No Please, After You: Detecting Fraud in Affiliate Marketing Networks

Peter Snyder <psnyde2@uic.edu> and Chris Kanich <ckanich@uic.edu>
University of Illinois at Chicago

Overview

1. **Problem Area:** affiliate marketing

2. **Data Set**: HTTP request records

3. **Methodology**: classification algorithm

4. **Findings**: numbers and stakeholder analysis

1. Problem Area

Affiliate Marketing

Online Retailers

Publishers

Web Users

Affiliate Marketing

- Common method for funding "free" content

- Largest programs include Amazon, GoDaddy, eBay and WalMart

- Both direct programs and networks / middle-parties

amazon Prime

All ▾ [search] 🔍

🔧 Father's Day Savings
Sponsored by DEWALT

Shop by **Department** ▾   Peter's Amazon.com   Today's Deals   Gift Cards   Sell   Help

Hello, Peter **Your Account** ▾   Your **Prime** ▾   Wish **List** ▾   🛒 0 **Cart** ▾

Amazon Fashion

Free One-Day Shipping
ON GIFTS FOR DAD

● ○ ○ ○ ○

Get Yourself a Little Something See more

OLIVER SACKS
The MAN Who MISTOOK

MELT-BANANA

×××××× ✕

PEOPLE WHO EAT DARKNESS

www.amazon.com/b/ref=br_imp_ara-1?_encoding=UTF8&field-enc-merchantbin=ATVPDKIKX0DER&node=11525499011&pf_rd_m=...&pf_rd_s=desktop-hero-softlines&pf_rd_r=045STG4DMW1ZNV81E8FM&pf_rd_t=36701&pf_rd_p=2100469162&pf_rd_i=desktop

HOMEPAGE > APPLIANCES, KITCHEN

The Best Food Processor

LAST UPDATED: JUNE 15, 2015   CHRISTINE CYR CLISSET

Tweet   Email

thesweethome.com/link/?l=B0000645TW&subtag=SH1777

thesweethome-20

`thesweethome-20`

Link to this page   Short Link To This Page   Add to Widget   Add to aStore   Share   Your Earnings Summary   What's New   Discussion Boards   amazon associates

Settings

amazon Prime

Home & Kitchen

Shop by Department

Peter's Amazon.com   Today's Deals   Gift Cards   Sell   Help

Hello, Peter
Your Account

Your Prime

Wish List

1 Cart

Father's Day Savings
Sponsored by DEWALT

Kitchen & Dining   Best Sellers   Wedding Registry   Small Appliances   Kitchen Tools   Cookware   Bakeware   Cutlery   Dining & Entertaining

Home & Kitchen › Kitchen & Dining › Small Appliances › Food Processors

Roll over image to zoom in

Cuisinart DFP-14BCN 14-Cup Food Processor, Brushed Stainless Steel

by Cuisinart

933 customer reviews

| 173 answered questions

List Price: $325.00
Price: **$141.75** Prime
You Save: **$183.25 (56%)**

In Stock.
Ships from and sold by Amazon.com in easy-to-open packaging. Gift-wrap available.

Want it tomorrow, June 16? Order within **1 hr 33 mins** and choose **One-Day Shipping** at checkout. Details

Color: Brushed Stainless

- 720-watt large-capacity food processor with intuitive "on"

Share

Qty: 1

Add to Cart

Turn on 1-Click ordering for this browser

Ship to:
IL 60290

Add to Wish List

Add to Wedding Registry

www.amazon.com

thesweethome.com/reviews/the-best-food-processor/

onId=AKIAJM4NKIC...GABP2PIRA&tag=thesweethome-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=B0000645TW&ascsubtag=SH1777

Link to this page   Short Link To This Page   Add to Widget   Add to aStore   Share   Your Earnings Summary   What's New   Discussion Boards   amazon associates

Settings

amazon Prime

Home & Kitchen

Father's Day Savings
Sponsored by DEWALT

Shop by
Department   Peter's Amazon.com   Today's Deals   Gift Cards   Sell   Help   Hello, Peter
Your Account   Your
Prime   Wish
List   1 Cart

Kitchen & Dining   Best Sellers   Wedding Registry   Small Appliances   Kitchen Tools   Cookware   Bakeware   Cutlery   Dining & Entertaining

Home & Kitchen › Kitchen & Dining › Small Appliances › Food Processors

HOMEP

The

LAS

Tweet

thesweethome.c

Roll over image to zoom in

Cuisinart DFP-14BCN 14-Cup Food Processor, Brushed Stainless Steel

by Cuisinart

★★★★½   933 customer reviews

| 173 answered questions

List Price: ~~$325.00~~
Price: **$141.75** ✓Prime
You Save: **$183.25 (56%)**

In Stock.

Ships from and sold by Amazon.com in easy-to-open
packaging. Gift-wrap available.

Want it tomorrow, June 16? Order within **1 hr 33 mins** and
choose **One-Day Shipping** at checkout. Details

Color: **Brushed Stainless**

- 720-watt large-capacity food processor with intuitive "on"

Share ✉ f ♥ 📌

Qty: 1

Add to Cart

Turn on 1-Click ordering for this browser

Ship to:
IL 60290

Add to Wish List

Add to Wedding Registry

www.amazon.com

The Best Food Processor | ...

thesweethome.com/reviews/the-best-food-processor/

Amazon.com: Cuisinart DF...

onId=AKIAJM4NKIQGABP2PIRA&tag=thesweethome-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=B0000645TW&ascsubtag=SH1777

Amazon.com Checkout

https://www.amazon.com/gp/buy/spc/handlers/display.html?hasWorkingJavascript=1

☐ Check this box to default to these delivery and payment options in the future.

Place your order

By placing your order, you agree to Amazon.com's privacy notice and conditions of use.

Get a **$1 credit for digital books, music, movies and TV shows**
Choose FREE No-Rush Shipping and receive a $1 credit towards your next eBook, digital music or Amazon Instant Video purchase or rental when this order ships.

Order Summary

Items:	$141.75
Shipping & handling:	$0.00
Total before tax:	$141.75
Estimated tax to be collected:	$8.86

Order total: **$150.61**

How are shipping costs calculated?

Prime shipping benefits have been applied to your order.

Guaranteed delivery date: June 17, 2015 If you order in the next 2 hours (Details)
Items shipped from Amazon.com

Cuisinart DFP-14BCN 14-Cup Food Processor, Brushed Stainless Steel
$141.75 ✓ *Prime*
Quantity: 1 Change
Sold by: Amazon.com LLC

🎁 Add a gift receipt

and see other gift options

Choose your Prime delivery option:

○ FREE No-Rush Shipping — get it Monday, June 22
Get a $1 credit for eBooks, digital music and videos. Details

○ FREE Standard Shipping — get it Friday, June 19

● FREE Two-Day Shipping — get it Wednesday, June 17

○ $11.99 One-Day Shipping — get it tomorrow, June 16

Place your order   **Order total: $150.61**

By placing your order, you agree to Amazon.com's privacy notice and conditions of use.

Need help? Check our Help pages or contact us

For an item sold by Amazon.com: When you click the "Place your order" button, we'll send you an email message acknowledging receipt of your order. Your contract to purchase an item will not be complete until we send you an email notifying you that the item has been shipped.

Important information about sales tax you may owe in your state

Shop by Department

Kitchen & D

Home & Kitc

HOMEP

The

📅 LAS

Tweet

thesweethome.c

Affiliate Marketing Terms

- **Affiliate Marketing ID:**
  Unique identifier that <u>Online Retailers</u> use to tie <u>Web Users</u> to <u>Publishers</u>

- **Affiliate Marketing Cookie:**
  Cookie set by <u>Online Retailer</u>, tying <u>Web User</u> to the "delivering" <u>Publisher</u>

- **Cookie Setting URL:**
  End points, controlled by <u>Online Retailers</u>, that set an affiliate marketing cookie on a <u>Web User</u>

Affiliate Marketing Fraud

- **Assumption**

  Having an affiliate marketing cookie →

      User intended to visit the online retailer →

          Retailer helped sale

- **Exploit**

  Get your affiliate marketing cookie on as many browsers as possible

- **Methods**

  hidden iframes, plugins, malware, automatic redirects, etc.

2. Data Set

Affiliate Marketing Programs

- 164 affiliate marketing programs

  - **Popular**: Amazon, GoDaddy

  - **Networks**: ClickCash, MoreNiche

- Selection methods

  - Predictable URLs

  - HTTP / no encryption

Affiliate Marketing Programs

Data	Amazon	GoDaddy
Domains	(www\.)amazon\.com	^godaddy\.*
Cookie Setting URLs	^/(?:.*(dp|gp)/.*)?	[&?]tag=(?:&|\?|^|;)isc=
Conversion URLs	*handle-buy-box*	*domains/domain-configuration\.aspx*
Affiliate ID Values	tag=(.*?)(?:&|$)	cvosrc=(.*?)(?:&|$)

HTTP Request Logs

- 660G of HTTP Requests (bro-log format)

- 2.3 billion records

- January and February 2014

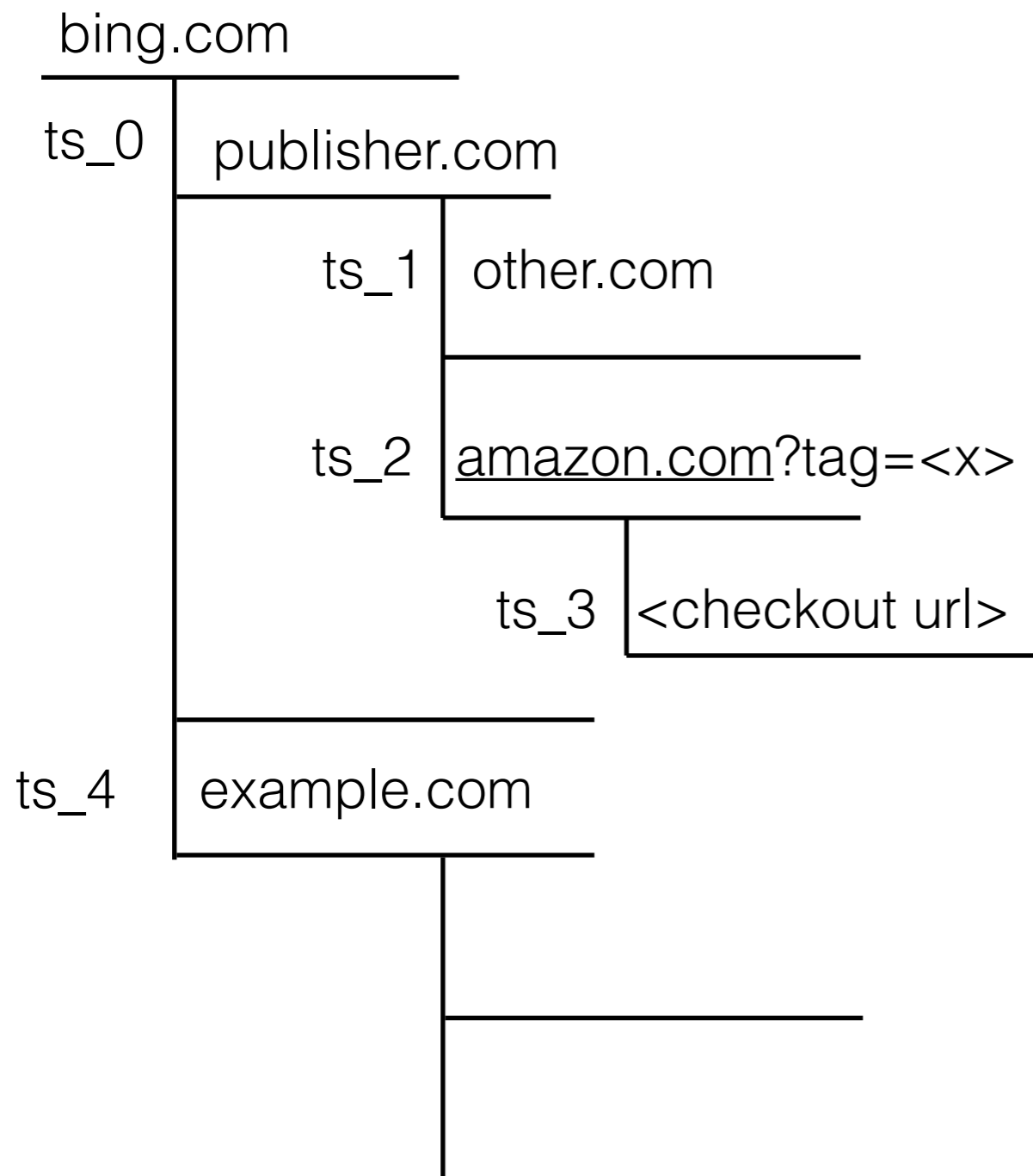| Request Information	Response Information
Sender and destination IP	Mime type
Domain and path	HTTP response code
Referrer	
Timestamp	
Cookies	
User agent	

3. Methodology

Data and Preprocessing
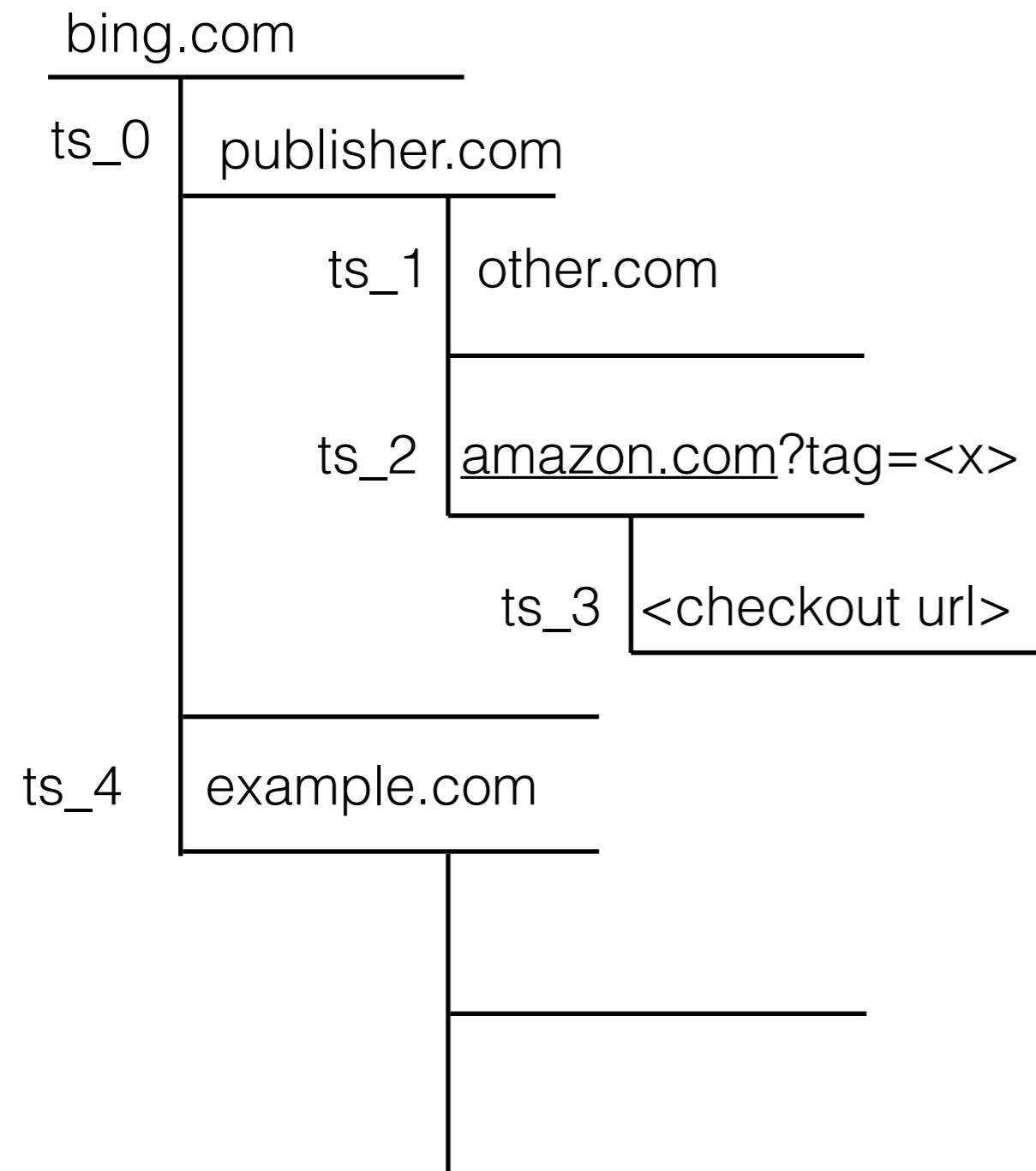
2014-01-01T18:00:00-0500   129.174.37.173   r14---sn-p5qlsnez.googlevideo.com   http://www.youtube.com/watch?v=TabUnV7xoHM   /videoplayback?algorithm
2014-01-01T17:59:59-0500   192.5.215.222   www.mmadsgadget.com   -   /t?id=98394efb-7f7c-4ce4-0152-dee1df6de648&size=728x90
2014-01-01T17:59:59-0500   192.5.215.222   www.mmadsgadget.com   http://www.twitbridge.com/ads/300x250.cshtml?d=http://www.mangahere.com/manga/akuma_na_e
2014-01-01T18:00:00-0500   124.149.235.78   wxmaps.org   -   /apple-touch-icon-72x72-precomposed.png
2014-01-01T17:59:55-0500   129.174.67.117   tracker.bitsoup.me   -   /announce.php?passkey=7776250d207e3bfce323b2107bbacf83&info_hash=b\xa2'M\x06;x\x
2014-01-01T18:00:00-0500   129.174.67.117   tracker.bitsoup.me   -   /scrape.php?passkey=7776250d207e3bfce323b2107bbacf83&info_hash=b\xa2'M\x06;x\xae
2014-01-01T18:00:00-0500   192.5.215.222   www.twitbridge.com   http://www.mangahere.com/manga/akuma_na_eros/c003.2/4.html   /ads/728x90.cshtml?d=htt
2014-01-01T18:00:00-0500   38.119.104.180   edge.sharethis.com   http://www.era.com/property/73716557-101024643/17600-e-fieldbrook-cir-e-boca-raton-fl-33
2014-01-01T18:00:00-0500   38.119.104.180   b.scorecardresearch.com   http://www.era.com/property/73716557-101024643/17600-e-fieldbrook-cir-e-boca-raton-fl-33
2014-01-01T17:59:59-0500   199.26.254.212   scores.espn.go.com   http://assets.espn.go.com/swf/caster/snapshotCasterClientv2.swf   /ncf/caster/snapshot?ses
2014-01-01T18:00:00-0500   192.5.215.222   www.twitbridge.com   http://www.mangahere.com/manga/akuma_na_eros/c003.2/4.html   /ads/300x250.cshtml?d=ht
2014-01-01T18:00:00-0500   199.26.254.69   whois.freeshell.ustc.edu.cn   -   /whois/data
2014-01-01T18:00:00-0500   192.5.215.222   www.twitbridge.com   http://www.mangahere.com/manga/akuma_na_eros/c003.2/4.html   /ads/160x600.cshtml?d=ht
2014-01-01T18:00:00-0500   129.174.0.60   updates.ironport.com   -   /case/1.0/pkg_version/default/1388616888070135
2014-01-01T18:00:00-0500   129.174.0.60   updates.ironport.com   -   /case/1.0/dfa/default/1388616888304604
2014-01-01T18:00:00-0500   5.10.83.24   911digitalarchive.org   -   /repository_object.php?object_id=111820
2014-01-01T18:00:00-0500   23.238.207.60   russianhistoryblog.org   http://russianhistoryblog.org/2011/02/russians-in-east-prussia-1914-pt-2/   /wp-comm
2014-01-01T18:00:00-0500   192.5.215.222   www.twitbridge.com   http://www.mangahere.com/manga/akuma_na_eros/c003.2/4.html   /ads/160x600.cshtml?d=ht
2014-01-01T18:00:00-0500   111.85.195.23   ipv4.tracker.harry.lu.nyud.net   -   /announce?info_hash=\xc2\xf1\x81\x9bb$\xa7Cv\x12Q\x9f*~\x02Gh\x15s\xe6&p
2014-01-01T18:00:00-0500   129.174.0.60   updates.ironport.com   -   /case/1.0/dfa_updates/default/1388616853644950
2014-01-01T17:59:59-0500   129.174.143.252   108.175.35.156   -   /1654089019.ismv?c=us&n=11279&v=3&e=1388644580&t=-1aKqXqPhD01POJZRz9VmZuRomQ&d=ps3&p=5.B
2014-01-01T18:00:00-0500   192.5.215.213   108.175.34.103   http://movies.netflix.com/WiPlayer?movieid=70122768&trkid=7882979   /1550506182.ismv/range/5
2014-01-01T18:00:00-0500   199.26.254.69   whois.freeshell.ustc.edu.cn   -   /whois/registrant
2014-01-01T17:59:56-0500   192.5.215.215   ns-streamingaudio.itunes.apple.com   -   /us/r1000/017/Music4/v4/45/8c/60/458c60b4-da3b-2b07-8975-8824e2b
2014-01-01T18:00:00-0500   124.120.221.183   ipv4.tracker.harry.lu.nyud.net   -   /announce?info_hash=nJ\xf4\xd1\x14\xd0\xbc|L\xd59\xb5\x93\x1f\x1eM\xabT
2014-01-01T18:00:00-0500   124.120.221.183   ipv4.tracker.harry.lu.nyud.net   -   /announce?info_hash=*o\xc7\xed\x1c\xd7\x192\x85qW\xca\x99^"\xcdO\x18m\x1
2014-01-01T18:00:01-0500   38.119.104.180   media.newjobs.com   http://jobsearch.monster.com/search/Systems-Engineer_5?where=Massachusetts   /nmy/use
2014-01-01T18:00:01-0500   129.174.67.117   tracker.bitsoup.me   -   /announce.php?passkey=7776250d207e3bfce323b2107bbacf83&info_hash=dl\xa2\xd2\x2\xcaB
2014-01-01T18:00:01-0500   192.5.215.222   www.mmadsgadget.com   -   /t?id=98394efb-7f7c-4ce4-0152-dee1df6de648&size=728x90
2014-01-01T17:59:58-0500   129.174.143.252   108.175.35.156   -   /1654089019.ismv?c=us&n=11279&v=3&e=1388644580&t=-1aKqXqPhD01POJZRz9VmZuRomQ&d=ps3&p=5.B
2014-01-01T18:00:01-0500   129.174.138.174   ac.search.naver.com   http://www.naver.com/   /nx/ac?_callback=window.__jindo_callback._$3361_0&q=are counterp
2014-01-01T18:00:00-0500   173.64.103.239   calu-atrack.appspot.com.nyud.net   -   /announce?info_hash=\x057\x14,-\x7f\xbbp\xab\x12/\xc8\x0dC\xb5k]
2014-01-01T17:59:58-0500   192.5.215.217   valve205.cs.steampowered.com   -   /depot/206421/chunk/c274f93a6b4428922bba93b3d727f8d2def1983b
2014-01-01T17:59:01-0500   199.26.254.69   whois.freeshell.ustc.edu.cn   -   /ip/getnext
2014-01-01T18:00:01-0500   192.5.215.211   www.google.com   -   /url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CDsQFjAB&url=http://engineering.stanford.
2014-01-01T18:00:01-0500   131.253.24.130   publichistorycommons.org   -   /thoughts-on-completing-my-first-major-public-history-project/
2014-01-01T18:00:01-0500   129.174.70.208   crl.microsoft.com   -   /pki/crl/products/CodeSignPCA2.crl
2014-01-01T18:00:01-0500   38.119.104.180   jdn.monster.com   http://oas.monster.com/RealMedia/ads/adstream_sx.ads/us.monster.en/js20/powersearch.aspx/1860996
2014-01-01T18:00:01-0500   157.55.34.29   teachinghistory.org   -   /history-content/historical-site/10448
2014-01-01T18:00:01-0500   192.5.215.219   googleads.g.doubleclick.net   -   /mads/static/sdk/native/sdk-core-v40.js

Browsing Session Trees

bing.com

ts_0   publisher.com

ts_1   other.com

ts_2   amazon.com?tag=<x>

ts_3   <checkout url>

ts_4   example.com

Xie, Guowu, et al. "Resurf: Reconstructing web-surfing activity from network traffic." *IFIP Networking Conference, 2013*. IEEE, 2013.
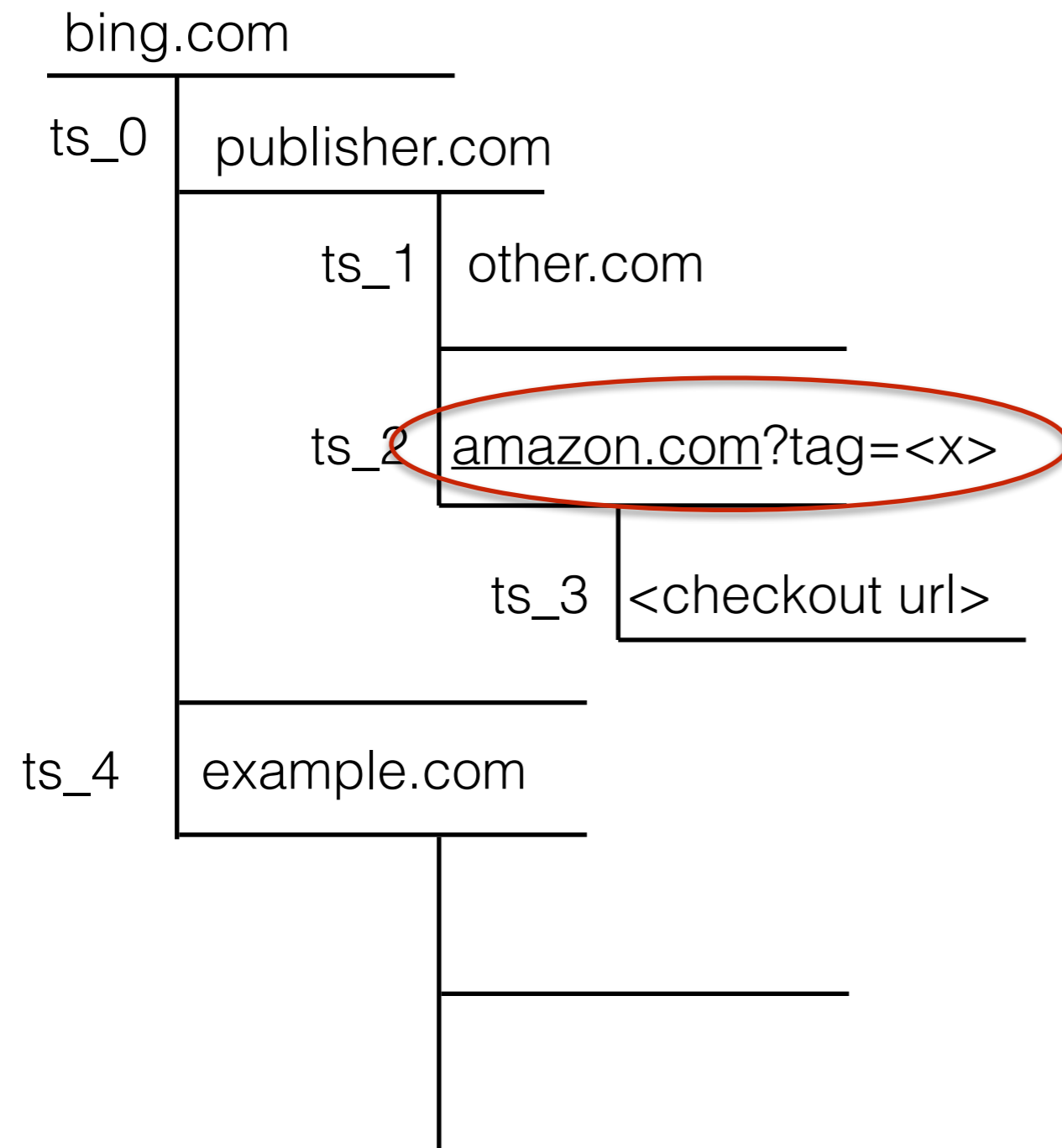
Browsing Session Trees

bing.com

ts_0 | publisher.com

ts_1 | other.com

ts_2 | amazon.com?tag=<x>

ts_3 | <checkout url>

ts_4 | example.com

Simple Measurements

- Number of referrals in each program

- Number of publishers in each program

- Number of conversions / purchases in each program

- How long a user takes to be referred

- How long a user spent on site after being referred

Browsing Session Trees

bing.com

ts_0 | publisher.com

ts_1 | other.com

ts_2 | amazon.com?tag=<x>

ts_3 | <checkout url>

ts_4 | example.com

Simple Measurements

- Number of referrals in each program

- Number of publishers in each program

- Number of conversions / purchases in each program

- How long a user takes to be referred

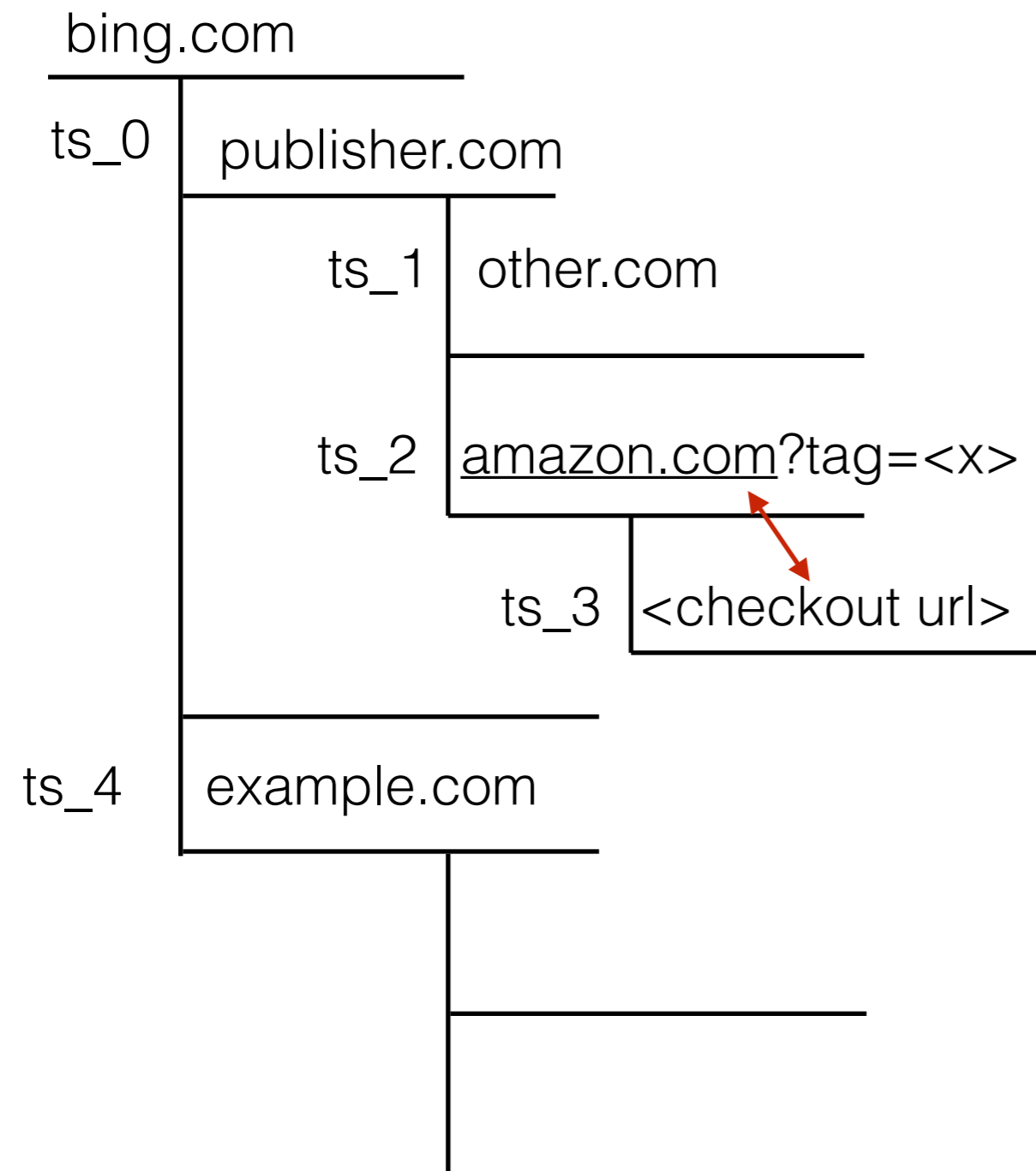- How long a user spent on site after being referred

Browsing Session Trees

bing.com

ts_0 | publisher.com

ts_1 | other.com

ts_2 | amazon.com?tag=<x>

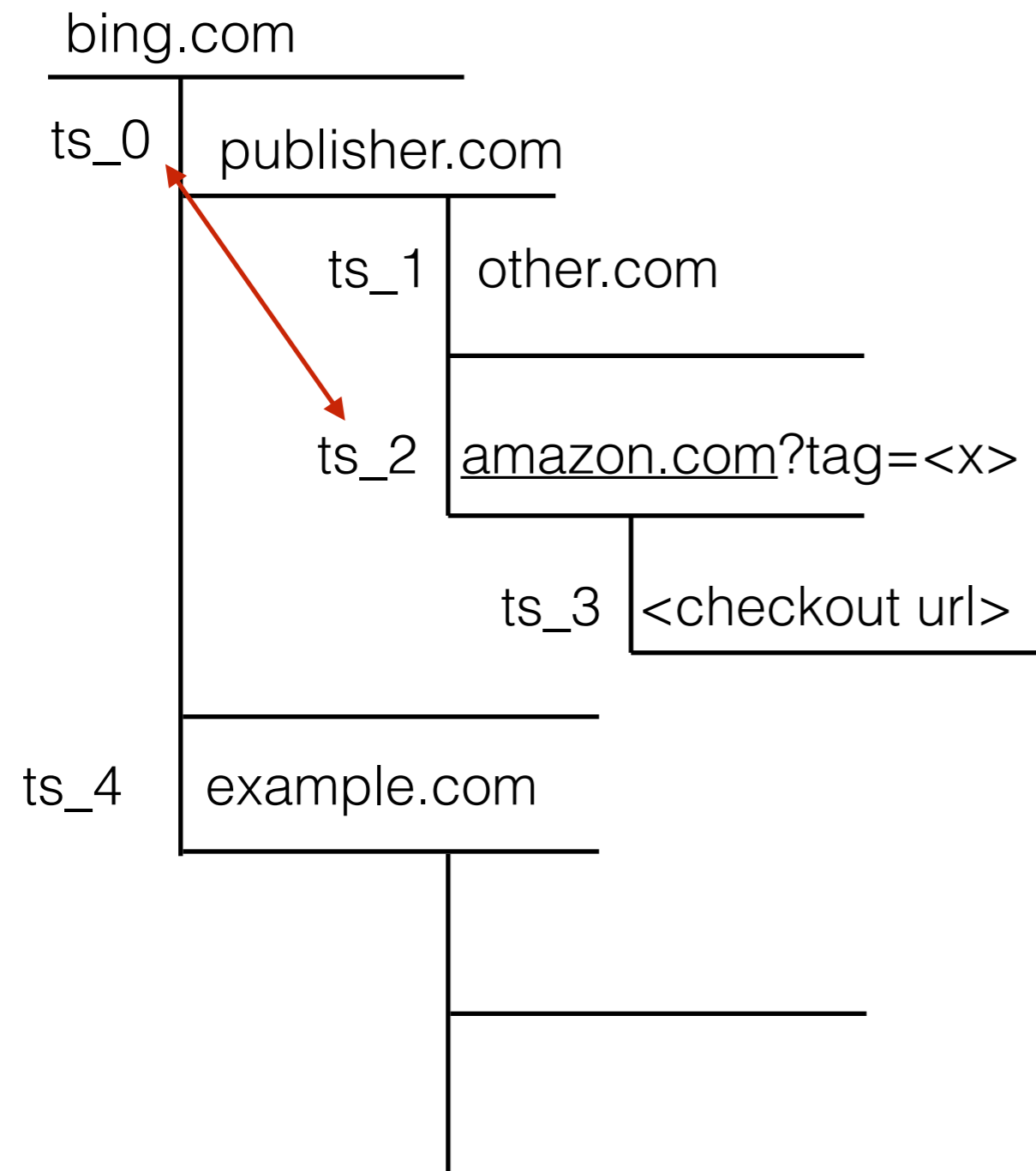ts_3 | <checkout url>

ts_4 | example.com

Simple Measurements

- Number of referrals in each program

- Number of publishers in each program

- Number of conversions / purchases in each program

- How long a user takes to be referred

- How long a user spent on site after being referred

Browsing Session Trees

bing.com

ts_0 publisher.com

ts_1 other.com

ts_2 amazon.com?tag=<x>

ts_3 <checkout url>

ts_4 example.com

Simple Measurements

- Number of referrals in each program

- Number of publishers in each program

- Number of conversions / purchases in each program

- How long a user takes to be referred

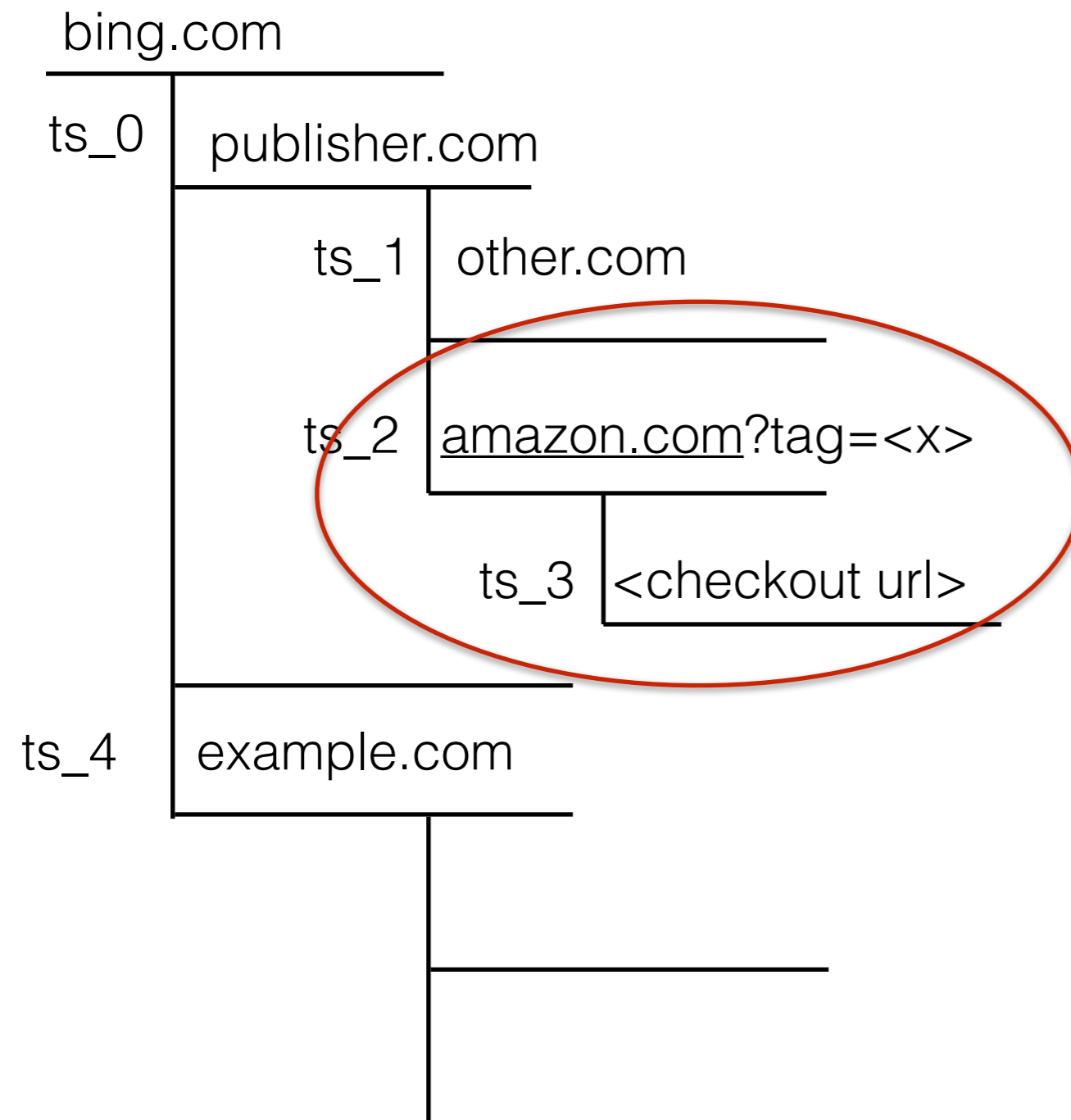- How long a user spent on site after being referred

Browsing Session Trees

bing.com

ts_0 | publisher.com

ts_1 | other.com

ts_2 | amazon.com?tag=<x>

ts_3 | <checkout url>

ts_4 | example.com

Simple Measurements

- Number of referrals in each program

- Number of publishers in each program

- Number of conversions / purchases in each program

- How long a user takes to be referred

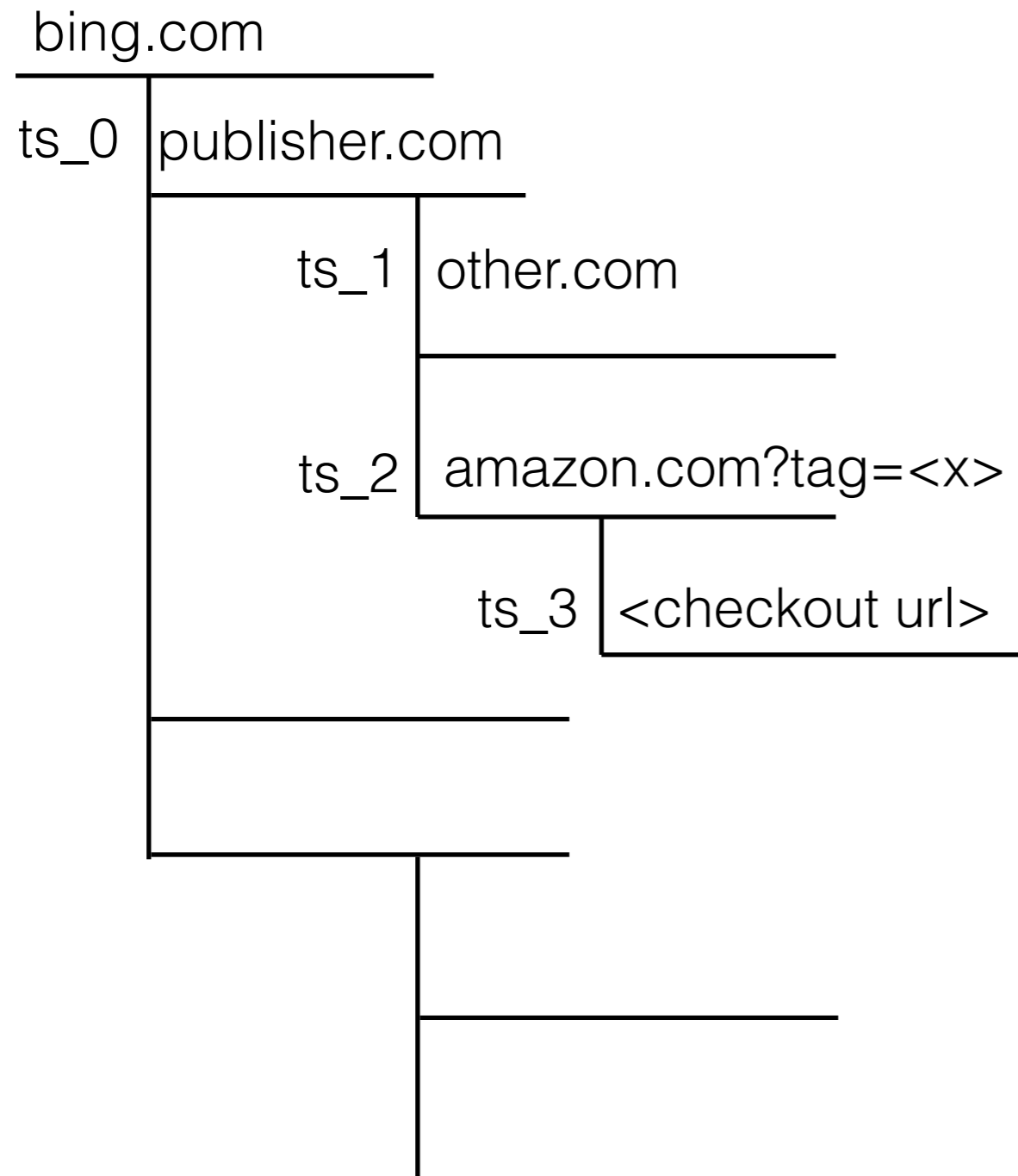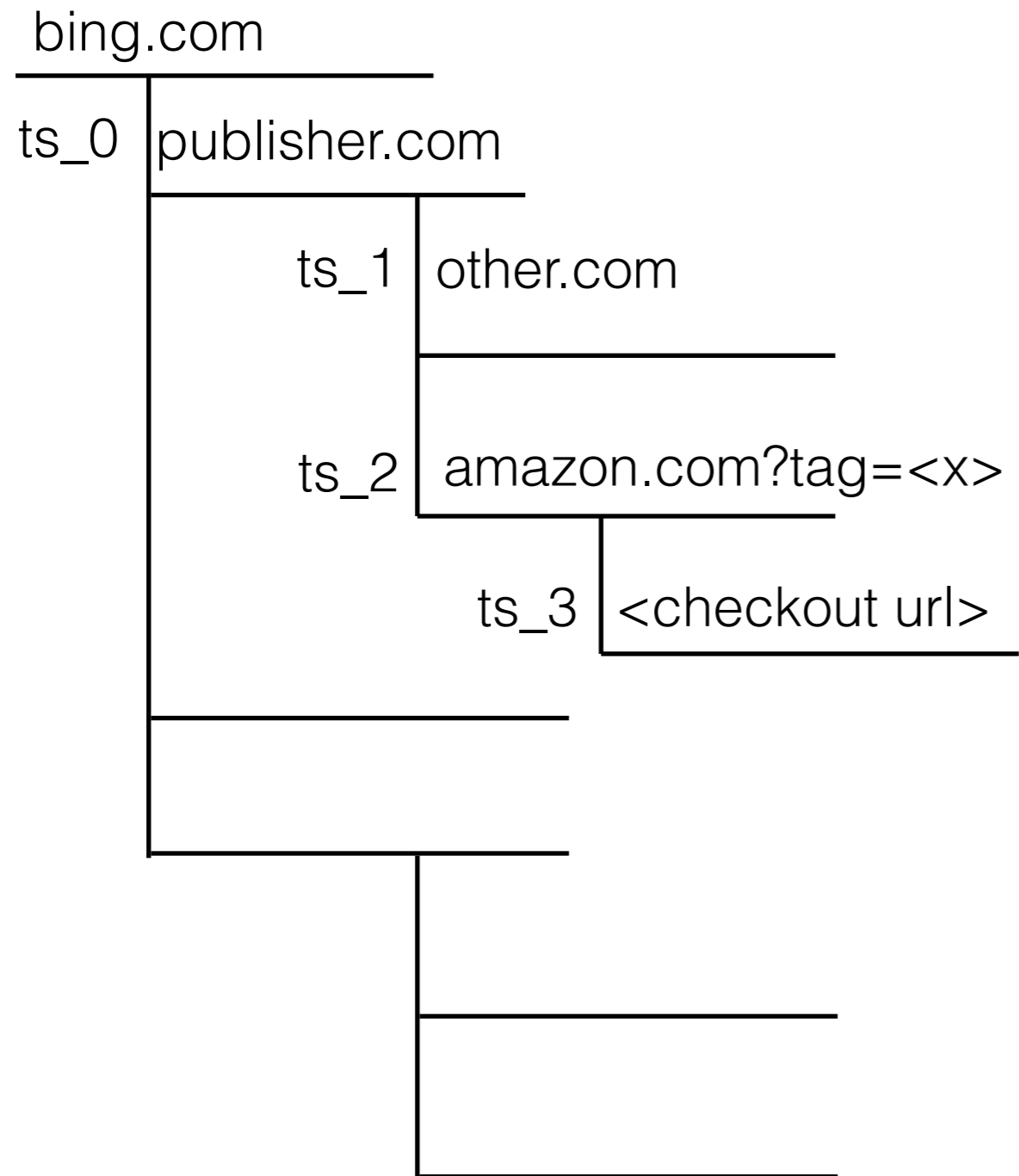- How long a user spent on site after being referred
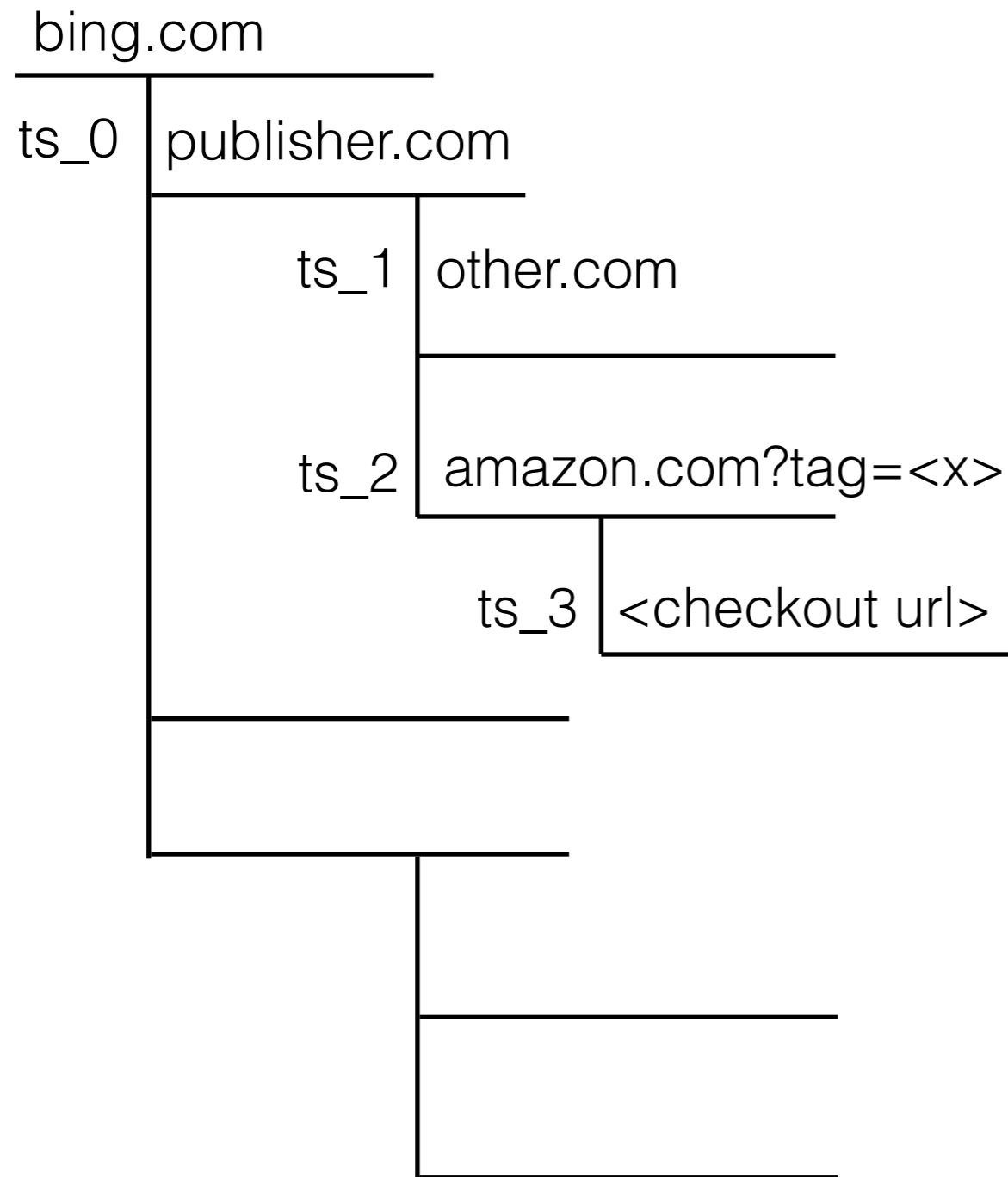
Classifier: Training Set

bing.com

ts_0 publisher.com

ts_1 other.com

ts_2 amazon.com?tag=<x>

ts_3 <checkout url>

- Did the user intend to travel from <u>some-referrer.com</u> to <u>amazon.com</u>?

- Built training set of 1141 relevant trees (subset of January data)

- If referrer was still available, **direct test**

- If referrer was not available, **infer from graph** (log data)

Classifier: Features

bing.com

ts_0 publisher.com

ts_1 other.com

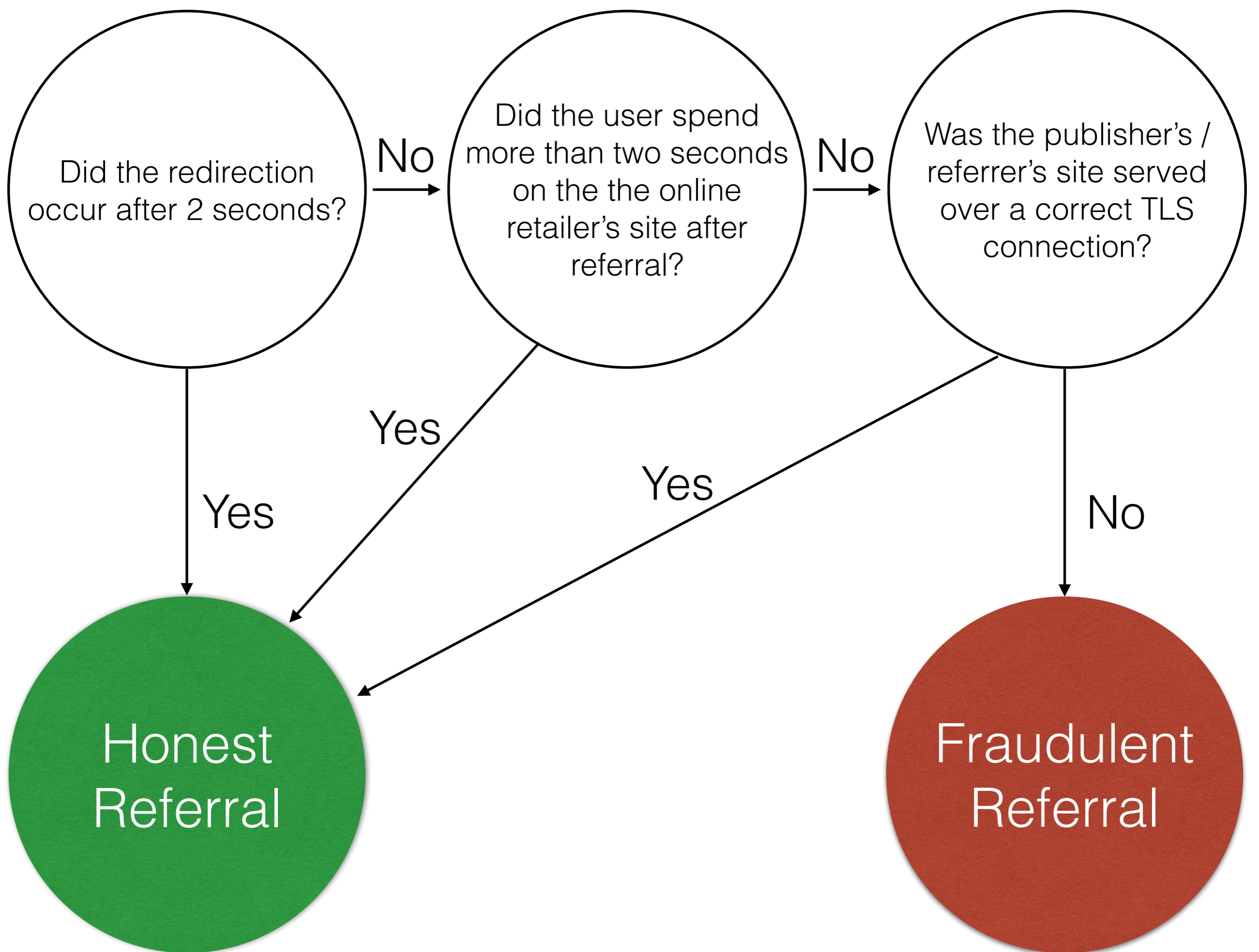ts_2 amazon.com?tag=<x>

ts_3 <checkout url>

1. Time before referral
2. Time after referral
3. Is referrer SSL?
4. Graph size
5. Is referrer reachable?
6. Google page rank of referrer
7. Alexia traffic rank
8. Is referrer domain registered?
9. # years domain is registered
10. Tag count

Classifier: Features

bing.com

ts_0 | publisher.com

ts_1 | other.com

ts_2 | amazon.com?tag=<x>

ts_3 | <checkout url>

1. **Time before referral**
2. **Time after referral**
3. **Is referrer SSL?**
4. Graph size
5. Is referrer reachable?
6. Google page rank of referrer
7. Alexia traffic rank
8. Is referrer domain registered?
9. # years domain is registered
10. Tag count

93.3% accuracy

4. Findings

Online Retailer Popularity

Retailer	Requests	Unique Sessions
Amazon.com	2,663,574	87,654
GoDaddy	7,320	364
ImLive.com	731	194
wildmatch.com	3	1
Total (166 programs)	2,671,808	88,257

Publishers

Retailer	Honest	Fraudulent	Total
Amazon.com	2,268	1,396	3,664
GoDaddy	5	19	24
ImLive.com	4	7	11
wildmatch.com	0	1	1
Total (166 programs)	2,281	1,426	3,707

Affiliate Marketer Referrals

Retailer	Honest	Fraudulent	Total
Amazon.com	12,870	2,782	15,652
GoDaddy	399	98	497
ImLive.com	9	13	22
wildmatch.com	0	1	1
Total (166 programs)	13,283	2,897	16,180

Conversion Events

Retailer	Amazon.com	GoDaddy	Total (166 programs)
Conversion Events	15,624	26	15,650
Affiliate Conversions	955	8	693
Honest	781	8	789
Fradulent	174	0	174
"Stolen"	0	0	0

In The Paper…

- Session tree building algorithm

- Details of how we generated the classifier

- Stakeholder analysis of affiliate marketing fraud

- More numbers…

Thanks!

Peter Snyder – psnyde2@uic.edu
Chris Kanich – ckanich@uic.edu
University of Illinois at Chicago

UIC
UNIVERSITY
OF ILLINOIS
AT CHICAGO