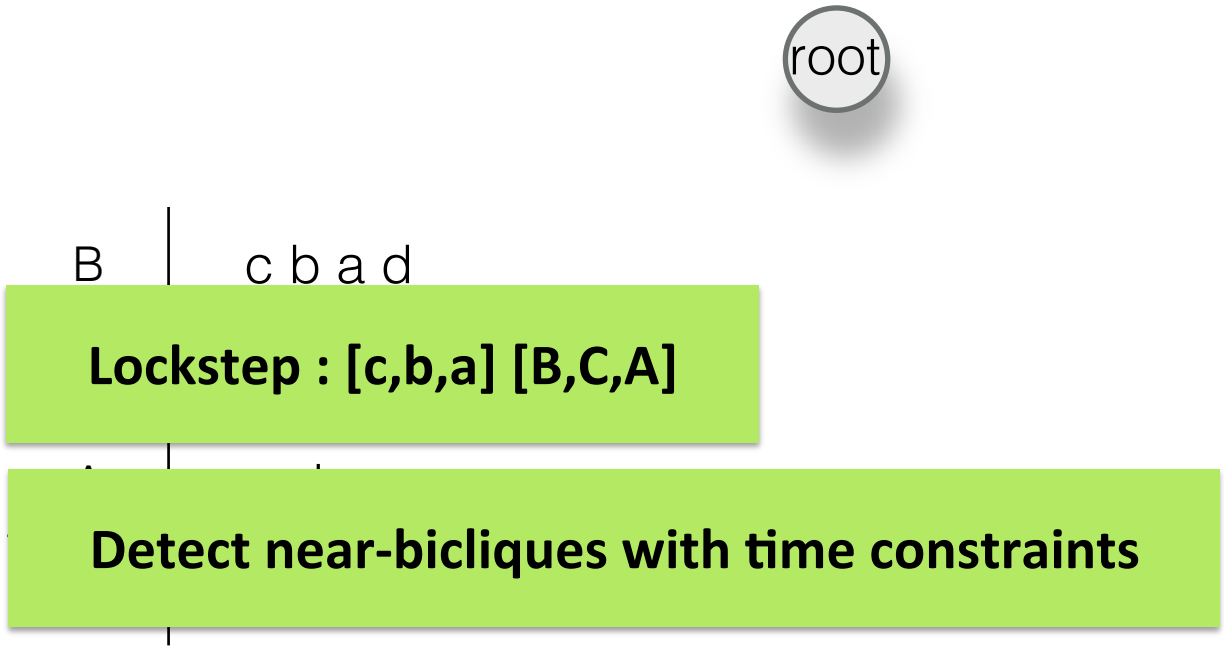


- Beewolf
  - Detect **lockstep patterns**
    - Offline: from the entire input dataset
    - Streaming: from the stream of data
  - Four core components
    - Star Detection, Galaxy graph, FP tree, Lockstep Detection

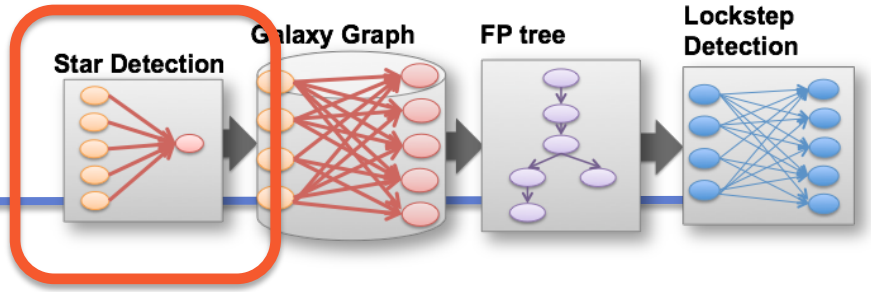
Goal

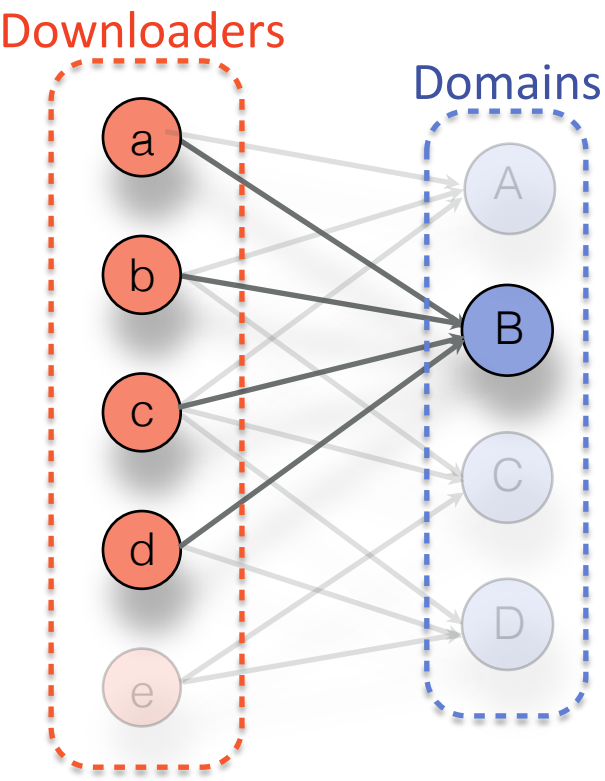

Star Detection
Galaxy Graph
FP tree
Lockstep Detection


Downloaders
Domains
a
b
c
d
e
A
B
C
D


root
B
c b a d
Lockstep : [c,b,a] [B,C,A]
Detect near-bicliques with time constraints

Star Detection

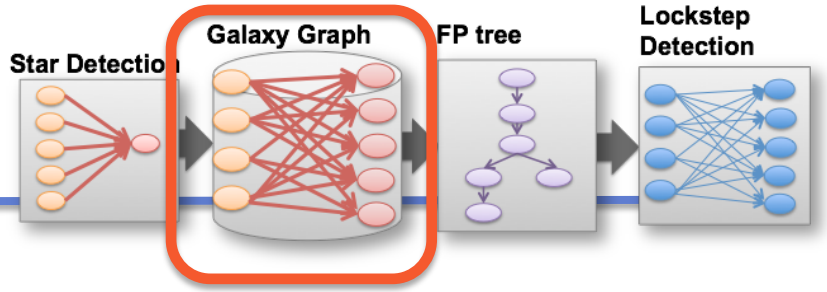


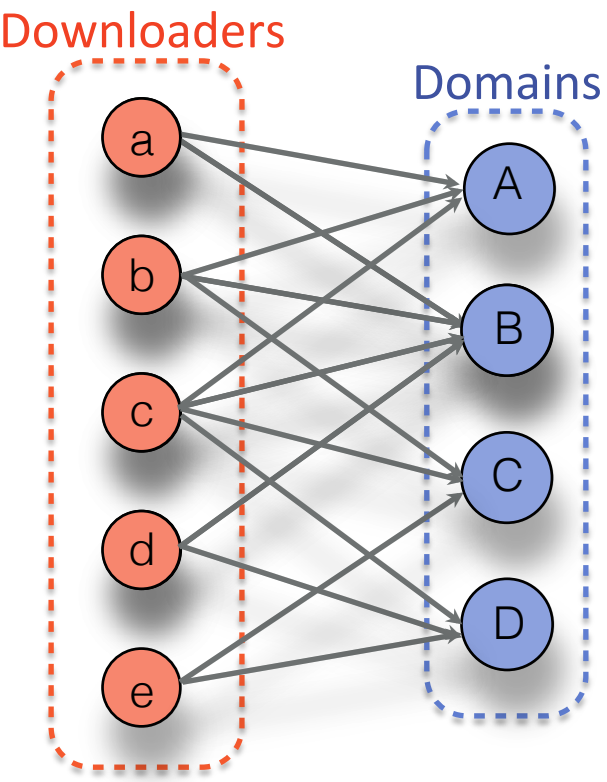


root

B	c b a d
C	c b e
A	c b a
D	c d e

Galaxy Graph

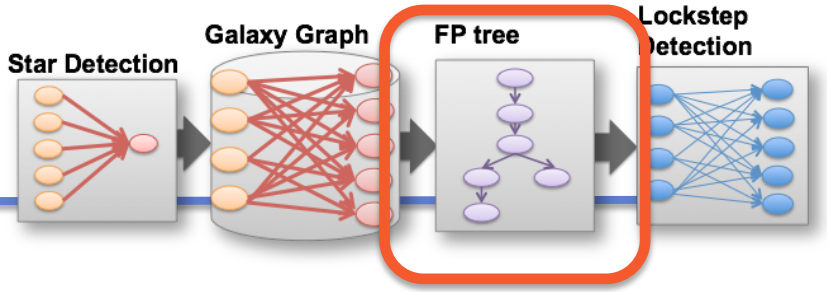


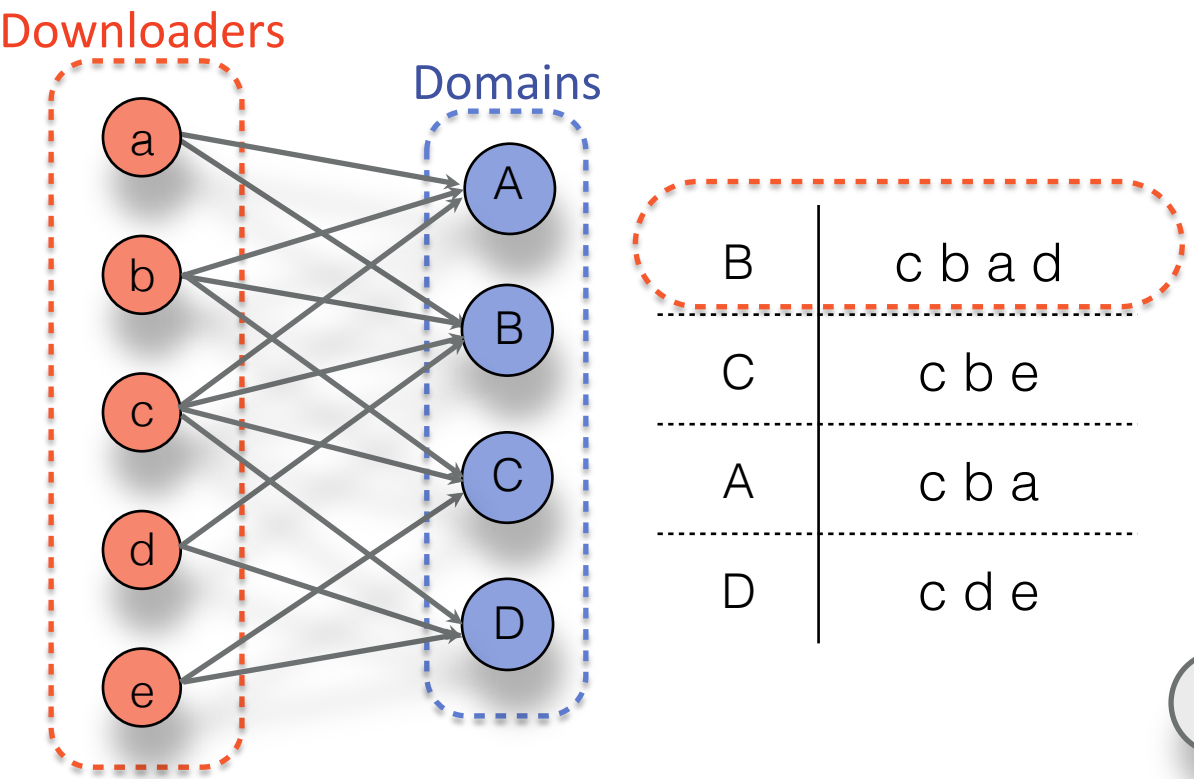


root

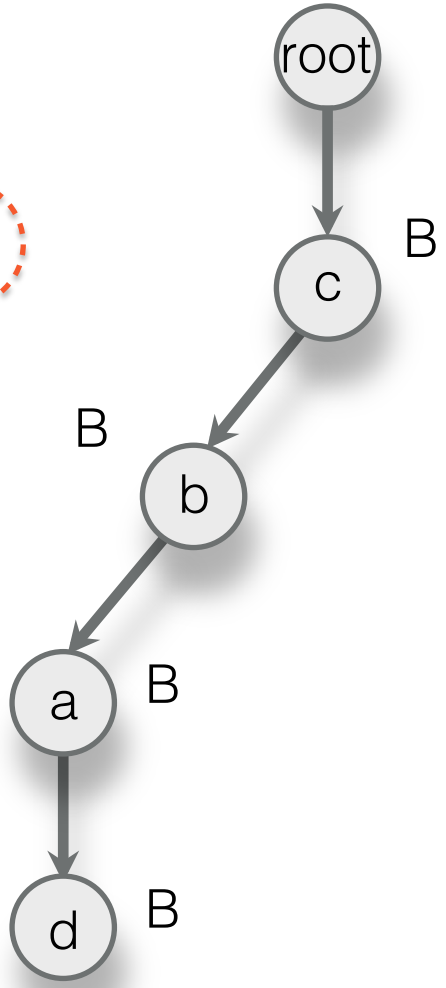
B	c b a d
C	c b e
A	c b a
D	c d e

Frequent Pattern Tree

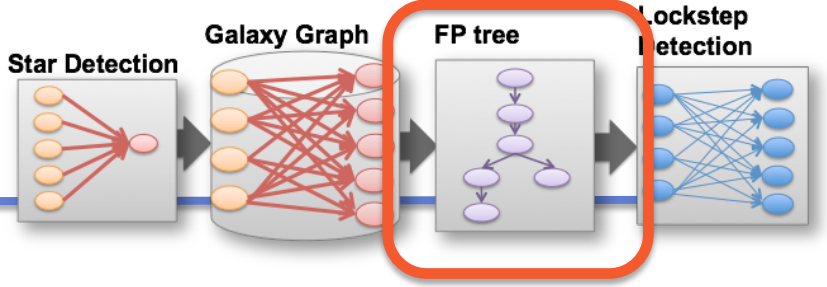


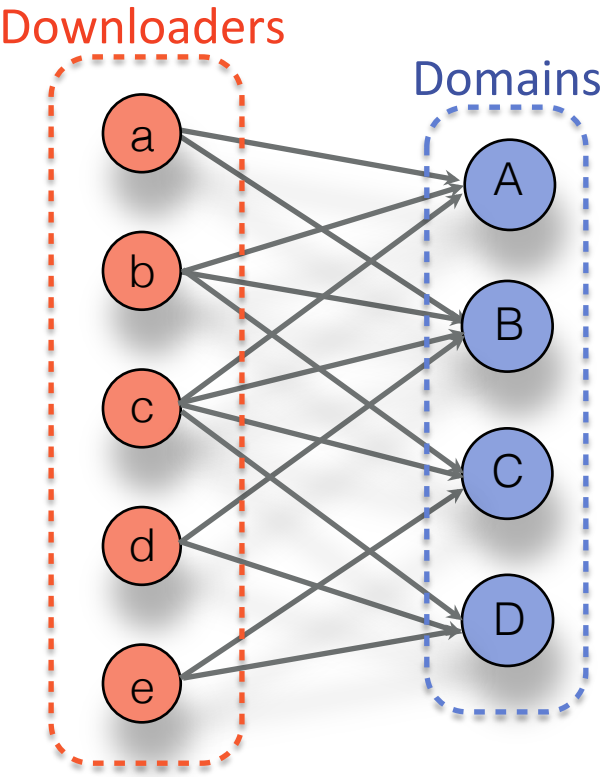


B	c b a d
C	c b e
A	c b a
D	c d e

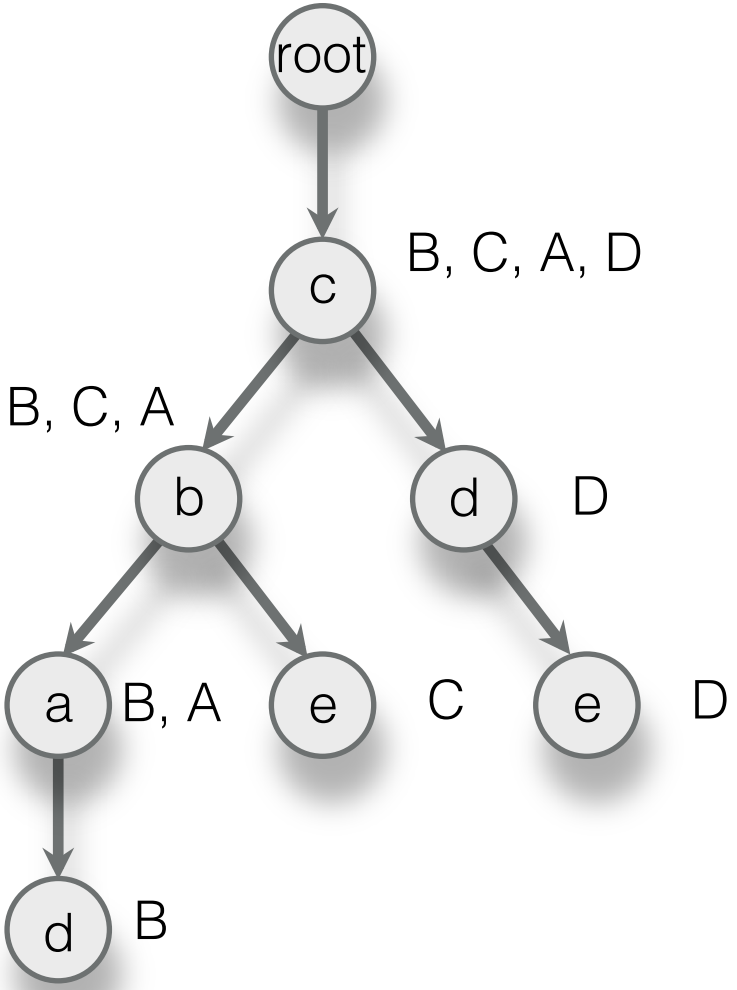


Frequent Pattern Tree

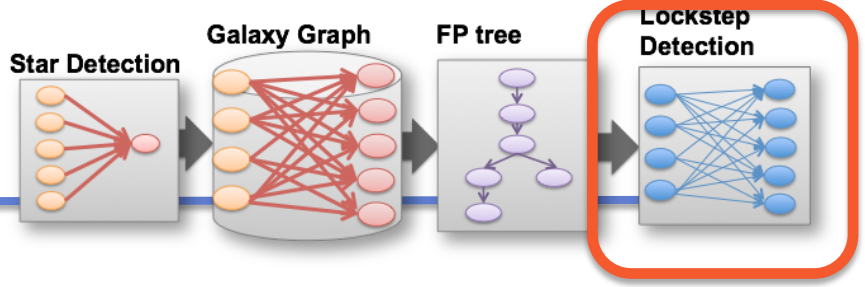


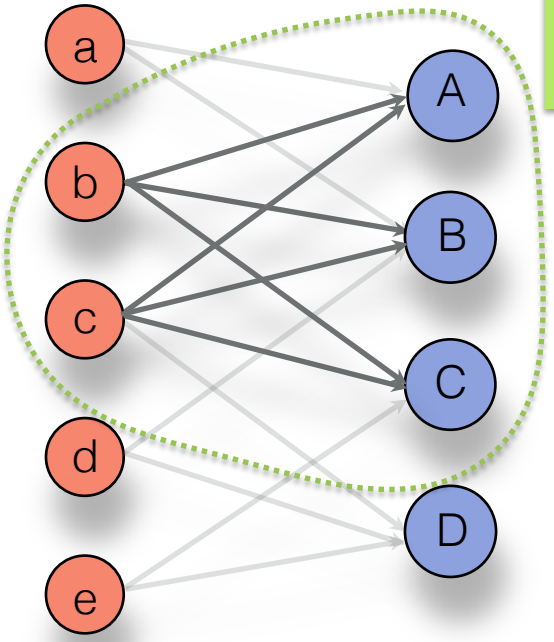


B	c b a d
C	c b e
A	c b a
D	c d e



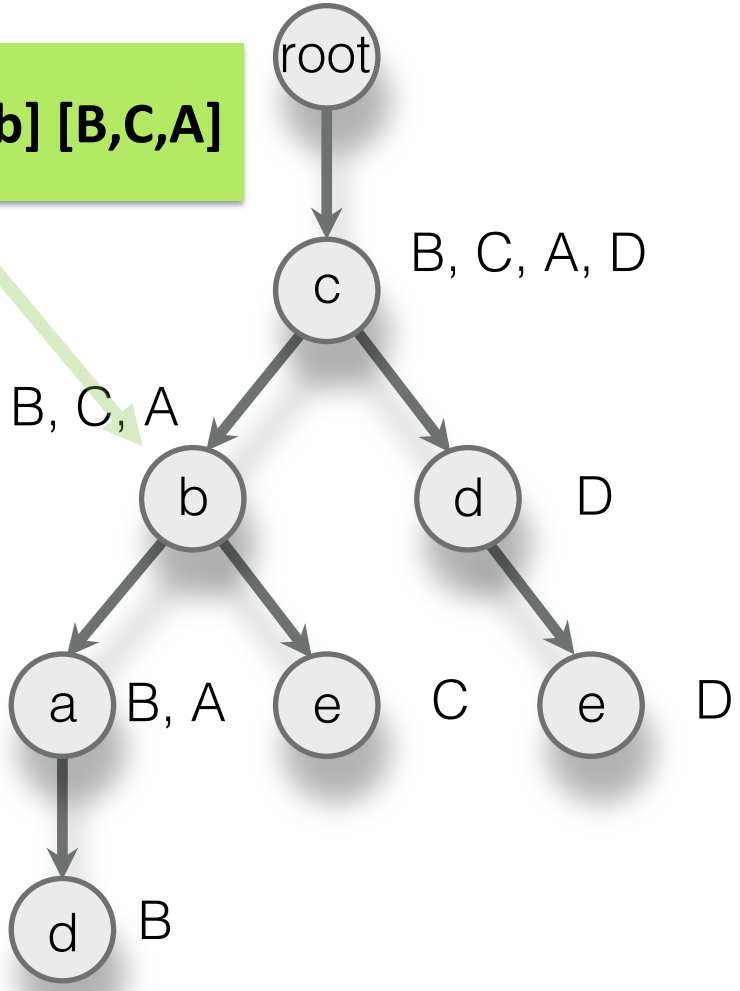
Lockstep Detection



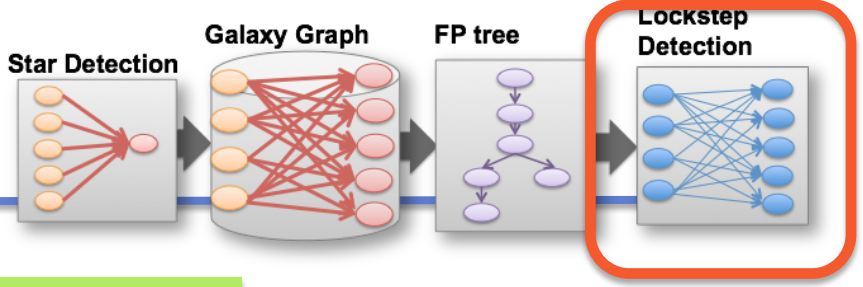


Complete Biclique: [c,b] [B,C,A]

B	c b a d
C	c b e
A	c b a
D	c d e

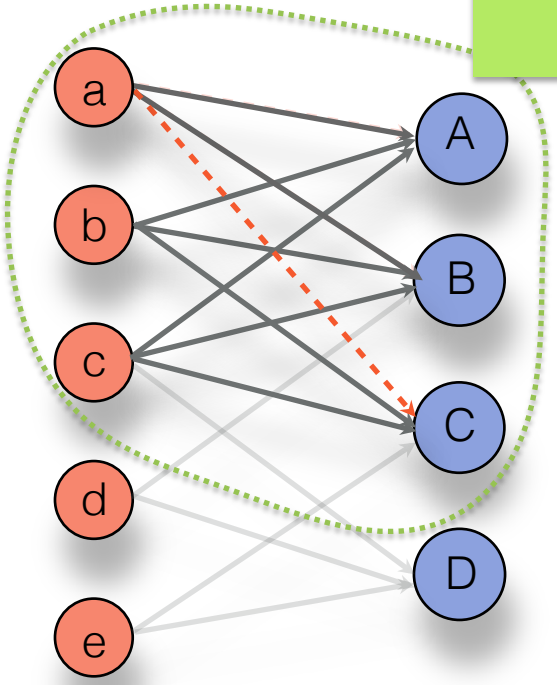


Addressing Limitations (1)

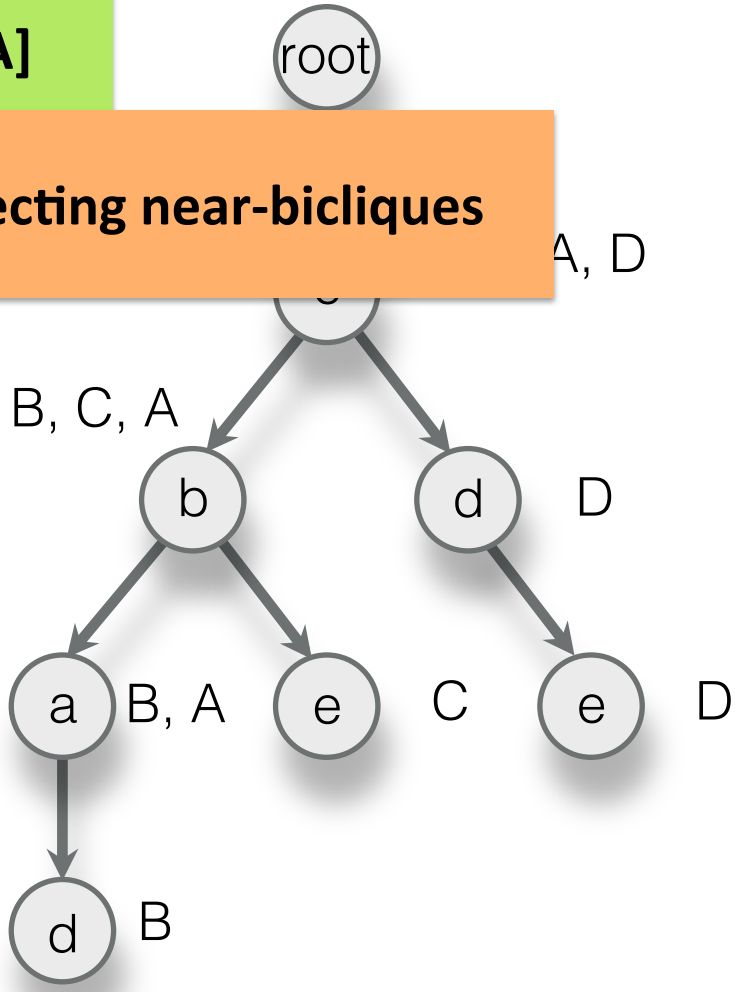


Lockstep : [c,b,a] [B,C,A]

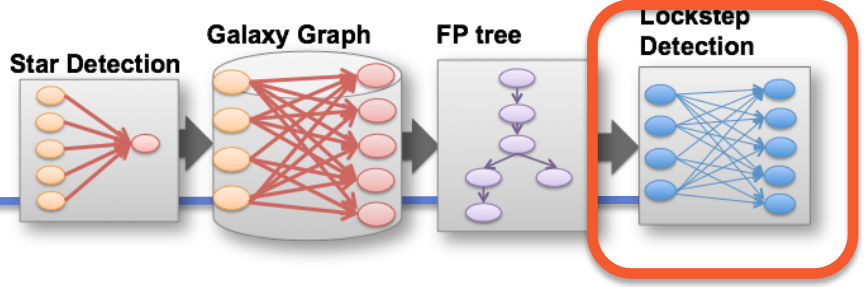
Heuristic for detecting near-bicliques

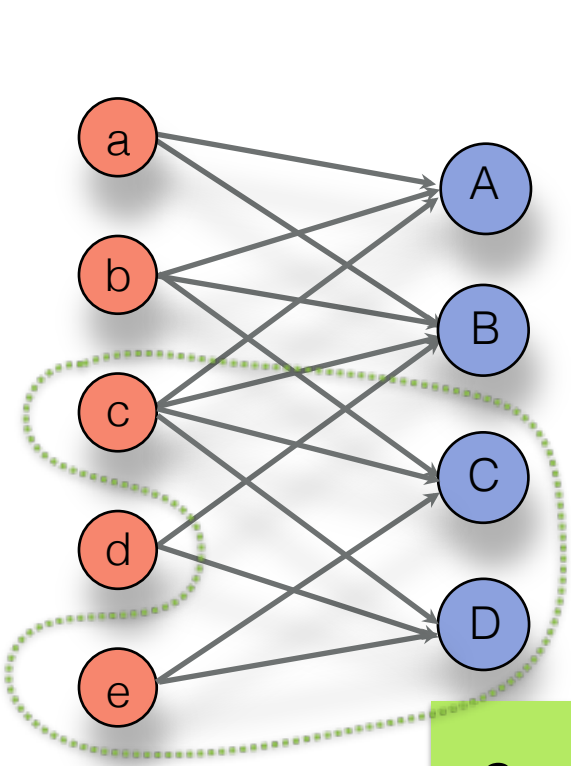


C	c b e
A	c b a
D	c d e

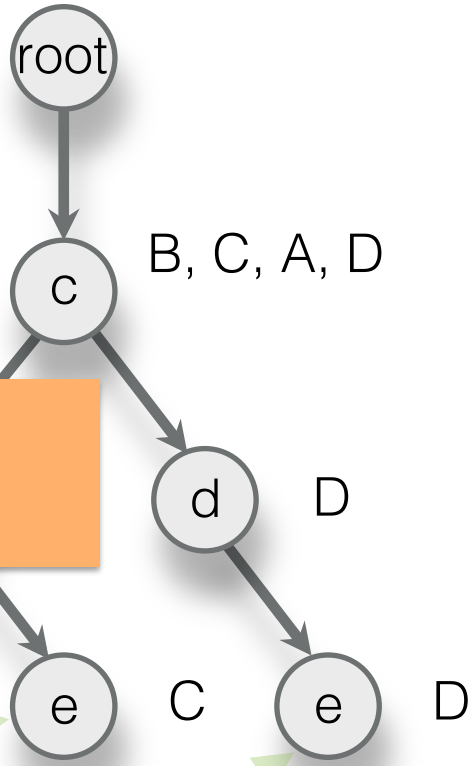


Addressing Limitations (2)





B	c b a d
C	c b e
D	c d e



Supplementation phase

Complete Biclique: [c,b,e] [C]

Complete Biclique: [c,d,e] [D]

Outline

- System overview
- Lockstep analysis
  - Attribution
  - Observations
- Evaluation
  - Streaming
- Conclusion

Lockstep Analysis

- Beewolf in offline mode
- Time window Δt of 3 days
  - Shorter than the typical reaction time of domain blacklist
- Summary
  - Locksteps: 67,094

Label by Publisher

- Identify the organization
- **Representative publisher (rep-pub)**
  - A publisher that accounts more than 50% of the signed downloaders in the lockstep

    ex) [OutBrowse, OutBrowse, MindAd LTD]

    OutBrowse
  - Cannot identify rep-pub: mixed
- Categorization (rep-pub)
  - **PUP**, **PPI**, **benign(BN)**, **other**, mixed, unknown(UK)

Label by Publisher Result

- Identified 335 rep-pubs
- Investigate the top 50 rep-pubs
- Large portion of the locksteps correspond to the Mixed category followed by PUP

Difficult to place in a specific category

Label by Payload

- Understand the purpose of the lockstep
- Detection performance evaluation
- First, label the **downloader** by the payload they distribute
  - Malware downloader (MD)
  - PUP downloader (PD)
  - Benign downloader (BD)
  - Unknown downloader (UD)

Label by Payload Cont'

Suspicious

- **Malware downloader lockstep (MDL)**: lockstep that include at least one MD
- **PUP downloader lockstep (PDL)**: contains PD but no MD
- **Unknown downloader lockstep (UDL)**: no suspicious downloader
- **Benign downloader lockstep (BDL)**: no suspicious downloader, contain BD

Benign

Label by Payload Result

- Higher success rate in labeling (2.33% UDLs)
- MDL occupy more than 80% of the total lockstep while BDL are low (4.82%)

Overlap Between Malware and PUP Delivery Ecosystems

- Overlap of downloaders
  - **Large overlap**
    - 36.7% of the downloaders are present in both MDLs and PDLs
    - Associated with 97.8% of all the PDLs

- Malsign blacklist
  - 1,926 downloaders signed by 212 publishers in locksteps
  - Involved in 66.8% of MDLs and 37.2% of PDLs

Many PUP publishers are likely involved in malware delivery

Overlap Between Malware and PUP Delivery Ecosystems Cont'

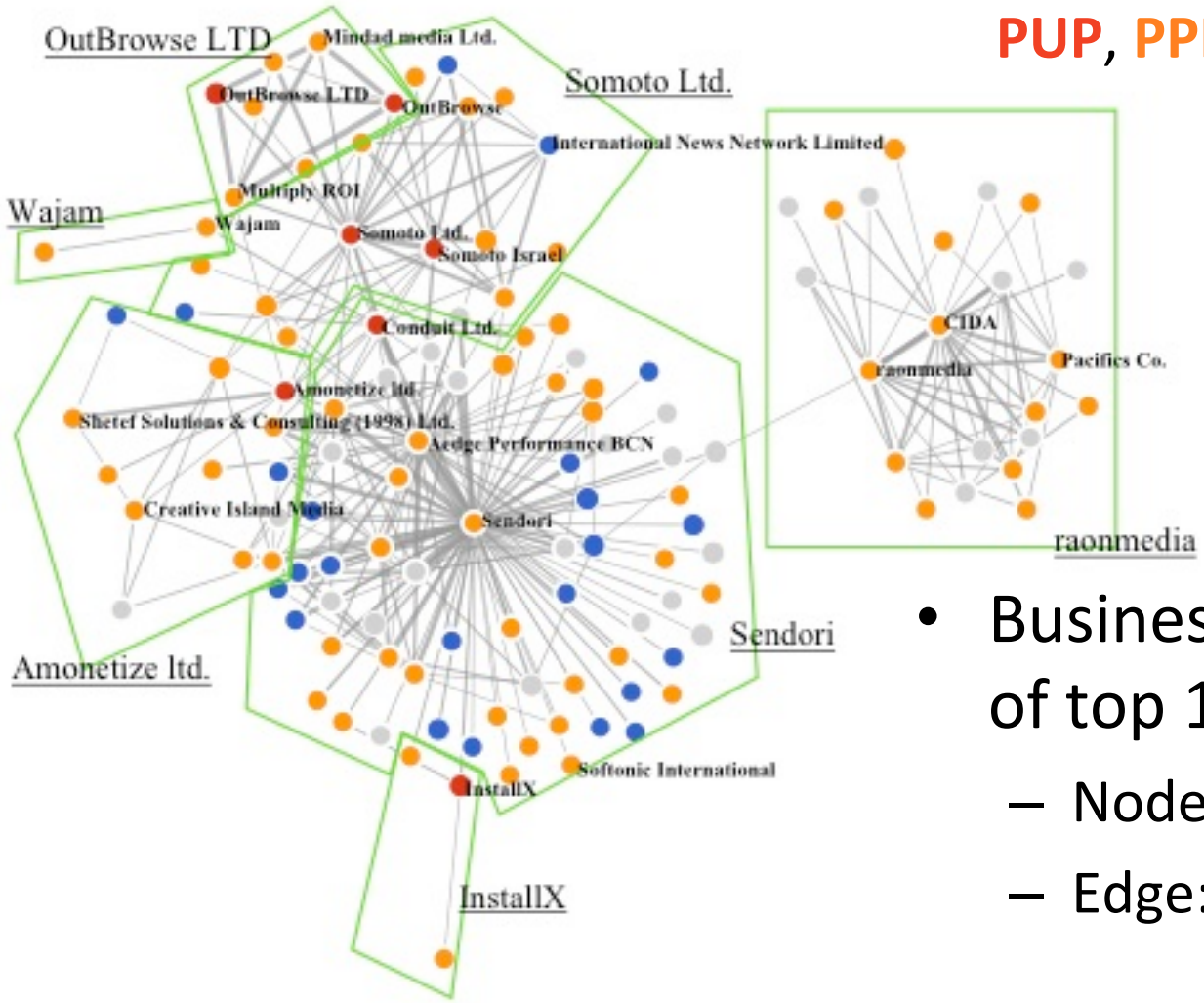
- Recent measurements of commercial PPIs (Kotzias+ 2016, Thomas+ 2016)
  - Did not find substantial overlap

- Key distinction
  - Geographical distribution
    - Hosts from 72 different countries
  - Different observation period / malware set
  - Locksteps detect indirect relationships
    - Utilize unsigned downloaders for malicious payloads

Business Relationships

- Publishers appearing together in locksteps
  - Utilize the same server side infrastructure
    - Reflects a relationship among the corresponding distribution networks
  - Two different publisher relationships
    - Partner: downloaders in downloaded-by relationship
    - Neighbor: No direct download relationship
      - Organization that use multiple code signing certificate
      - Relationships with a common third party

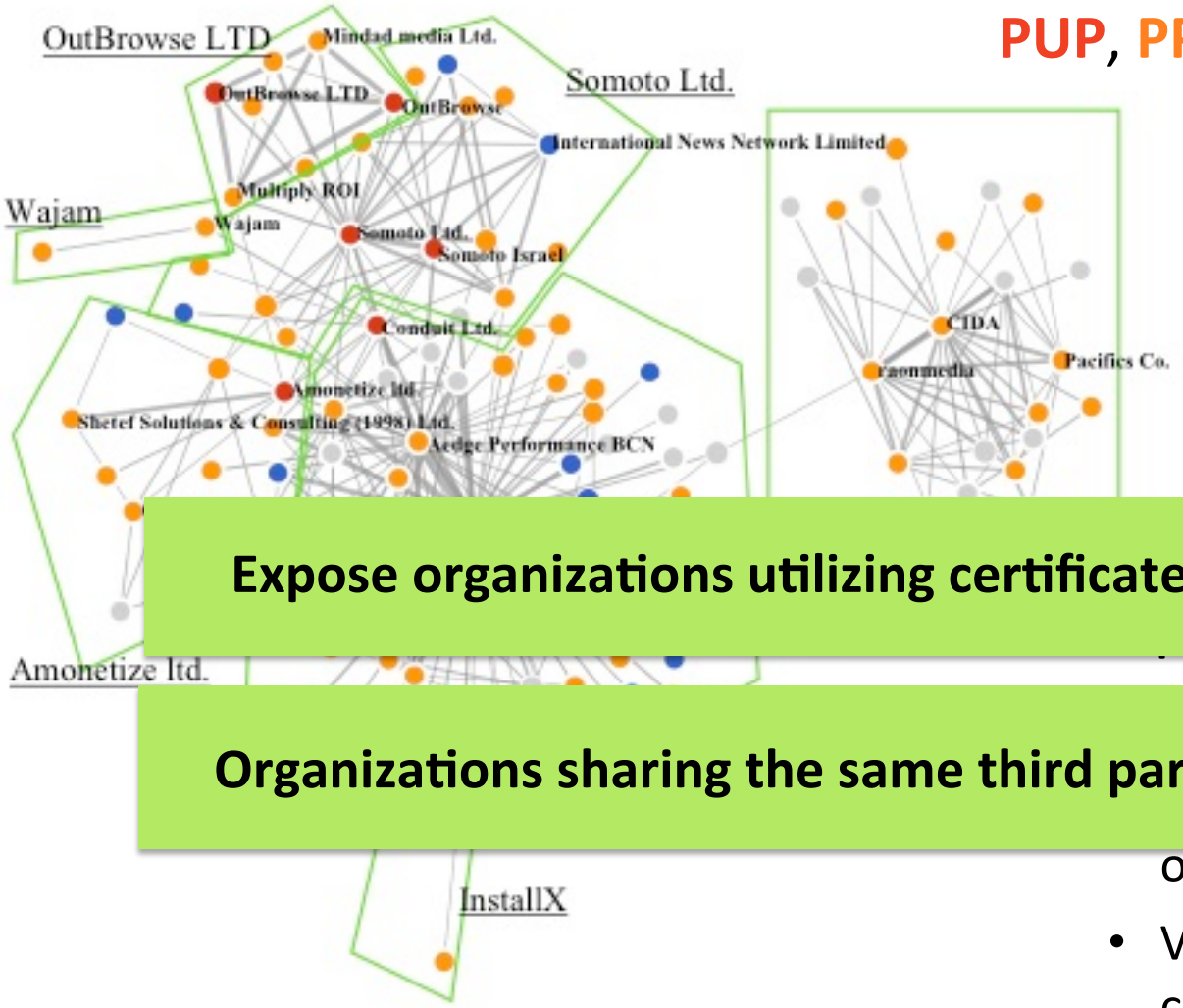
Business Relationships Cont'



PUP, PPI, benign(BN), other

- Business relationship graph of top 13 rep-pubs
  - Node: publisher
  - Edge: business relationship

Business Relationships Cont'



PUP, PPI, benign(BN), other

Expose organizations utilizing certificate polymorphism

Organizations sharing the same third party infrastructure

ates

of the Outbrowse PPI

- Variants of the rep-pub's certificate

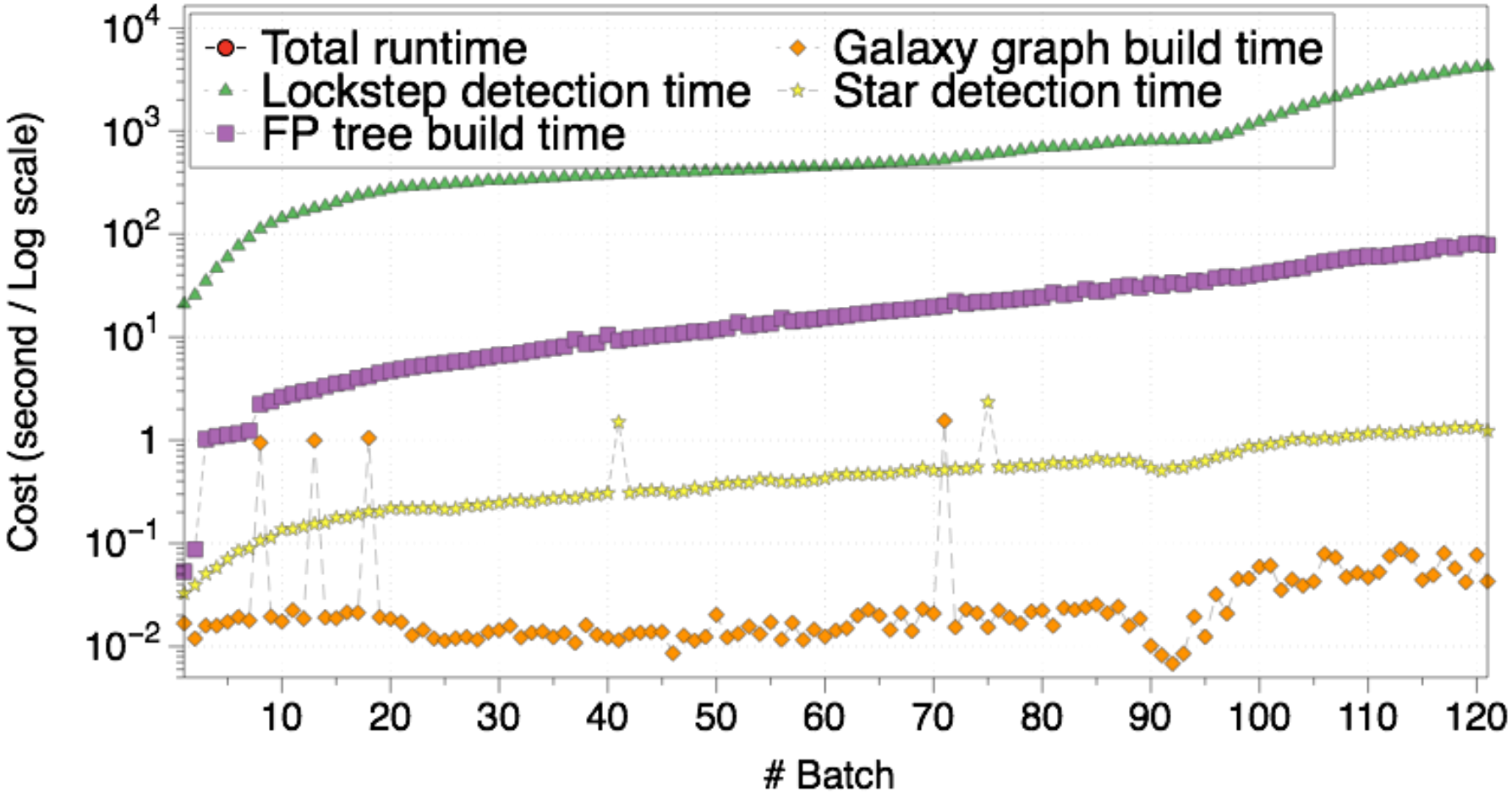
Outline

- System overview
- Lockstep analysis
  - Attribution
  - Observations
- Evaluation
  - Streaming
- Conclusion

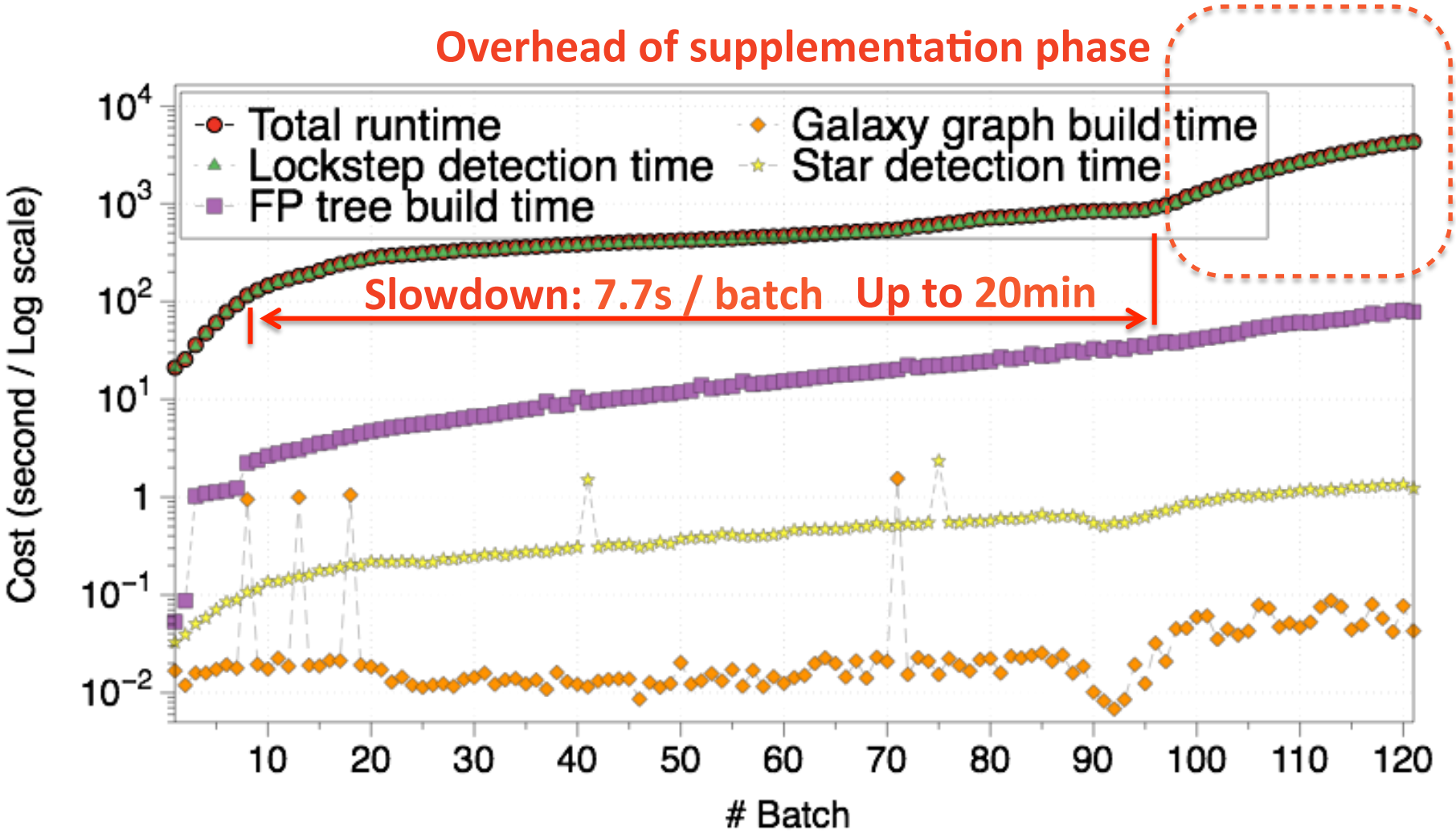
Streaming Setup

- Batch of Download events from the year 2013
  - Download events in time window $\Delta t = 3$ days per batch
  - 122 batch in total
  - Check the computation cost (time) growth

Streaming Performance: Serial

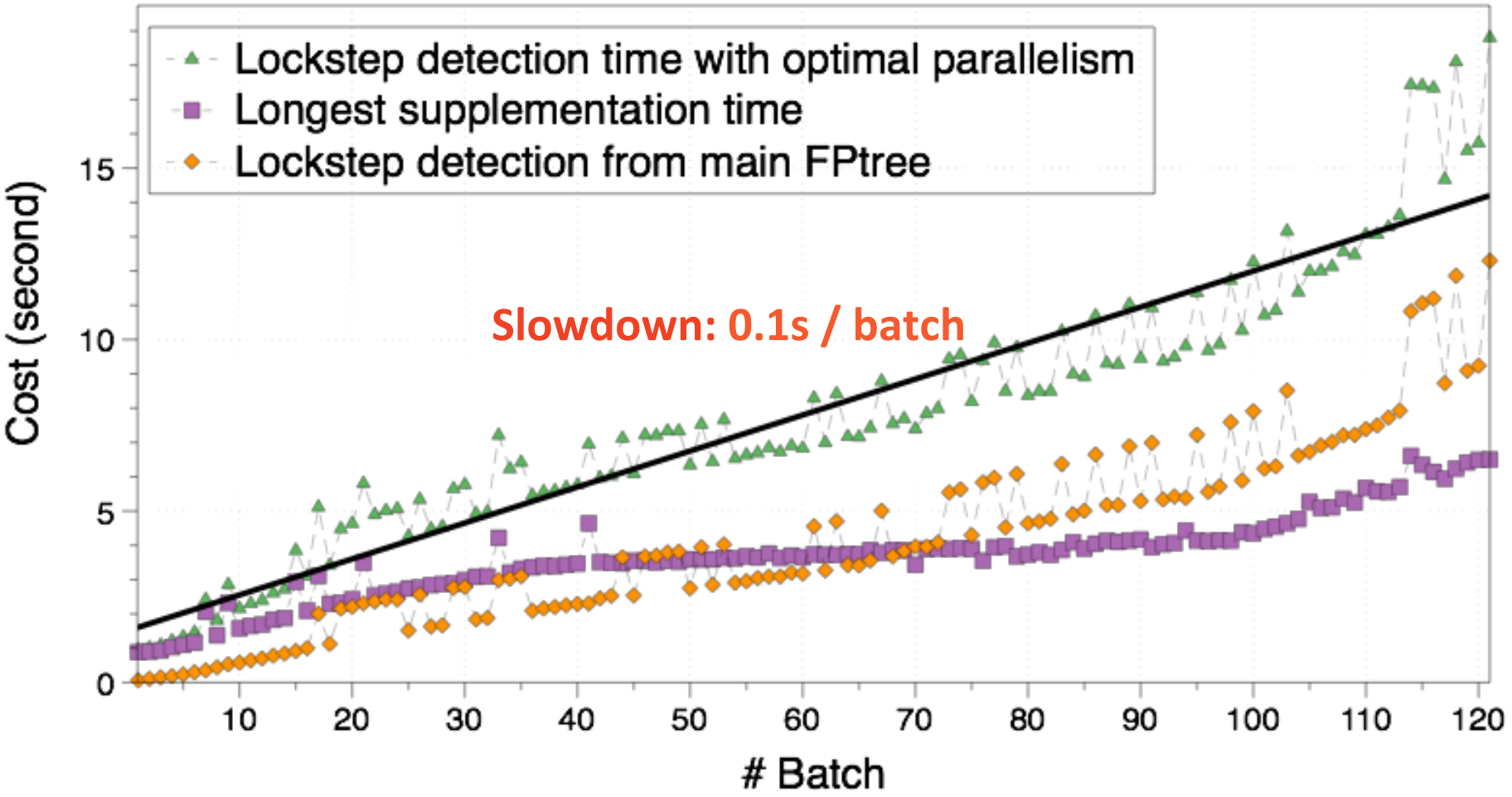


Streaming Performance: Serial



Streaming Performance: Optimal Parallelism

Supplementation processes are independent => Run in parallel



Outline

- System overview
- Lockstep analysis
  - Attribution
  - Observations
- Evaluation
  - Detection performance
  - Streaming
- Conclusion

Conclusion

- We introduce Beewolf
  - Unsupervised and deterministic system, operates on stream of data
  - Discover indirect relationships (reflect PUP/malware overlap)

- Implication beyond malware detection
  - Beewolf can detect other kinds of coordinated actions (Beaconing, C&C commucation, posting in SNS)

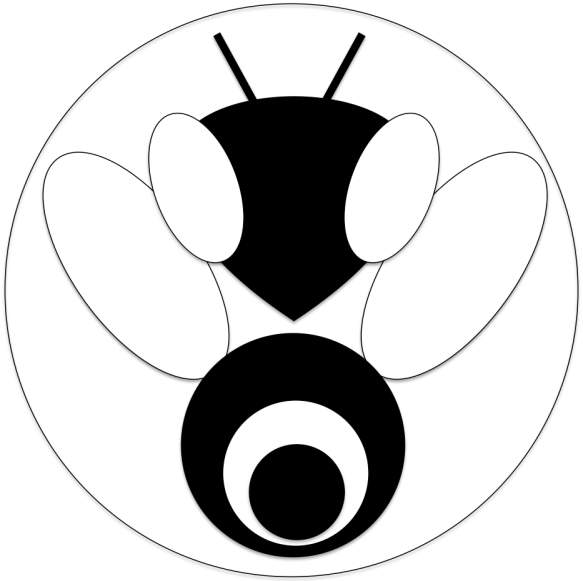
- Data release
  - http://www.beewolf.org

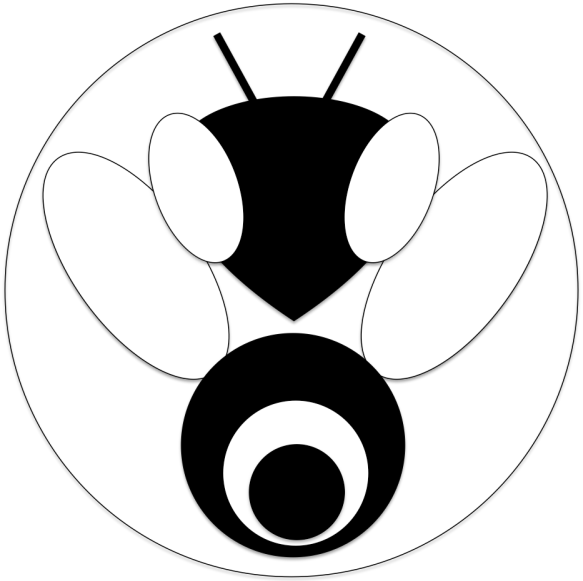
Thank you!

BumJun Kwon

bkwon@umd.edu

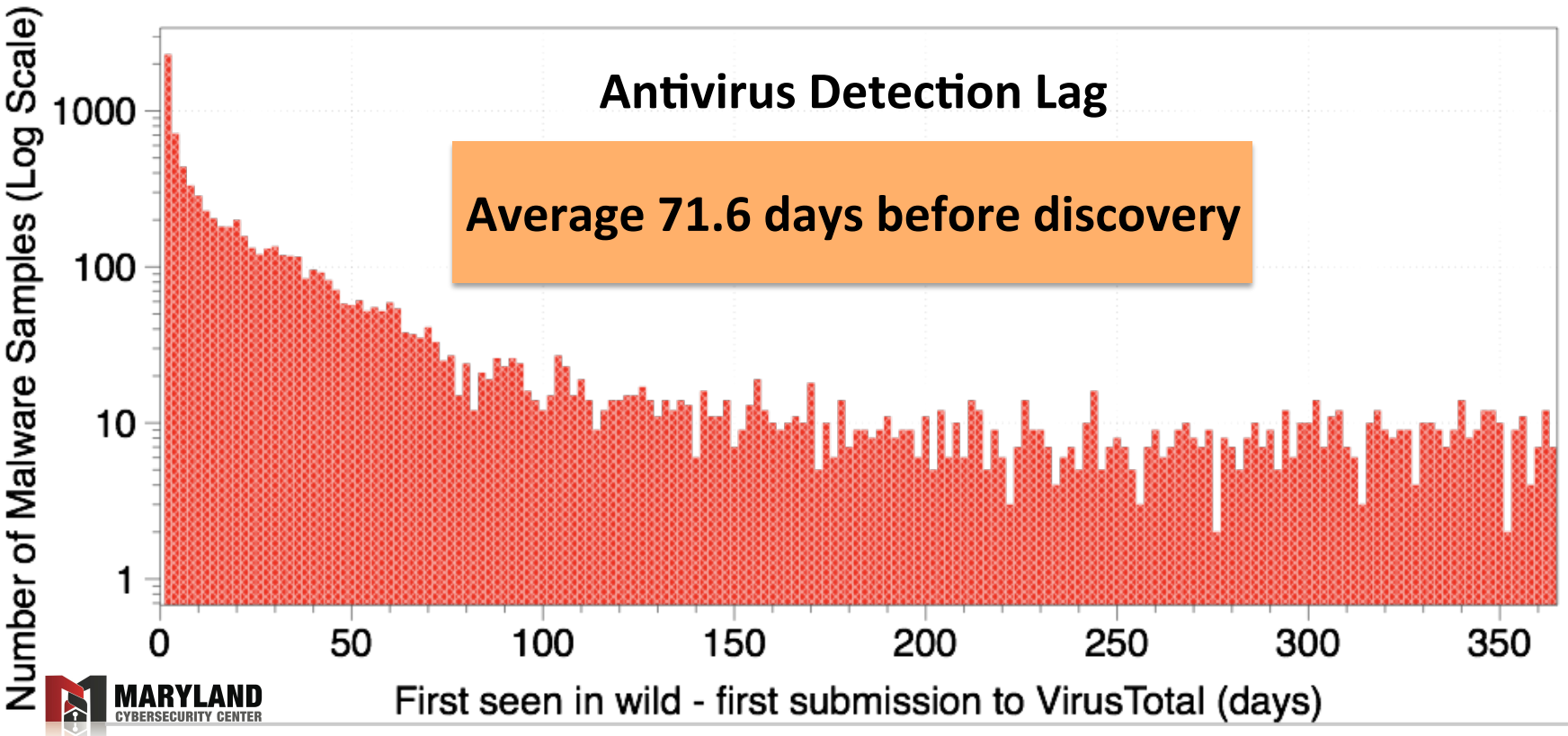
http://www.beewolf.org



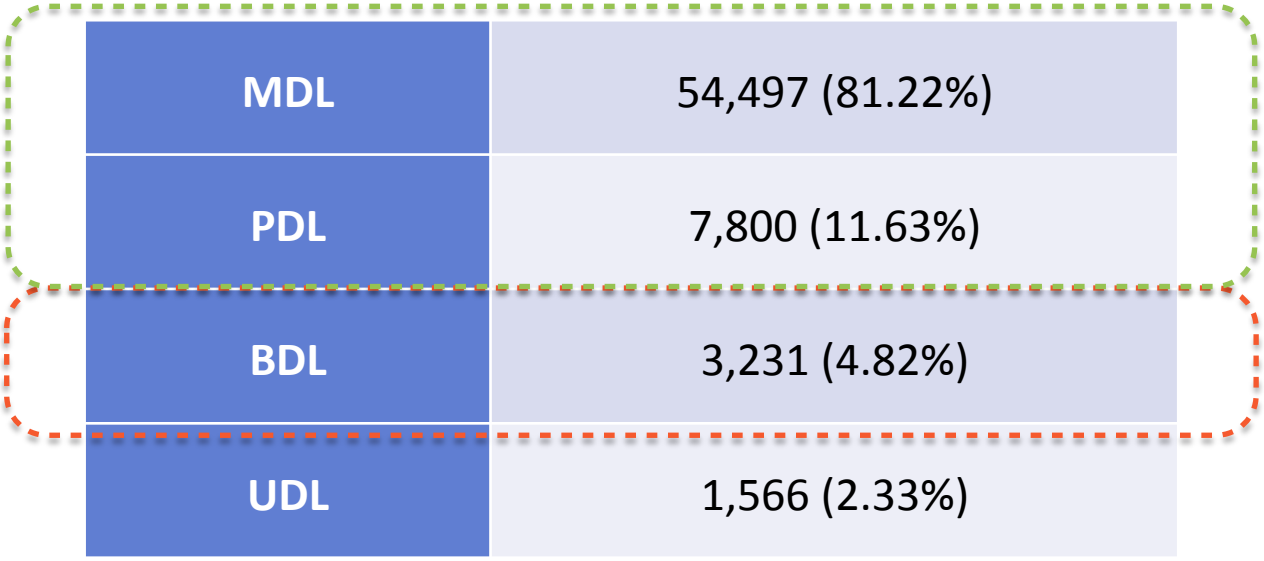


The Detection Lag

- Downloaders
  - Downloading is not a sign of inherently malicious intent
  - Signed downloaders



Detection Performance



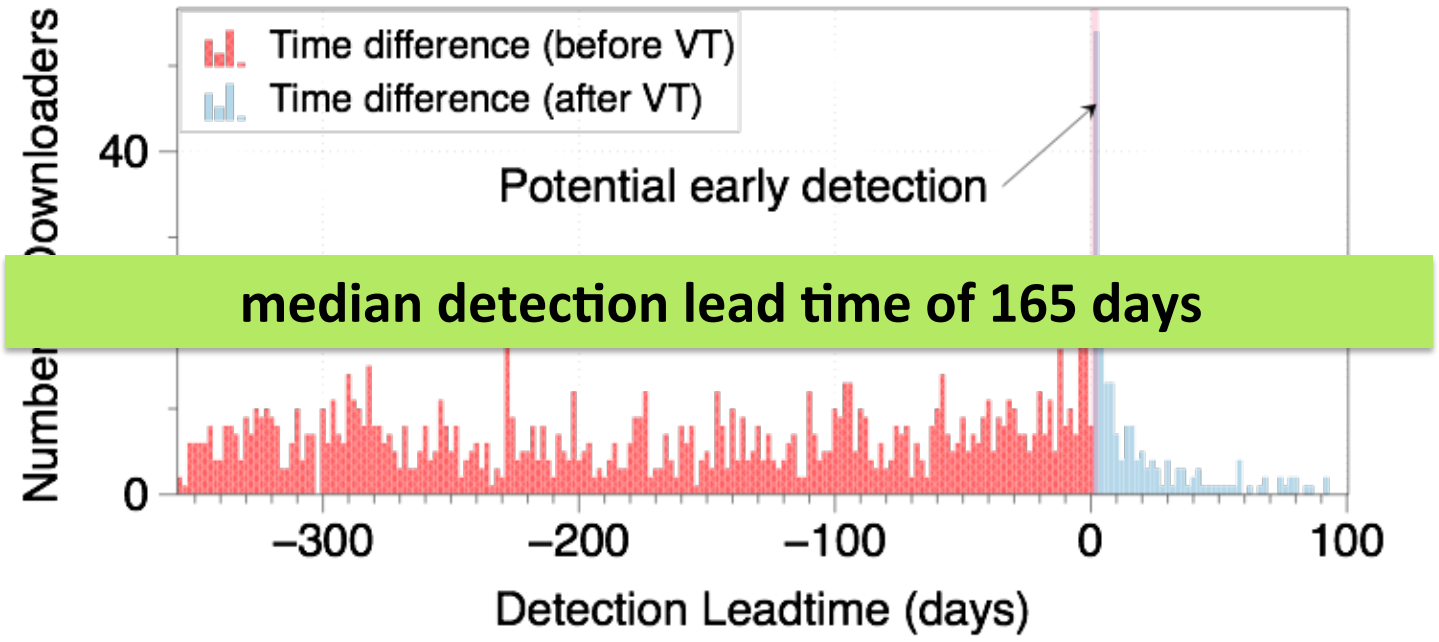
MDL	54,497 (81.22%)
PDL	7,800 (11.63%)
BDL	3,231 (4.82%)
UDL	1,566 (2.33%)

False positive fewer than 5%

True positive (suspicious locksteps) account for 92.85% of locksteps

Detection Lead Time

- How early we can detect suspicious downloaders or domains that are previously unknown?
  - Downloaders: detect unknown executables in lockstep before their first submission to VirusTotal

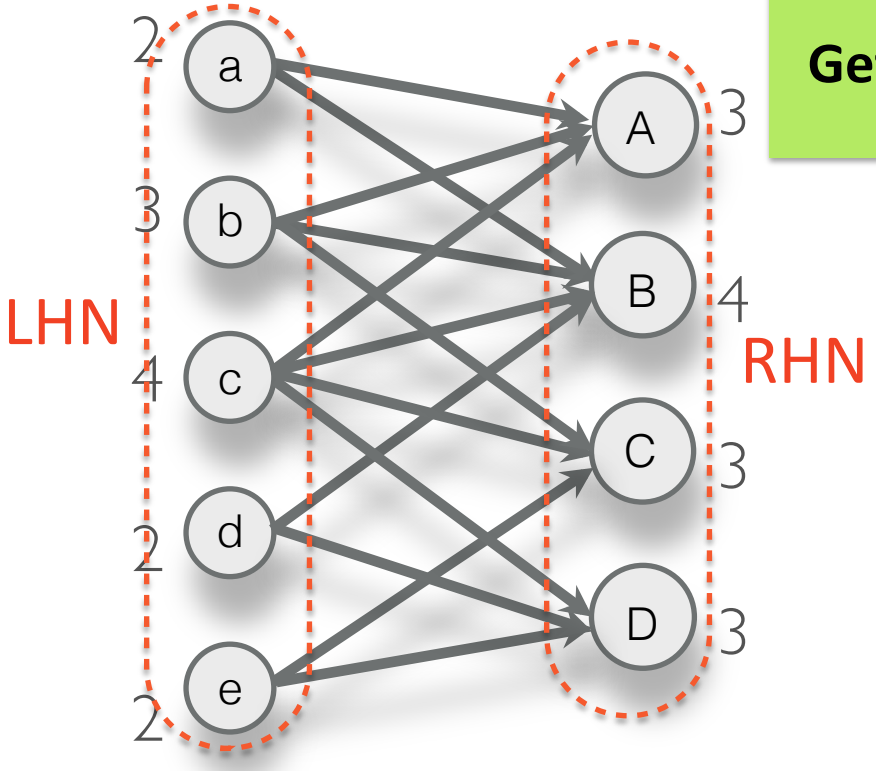


Detection Lead Time Cont'

- How early we can detect suspicious downloaders or domains that are previously unknown?
  - Downloaders: detect unknown executables in lockstep before their first submission to VirusTotal
  - Domains: flag unknown domains in lockstep before listed to public URL blacklists

median detection lead time of 196 days

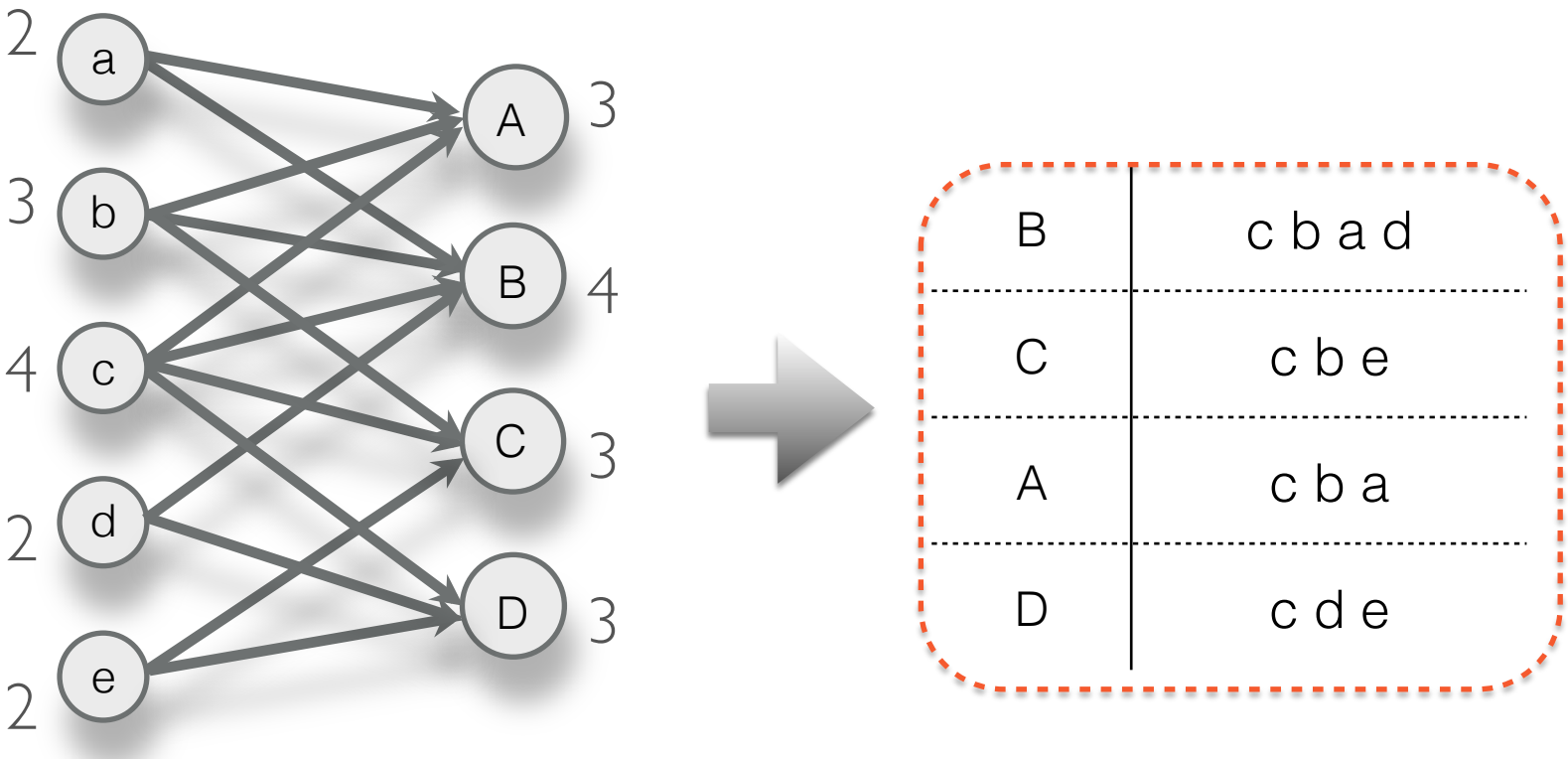
Frequent Pattern Tree (1)



Get the degree for the nodes

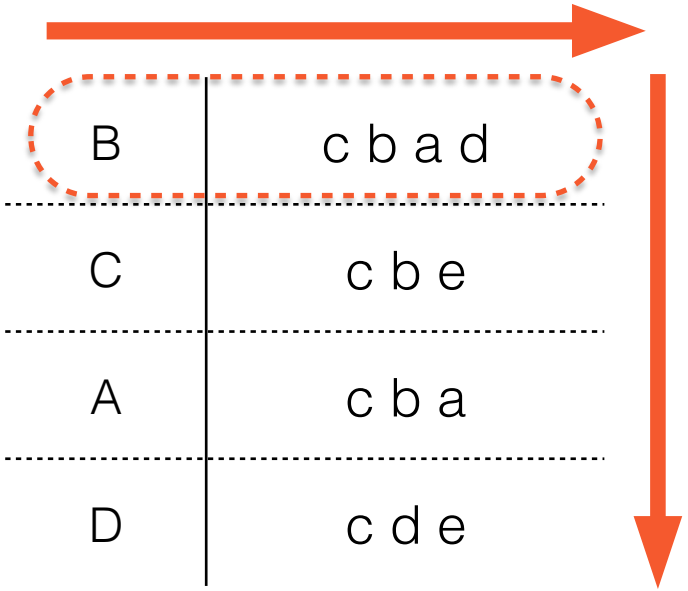
- Pre-setup
  - Bipartite graph of downloaders and second level domain names (domains)

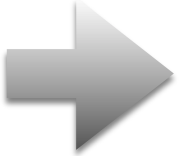
Frequent Pattern Tree (1)

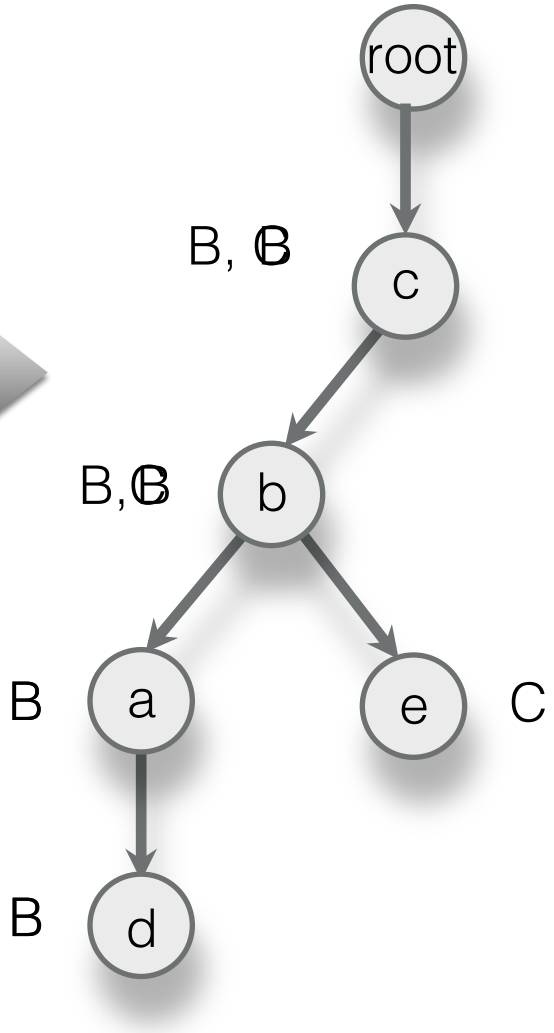


- Adjacency list
  - Sorted in degree-descending order (First sort RHNs, then for each RHN sort its neighbor LHNs)

Frequent Pattern Tree (2)

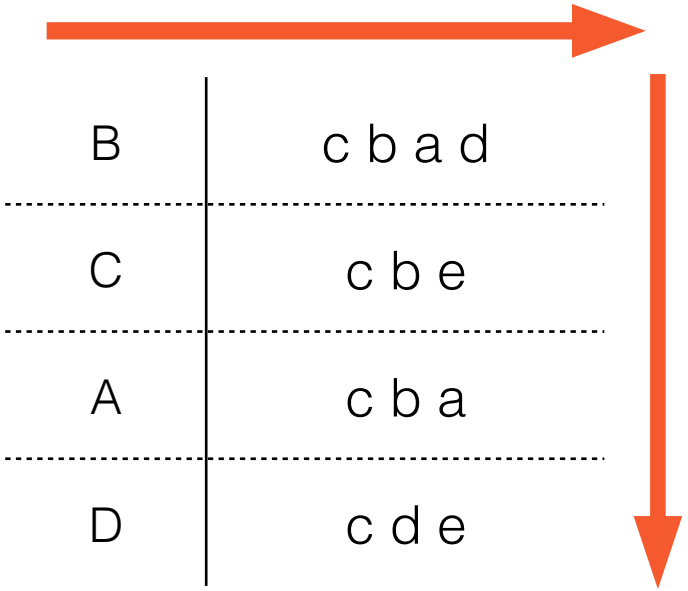


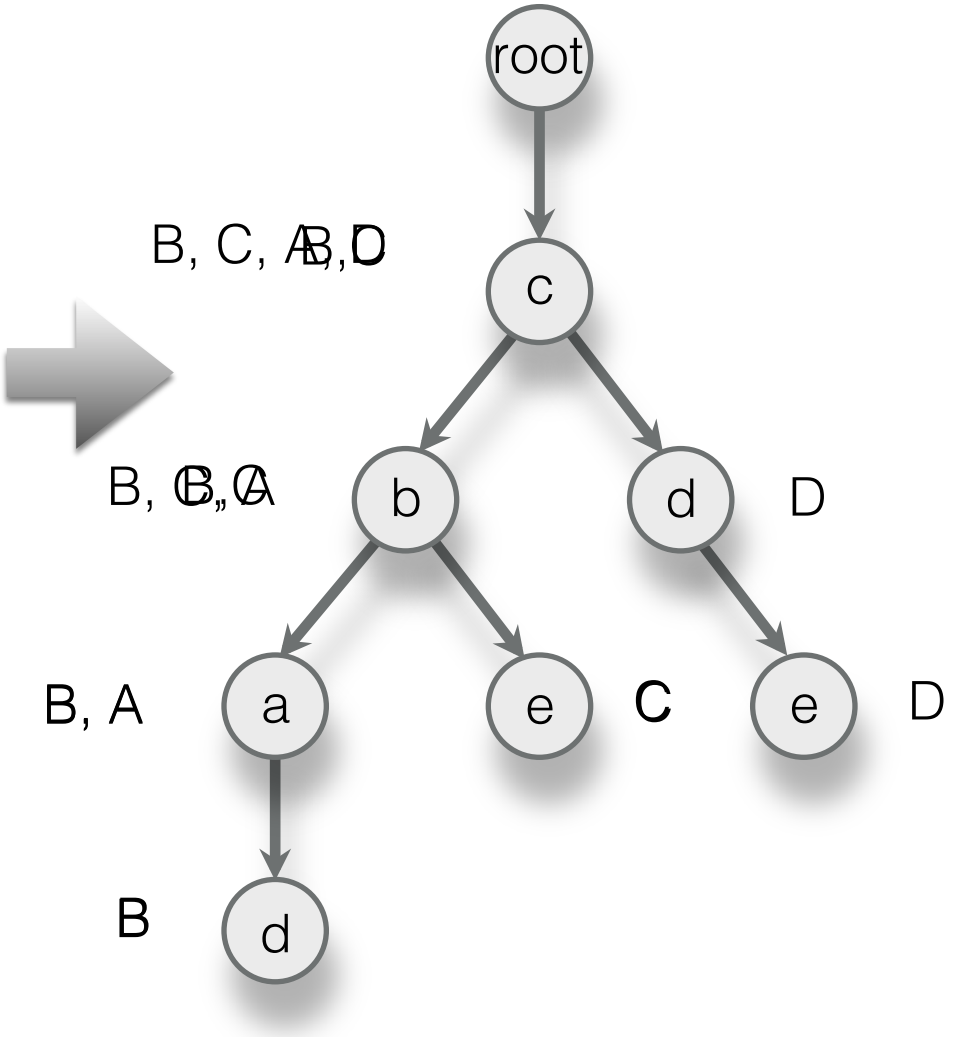




Perform insertion (node: LHN)
1) Not the child: insert as child
2) Add the RHN to the visited list

Frequent Pattern Tree (2)

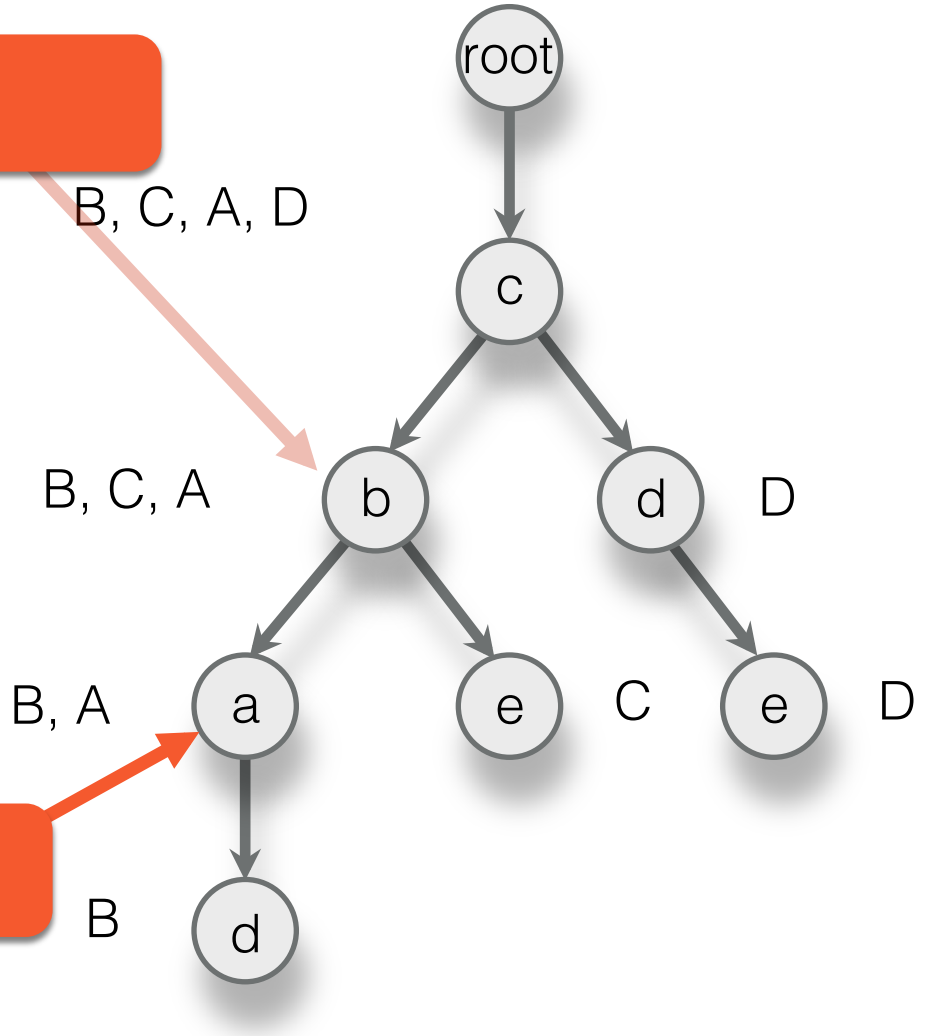




Perform insertion (node: LHN)
1) Not the child: insert as child
2) Add the RHN to the visited list

Frequent Pattern Tree (3)

Lockstep: [c,b] [B,C,A]

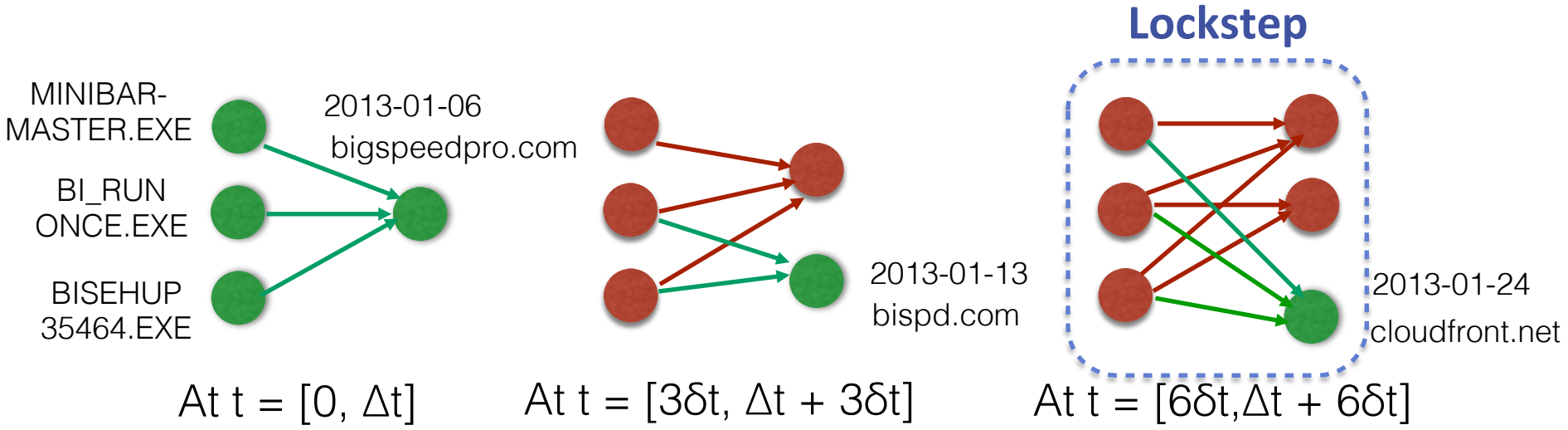


Lockstep: [c,b,a] [B,A]

Outline

- Detecting silent delivery campaigns
  - Lockstep behavior
  - How to detect locksteps: Frequent pattern tree
  - Dataset
  - Lockstep attribution
- System
- Silent distribution campaigns
  - Properties of locksteps
  - Overlap between malware and PUP delivery ecosystems
  - Business relationships
- Evaluation
- Conclusion

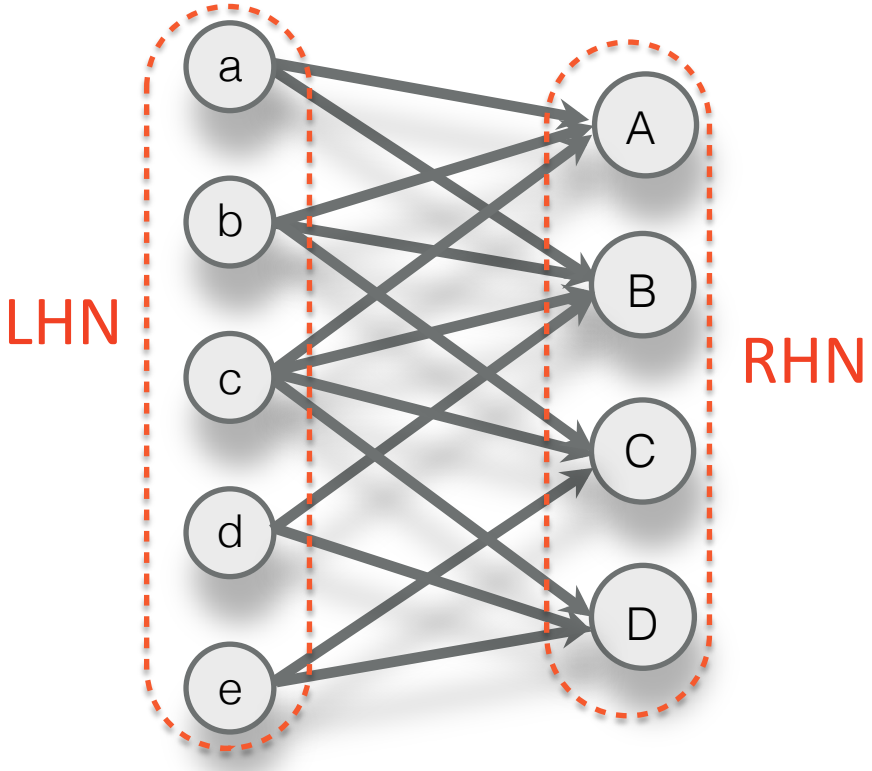
Lockstep Behaviors



- Lockstep behavior
  - Downloader - Domain interaction
  - Temporal pattern: access the same domain within a bounded time period Δt
  - Coordinated downloads that do not experience random delays

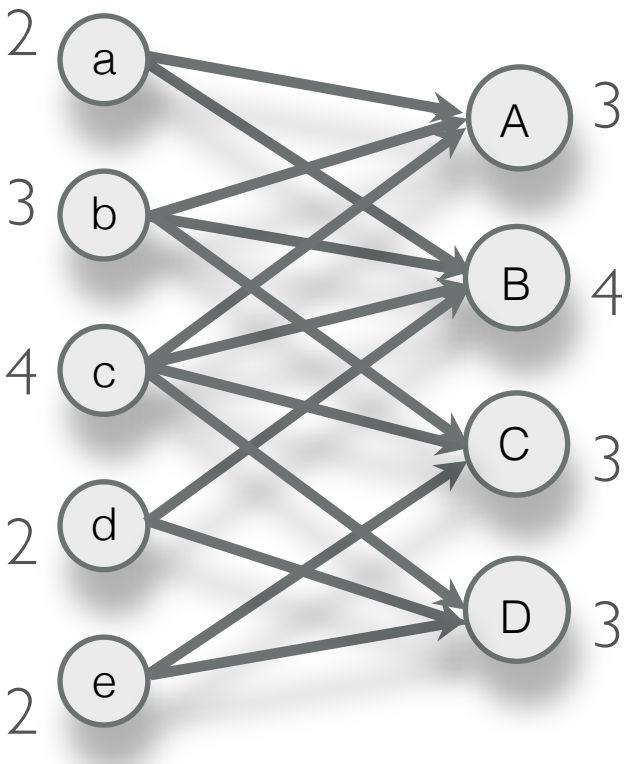
Lockstep behavior exposes remotely controlled downloaders and reveals the domains involved in subsequent campaigns

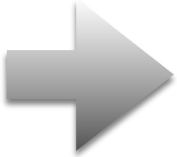
Frequent Pattern Tree (1)



- Pre-setup
  - Bipartite graph of downloaders and second level domain names (domains)

Frequent Pattern Tree (1)



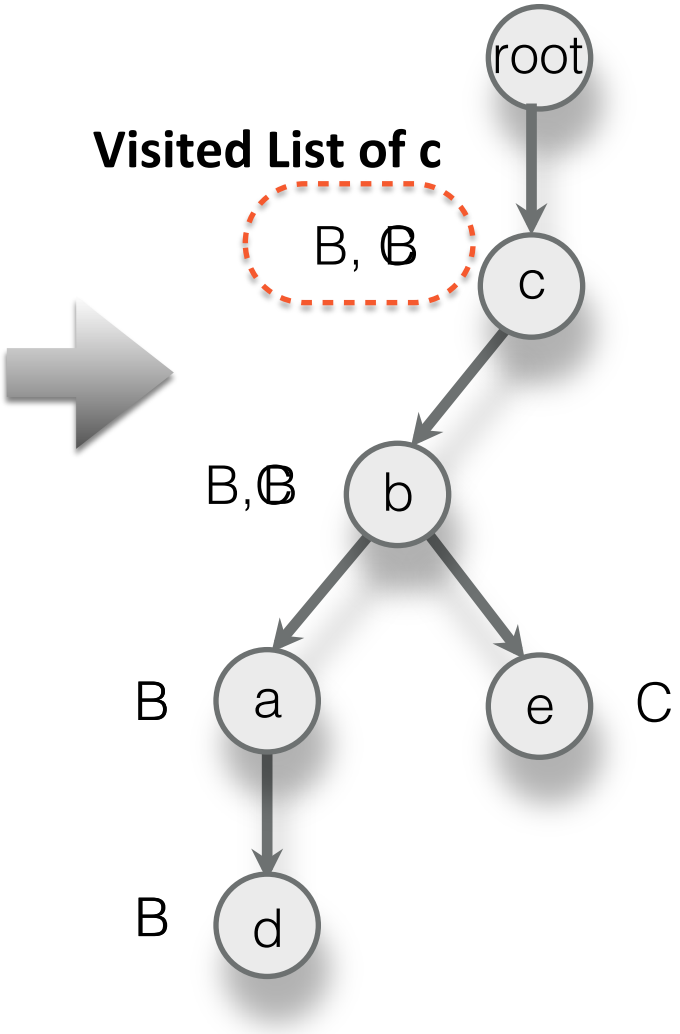


B	c b a d
C	c b e
A	c b a
D	c d e

- Adjacency list
  - Sorted in degree-descending order

Frequent Pattern Tree (2)

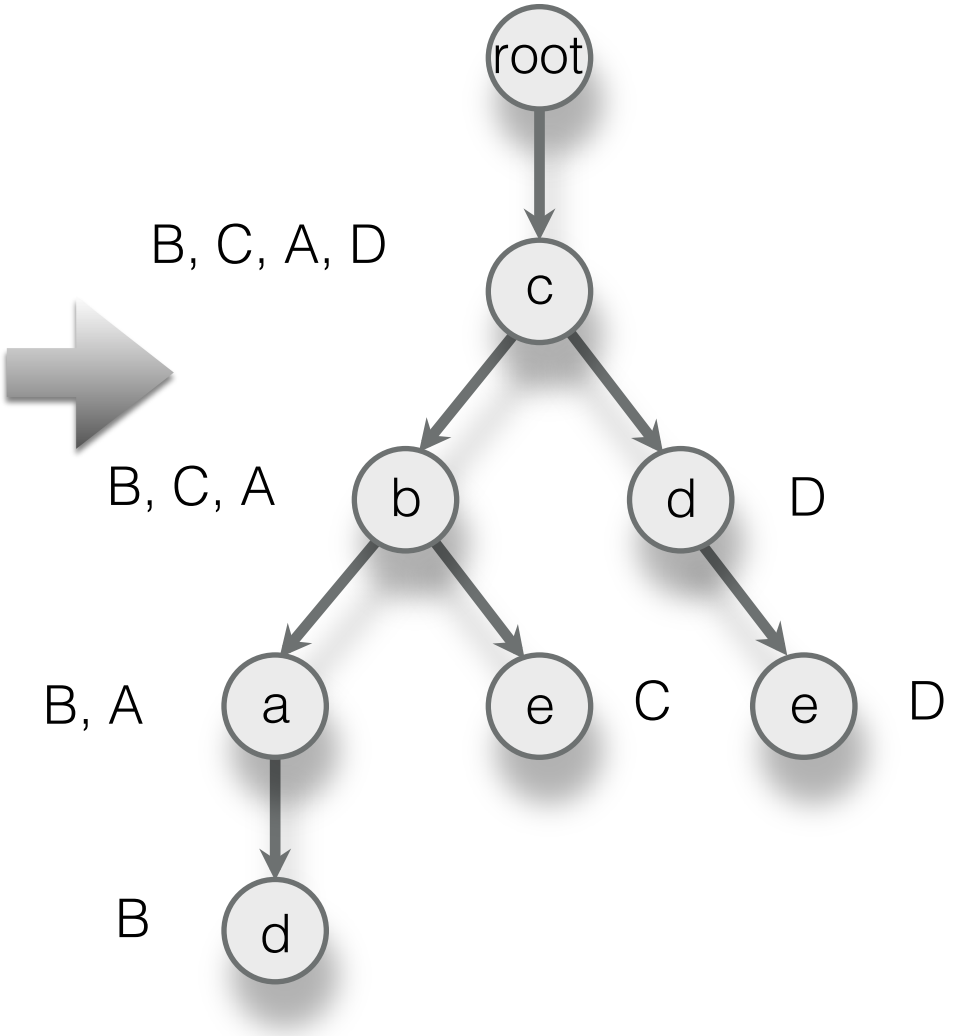
B	c b a d
C	c b e
A	c b a
D	c d e



Perform insertion (node: LHN)
1) Not the child: insert as child
2) Add the RHN to the visited list

Frequent Pattern Tree (2)

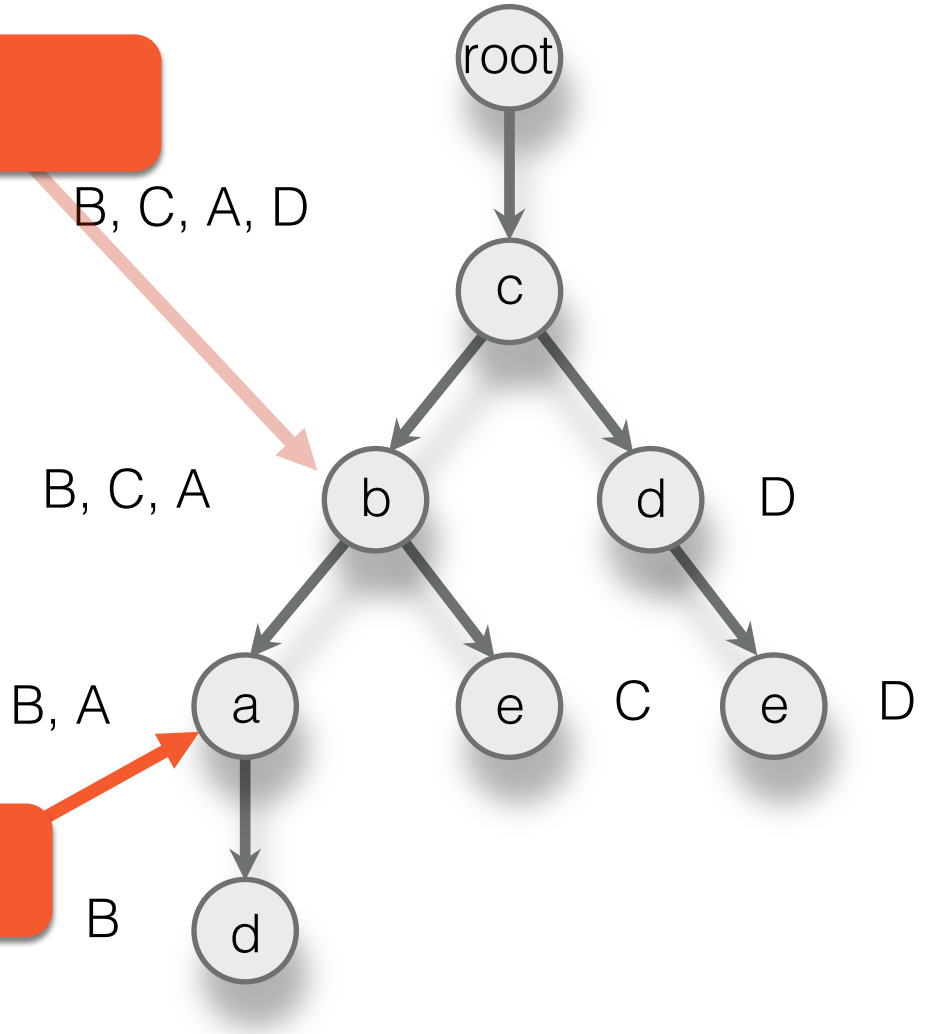
B	c b a d
C	c b e
A	c b a
D	c d e



Perform insertion (node: LHN)
1) Not the child: insert as child
2) Add the RHN to the visited list

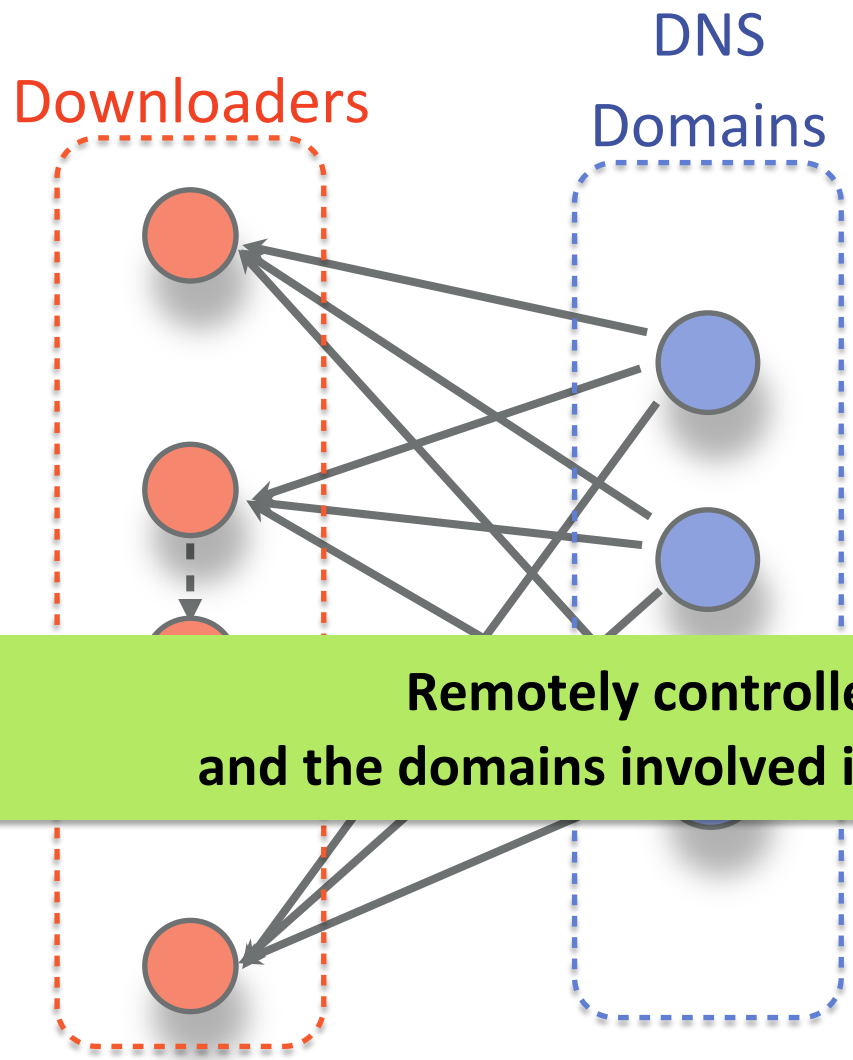
Frequent Pattern Tree (3)

Lockstep: [c,b] [B,C,A]



Lockstep: [c,b,a] [B,A]

How to Detect Silent Delivery Campaigns Cont'

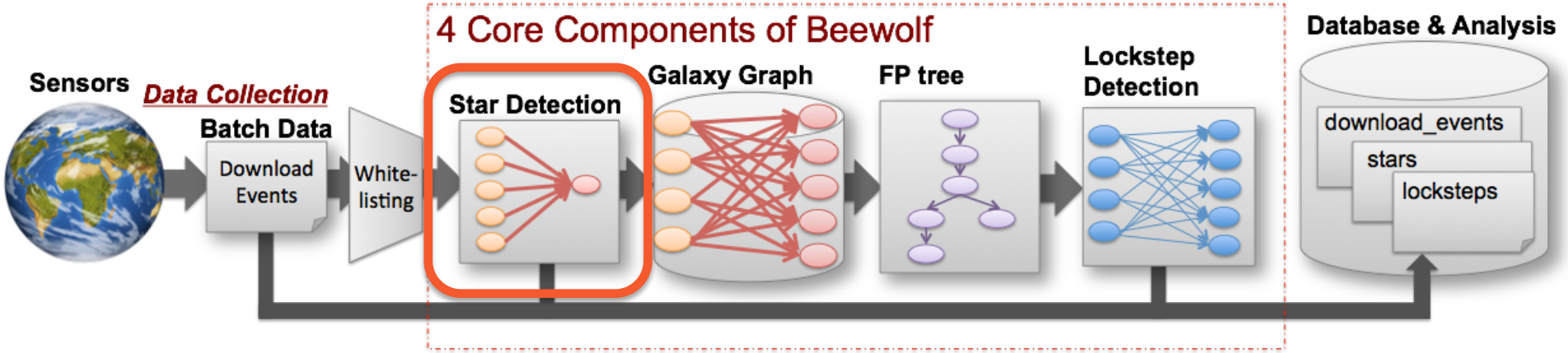


Lockstep behavior:

- **Coordinated** downloads without random delays
- Downloaders-domains in near-bicliques

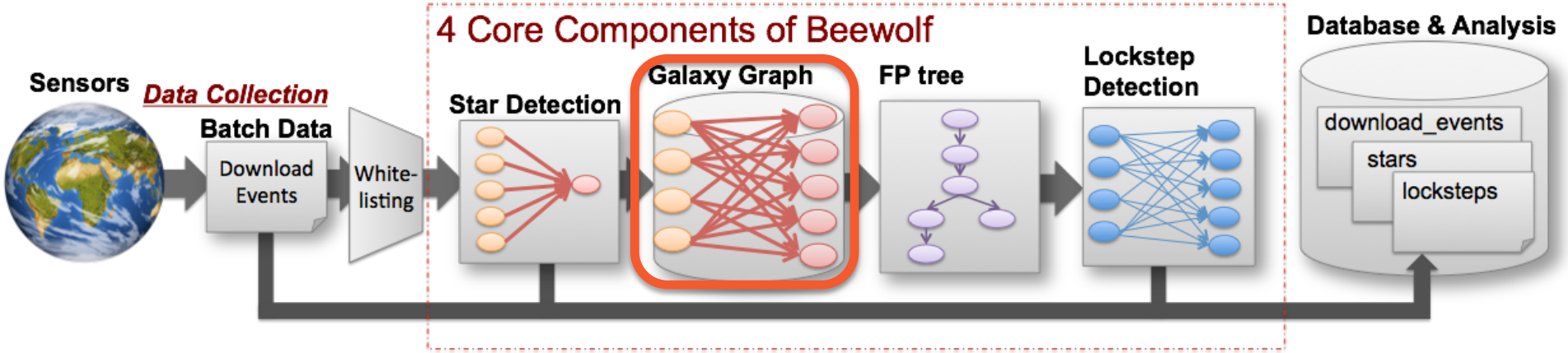
Remotely controlled downloaders and the domains involved in subsequent campaigns

Star Detection



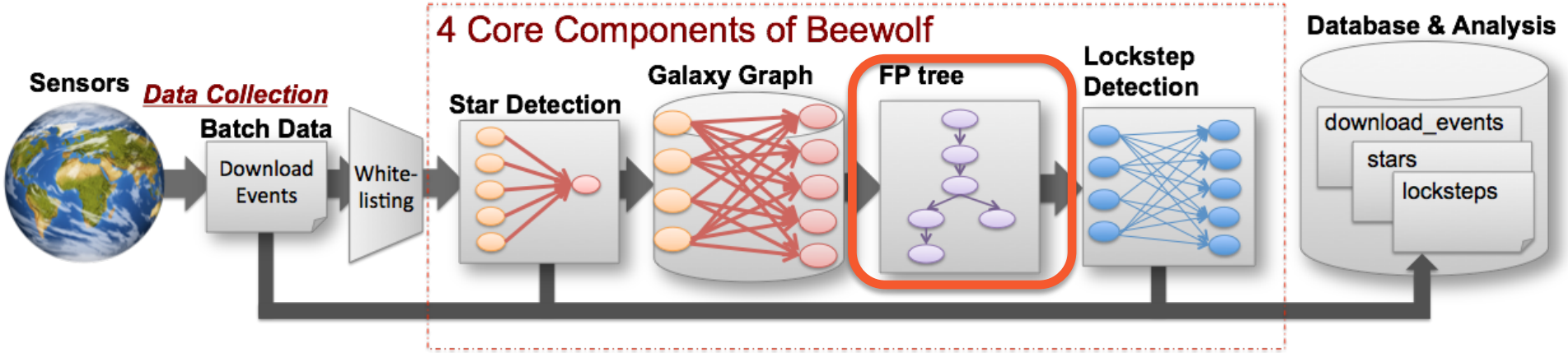
- Detect **Stars**
  - Complete bipartite graph of a single domain and at least 2 downloaders
  - Star corresponds to the row of the adjacency list
- Collect all stars within time window Δt
  - For each domain, aggregate the adjacent downloaders

Galaxy Graph



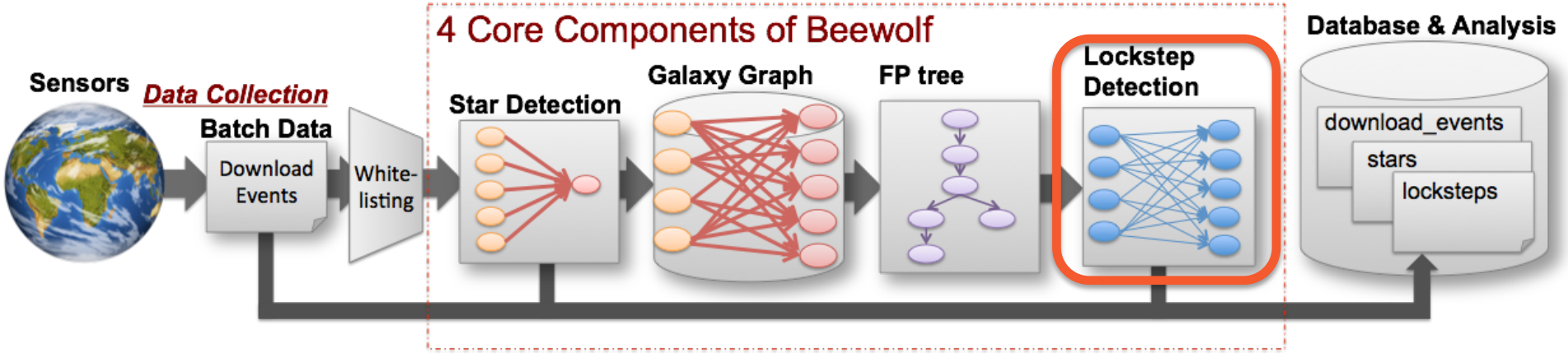
- Bipartite graph of set of stars
- Update the galaxy graph **incrementally**
  - For each star, add the central node and its adjacent nodes to the graph
  - Discard if the star is a subset of some existing star

FP Tree



- Limitations
  - Does not return near-bicliques
    - Heuristic for detecting near-bicliques
  - Misses part of complete bicliques
    - Independent **supplementation phase**

Lockstep Detection



- Traverse the FP tree from the root and collect all the locksteps
- Assign identifiers to the detected locksteps