

Automated Black-Box Detection of Side-Channel Vulnerabilities in Web Applications

Peter Chapman

David Evans

University of Virginia

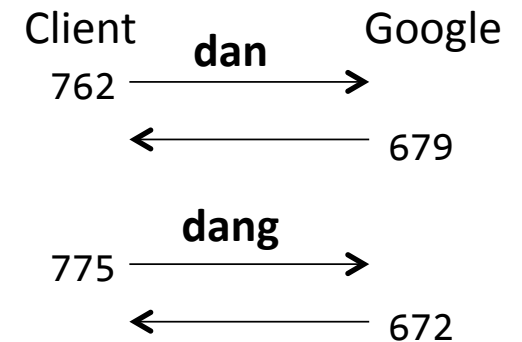
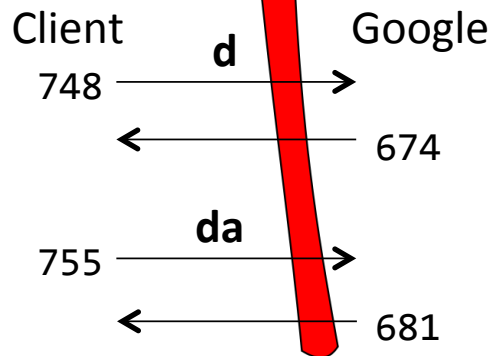
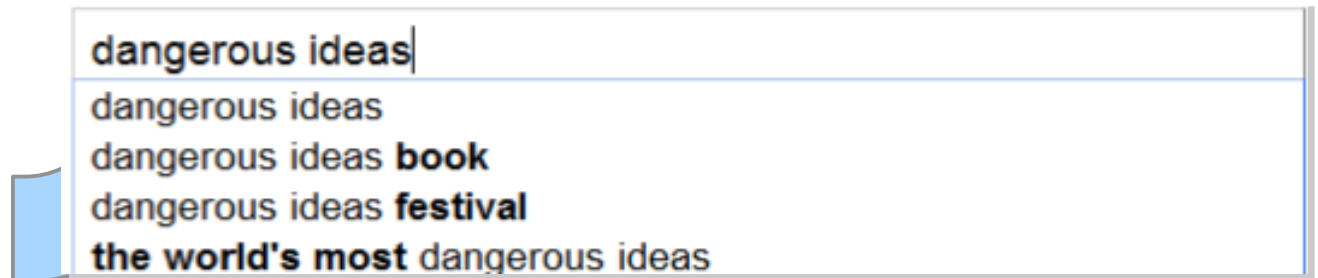
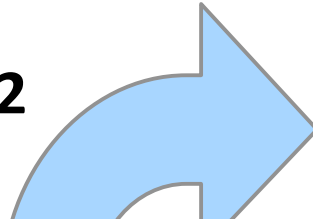
<http://www.cs.virginia.edu/sca/>

CCS '11

October 19, 2011

Side-Channel Leaks in Web Apps

HTTPS over WPA2



Modern Web Apps

Dynamic and Responsive
Browsing Experience



On-Demand Content

HTML



Traffic



Latency



Responsiveness

Traffic is now closely associated
with the demanded content.

Motivation: Detect Vulnerabilities

doctor's office

About 4,750,000 results (0.22 seconds)

Sentara Physicians - We Provide Quality Family Care
www.sentara.com/sentara_medical_group
Request an Appointment Online.

Places for doctor's office near Charlottesville, VA 22903

- University of Virginia Health System : Northridge Internal Medicine** - 1 Google review
uvahealth.com - 2955 Ivy Road, Charlottesville - (434) 243-4500
- Custer M MD** - 1 Google review
rychiropractic.com - 4422 Ivy Commons, Charlottesville - (434) 817-3666
- U Va Hematology/Oncology: Brenin Christiana MD** - Place page
maps.google.com - 3135 Rocks Farm Ct, Charlottesville - (434) 924-8552
- Pediatric Associates of Charlottesville, PLC** - Place page
www.charlottesvillepeds.com - 2411 Ivy Road, Charlottesville - (434) 296-8300
- University of Virginia Health System : Head and Neck Surgery ...** - Place page
www.uvaotolaryngology.com - 415 Ray C Hunt Drive, Charlottesville - (434) 924-2153
- UVA Psychiatry & Nhhvnl** - Place page
maps.google.com - 2955 Ivy Rd # 210, Charlottesville - (434) 243-4646

Cville Family Medicine
www.charlottesvillemedicine.com
Excellent health care...on-site

Basic Info Life Events Federal Q&A State Q&A Review Filing Next Year

Income Deductions Credits Miscellaneous Summary **Jump To Forms & Topics**

Wages

Did you receive any wages or a salary from an employer in 2011?

Yes No

You should receive a Form W-2 from each employer by February 1, 2012. Enter each Form W-2 **separately**. After you complete each Form W-2, the program will ask if you have additional copies of Form W-2 to enter and allow you the opportunity to add new wage information.

Back Continue

Find your symptom checker

1 Introduction 2 Who is the checker for? 3 Symptom questions 4 Results

Find your symptom checker

What happens now?
Work through the questions on the following pages to find the most suitable symptom checker for you.

Find your symptom checker

Once you've found the most appropriate symptom checker, you can work through the assessment and receive advice which may include one of the following:

- Self care - advice on how to look after yourself and manage your symptoms

Please note, this service is provided for people living in England only. For health advice and information in other parts of the UK please visit NHS Northern Ireland, NHS 24 (Scotland) or NHS Direct Wales.

dangerous ideas

dangerous ideas

dangerous ideas **book**

dangerous ideas **festival**

the world's most dangerous ideas

darwin's dangerous ideas

most dangerous ideas

in defense of dangerous ideas

Motivation: Evaluate Defenses

Randomized or Uniform
Communication Attributes

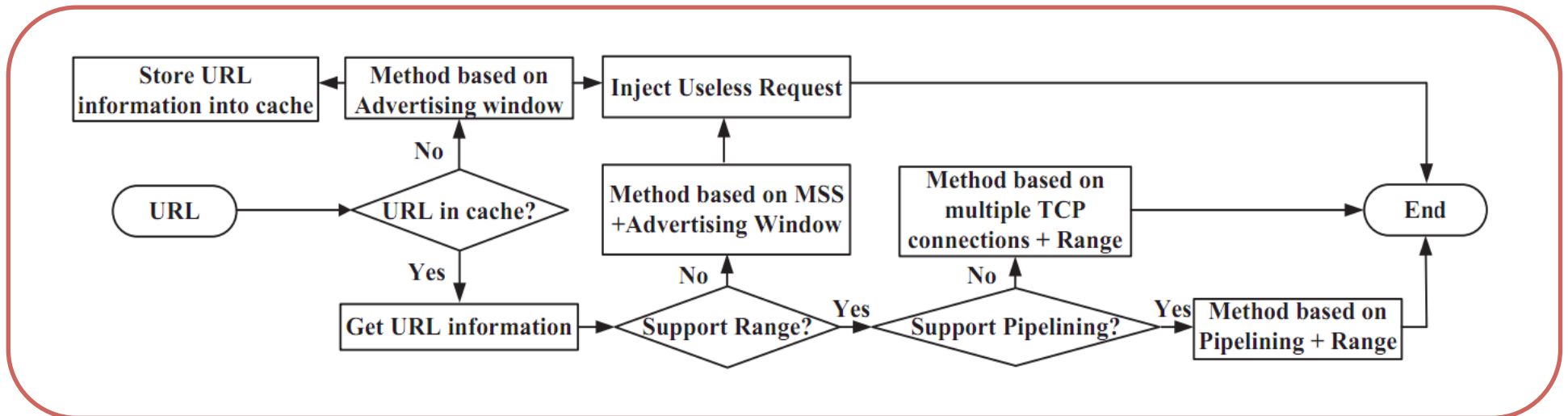
Packet Sizes

Transfer
Control Flow

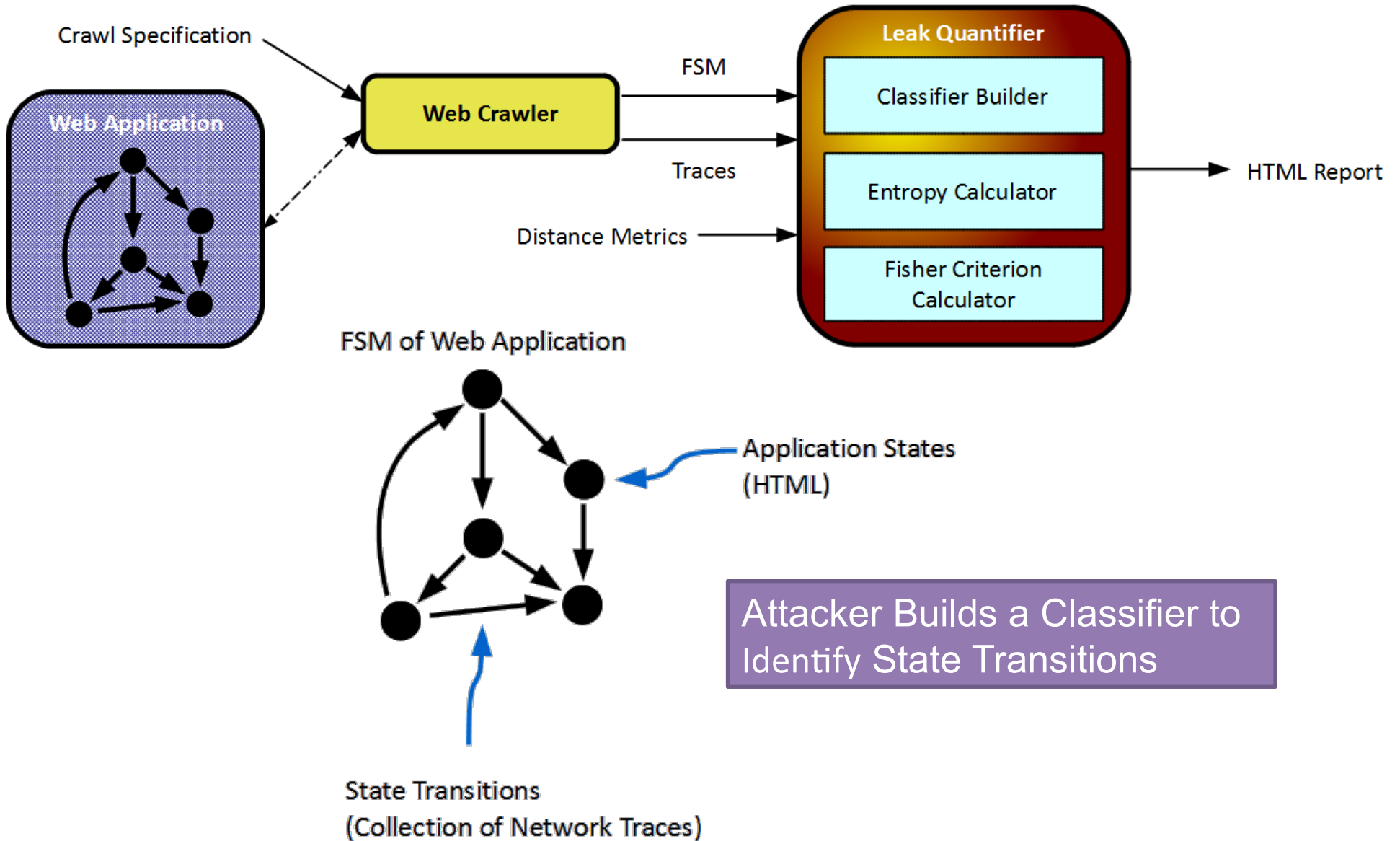
Inter-Packet
Timings

Requests and
Responses

HTTPOS [Luo+, NDSS 2011]

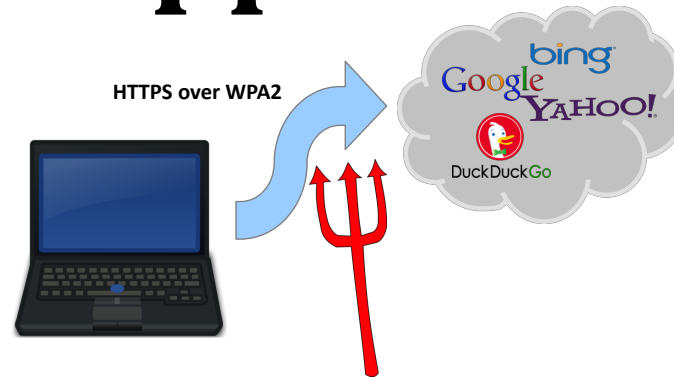


Approach



A Black-Box Approach

Similar to Real Attack Scenario



Applicable to Most Web Applications

Full Browser Analysis

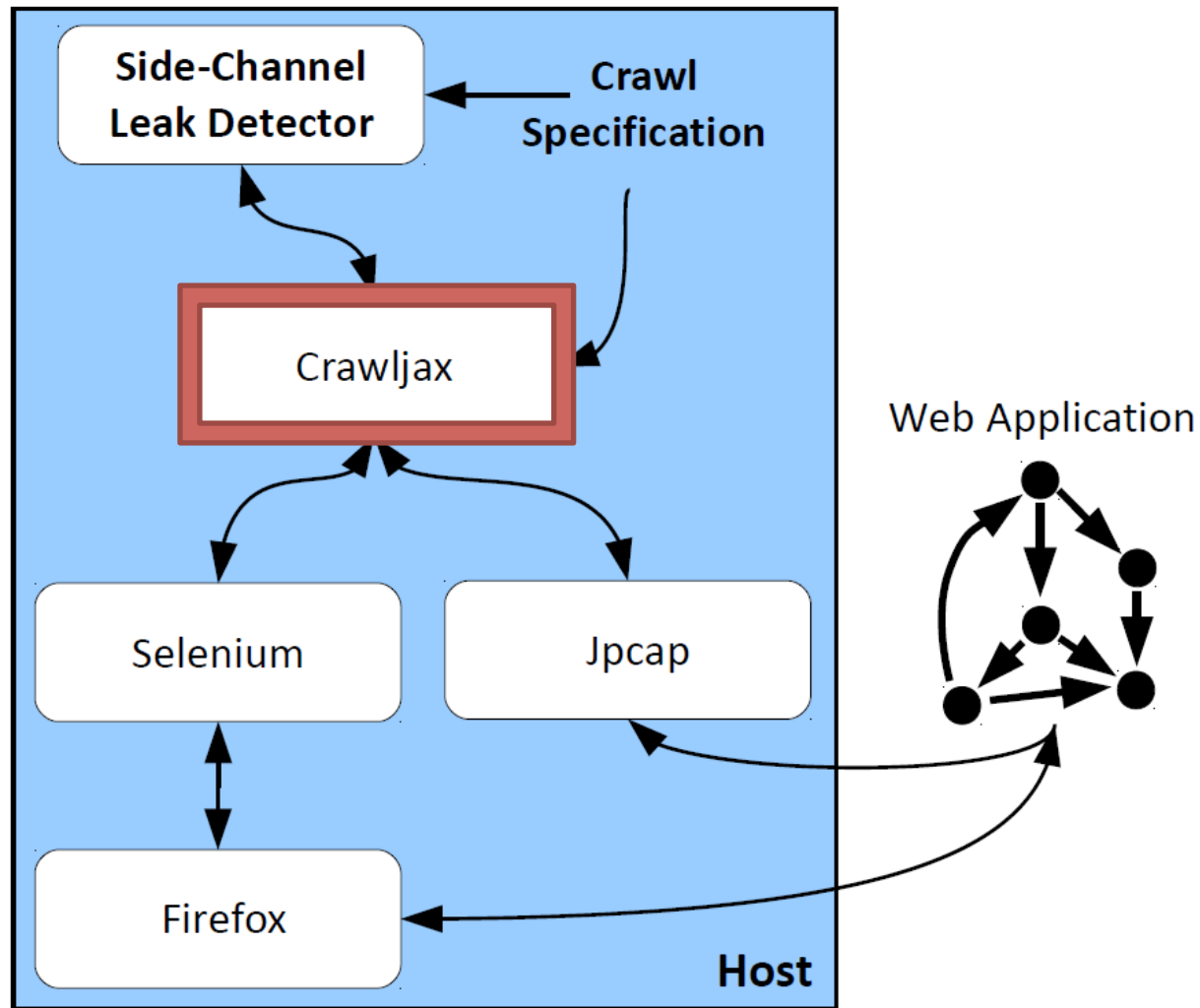
Google
bing
YAHOO!

NHS

Google health
beta



Black-Box Web Application Crawling



Crawljax

The screenshot shows the Crawljax website. At the top left, the logo "Crawljax" is displayed with the tagline "Automate Ajax Crawling and Testing". A navigation menu includes "Home", "About", "Download", "Documentation", and "Plugins". The main content area features a "Home" section with a paragraph about the "crawljax-2.0" release, a description of Crawljax as an open-source Java tool, and a list of capabilities. A "Documentation" link with a notebook icon is visible on the right. A "Log in" link is at the bottom right. Two callout boxes are overlaid: an orange one in the upper middle and a blue one in the lower right.

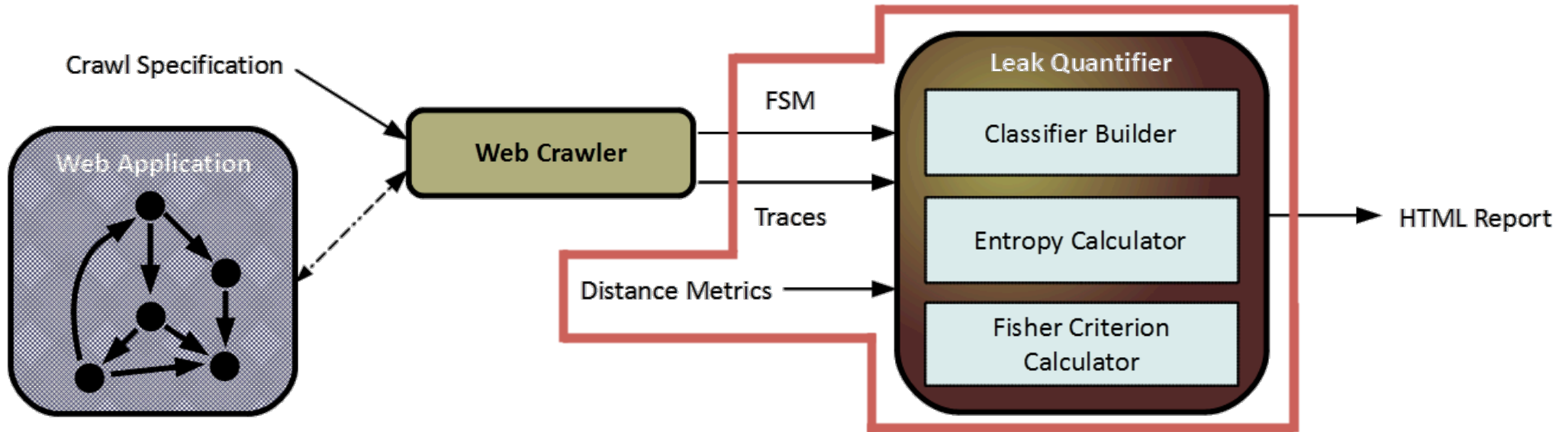
Web crawling back-end drives Firefox instance via Selenium

Designed to build Finite-State Machines of AJAX Applications

<http://crawljax.com/>

<http://crawljax.com/>

Approach



Threat Models and Assumptions

Both: Victim begins at root of application

WiFi

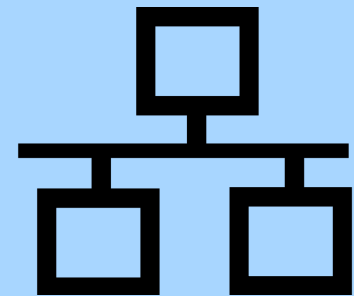
No disruptive traffic

Distinguish incoming and outgoing



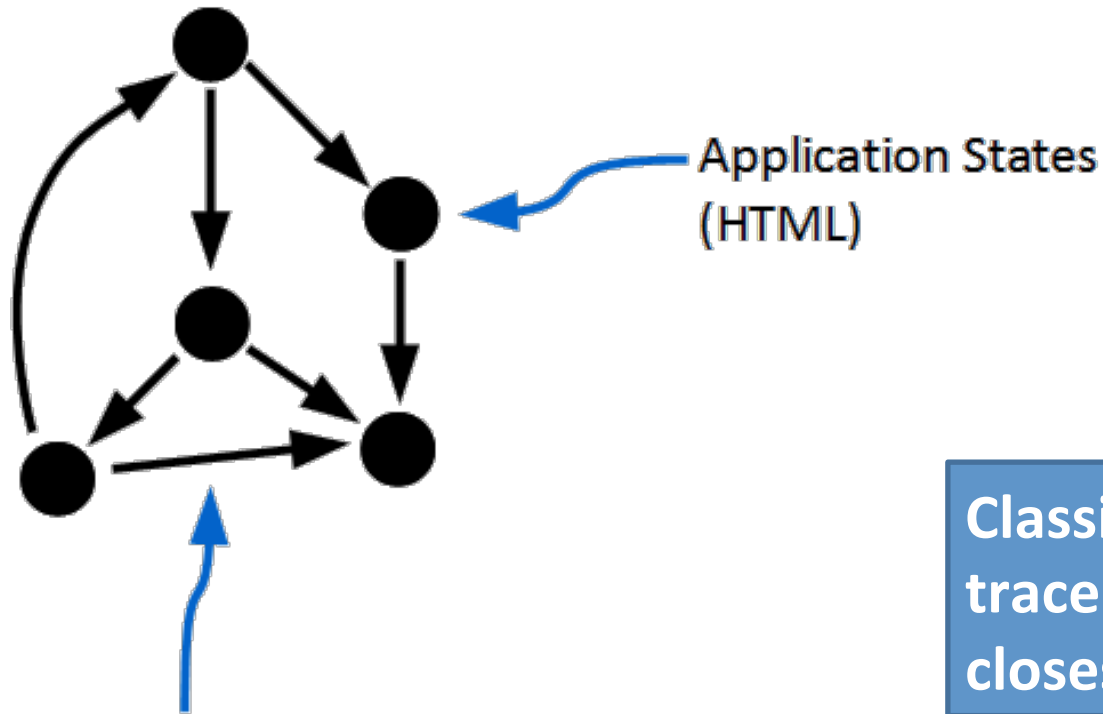
ISP

Access to TCP header



Nearest-Centroid Classifier

FSM of Web Application



Given an unknown network trace, we want to determine to which state transition it belongs

Classify unknown trace as one with the closest centroid

State Transitions
(Collection of Network Traces)

Distance Metrics

Metrics to determine similarity between two traces

Size-Weighted-Edit-Distance

Convert to string, weighted edit distance based on size

Edit-Distance

Unweighted edit distance

Direction

Size of bytes on each party

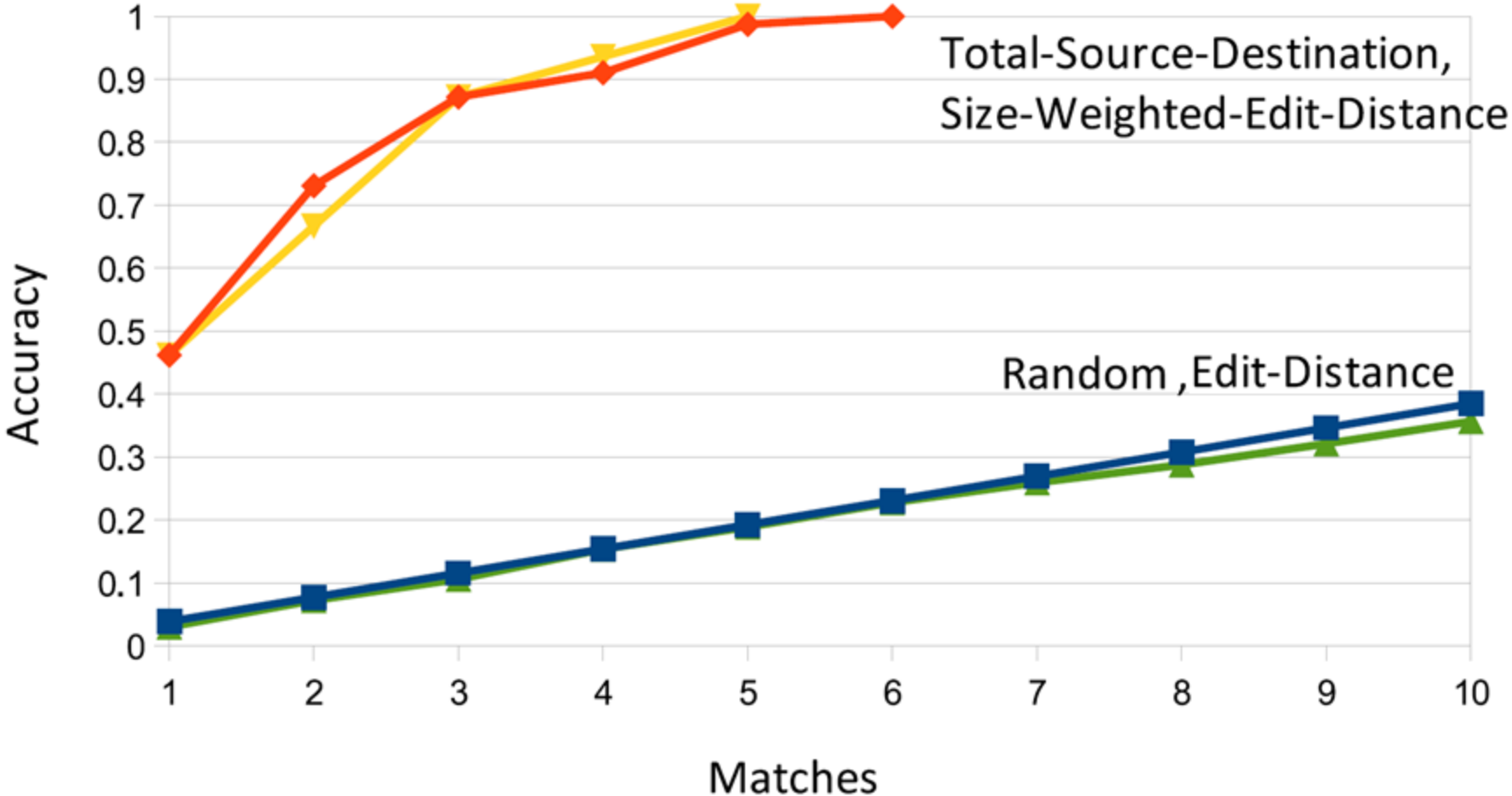
72.14.204	->	192.168.1	62 bytes	
192.168.1	->	72.14.204	281 bytes	A 62 bytes
72.14.204	->	192.168.1	1860 bytes	B 62 bytes
192.168.1	->	72.14.204	294 bytes	A 482 bytes
72.14.204	->	192.168.1	296 bytes	B 693 bytes
192.168.1	->	72.14.204	453 bytes	A 62 bytes
72.14.204	->	192.168.1	2828 bytes	B 62 bytes
192.168.1	->	72.14.204	281 bytes	A 281 bytes
72.14.204	->	192.168.1	1860 bytes	B 1860 bytes
192.168.1	->	72.14.204	294 bytes	A 294 bytes
72.14.204	->	192.168.1	296 bytes	B 296 bytes
192.168.1	->	72.14.204	453 bytes	A 453 bytes
72.14.204	->	192.168.1	2828 bytes	B 2828 bytes



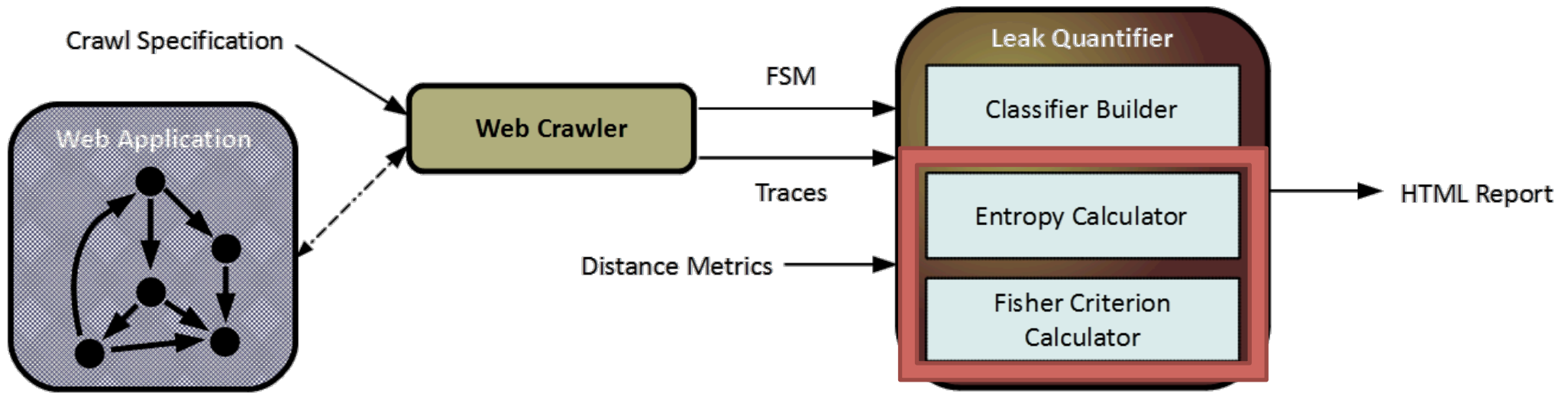
Classifier Performance - Google Search

First character typed, ISP threat model

dangerous ideas|
dangerous ideas
dangerous ideas **book**
dangerous ideas **festival**
the world's most dangerous ideas
darwin's dangerous ideas
most dangerous ideas
in defense of dangerous ideas



Quantifying Leaks



Leak quantification should be independent of a specific classifier implementation

Entropy Measurements

Entropy measurements are a function of the average size of an attacker's uncertainty set given a network trace

Problems

The same network trace can be the result of multiple classifications

Every possible network trace is unknown

$$H(X) = \frac{\sum_{i=0}^n \log_2 p(\bar{x}_i)}{n}$$

Size of uncertainty set

Centroid for class

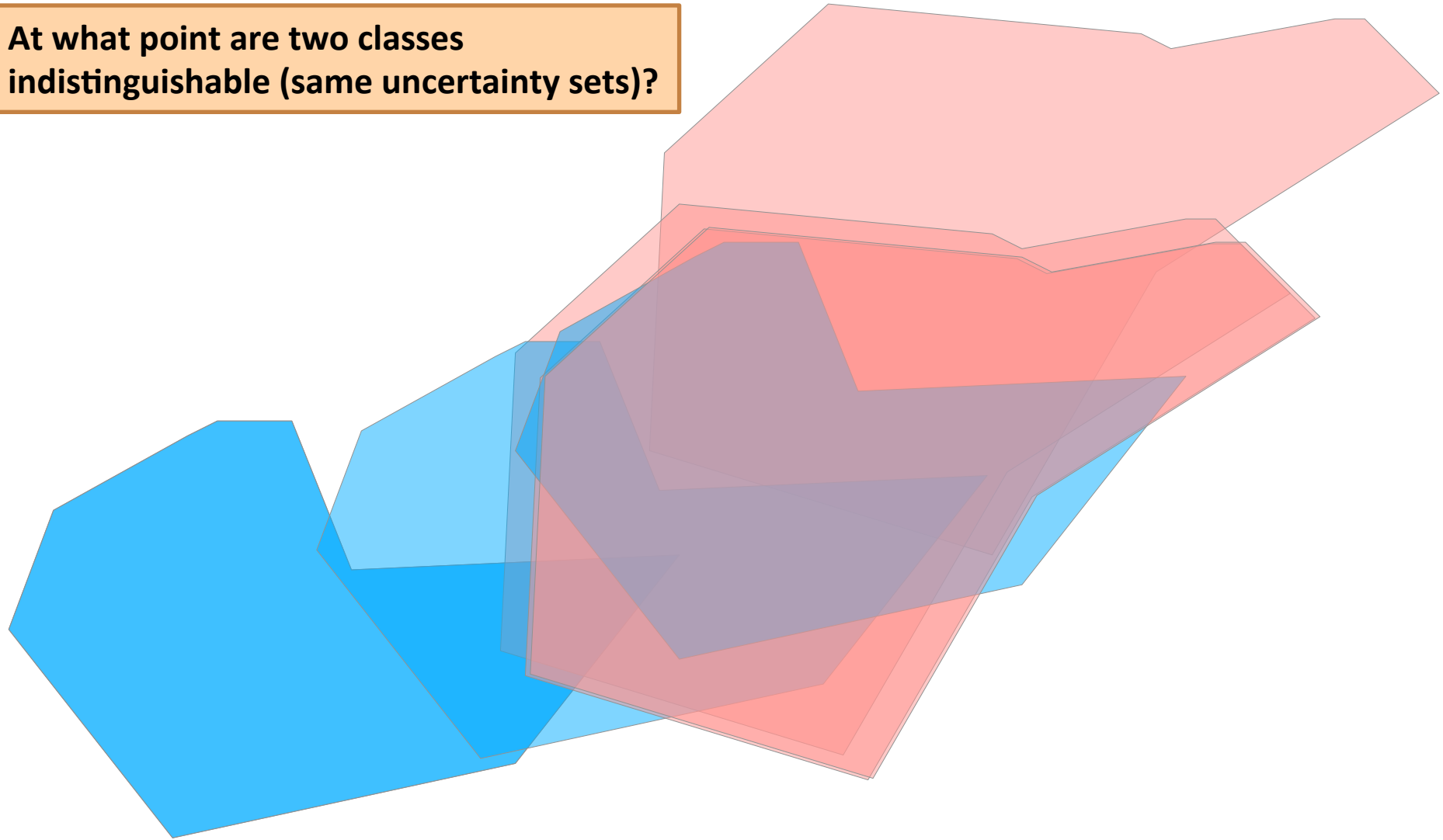
Use the centroids

Number of classes

Traditional Entropy Measurement

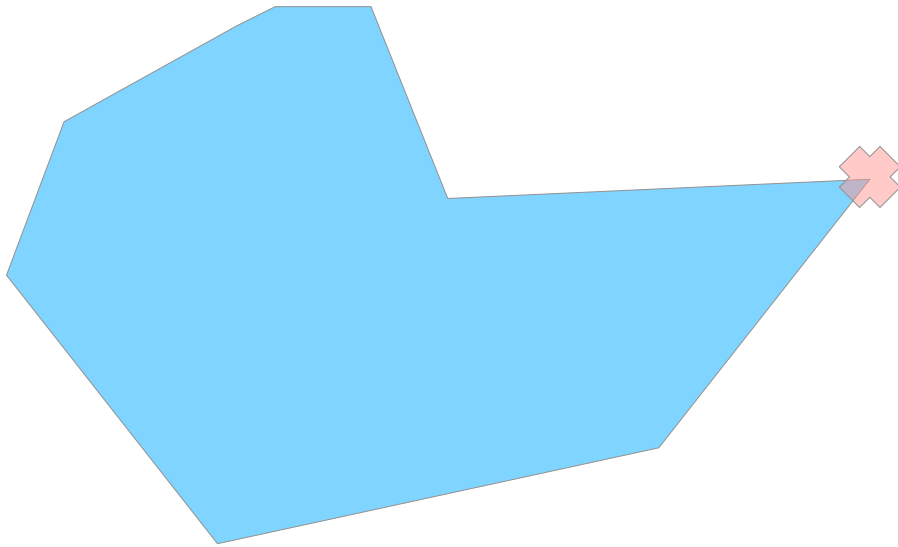
Determining Indistinguishability

At what point are two classes indistinguishable (same uncertainty sets)?



Determining Indistinguishability

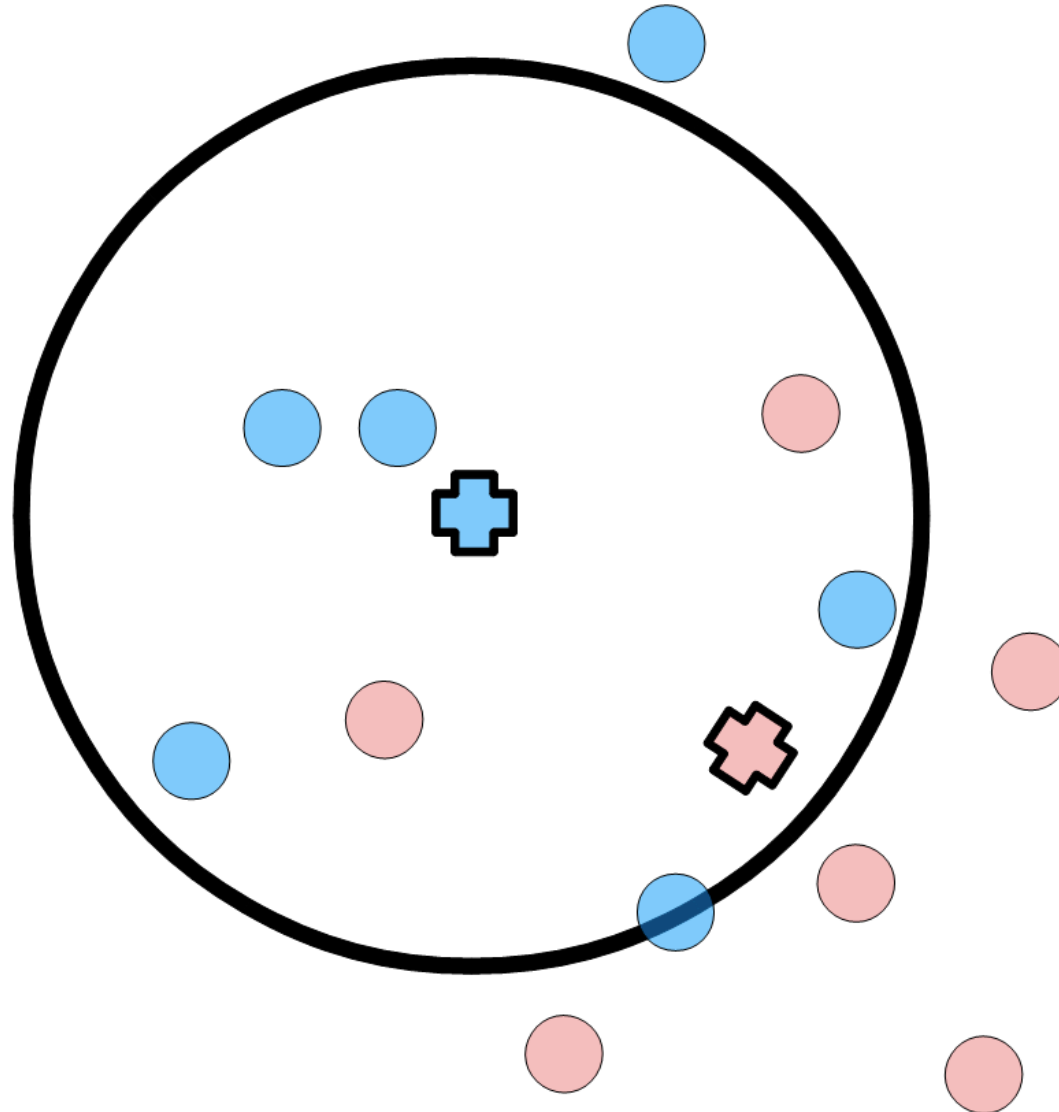
Compare points to centroids?



Same issue with individual points.

In practice the area can be very large due to high variance in network conditions

Entropy Distinguishability Threshold



Threshold of 75%

Google Search Entropy Calculations

	Threshold		
	100%	75%	50%
Desired	4.70	4.70	4.70
Total-Source-Destination	2.95	2.40	0.44
Size-Weighted-Edit-Distance	1.13	0.56	0.44
Edit-Distance	4.70	4.70	4.70

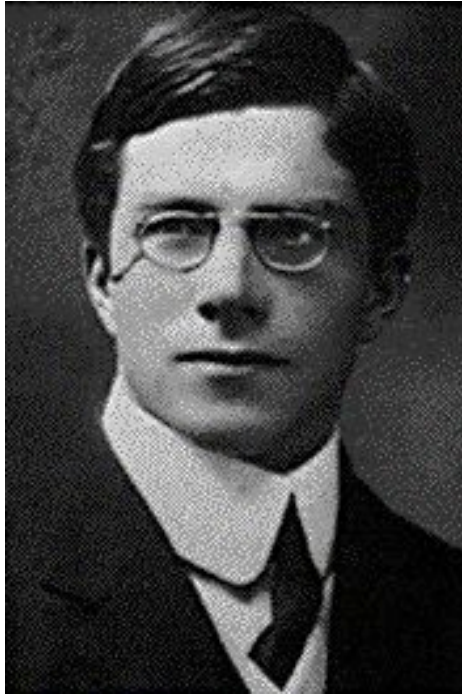
(measured in bits of entropy)

We'd rather not use something with an arbitrary parameter

Fisher Criterion

[11] Ronald A. Fisher. The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, 1936.

Fisher Criterion



Ronald Fisher (1890-1962)

Developed many statistical tools as a part of his prominent role in the eugenics community

Married Arthur Guinness' daughter, secret wedding (she was 17) in 1917



Arthur Guinness (1835-1910)

Fisher Criterion

Like all good stories, this one starts with a Guinness.

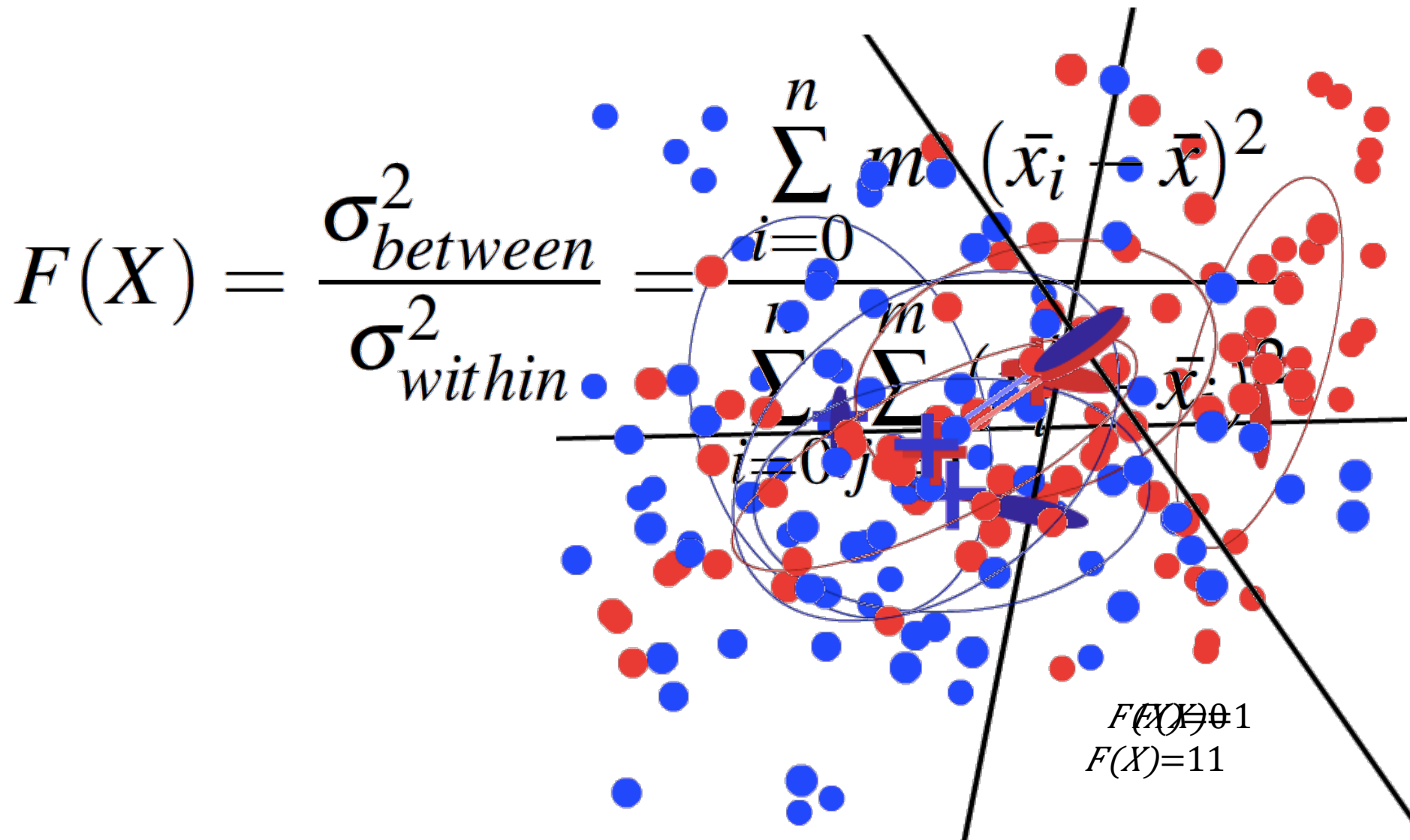


Arthur Guinness (1725-1803)



"Guinness is Good for You"

Fisher Criterion



Google Search Fisher Calculations

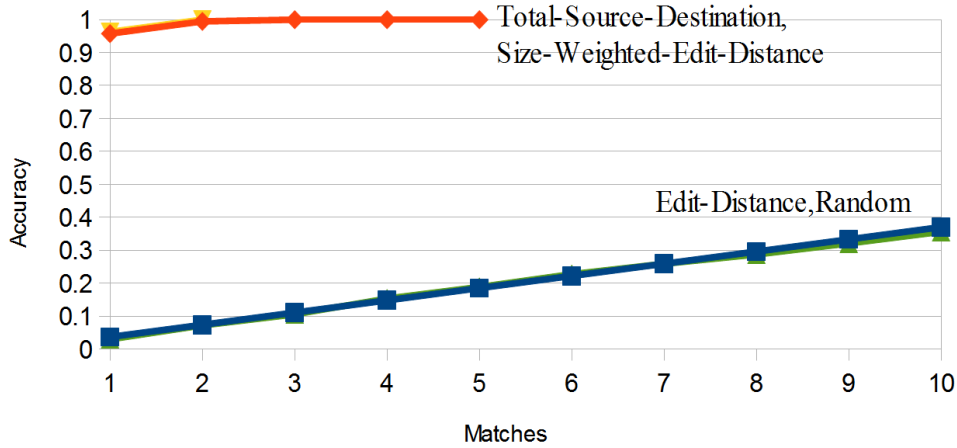
Fisher Criterion Calculations	
Total-Source-Destination	4.13
Size-Weighted-Edit-Distance	41.7
Edit-Distance	0.00

Entropy Calculations

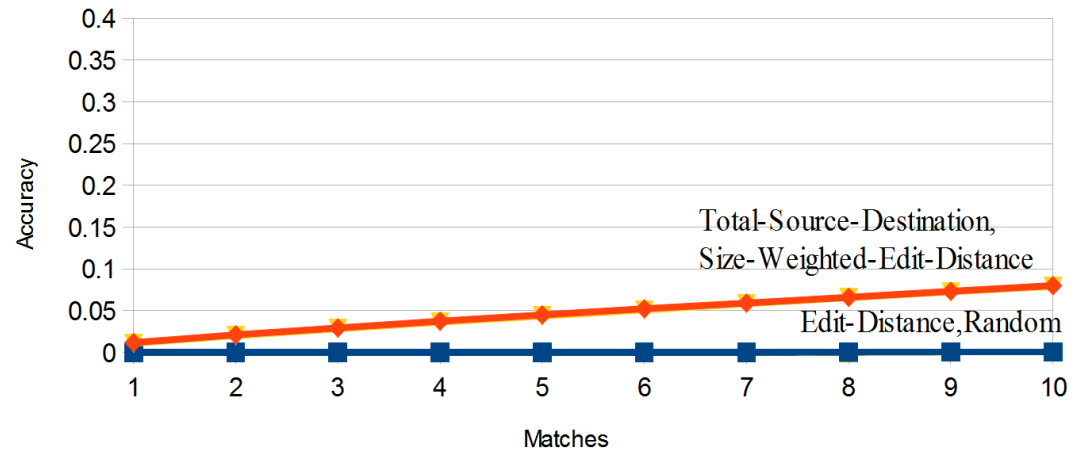
	100%	75%	50%
Desired	4.70	4.70	4.70
Total-Source-Destination	2.95	2.40	0.44
Size-Weighted-Edit-Distance	1.13	0.56	0.44
Edit-Distance	4.70	4.70	4.70

Other Applications

Bing Search Suggestions

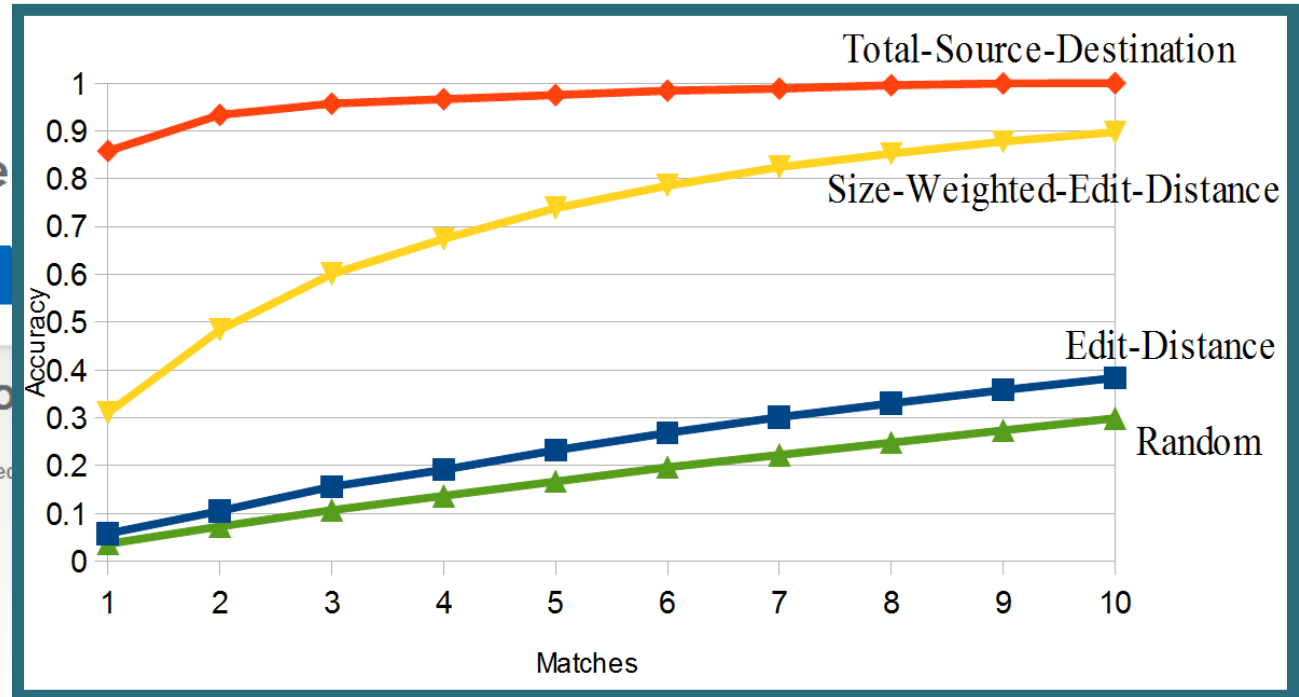


Yahoo Search Suggestions



Other Applications

NHS Symptom Checker



Find your symptom checker

1 Introduction > 2 Who is the checker for?

Please select the general area of concern

Some symptoms are covered by more than one health and symptom checker or enquiry:

- Rashes or skin problems [Help]
- Pregnancy problems [Help]
- Accident, wound or injury [Help]
- Stomach, bowel and bladder [Help]
- Head and neck [Help]
- General health [Help]

- Bones and muscles [Help]
- Children's health [Help]
- Ear, nose and throat [Help]

◀ Previous

Next ▶

See paper for Google Health Find-A-Doctor

Evaluating Defenses

With black-box approach,
evaluating defenses is easy!

HTTPOS: Sealing Information Leaks with Browser-side Obfuscation of Encrypted Flows

Xiapu Luo^{§*}, Peng Zhou[§], Edmond W. W. Chan[§], Wenke Lee[†], Rocky K. C. Chang[§], Roberto Perdisci[‡]
The Hong Kong Polytechnic University[§], Georgia Institute of Technology[†], University of Georgia[‡]
{csxluo, cspzhouroc, cswchan, csrchang}@comp.polyu.edu.hk, wenke@cc.gatech.edu, perdisci@cs.uga.edu

Abstract

Leakage of private information from web applications—even when the traffic is encrypted—is a major security threat to many applications that use HTTP for data deliv-

be profiled from traffic features [29]. A common approach to preventing leaks is to obfuscate the encrypted traffic by changing the statistical features of traffic, such as packet size and packet timing information.

Existing methods for defending against information

NDSS 2011

HTTPOS Search Suggestions

Before HTTPOS

(matches)

	1	10
Random	2.9%	35.6%
Total-Source-Destination	46.1%	100%
Size-Weighted-Edit-Distance	46.1%	100%
Edit-Distance	3.8%	39.5%

(matches)

After HTTPOS

	1	10
Random	2.9%	35.6%
Total-Source-Destination	3.4%	38.0%
Size-Weighted-Edit-Distance	3.8%	38.0%
Edit-Distance	3.4%	35.5%

HTTPOS Search Suggestions

Before HTTPOS

Fisher Criterion Calculations	
Total-Source-Destination	4.13
Size-Weighted-Edit-Distance	41.7
Edit-Distance	0.00

After HTTPOS

Fisher Criterion Calculations	
Total-Source-Destination	0.28
Size-Weighted-Edit-Distance	0.43
Edit-Distance	0.14

HTTPOS works well with search suggestions

HTTPOS Google Instant

Before HTTPOS

(matches)

	1	10
Random	2.9%	35.6%
Total-Source-Destination	47.5%	88.3%
Size-Weighted-Edit-Distance	7.3%	52.6%
Edit-Distance	7.7%	56.0%

(matches)

After HTTPOS

	1	10
Random	2.9%	35.6%
Total-Source-Destination	43.7%	87.6%
Size-Weighted-Edit-Distance	8.2%	51.4%
Edit-Distance	8.7%	55.0%

HTTPOS Google Instant

Before HTTPOS

Fisher Criterion Calculations	
Total-Source-Destination	1.13
Size-Weighted-Edit-Distance	0.34
Edit-Distance	0.22

After HTTPOS

Fisher Criterion Calculations	
Total-Source-Destination	0.60
Size-Weighted-Edit-Distance	0.55
Edit-Distance	0.47

No training phase, so HTTPOS works well with search suggestions, but not entire pages

Summary

Evaluated real web apps and a proposed defense system

Developed Fisher Criterion as an alternative measurement for information leaks in this domain

With a tutorial

Code available now: <http://www.cs.virginia.edu/sca>