2019 IEEE Conference on Communications and Network Security

BotFlowMon

Learning-Based, Content-Agnostic Identification of Social Bot Traffic Flows

Yebo Feng, Jun Li, Lei Jiao
University of Oregon
{yebof, lijun, jiao}@cs.uoregon.edu

Xintao Wu
University of Arkansas
xintaowu@uark.edu

Online Social Network with Social Bots

- Online social networks (OSNs) are increasingly threatened by software-controlled **social bots**.
    - They impersonate real OSN users.
    - Majority of OSN malicious activities come from social bots (80% - 90%).
- They can, for example:
    - infiltrate an OSN
    - launch spam campaigns
    - spread fraud information
    - collect private data, and
    - perform financial fraud.

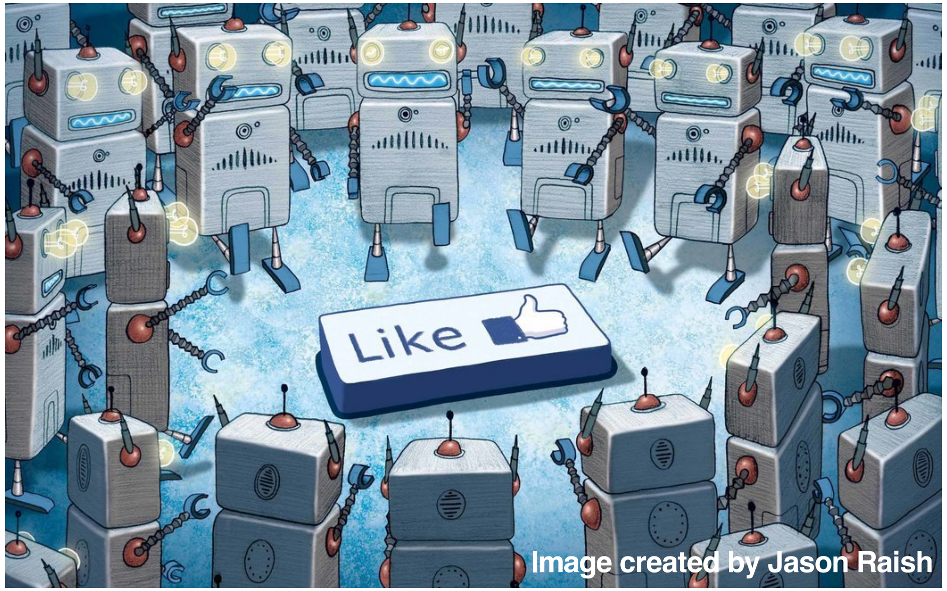
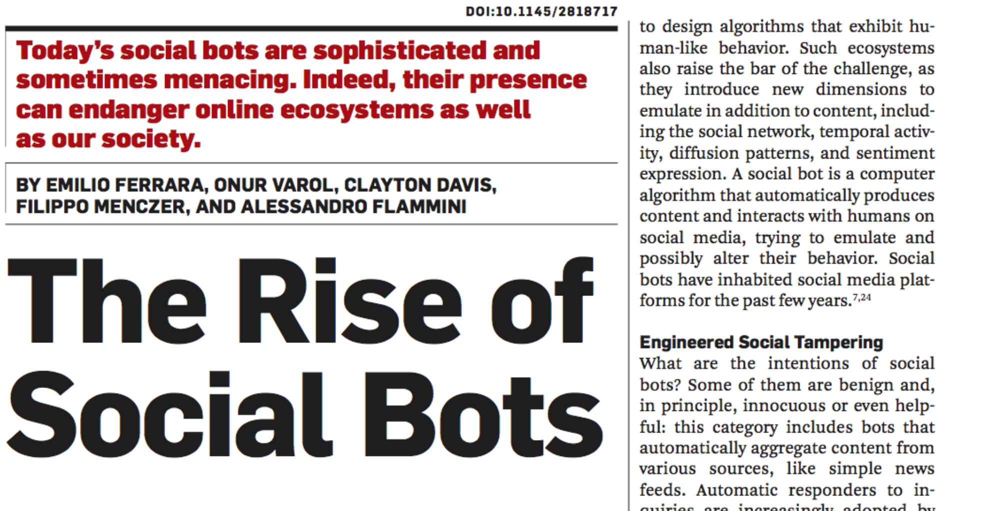


Image created by Jason Raish



DOI:10.1145/2818717

Today's social bots are sophisticated and sometimes menacing. Indeed, their presence can endanger online ecosystems as well as our society.

BY EMILIO FERRARA, ONUR VAROL, CLAYTON DAVIS, FILIPPO MENCZER, AND ALESSANDRO FLAMMINI

The Rise of Social Bots

to design algorithms that exhibit human-like behavior. Such ecosystems also raise the bar of the challenge, as they introduce new dimensions to emulate in addition to content, including the social network, temporal activity, diffusion patterns, and sentiment expression. A social bot is a computer algorithm that automatically produces content and interacts with humans on social media, trying to emulate and possibly alter their behavior. Social bots have inhabited social media platforms for the past few years.[7,24]

Engineered Social Tampering

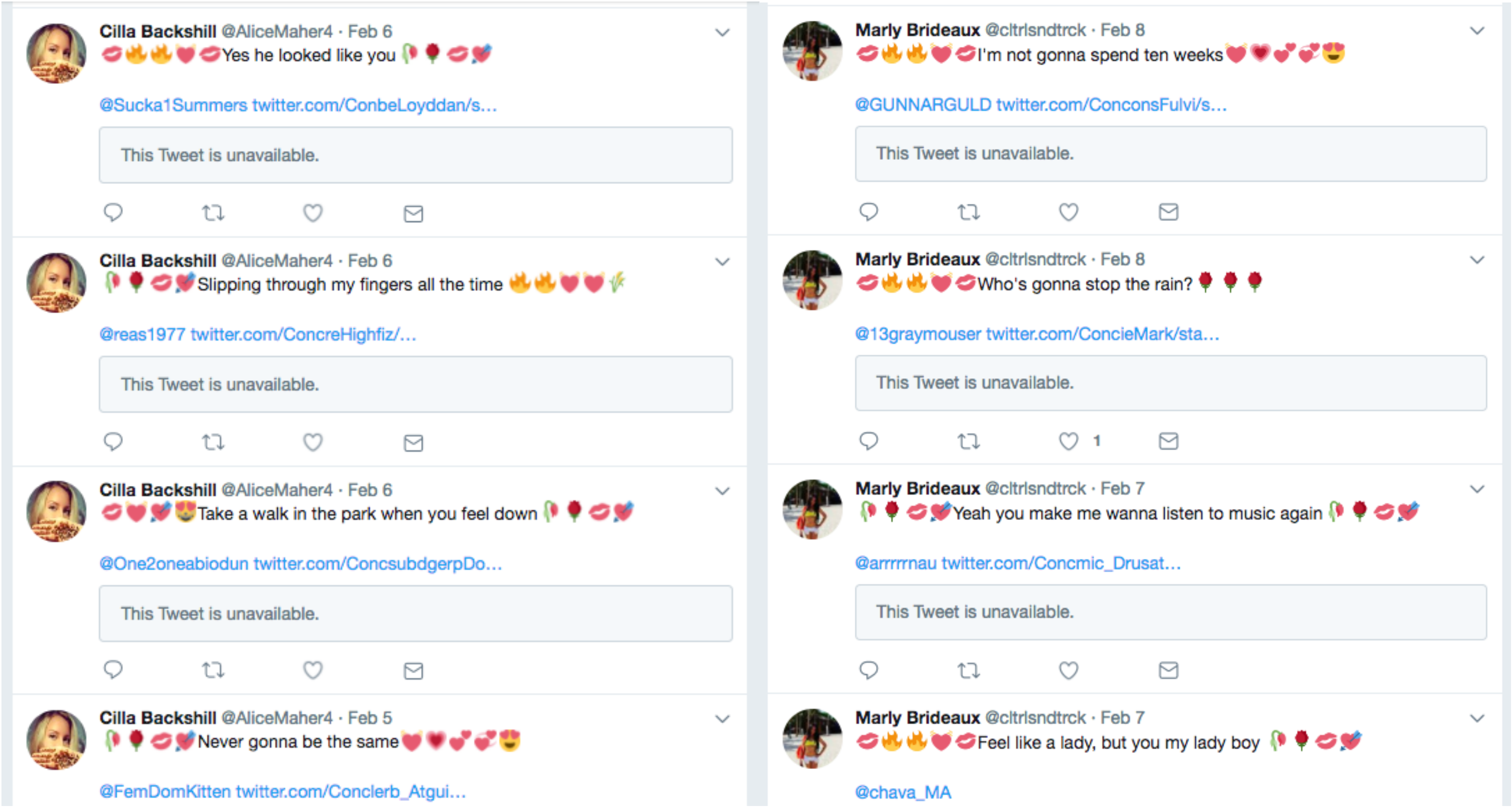
What are the intentions of social bots? Some of them are benign and, in principle, innocuous or even helpful: this category includes bots that automatically aggregate content from various sources, like simple news feeds. Automatic responders to inquiries are increasingly adopted by

Social Bot Examples

- Post bot
- Chatbot
- Amplification bot
- OSN crawler bot
- Hybrid bot

Post Bot

- Automatically post spam tweets, Facebook posts, etc.
- Can contain malicious URLs or texts
- The most common social bots



Cilla Backshill @AliceMaher4 · Feb 6
💋🔥🔥💓💋Yes he looked like you 🥀🌹💋💘

@Sucka1Summers twitter.com/ConbeLoyddan/s...

This Tweet is unavailable.

Cilla Backshill @AliceMaher4 · Feb 6
🥀🌹💋💘Slipping through my fingers all the time 🔥🔥💓💓🌾

@reas1977 twitter.com/ConcreHighfiz/...

This Tweet is unavailable.

Cilla Backshill @AliceMaher4 · Feb 6
💋💓💘😍Take a walk in the park when you feel down 🥀🌹💋💘

@One2oneabiodun twitter.com/ConcsubdgerpDo...

This Tweet is unavailable.

Cilla Backshill @AliceMaher4 · Feb 5
🥀🌹💋💘Never gonna be the same💓💗💕💞😍

@FemDomKitten twitter.com/Conclerb_Atgui...

Marly Brideaux @cltrlsndtrck · Feb 8
💋🔥🔥💓💋I'm not gonna spend ten weeks💓💗💕💞😍

@GUNNARGULD twitter.com/ConconsFulvi/s...

This Tweet is unavailable.

Marly Brideaux @cltrlsndtrck · Feb 8
💋🔥🔥💓💋Who's gonna stop the rain? 🌹🌹🌹

@13graymouser twitter.com/ConcieMark/sta...

This Tweet is unavailable.

1

Marly Brideaux @cltrlsndtrck · Feb 7
🥀🌹💋💘Yeah you make me wanna listen to music again 🥀🌹💋💘

@arrrrrnau twitter.com/Concmic_Drusat...

This Tweet is unavailable.

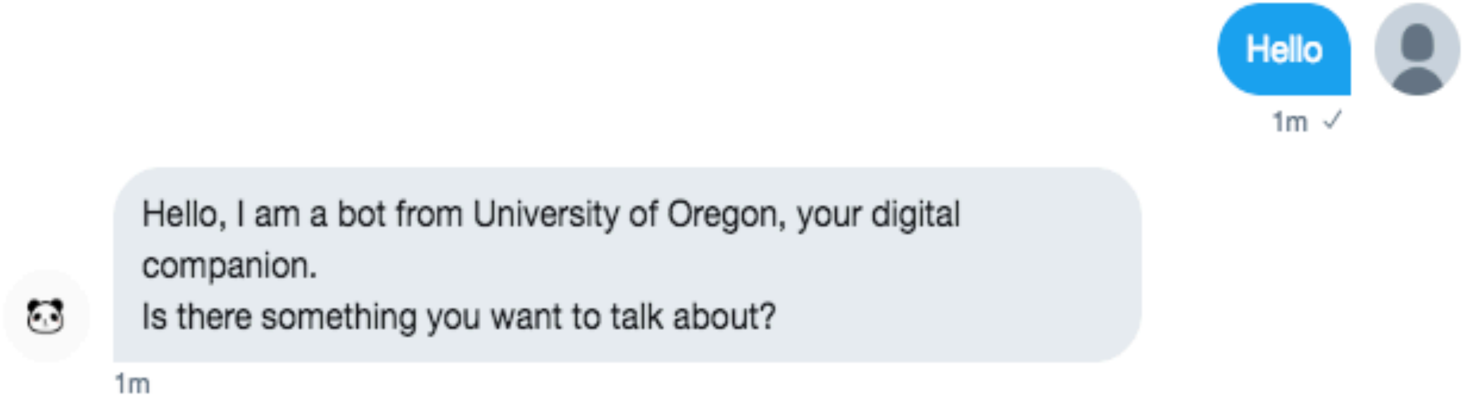
Marly Brideaux @cltrlsndtrck · Feb 7
💋🔥🔥💓💋Feel like a lady, but you my lady boy 🥀🌹💋💘

@chava_MA

Examples are from medium.com - *Analyzing a bot attack on a news site in the United States.*

Chatbot

- Automatically converse with regular users



Hello

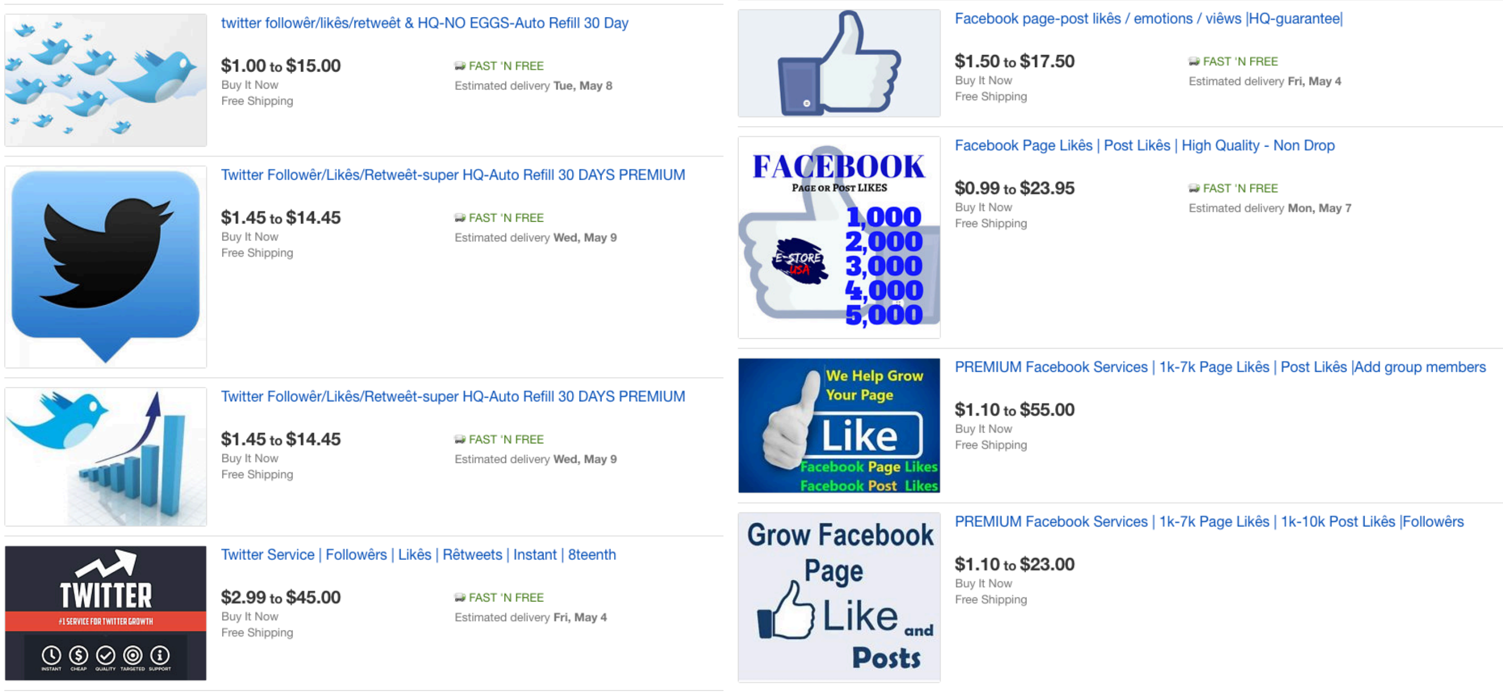
1m ✓

Hello, I am a bot from University of Oregon, your digital companion.
Is there something you want to talk about?

1m

Amplification bot

- Amplify OSN accounts via fake followers, or
- Amplify content by, for example, artificial retweets and likes



twitter followêr/likês/retweêt & HQ-NO EGGS-Auto Refill 30 Day
$1.00 to $15.00
Buy It Now
Free Shipping
FAST 'N FREE
Estimated delivery **Tue, May 8**

Twitter Followêr/Likês/Retweêt-super HQ-Auto Refill 30 DAYS PREMIUM
$1.45 to $14.45
Buy It Now
Free Shipping
FAST 'N FREE
Estimated delivery **Wed, May 9**

Twitter Followêr/Likês/Retweêt-super HQ-Auto Refill 30 DAYS PREMIUM
$1.45 to $14.45
Buy It Now
Free Shipping
FAST 'N FREE
Estimated delivery **Wed, May 9**

Twitter Service | Followêrs | Likês | Rêtweets | Instant | 8teenth
$2.99 to $45.00
Buy It Now
Free Shipping
FAST 'N FREE
Estimated delivery **Fri, May 4**

Facebook page-post likês / emotions / viêws |HQ-guarantee|
$1.50 to $17.50
Buy It Now
Free Shipping
FAST 'N FREE
Estimated delivery **Fri, May 4**

Facebook Page Likês | Post Likês | High Quality - Non Drop
$0.99 to $23.95
Buy It Now
Free Shipping
FAST 'N FREE
Estimated delivery **Mon, May 7**

PREMIUM Facebook Services | 1k-7k Page Likês | Post Likês |Add group members
$1.10 to $55.00
Buy It Now
Free Shipping

PREMIUM Facebook Services | 1k-7k Page Likês | 1k-10k Post Likês |Followêrs
$1.10 to $23.00
Buy It Now
Free Shipping

Examples are from the searching results of ebay.com.

OSN Crawler Bot

- Page crawler
  - read the HTML files of OSN users
- API crawler
  - become friends of OSN users and fetch their information via API calls

API sample code:

```
api.GetUser(user)
api.GetReplies()
api.GetUserTimeline(user)
api.GetHomeTimeline()
api.GetStatus(status_id)
api.GetStatuses(status_ids)
api.GetFriends(user)
api.GetFollowers()
api.GetFeatured()
```

Outline

- **Related work**
- BotFlowMon Overview
- Five Modules of BotFlowMon
- Evaluations
- Conclusions

Related Work

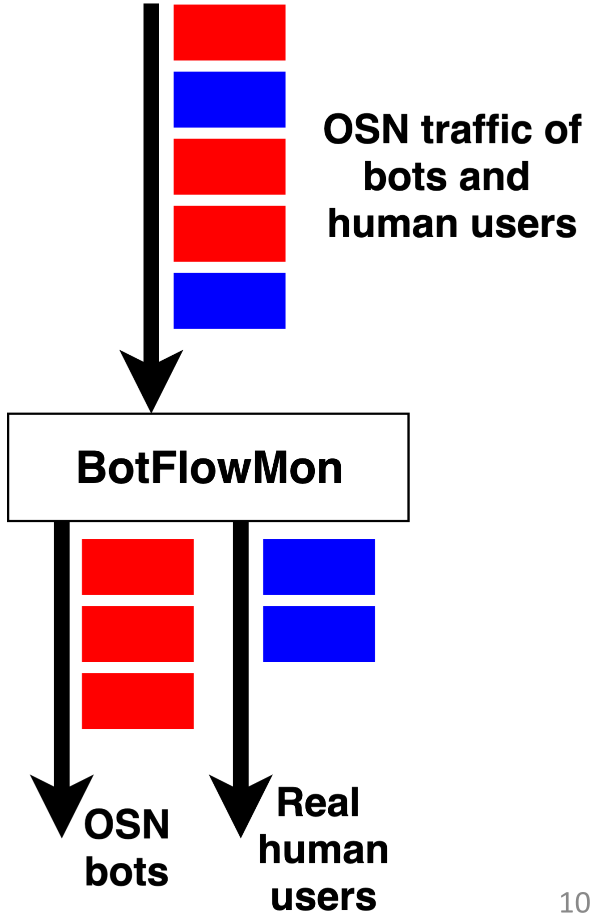
- **Content-based approaches**
  - Rely on post syntax, content, account activity, post linguistic features, etc.
  - Only executable by OSN providers
  - Could incur severe privacy concerns
- **Topology-based approaches**
  - Use topology structure of an online social network
  - Also only executable by OSN providers
  - Could incur moderate privacy concerns
- **Crowdsourcing-based approaches**
  - Ask participants to judge whether an account is a bot or not
  - incur a long running time, a high cost, and
  - privacy risk

Outline

- Related work
- **BotFlowMon Overview**
- Five modules of BotFlowMon
- Evaluations
- Conclusions

BotFlowMon: Content-Agnostic Social Bot Detection

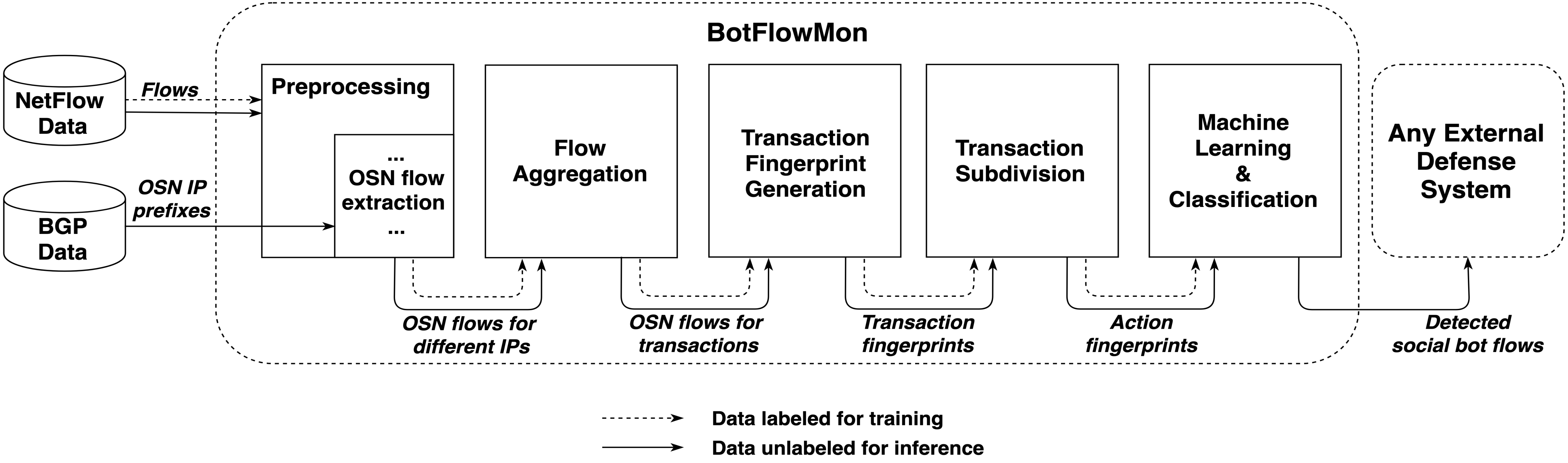
- BotFlowMon can process traffic flow data to distinguish social bot traffic flows from real OSN user flows.
- Any network service provider can deploy BotFlowMon
- Only need metadata of traffic flows
  - No IP packet payload data is needed
  - Fast and scalable
  - Privacy-preserving



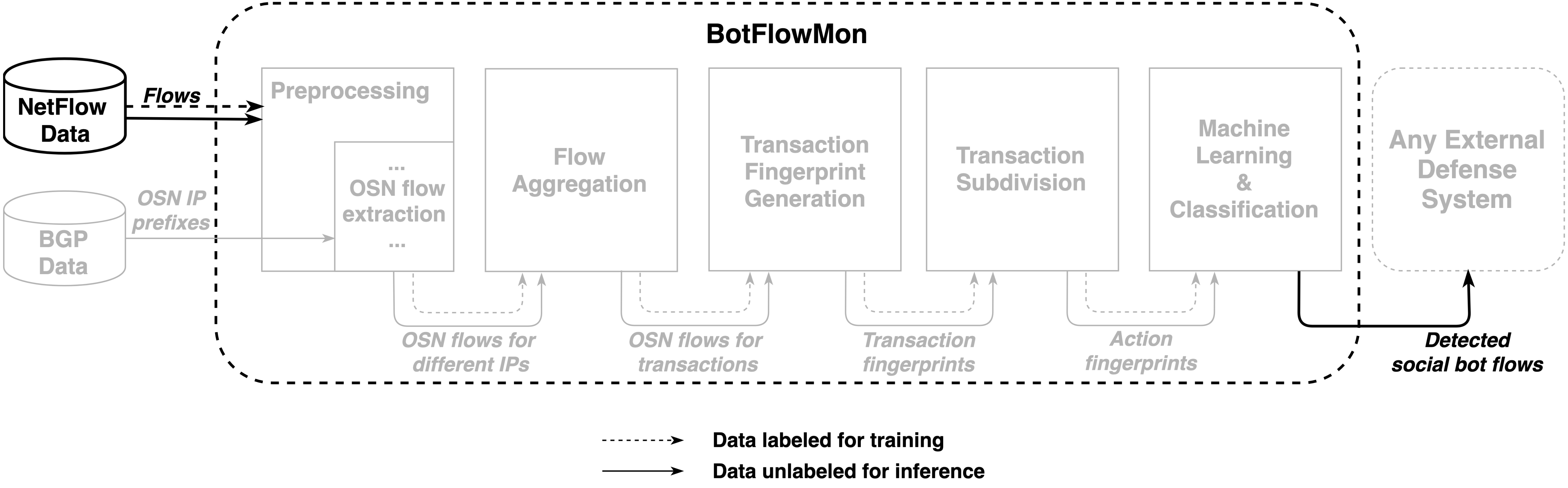
BotFlowMon Architecture

- **BotFlowMon has two modes:**
  - training mode: which uses labeled NetFlow data to derive a classification model.
  - detection mode: which uses the classification model to detect social bot flows from the input traffic flows.
- **With five modules:**
  - Preprocessing
  - Flow aggregation
  - Transaction fingerprint generation
  - Transaction subdivision
  - Machine learning & classification

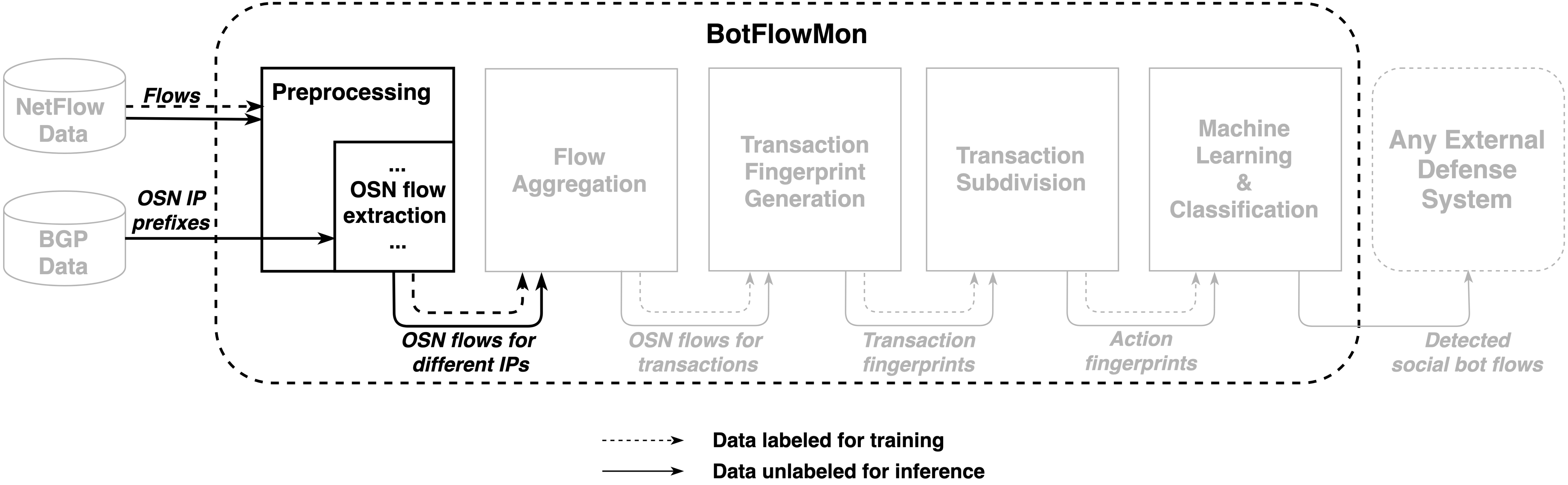
BotFlowMon Architecture


BotFlowMon
NetFlow Data
Flows
BGP Data
OSN IP prefixes
Preprocessing
...
OSN flow extraction
...
Flow Aggregation
Transaction Fingerprint Generation
Transaction Subdivision
Machine Learning & Classification
Any External Defense System
OSN flows for different IPs
OSN flows for transactions
Transaction fingerprints
Action fingerprints
Detected social bot flows
Data labeled for training
Data unlabeled for inference

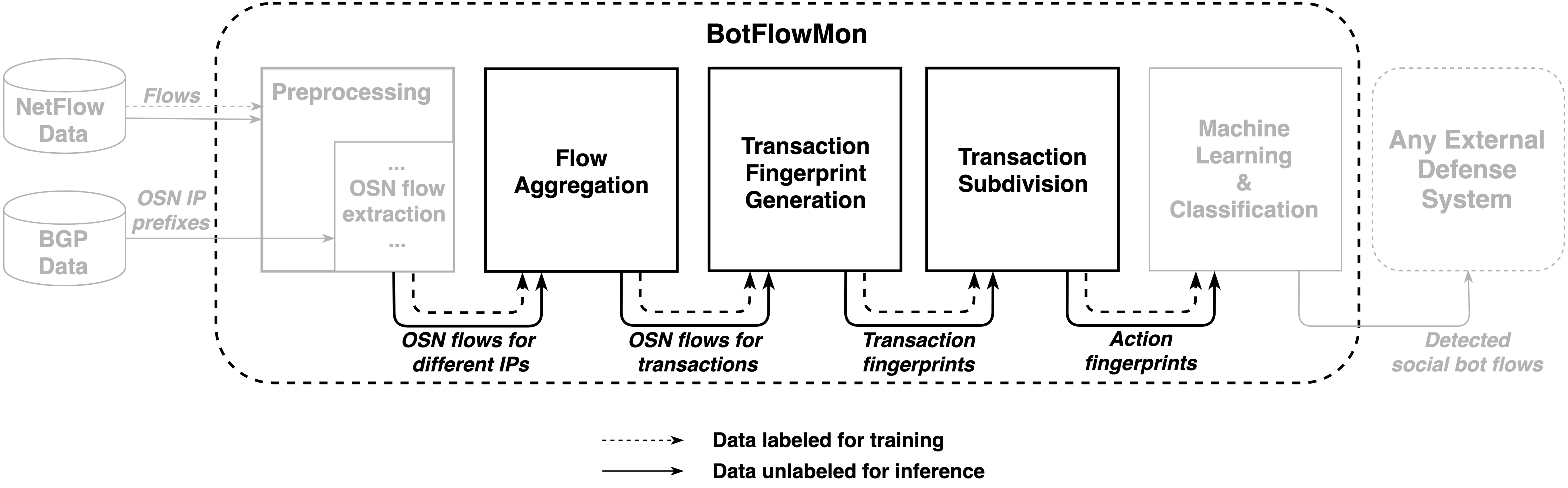
BotFlowMon Architecture


BotFlowMon
NetFlow Data
Flows
BGP Data
OSN IP prefixes
Preprocessing
...
OSN flow extraction
...
Flow Aggregation
Transaction Fingerprint Generation
Transaction Subdivision
Machine Learning & Classification
Any External Defense System
OSN flows for different IPs
OSN flows for transactions
Transaction fingerprints
Action fingerprints
Detected social bot flows
Data labeled for training
Data unlabeled for inference

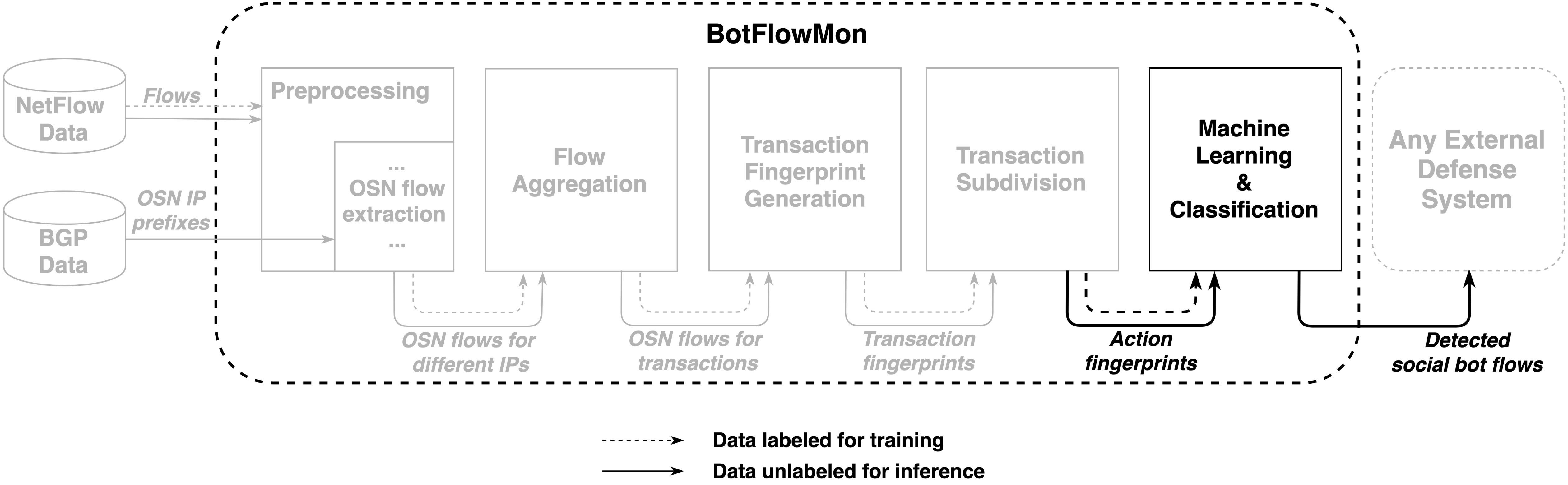
BotFlowMon Architecture


BotFlowMon
NetFlow Data
BGP Data
Flows
OSN IP prefixes
Preprocessing
...
OSN flow extraction
...
Flow Aggregation
Transaction Fingerprint Generation
Transaction Subdivision
Machine Learning & Classification
Any External Defense System
OSN flows for different IPs
OSN flows for transactions
Transaction fingerprints
Action fingerprints
Detected social bot flows
Data labeled for training
Data unlabeled for inference

BotFlowMon Architecture


BotFlowMon
NetFlow Data
Flows
BGP Data
OSN IP prefixes
Preprocessing
...
OSN flow extraction
...
Flow Aggregation
Transaction Fingerprint Generation
Transaction Subdivision
Machine Learning & Classification
Any External Defense System
OSN flows for different IPs
OSN flows for transactions
Transaction fingerprints
Action fingerprints
Detected social bot flows
Data labeled for training
Data unlabeled for inference

BotFlowMon Architecture


BotFlowMon
NetFlow Data
Flows
BGP Data
OSN IP prefixes
Preprocessing
...
OSN flow extraction
...
Flow Aggregation
Transaction Fingerprint Generation
Transaction Subdivision
Machine Learning & Classification
Any External Defense System
OSN flows for different IPs
OSN flows for transactions
Transaction fingerprints
Action fingerprints
Detected social bot flows
Data labeled for training
Data unlabeled for inference

NetFlow Data Format

- The information to leverage from NetFlow data is simple and straightforward.
- Content-agnostic, highly summarized information from packet headers.

Tag	Description	Tag	Description
%ts	Start time – first seen	%das	Destination AS number
%te	End time – last seen	%in	Input interface number
%td	Duration	%out	Output interface number
%pr	Protocol	%pkt	Number of packets
%sa	Source address	%byt	Number of bytes
%da	Destination address	%fl	Number of flows
%sap	Source address port	%flg	TCP flag
%dap	Destination address port	%tos	Type of service
%sp	Source port	%bps	Bits per second
%dp	Destination port	%pps	Packets per second
%sas	Source AS number	%bpp	Bytes per packet

Transactions & Actions

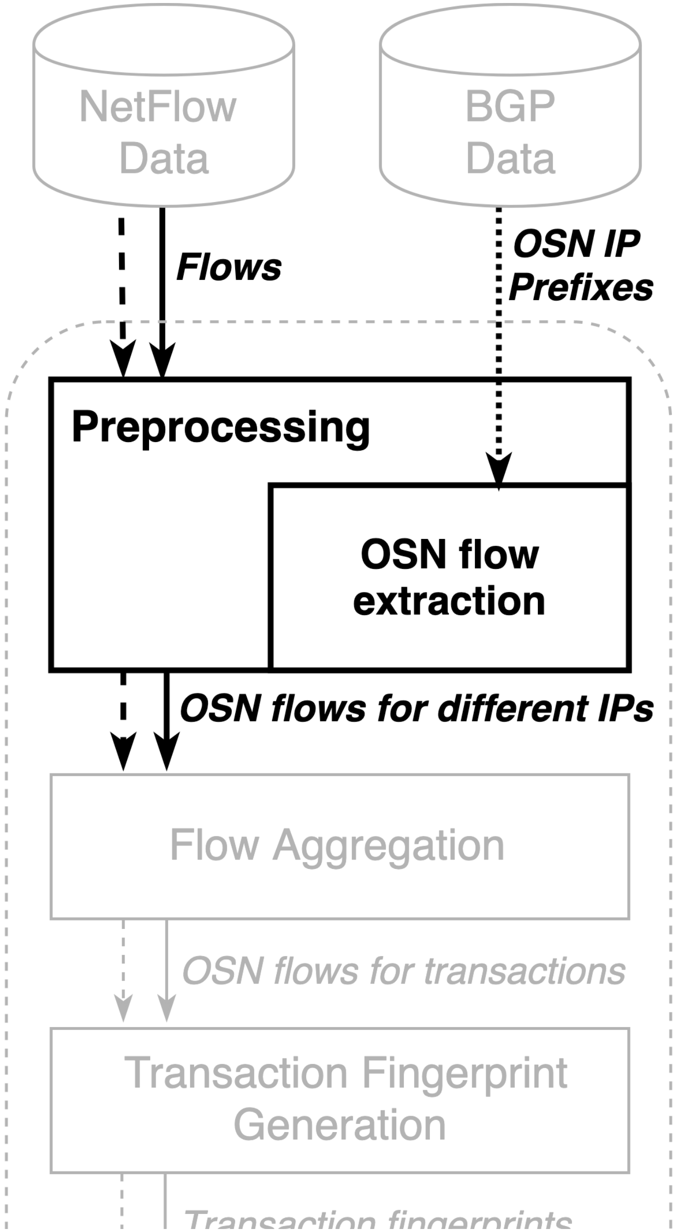
- BotFlowMon introduces two key concepts to study OSN traffic flows:
  - **transactions**
  - **actions**
- BotFlowMon aggregate flows into transactions
- Every transaction is composed of actions
- It can classify actions into bot actions and real user actions, and then classify the transactions based on how their actions are classified.

Outline

- Related work
- BotFlowMon Overview
- **Five Modules of BotFlowMon**
- Evaluations
- Conclusions

Preprocessing

- Raw NetFlow data is noisy and messy.
- This module:
  - denoise traffic flows,
  - extract OSN related traffic flows,
  - filter out flows with irrelevant protocols,
  - group flows by IP addresses, and
  - sort flows by timestamps.



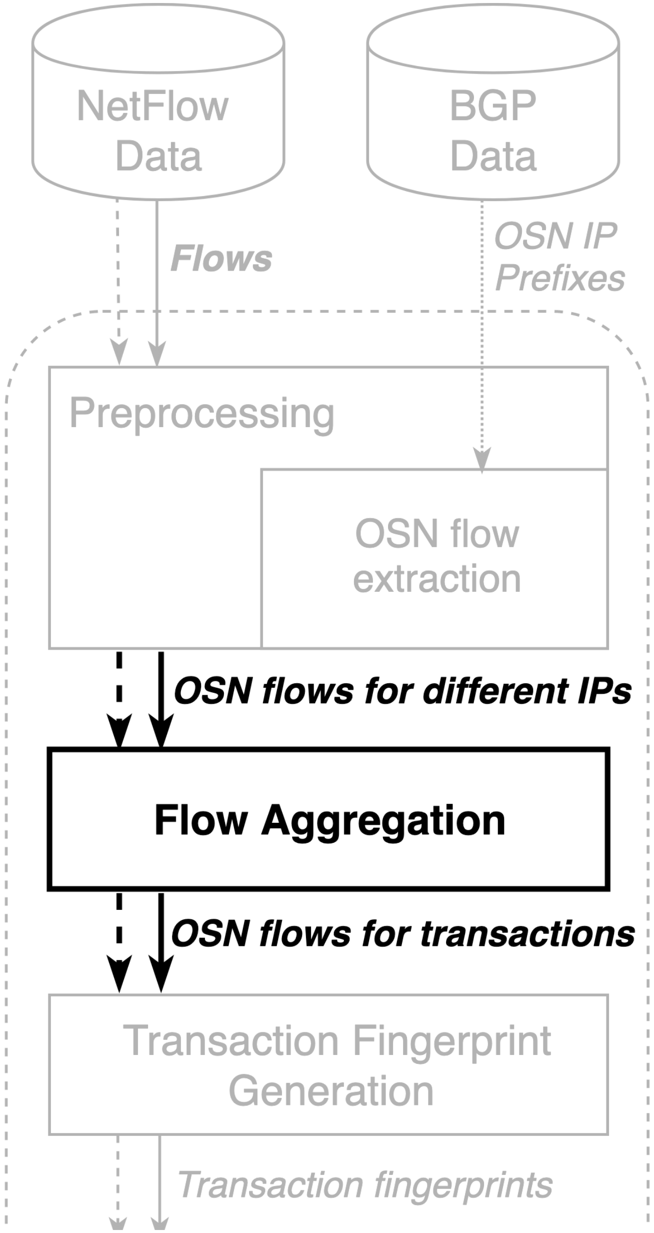
Flow Aggregation

Problem:

- No sufficient information from individual NetFlow records to detect social bot flows
- Both social bot and real OSN user behaviors are conducted at the application level.

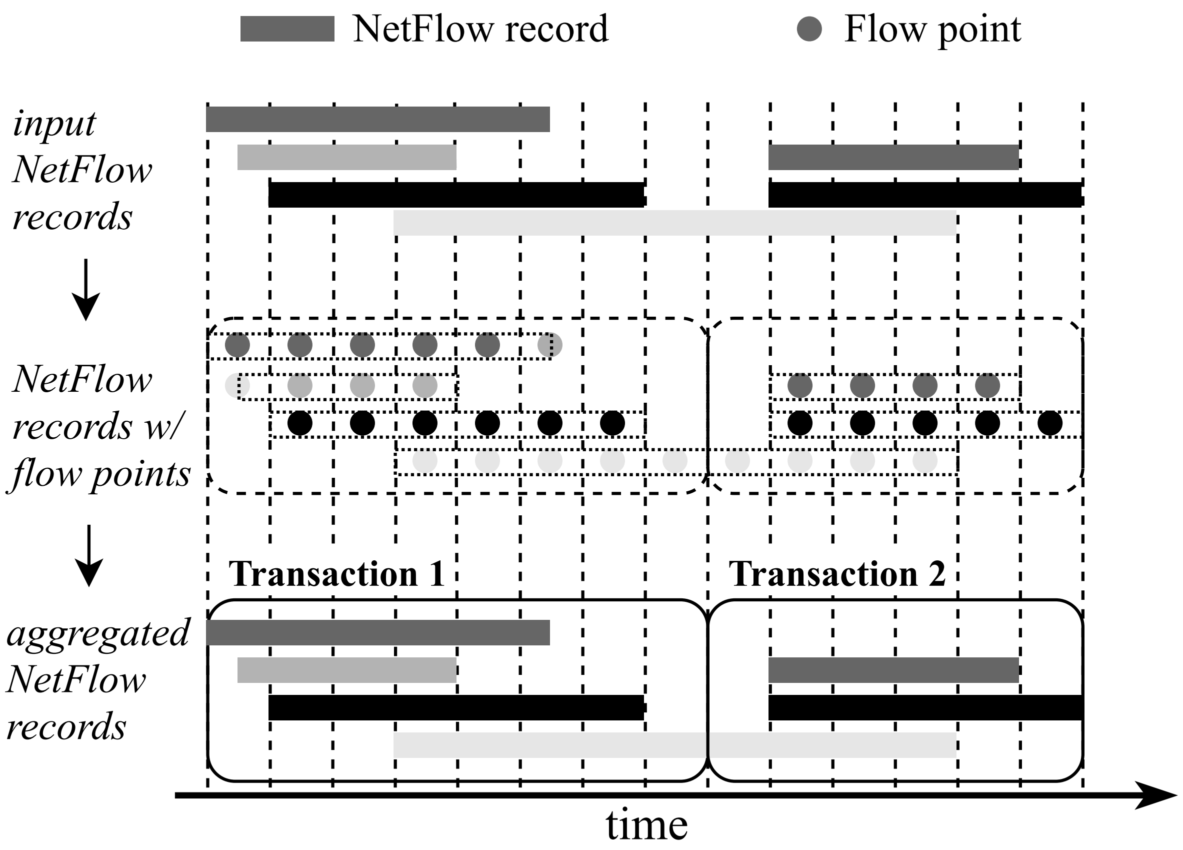
Aggregate adjacent flows into a transaction:

- More representative of application activities
- with more information to inspect collective OSN behaviors of flows



Flow Aggregation

1. Divide each flow’s duration into time bins of equal length, and define a flow point for each time bin.
2. Use modified DBSCAN to group flow points into clusters.
3. Inspect the time window of each cluster. The flows that fall within this window will belong to this transaction.



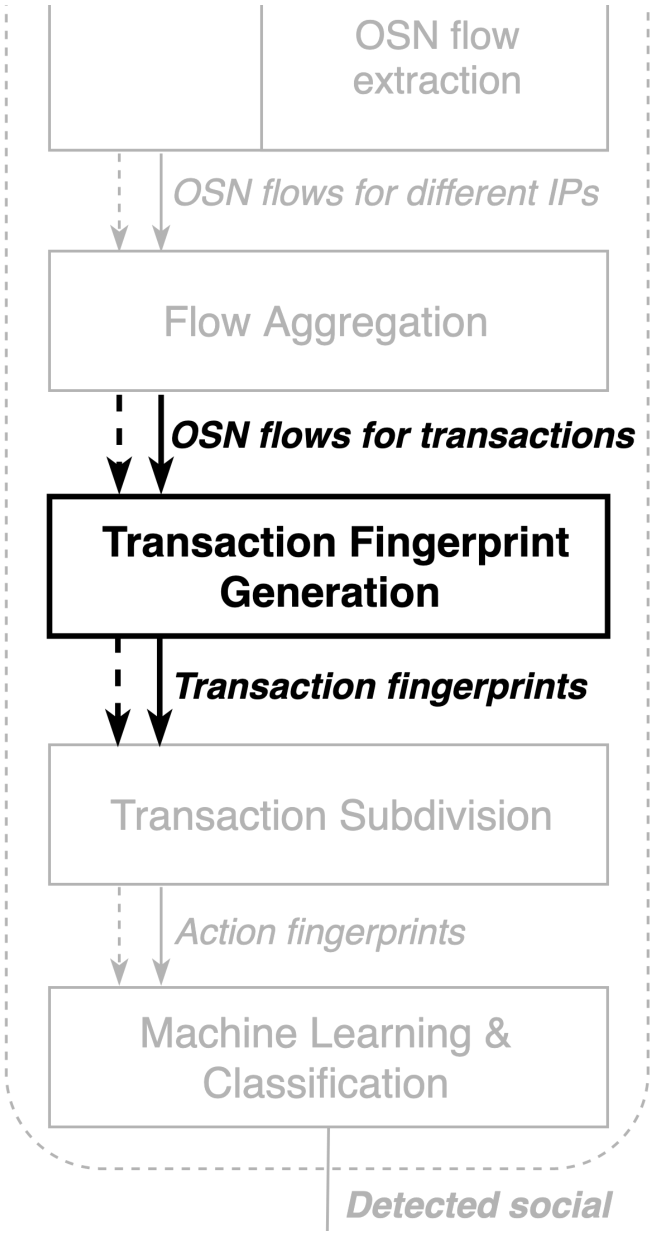
Transaction Fingerprint Generation

Problem:

- A transaction is a set of NetFlow records.
- Not regularized or normalized for comparison

Transaction fingerprint – a data fusion technique

- Derives an $f \times N$ matrix from every transaction
- Use this matrix as the fingerprint of the transaction
- Directly comparable and easy to visualize

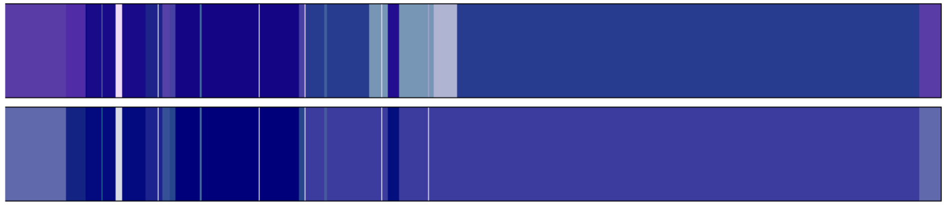

OSN flow extraction
OSN flows for different IPs
Flow Aggregation
OSN flows for transactions
Transaction Fingerprint Generation
Transaction fingerprints
Transaction Subdivision
Action fingerprints
Machine Learning & Classification
Detected social

Transaction Fingerprint Generation

Transaction Fingerprint:

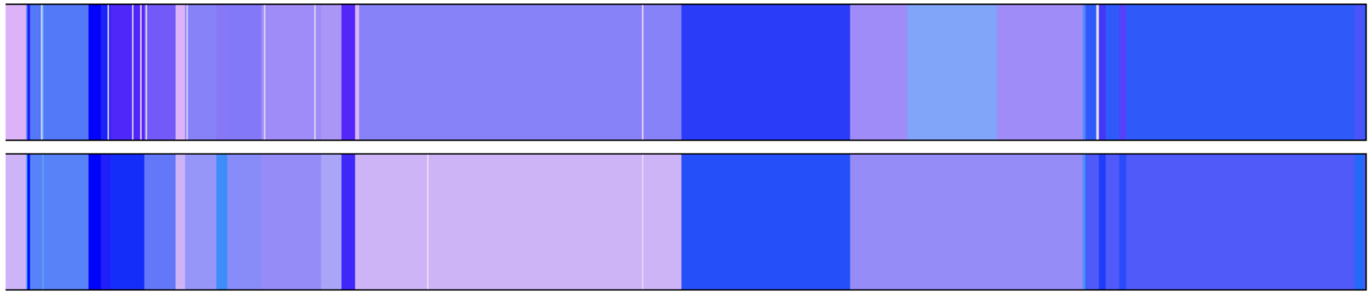
- $f \times N$ matrix
  - N is the number of time bins of equal length within the time window of the transaction
  - f is the number of features over each time bin
- E.g. 6 × 200
  - Incoming/outgoing bps, pps, ToS
- Visualizable

Features	Values			
1: outgoing bps	bps^o_{t1}	bps^o_{t2}	...	bps^o_{tN}
2: outgoing pps	pps^o_{t1}	pps^o_{t2}	...	pps^o_{tN}
3: outgoing ToS	tos^o_{t1}	tos^o_{t2}	...	tos^o_{tN}
4: incoming bps	bps^i_{t1}	bps^i_{i2}	...	bps^i_{tN}
5: incoming pps	pps^i_{t1}	pps^i_{i2}	...	pps^i_{tN}
6: incoming ToS	tos^i_{t1}	tos^i_{i2}	...	tos^i_{tN}

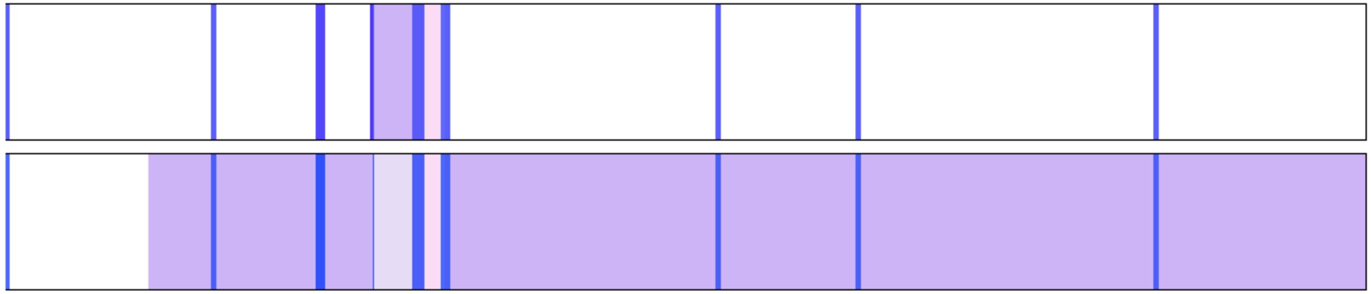


Transaction Fingerprint Example 1

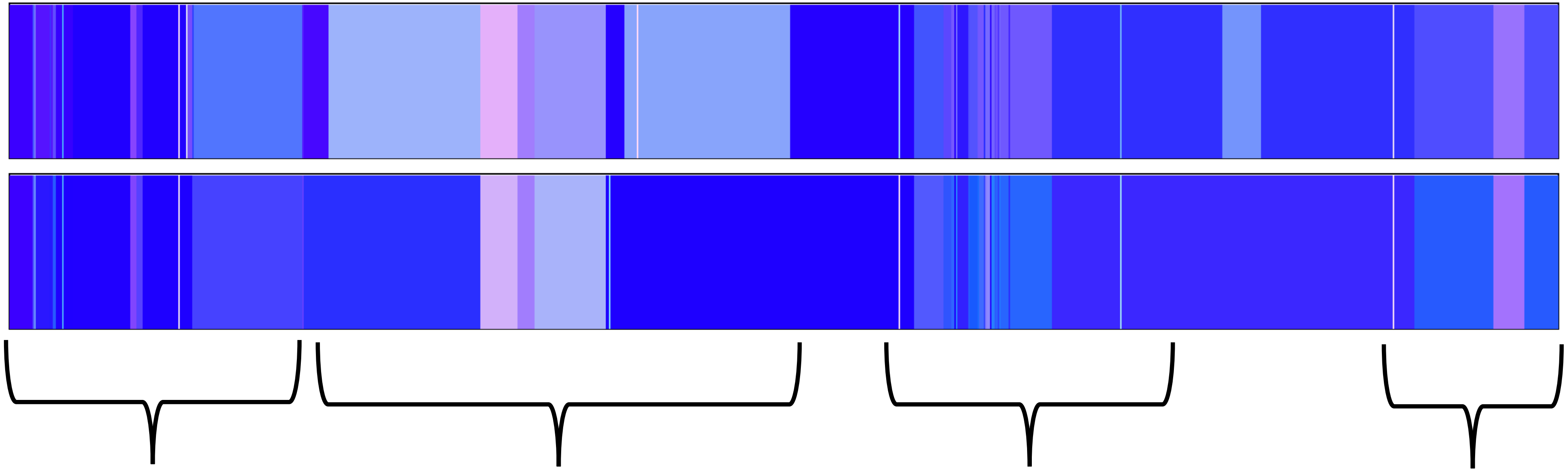
Human user



Chatbot



Transaction Fingerprint Example 2



Open the browser and load the page

Post 2 tweets that contain images

Reload the page

Post one tweet that contains image

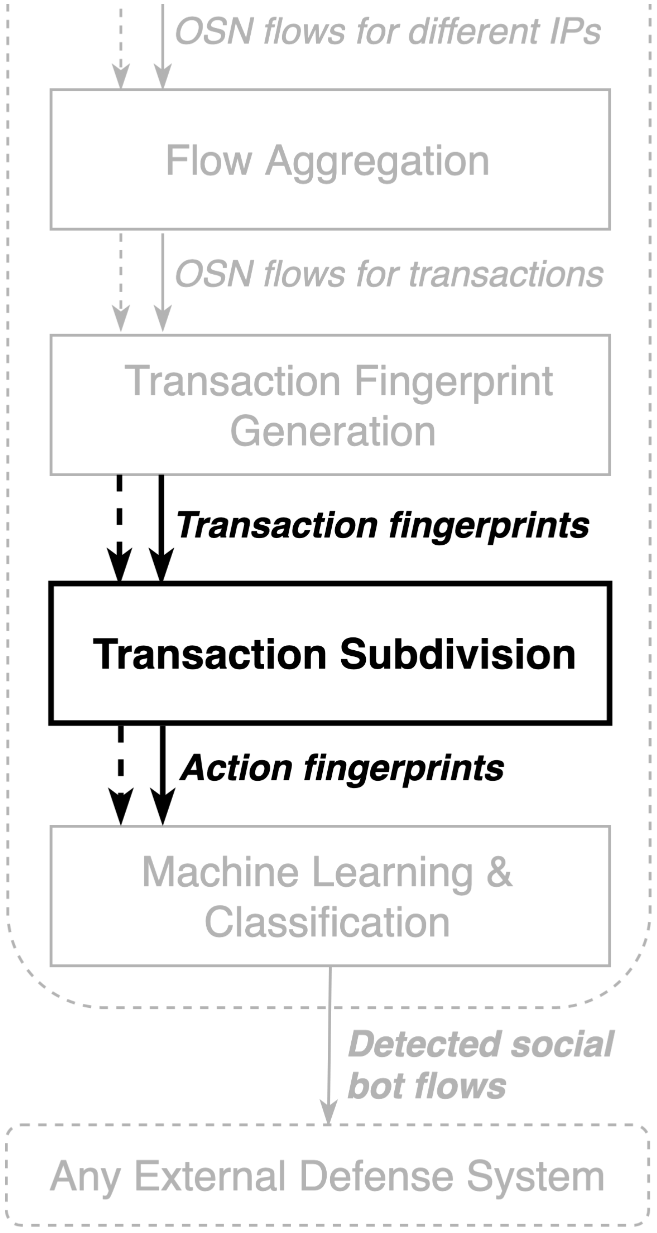
Transaction Subdivision

Problem:

- There can be countless types of transactions.
- Each transaction can be of an arbitrary duration.
- Size of training data is limited.

Subdivide a transaction into actions:

- Easier to differentiate bot actions from real user actions
- Reduce required training data size
- Increase training speed & detection accuracy

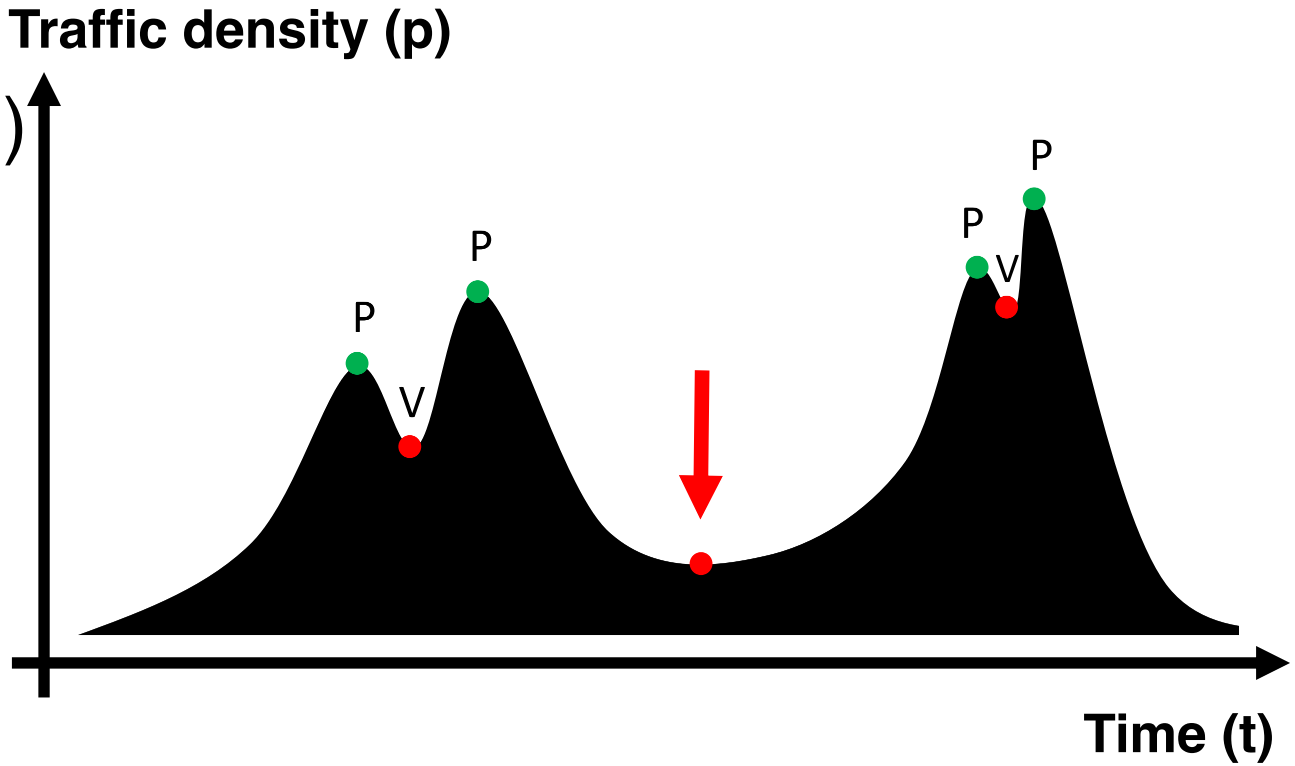

OSN flows for different IPs
Flow Aggregation
OSN flows for transactions
Transaction Fingerprint Generation
Transaction fingerprints
Transaction Subdivision
Action fingerprints
Machine Learning & Classification
Detected social bot flows
Any External Defense System

Transaction Subdivision

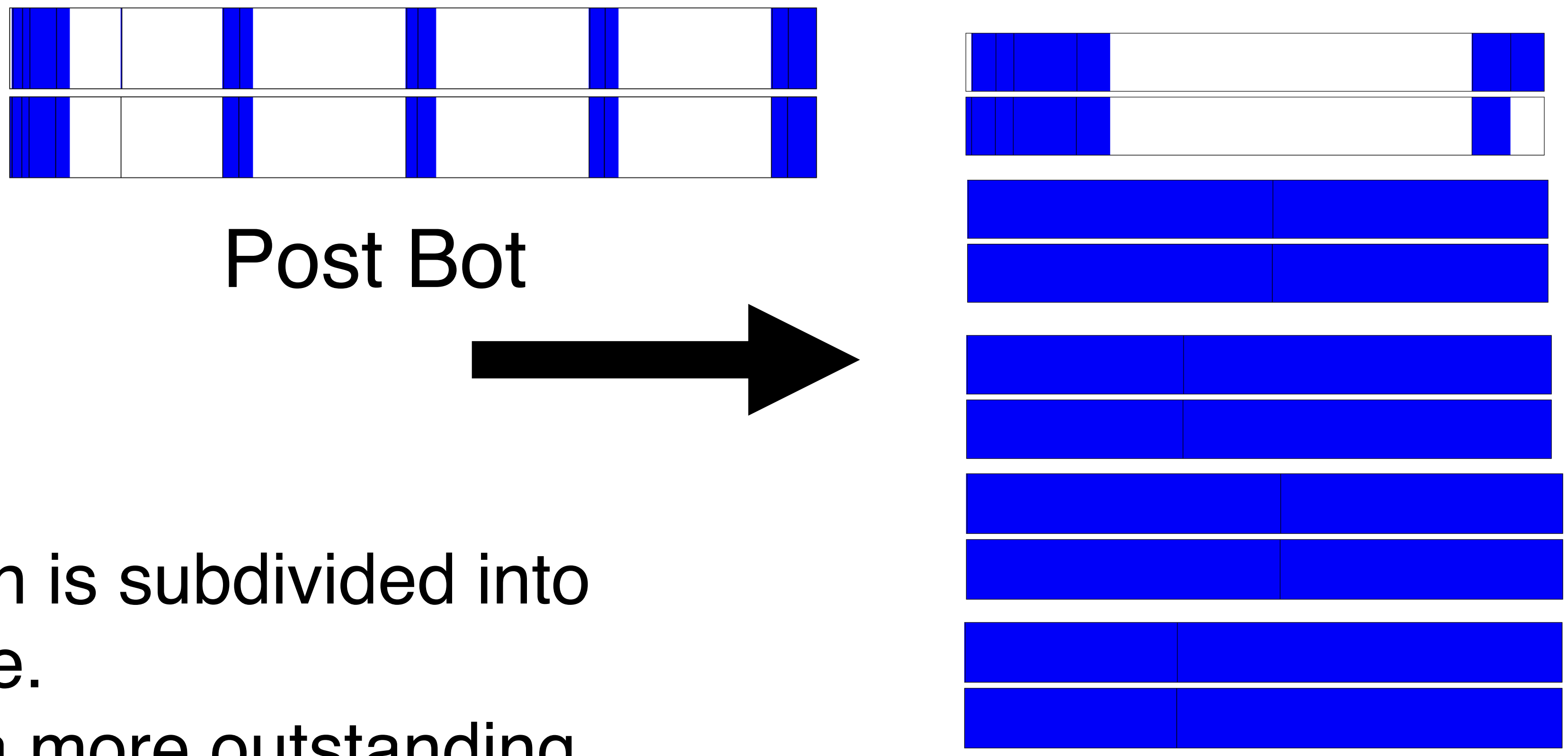
Subdivision Algorithm:

- A new density-valley-based clustering algorithm
  - Parameter r (duration threshold)
  - Find out all the density valley points
  - Choose the valley points with enough contrast with surroundings as subdivision moments of a transaction
- Output: a set of action fingerprints

$$Den(p) = \sum_f Byte_f / Duration$$

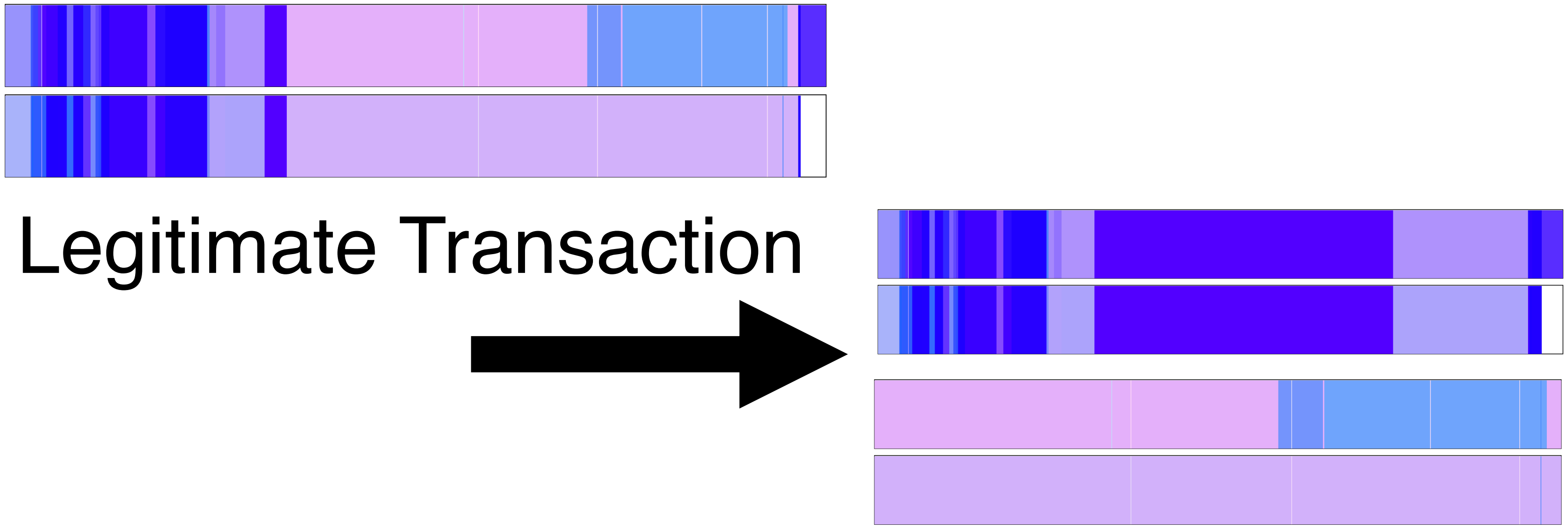


Transaction Subdivision Example 1



- A post bot’s transaction is subdivided into five actions in this case.
- Each action now has a more outstanding pattern than the original transaction fingerprint.

Transaction Subdivision Example 2



Legitimate Transaction

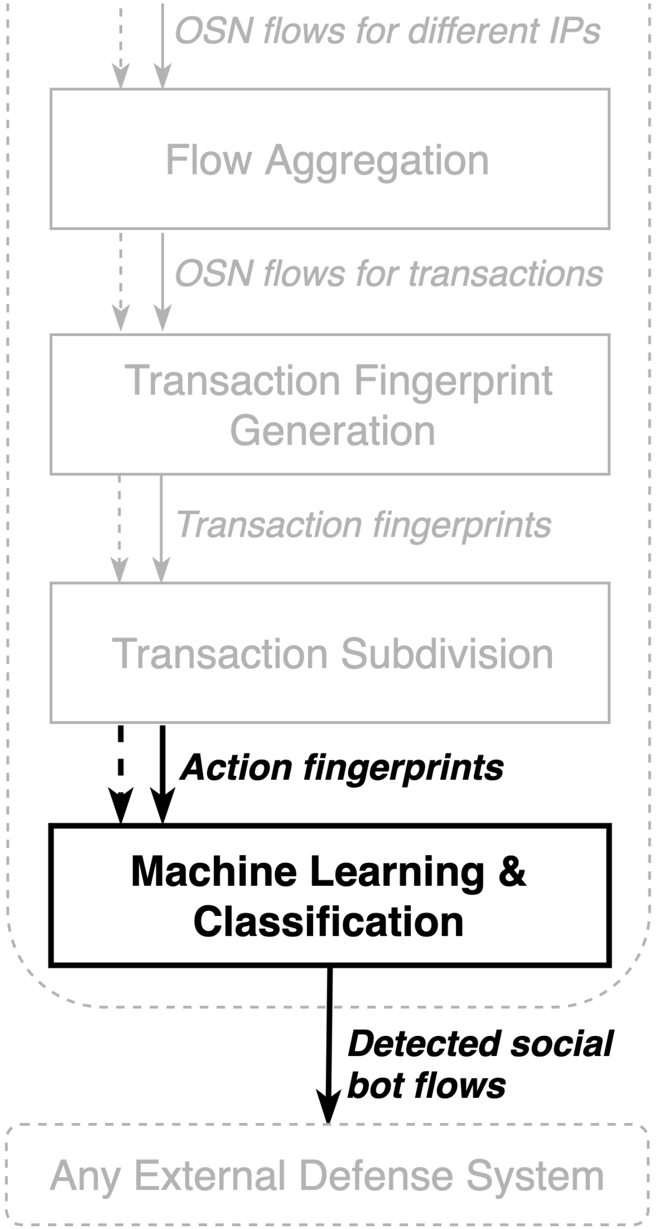
A transaction by a real user that is composed of two actions:

- one was opening an OSN site, and
- the other was scrolling down the page of the OSN site.

Machine Learning & Classification

BotFlowMon uses Multilayer Perceptron and Conventional Neural Network as its training approaches.

- Input: a set of action fingerprints.
- intermediary output: labeled action fingerprints.
- Then, use action fingerprints to vote for their transaction fingerprint's label.
- Final output: transaction fingerprints' labels.

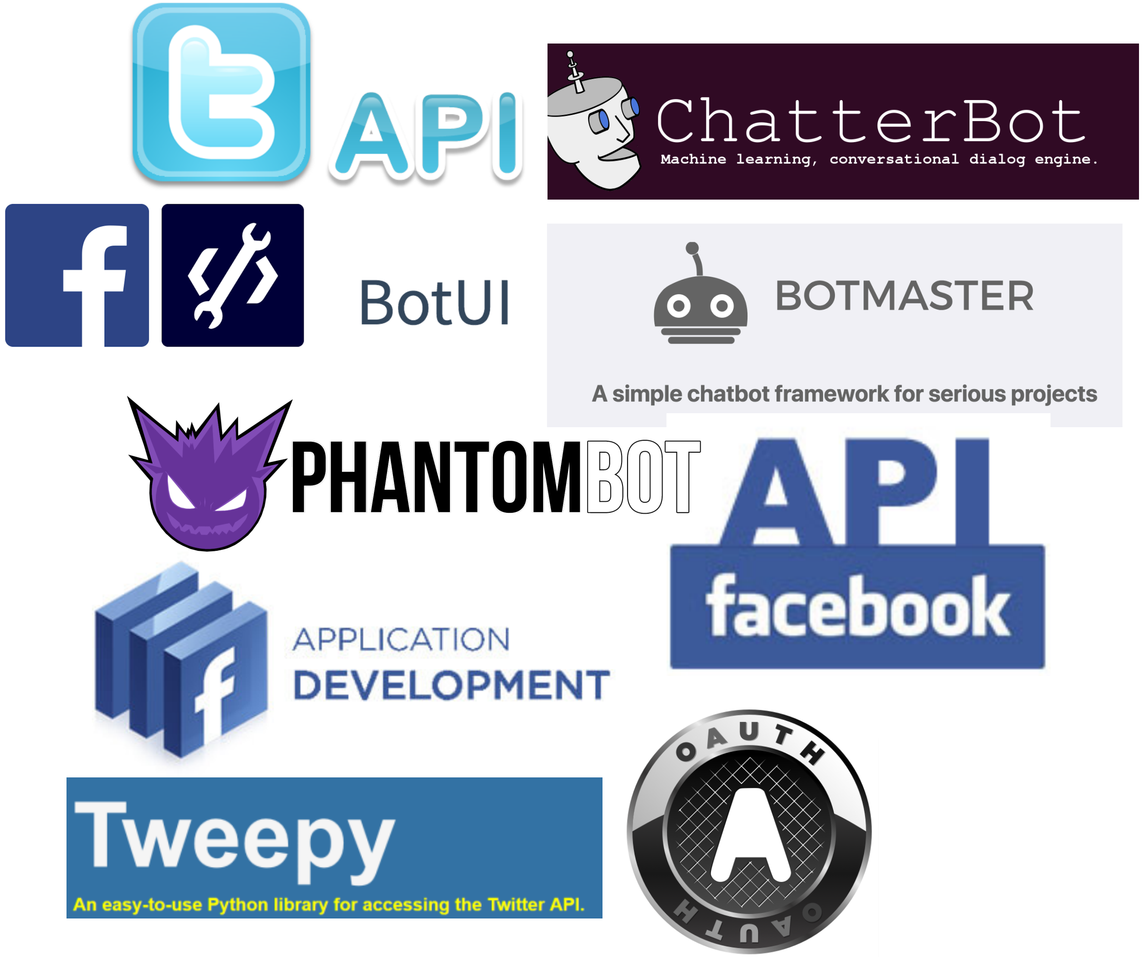

OSN flows for different IPs
Flow Aggregation
OSN flows for transactions
Transaction Fingerprint Generation
Transaction fingerprints
Transaction Subdivision
Action fingerprints
Machine Learning & Classification
Detected social bot flows
Any External Defense System

Outline

- Related work
- BotFlowMon Overview
- Five Modules of BotFlowMon
- **Evaluations**
- Conclusions

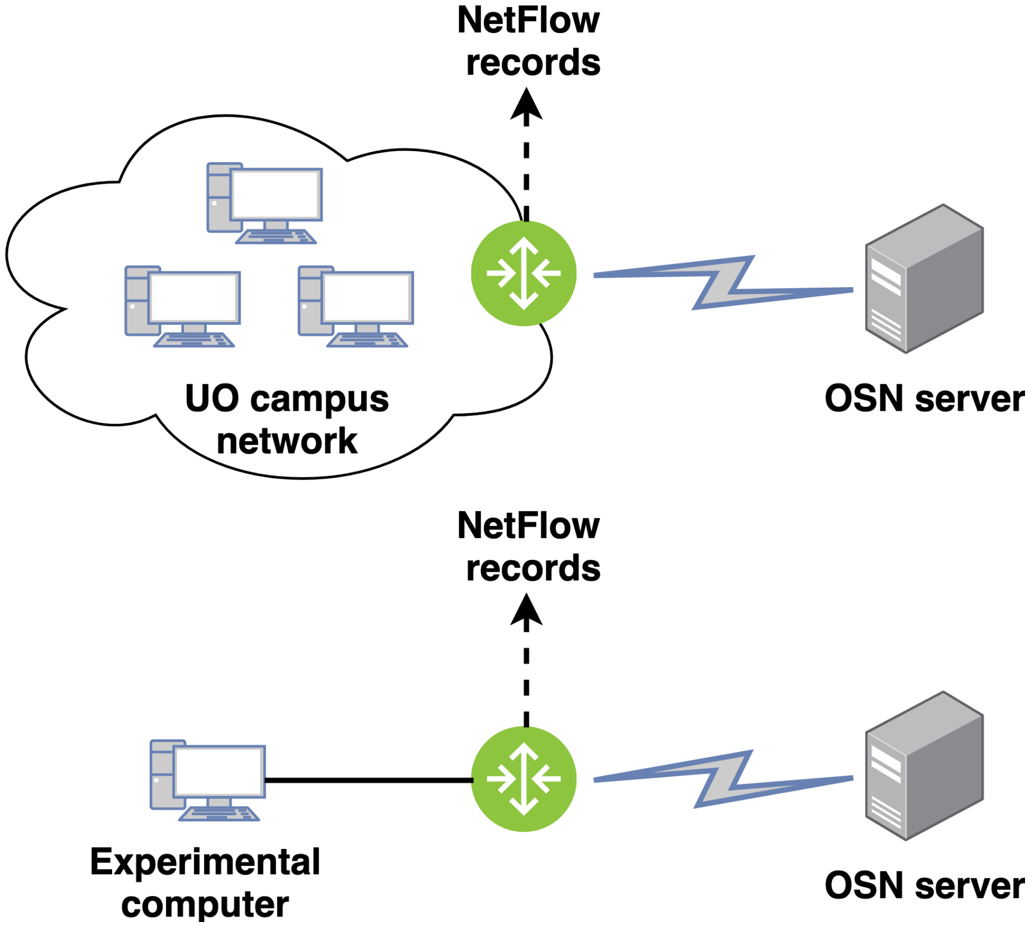
Bot Simulations

- Leverage existing open-sourced bot programs and frameworks
  - E.g., Botmaster, ChatterBot, PhantomBot
- Develop home-grown bot programs
- Five types of bots are simulated
  - post bot, chatbot, amplification bot, OSN crawler bot, hybrid bot


API
ChatterBot
Machine learning, conversational dialog engine.
BotUI
BOTMASTER
A simple chatbot framework for serious projects
PHANTOMBOT
API
facebook
APPLICATION
DEVELOPMENT
Tweepy
An easy-to-use Python library for accessing the Twitter API.
OAUTH

Data Collection

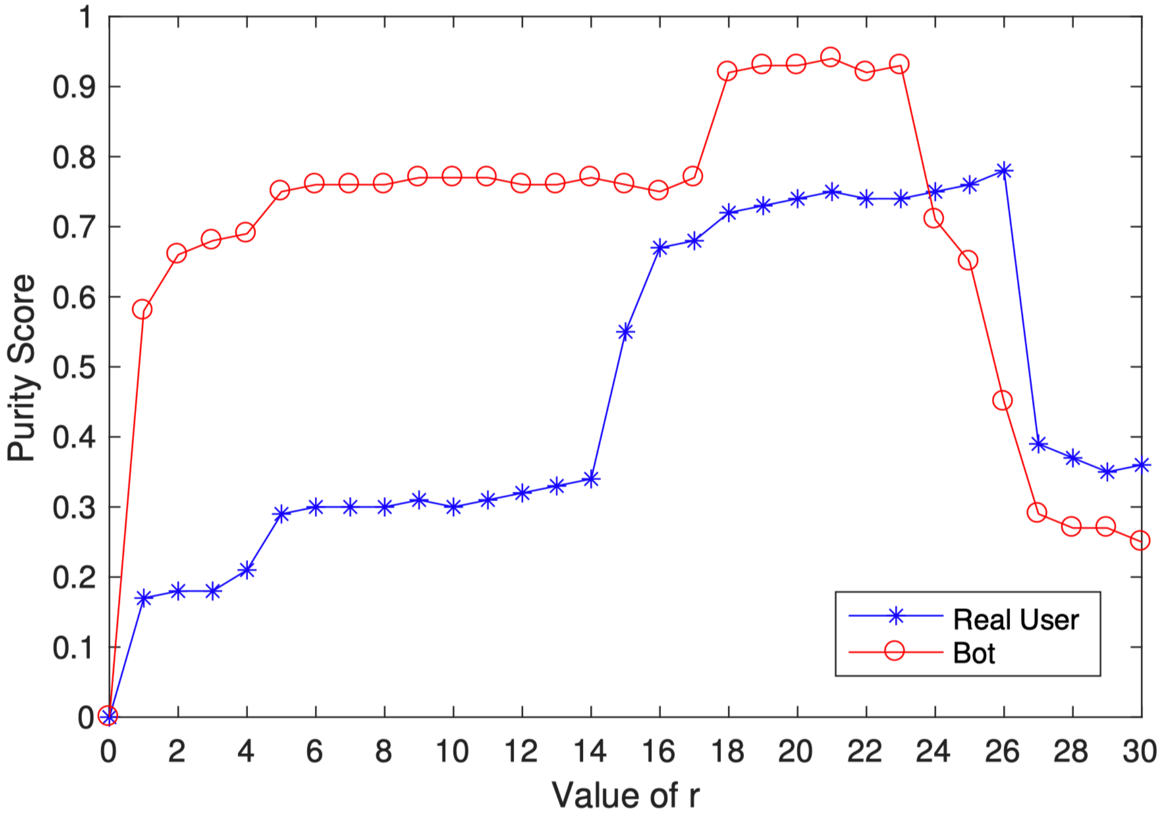
- NetFlow data collected from University of Oregon campus network
  - Data from realistic scenarios for analysis and verification
  - 507GBs are collected
- NetFlow data collected from experimental computers and routers
  - has superior flexibility and conveniences for simulation, data collection, and experiments
  - 28GBs are collected



Purity Score of Subdivision

- The subdivision algorithm’s purity scores with different r values.
- Optimal results: when r is in the range of 18 to 23.
- Doesn’t need to precisely partition all the transactions. It is designed to make data more friendly to the machine learning process.

Purity scores with different r values

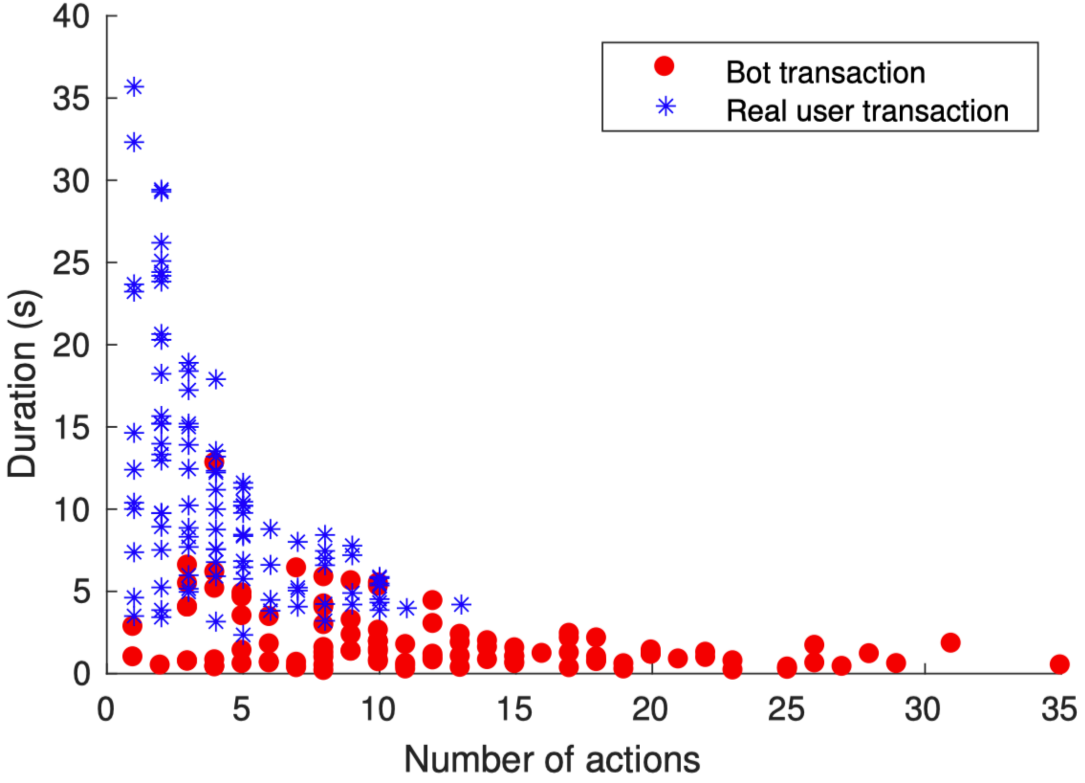


Transactions and Actions

1. Randomly select 100 bot and 100 real user transactions
2. Conduct subdivision
3. Record the number of resulted actions and average duration

- Bot transactions tend to have more actions
- Bot actions have shorter durations

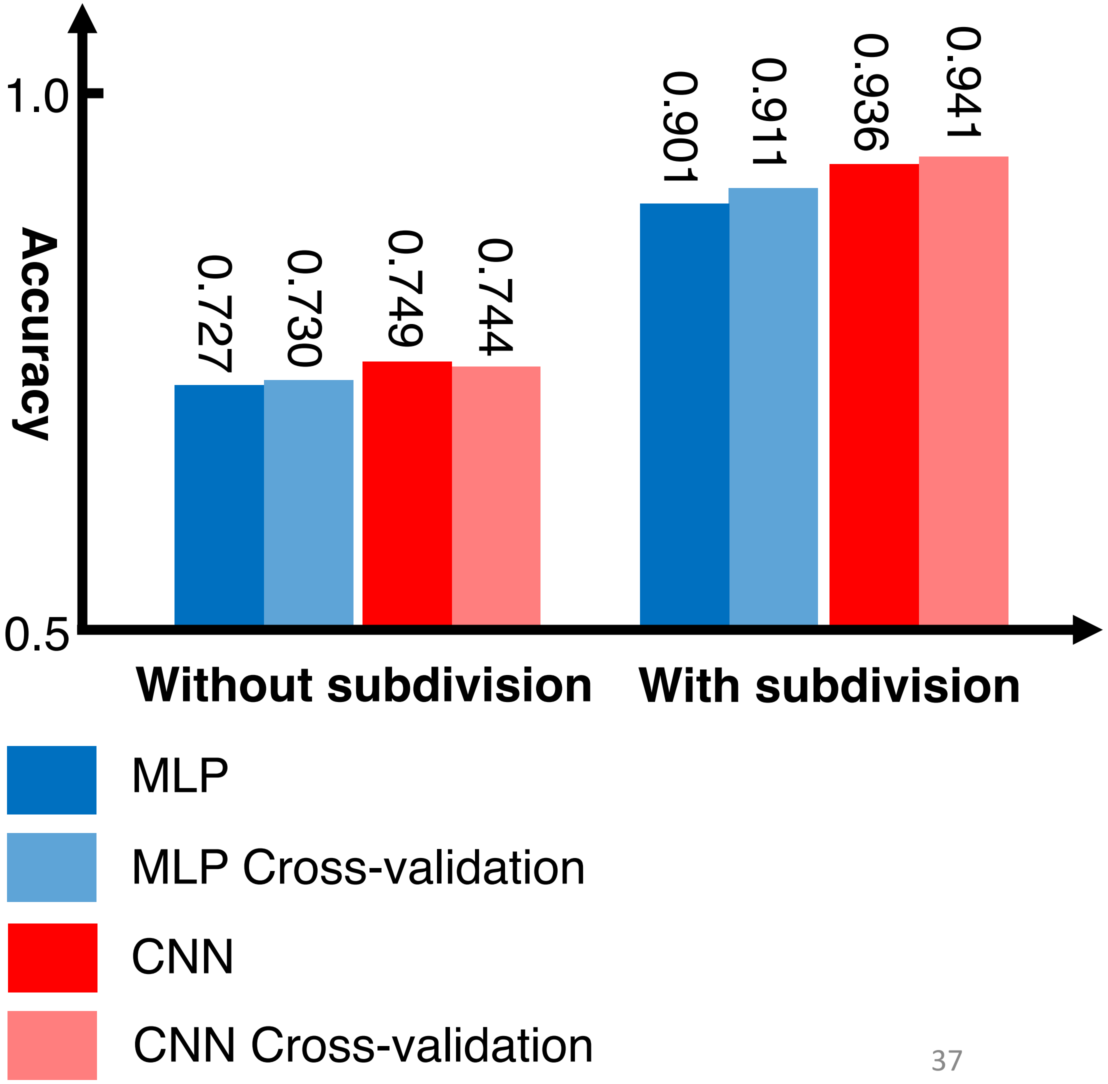
Scatter Diagram for Subdivision



Detection Accuracy

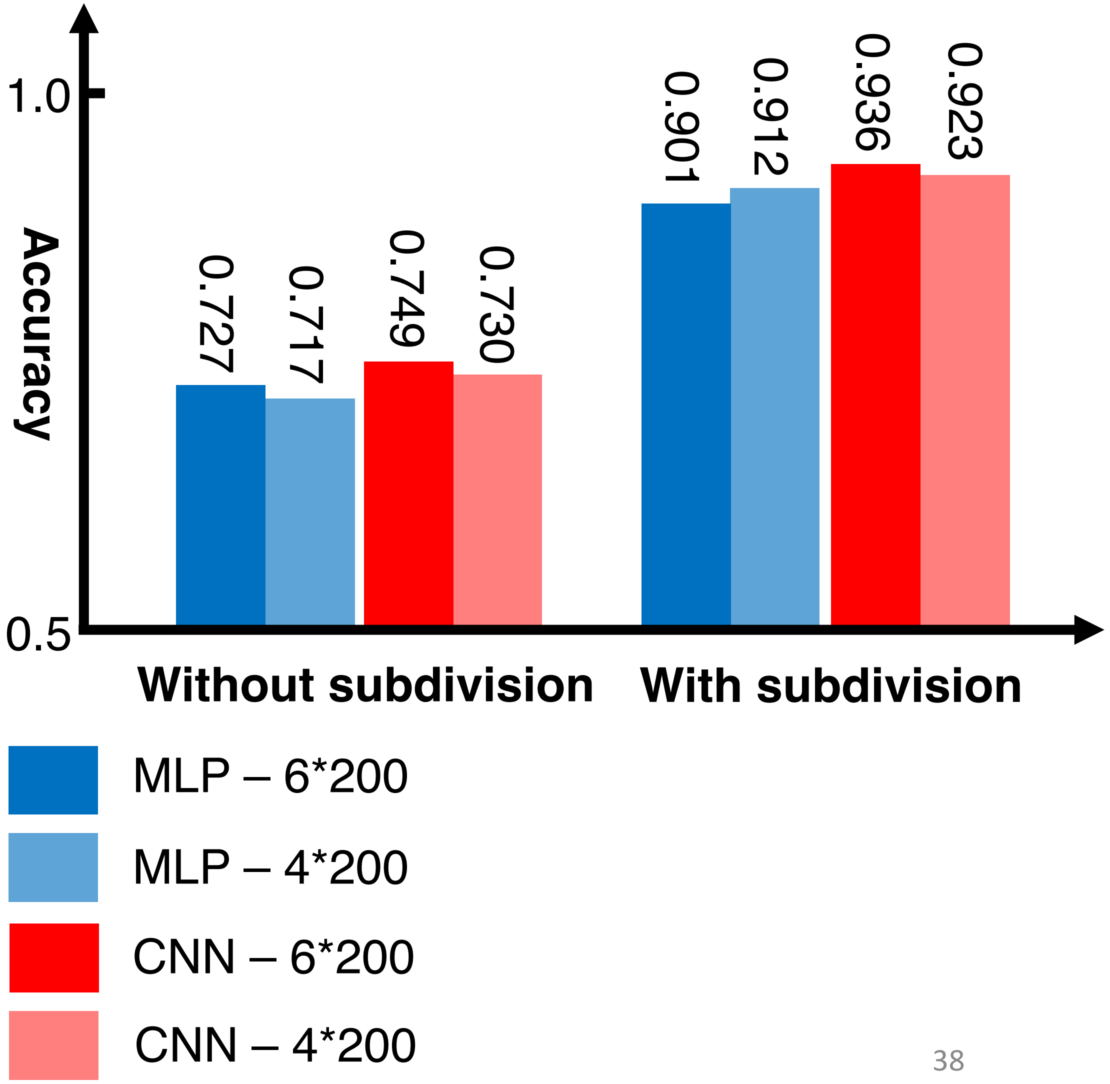
- **6 × 200 transaction fingerprint**
  - Incoming/outgoing bps, pps, ToS
- CNN Accuracy: 0.9361
  - Precision: 0.9887
  - Recall: 0.9067
  - F1 score: 0.9459

- CNN is slightly better than MLP
- The subdivision increases the accuracy significantly



Detection Accuracy

- Remove features from ToS field
  - ToS can be modified by third parties
  - Test the universality
- Accuracies almost remain the same
- Totally usable without ToS



Limitations

- Not all the social bots are malicious
    - The boundary between “good” bots and “bad” bots can be blurry.
    - Distinguishing social bots with malicious intentions from those that are innocent is hard to achieve without payload data.
- May not be able to detect zero-day social bots
    - Learning-based approach, fully depends on training dataset

Outline

- Related work
- BotFlowMon Overview
- Five Modules of BotFlowMon
- Evaluations
- **Conclusions**

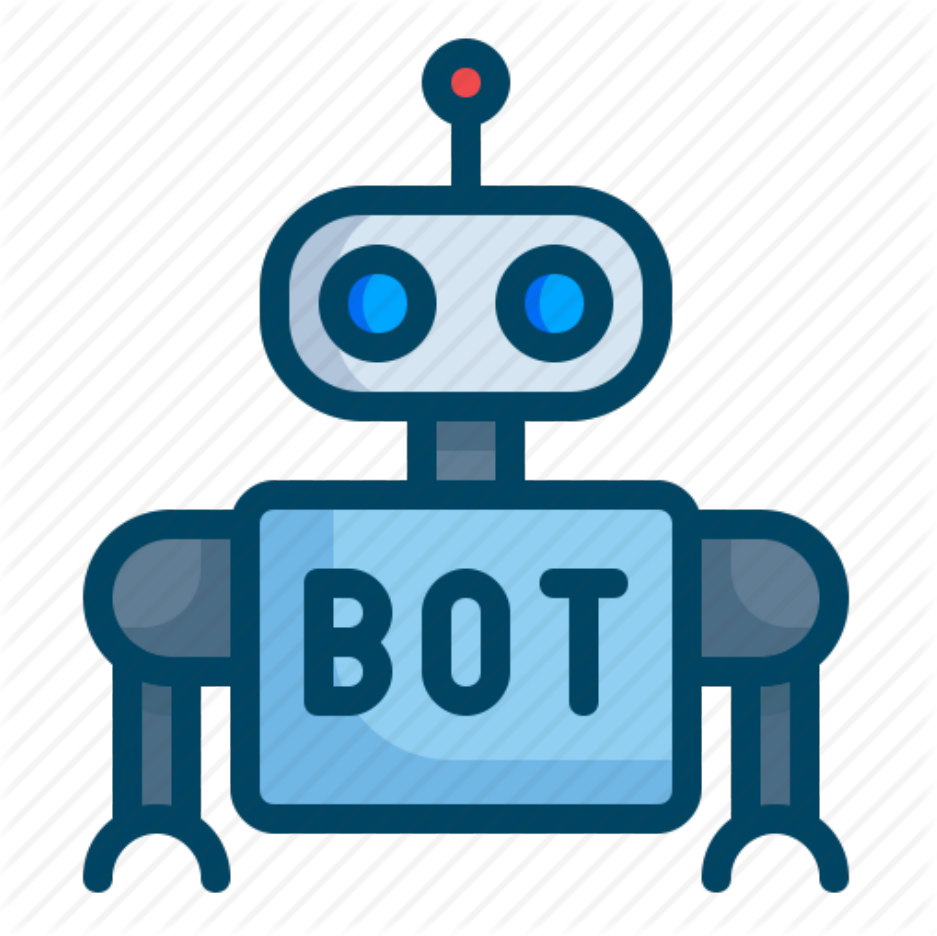
Conclusions

Social bots are becoming far more sophisticated and threatening than before.

Our contributions:

- BotFlowMon: flow-level social bot traffic identification
  - tackles big networking data to identify the traffic of OSN bots
  - content-agnostic, privacy-preserving and efficient
  - Easy to deploy by both OSN providers and ISPs
- Several new techniques and algorithms
  - an aggregation technique that derives transactions
  - a data fusion technique that extracts features from transactions and actions
  - a density-valley-based clustering algorithm

THANK YOU!



This material is based upon work supported by the National Science Foundation under Grant No.1564348. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.