TROGUARD:

Context-Aware Protection Against Web-Based Socially Engineered Trojans

Rui Han, Alejandro Mesa, University of Miami
Mihai Christodorescu, QualComm Research
Saman Zonouz, Rutgers University

UNIVERSITY OF MIAMI

QUALCOMM

RUTGERS NEW BRUNSWICK

Motivation

Eyecandy for your GNOME-Desktop - GNOME-Look.org - Mozilla Firefox

Eyecandy for your G...

gnome-look.org/index.php?xsortmode=high&page=0&xcontentmode=15x25x36x39x60

Google

GNOME-LOOK.ORG

EYECANDY FOR YOUR GNOME DESKTOP

Artwork | News | Groups | Knowledge | Events | Forum | People | Jobs

view | add artwork | my artwork | search content

ARTWORK

All
Wallpapers
GTK 1.x
GTK 2.x
GTK 3.x
Metacity
GNOME Shell
Compiz
Beryl

Latest | Alphabetical | Highest rated | Most downloaded

Faenza 1.3.1
GNOME Icon Theme
523 Comments 301 Fans
Score 93%
More

Conky Collection for Linux
Screenlet
Score 90%
More

```
done
```

Mac OS threats

Rank	Name	Percentage
1	**Trojan.OSX.FakeCo.a**	**52%**
2	**Trojan-Downloader.OSX.Jahlav.d**	**8%**
3	**Trojan-Downloader.OSX.Flashfake.ai**	**7%**
4	**Trojan-Downloader.OSX.FavDonw.c**	**5%**
5	**Trojan-Downloader.OSX.FavDonw.a**	**2%**
6	**Trojan-Downloader.OSX.Flashfake.ab**	**2%**
7	**Trojan-FakeAV.OSX.Defma.gen**	**2%**
8	**Trojan-FakeAV.OSX.Defma.f**	**1%**
9	Exploit.OSX.Smid.b	1%
10	**Trojan-Downloader.OSX.Flashfake.af**	**1%**

McAfee antivirus solution: http:www.securelist.com

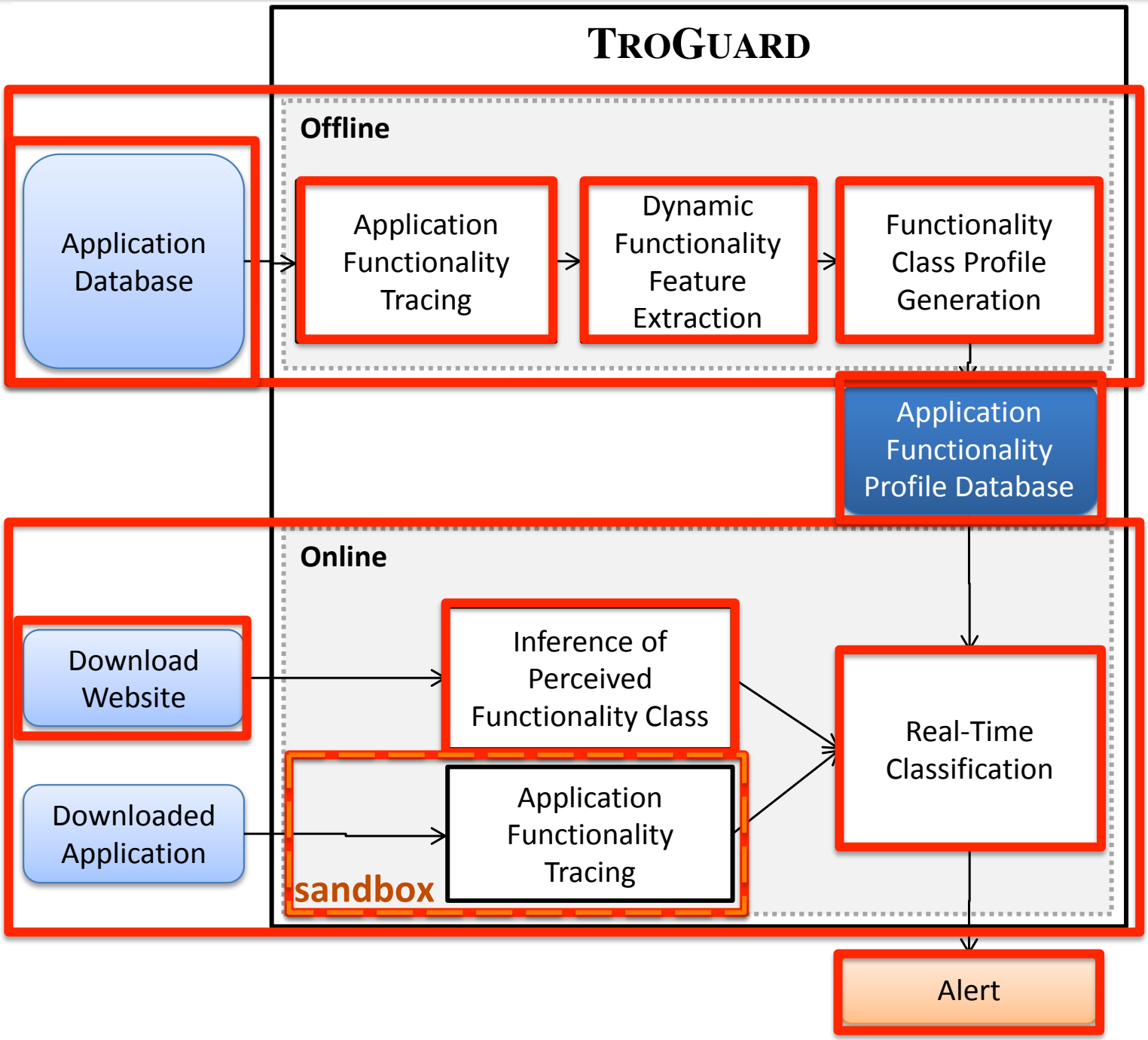
Example Malwares

Malware	Descriptions	Platform
TrojanClicker.VB .395	Trojan socially engineered as adobe flash update	Windows and Mac OS X
Faked Anti-Virus	Trojan or Adware socially engineered as anti-virus software	Windows, Mac OS X, and Linux
Opfake Browser	Malware socially engineered as Opera Browser	Android
WireLuker	Legitimate applications socially engineered with ad-wares and Trojan	Mac OS X and iOS

Contributions

- Answer the question: "Is this program doing what I expected it to do?"
- Bridge the semantic gap between functionality classes and low level behaviors
- Built on 100 Linux app profiles
- High detection rate on 50 Trojan apps

TROGUARD Architecture



Key Premise

- TROGUARD detects Trojans based on the premise that applications with **similar functionalities** expose **similar system-level behaviors**
- Applications with similar functionalities belong to a **functionality class**, they should exhibit common system level behaviors
  - *Learn web-browser behavior of well know instances(e.g., Firefox and Chrome)*
  - *Compare the web-browser profile with the behaviors of unknown downloaded web-browser app*

Functionality class

- It represents both user's understanding of software category and the system's observation of a software execution behavior

Functionality class	softpedia.com	download.cnet.com	tucows.com
Graphic Editor	Artistic software	Graphic Design SW	Design tools
Game	Games	Games	Games
Browser	Internet	Browsers	Internet
Instant Messenger	Communications	Communications	
Media Player	Multimedia	MP3/Audio Software	Audio/Video
Audio Editor			
Video Editor			
Text Editor	office	Productivity software	Business
IDE	Programming	Developer Tools	Develop/Web
Calculator	Utilities	Utilities	Home/Education

Applications

Class	Studied Applications
1. Graphic Editor	gimp, pinta, imagej, inkscape, kolourpaint, rawtherapee, mypaint, gpaint, gnome-paint, pencil
2. Games	sol, wesnoth, glchess, neverball, kmahjongg, supertuxkart, hedge- wars, pingus, frozen-bubble, eboard
3. Browser	chrome, firefox , opera, epiphany, midori, chromium, netsurf, arora, xxxterm, rekonq
4. Instant Messenger	skype, kmess, emesene, kopete, pidgin, psi, gajim, empathy, amsn, qutim
5. Media Player	smplayer, vlc, audacious, quodli- bet, gmusicbrowser, qmmp, abraca, amarok, guayadeque, aqualung
6. Audio Editor	audacity, avidemux, dvbcut, og- gconvert, kwave, wavbreaker, mp3splt-gtk, mhwaveedit, fillmore, soundconverter
7. Video Editor	openshot, lives, iriverter, kino, pitivi, videocut, winff, arista-gtk, kdenlive, curlew
8. Text Editor	kile, geany, texmaker, calligra- words, soffice.bin, lyx, tea, jed, emacs, vi
9. IDE	anjuta, codelite, codeblocks, net- beans, monodevelop, kdevelop, spyder, monkeystudio, drracket, idle
10. Calculator	grpn, gcalctool, EdenMath, speed- crunch, kcalc, keurocalc, extcalc, gip, galculator, gnome-genius

Functionality Tracing

- Manual testing
- Run 60 seconds for each application
- System call trace
- User-space information

  User interactivity

  Resource consumption

  IP addresses and port number

Feature Extraction

- Processing tracing data
- Four groups of feature

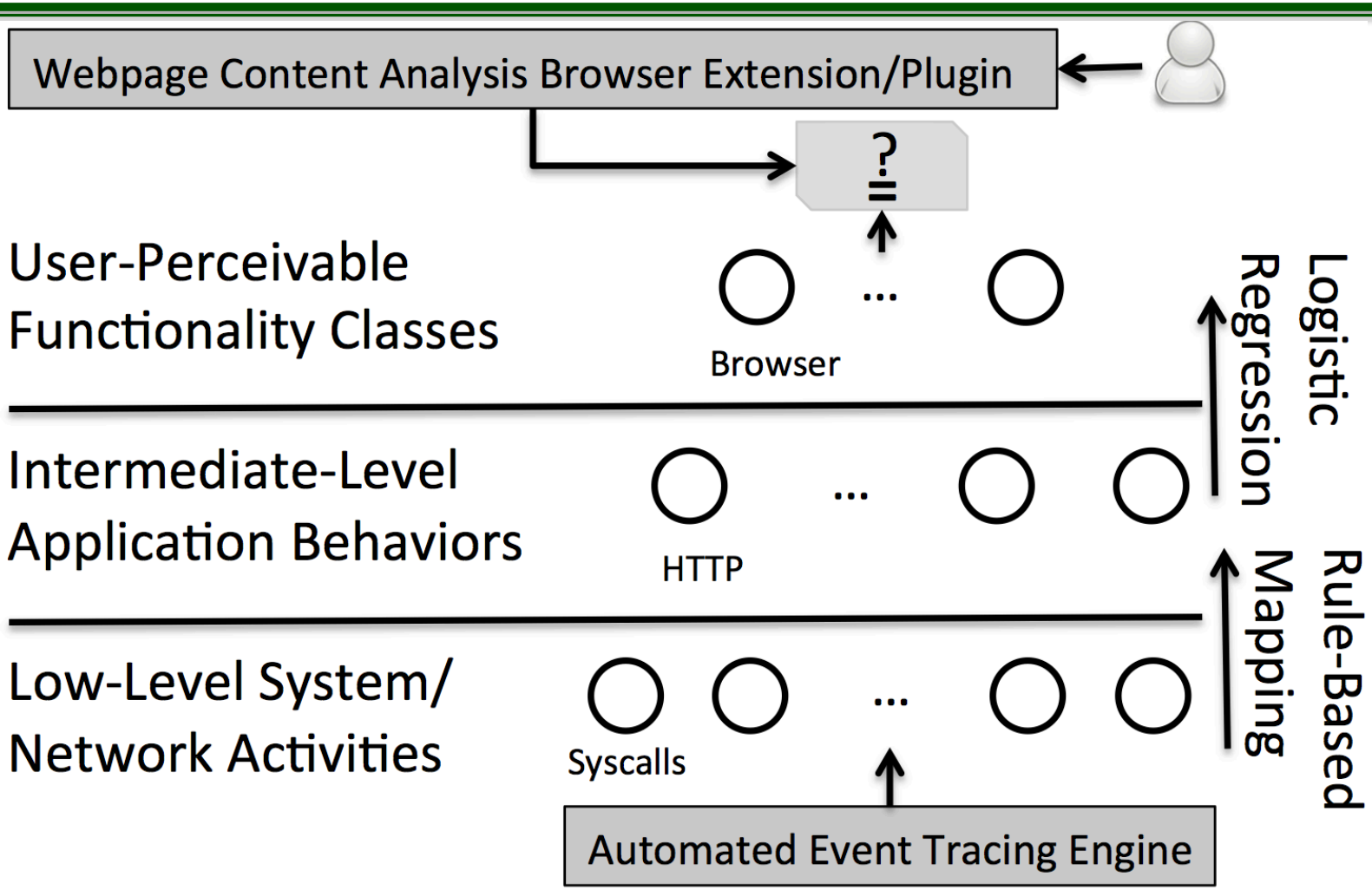
  file system

  Network

  resource usage

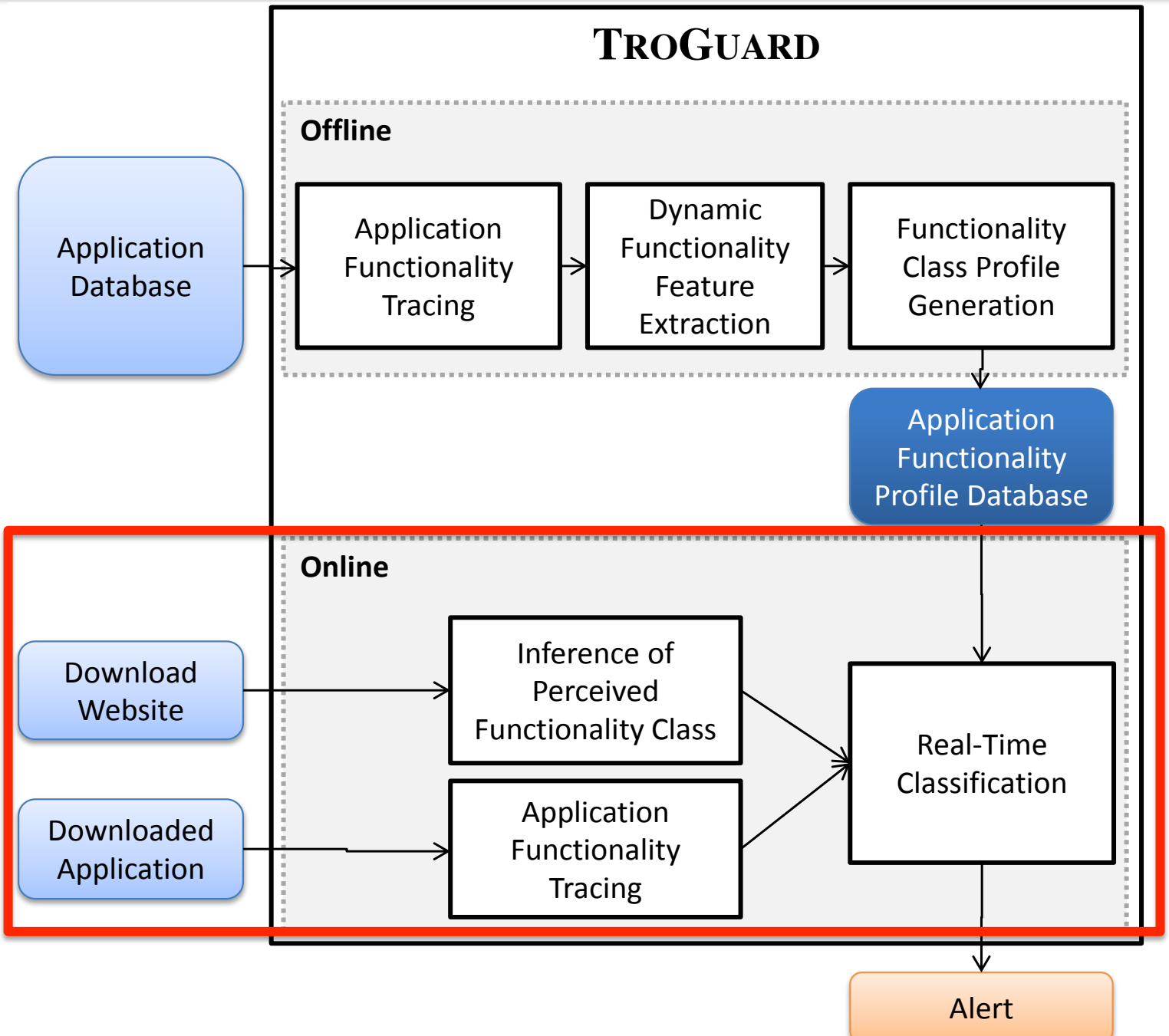
  user interactivity
- Intermediate feature

Intermediate Features



Example: **if(libssl3.so & fd = sys_socket(AF_INET, ..)**
&sys_write(fd, ..) & sys_read(fd, ..))
HTTP = true

TROGUARD Architecture

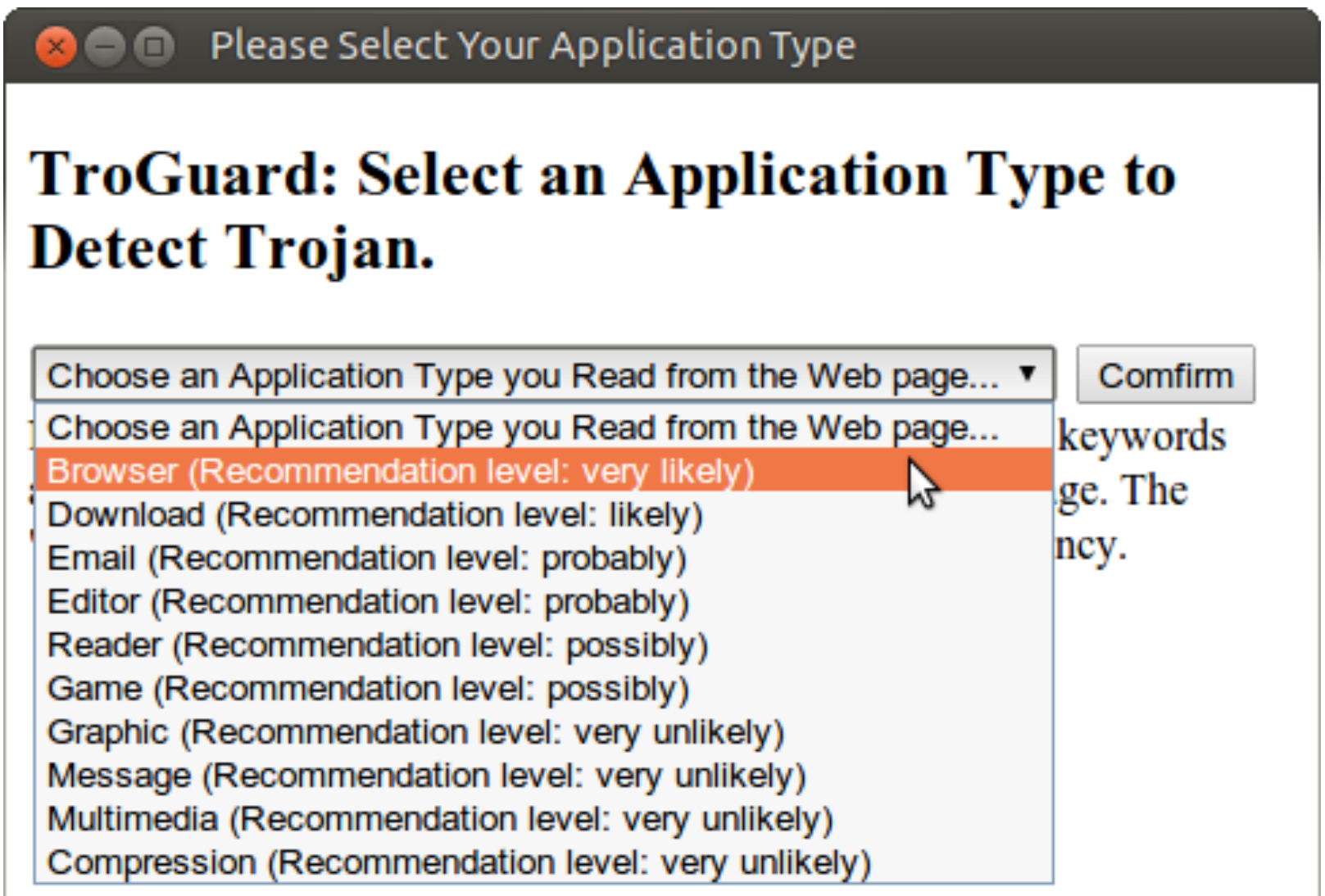


Web Page Analysis

- Give the explicit functionality class
- Web page contents analysis
- OCR to extract the texts in the images
- Analysis based on keywords

User Interface

- Browser extension for web page analysis



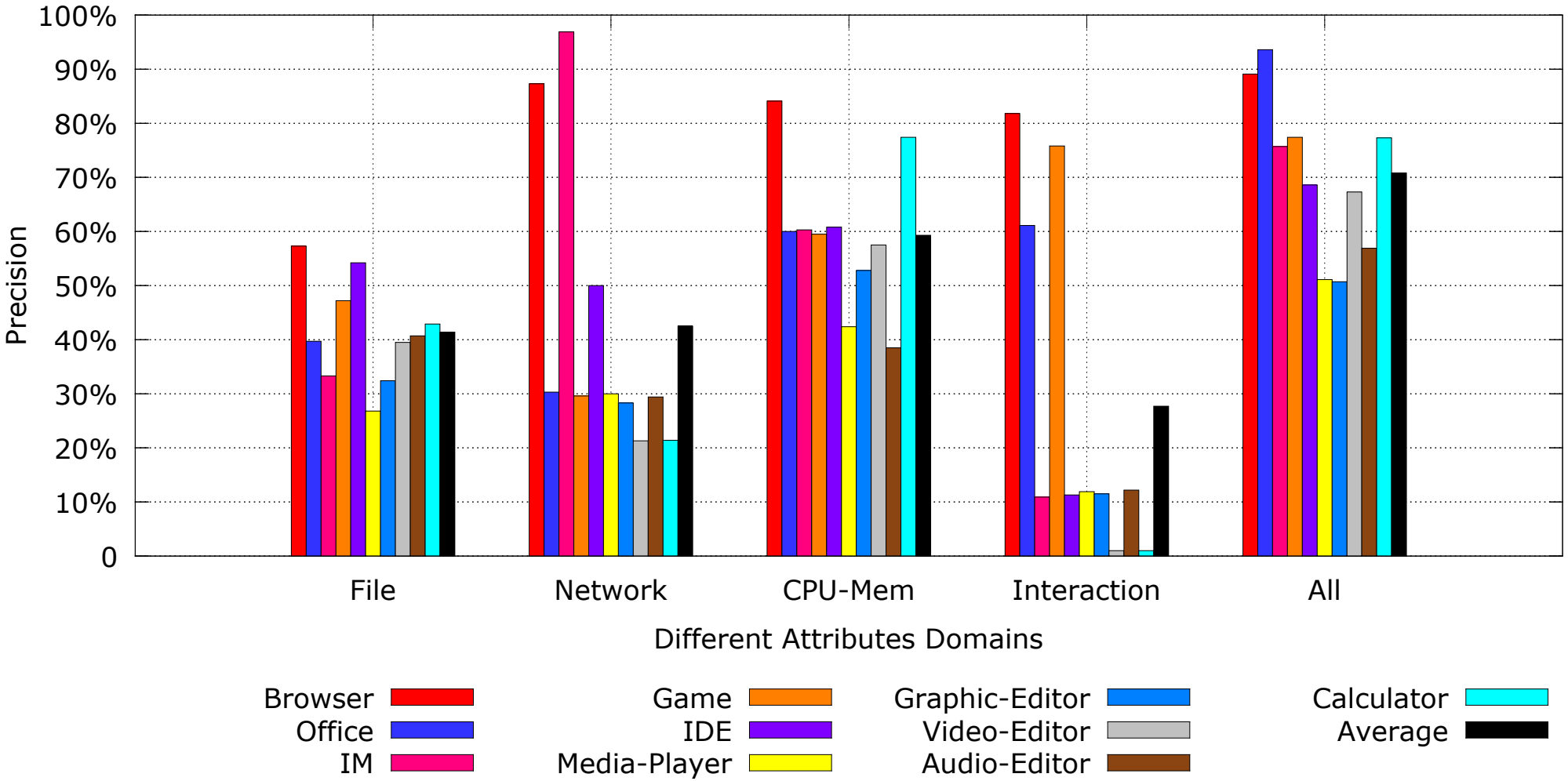
Sandboxing

- SELinux sandbox
- One policy for each app class
- Automatically generated by parsing all the logs from an app class

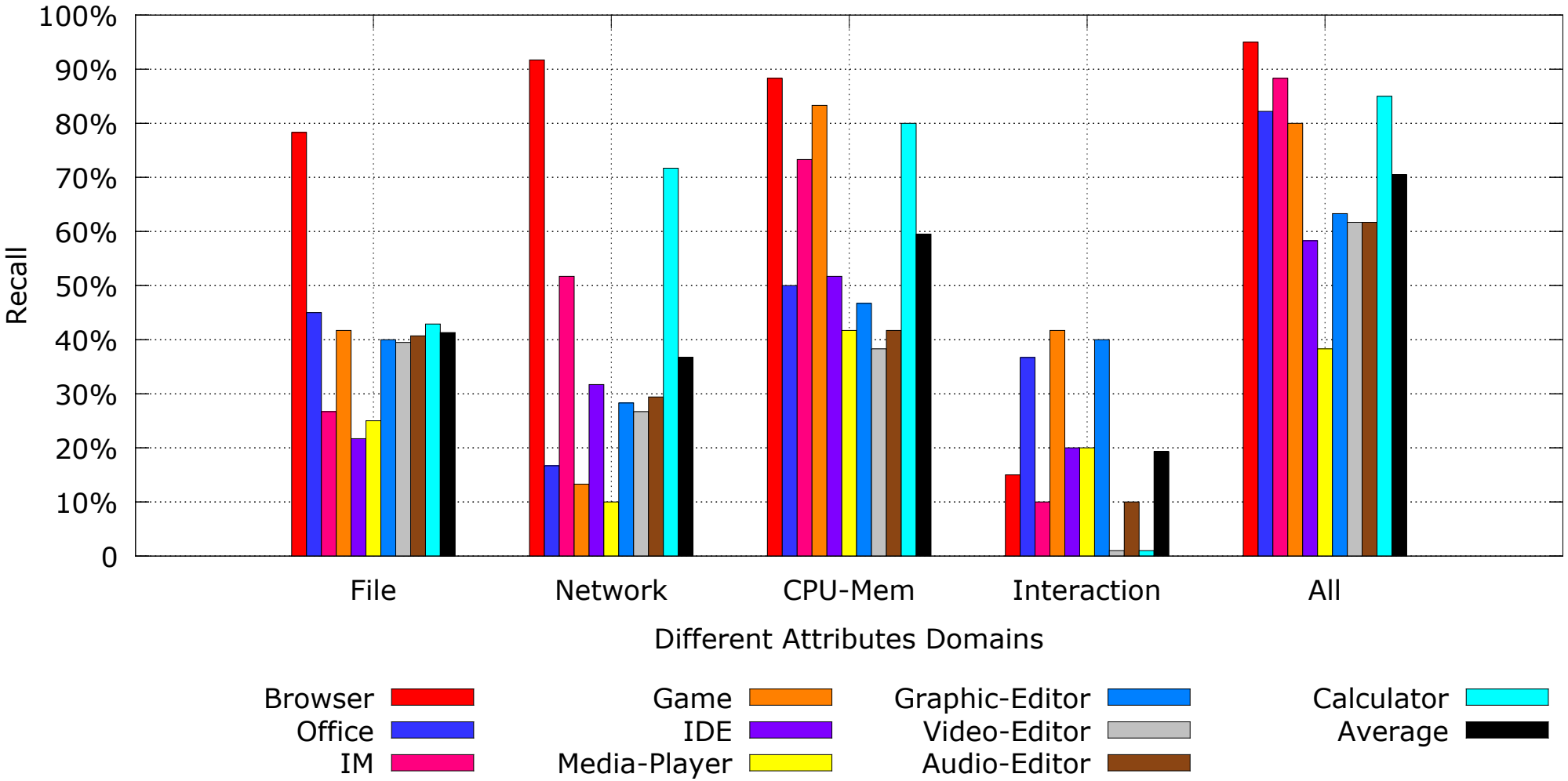
Classifier Evaluation

- 600 data points (10 second each)
- 10 fold cross validation
- 5 classifiers with different feature group
- Precision
- Recall
- Confusion Matrix

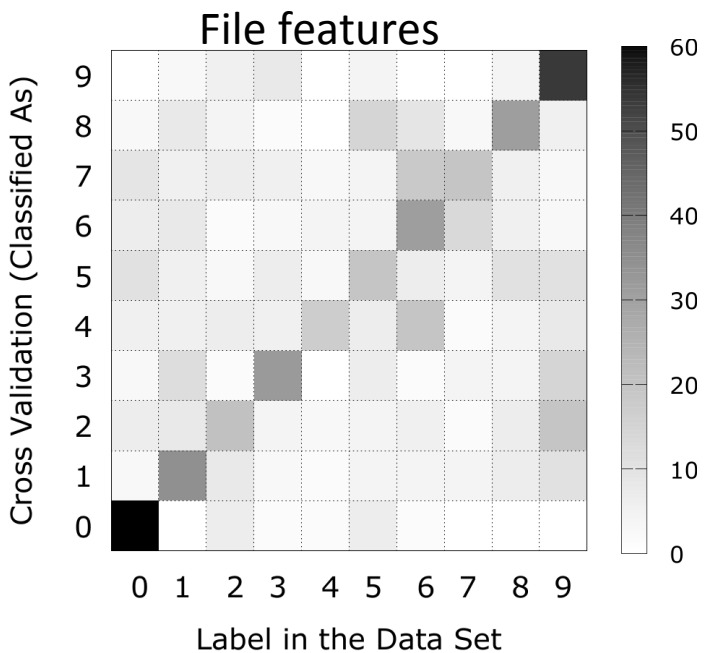
Precision

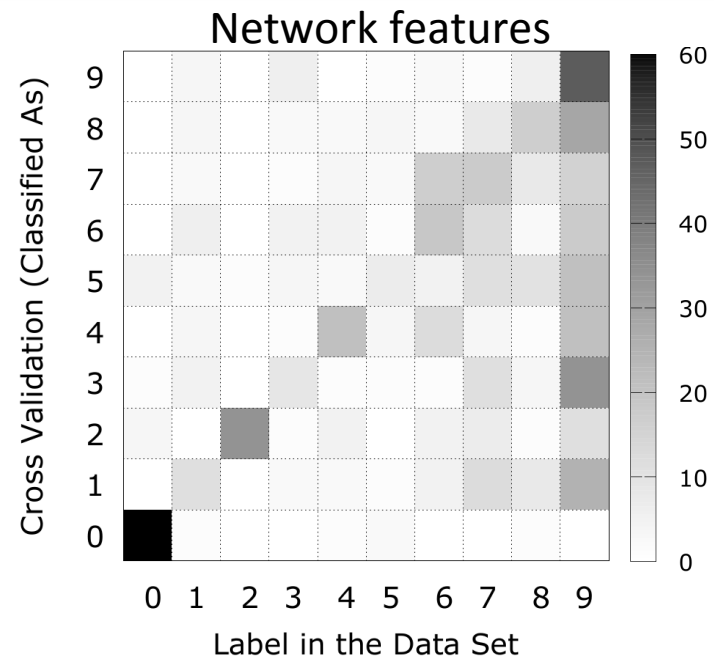

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0
Precision
File
Network
CPU-Mem
Interaction
All
Different Attributes Domains
Browser
Office
IM
Game
IDE
Media-Player
Graphic-Editor
Video-Editor
Audio-Editor
Calculator
Average

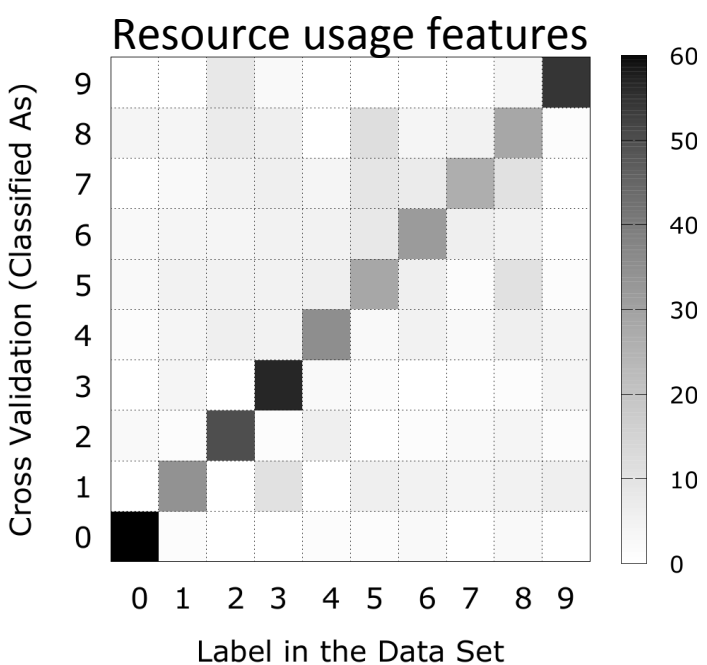
Recall

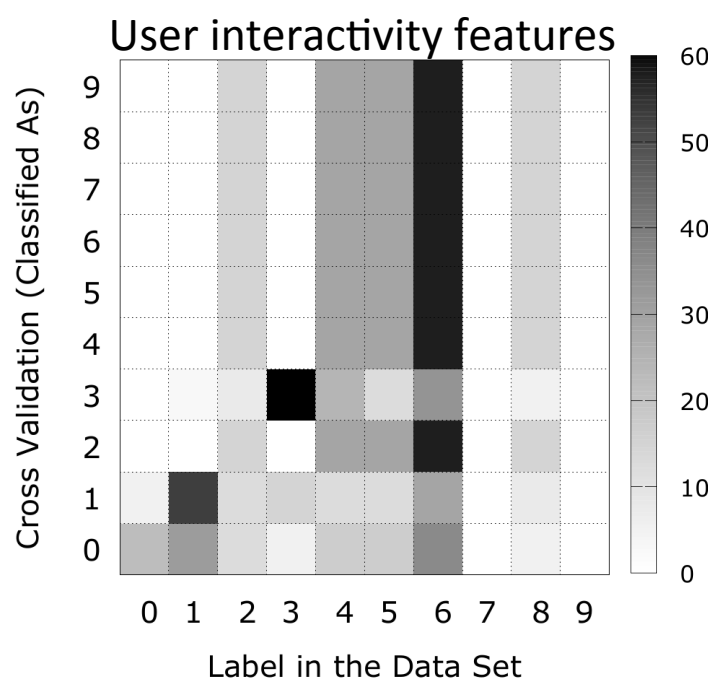

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0
Recall
File
Network
CPU-Mem
Interaction
All
Different Attributes Domains
Browser
Office
IM
Game
IDE
Media-Player
Graphic-Editor
Video-Editor
Audio-Editor
Calculator
Average

Confusion Matrices

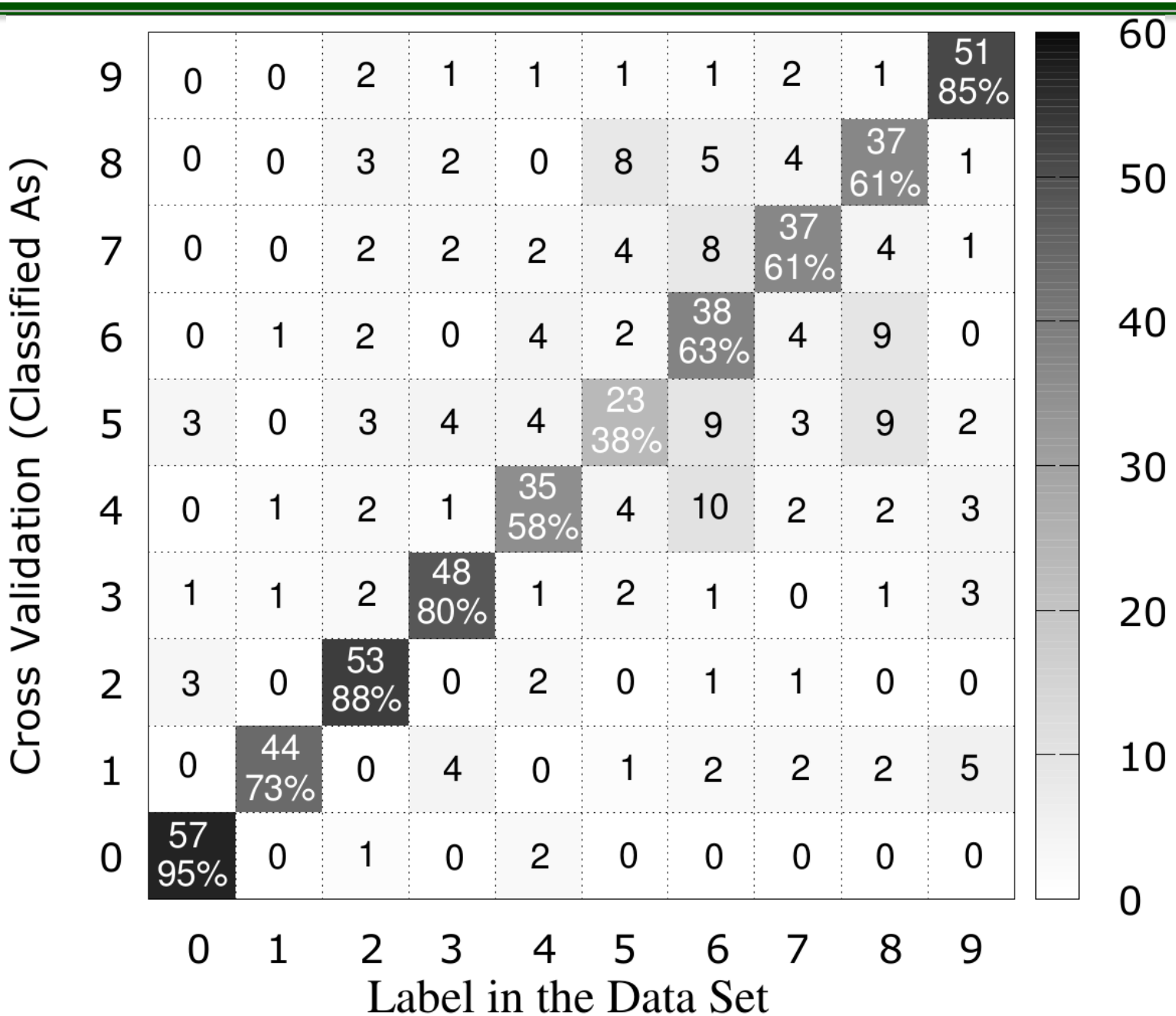




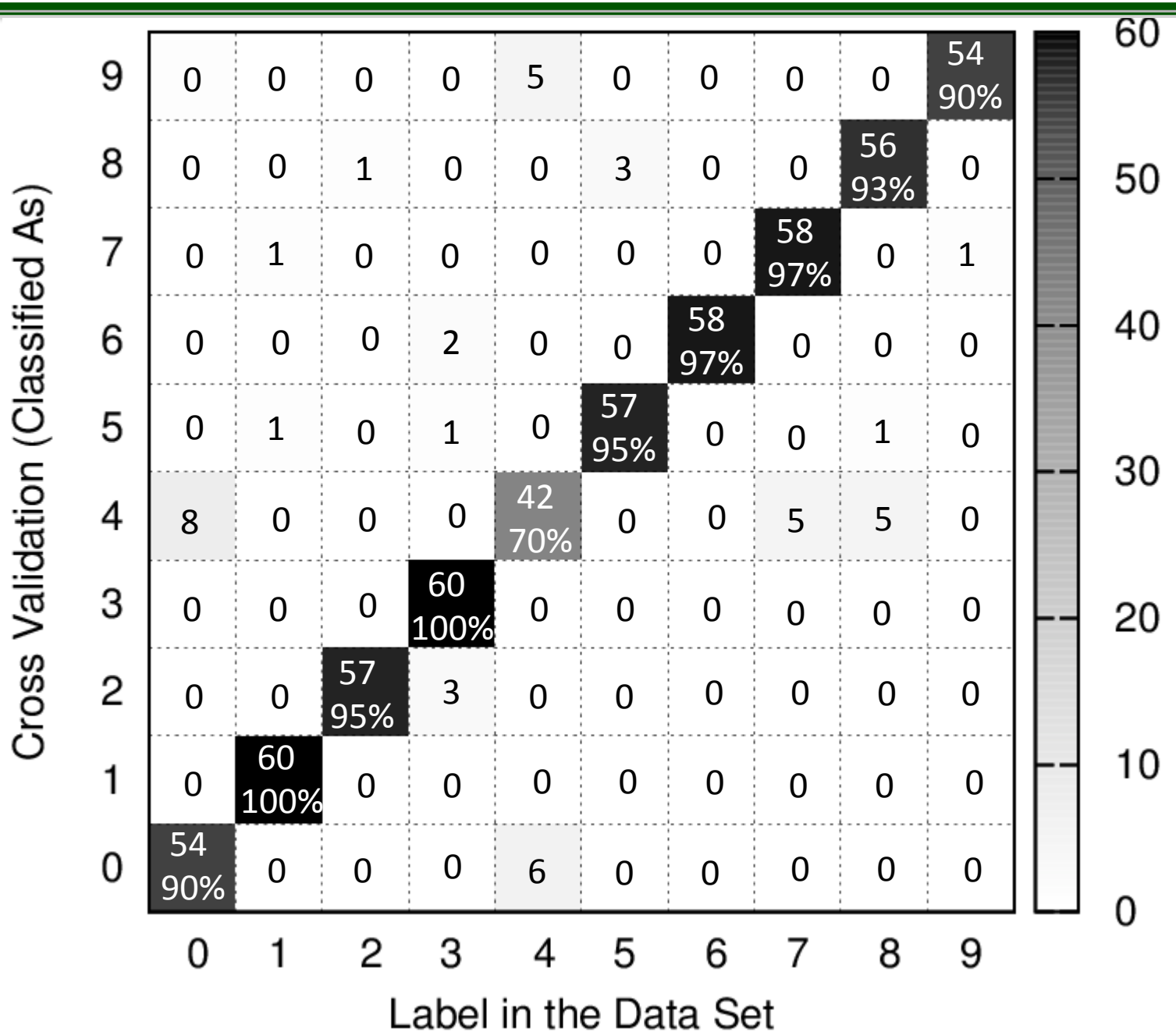




Confusion Matrices

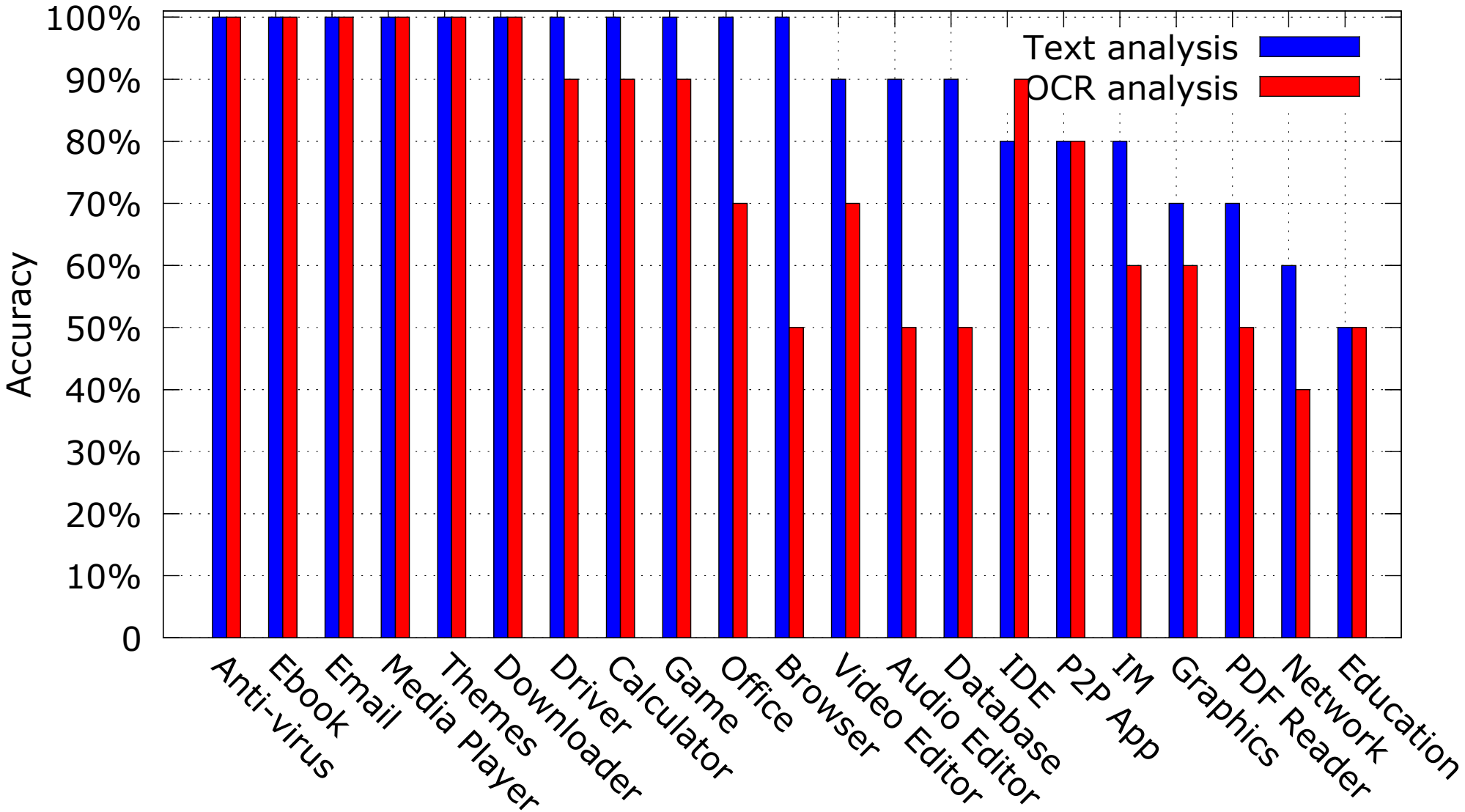


Intermediate Feature Results



Web Page Analysis Accuracy

- 100 Web page, 20 categories



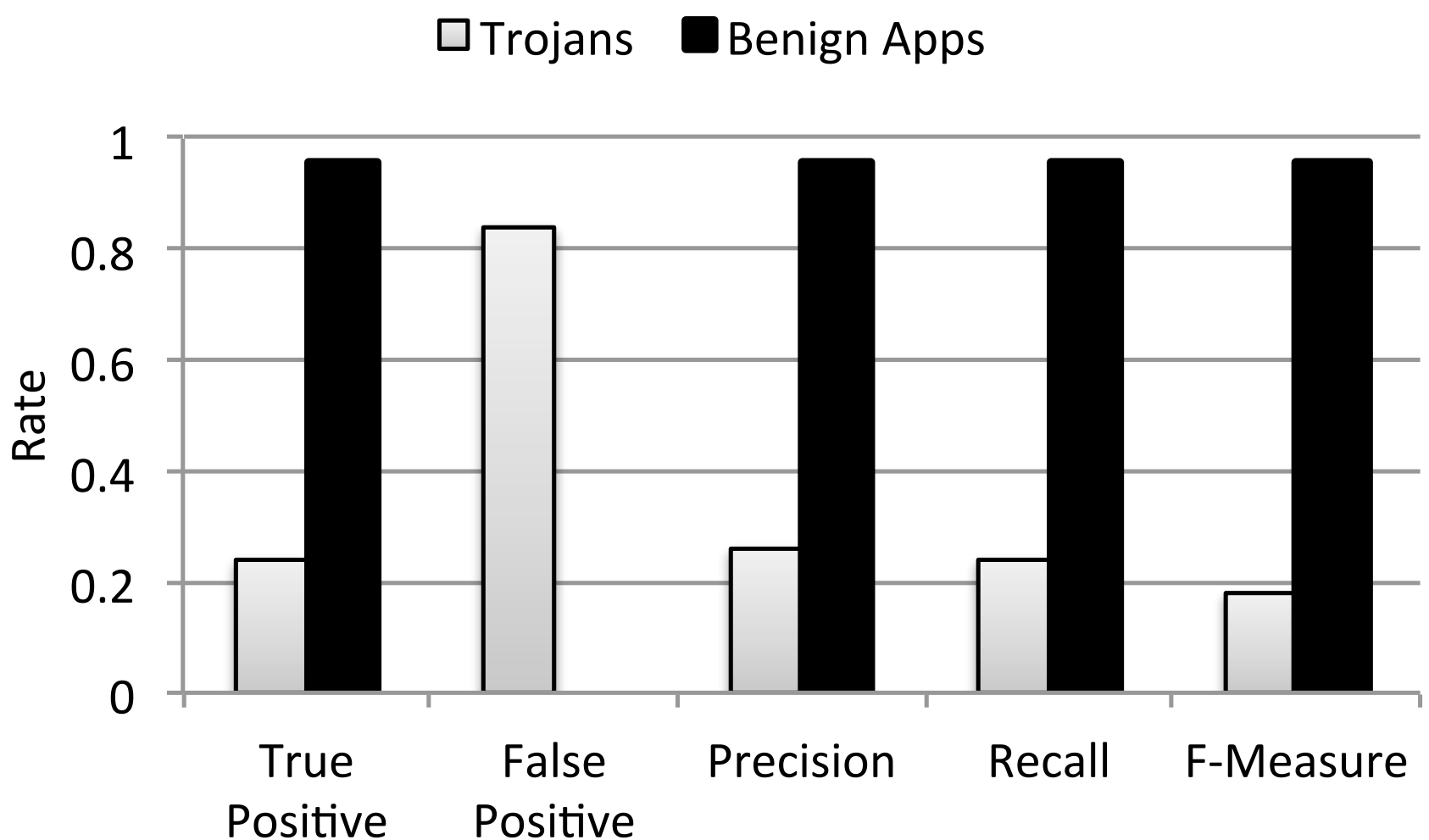
Case Study

- 10 benign apps ×5 payload = 50 Trojans

Functionality Class	Application	Metasploit Payload
1. Graphic Editor	gpaint	linux/x86/shell_bind_tcp linux/x86/shell/reverse_tcp linux/x86/vncinject/bind_tcp linux/x86/meterpreter/bind_tcp linux/x86/download_exec
2. Games	eboard	
3. Browser	xxxterm	
4. Instant Messenger	psi	
5. Media Player	qmmp	
6. Audio Editor	winff	
7. Video Editor	fillmore	
8. Text Editor	tea	
9. IDE	spyder	
10. Calculator	gnome-genius	

Case Study

- Predefined acceptance rate 0.8



Symbolic Execution

- Tested Core Utilities (four functionality classes)

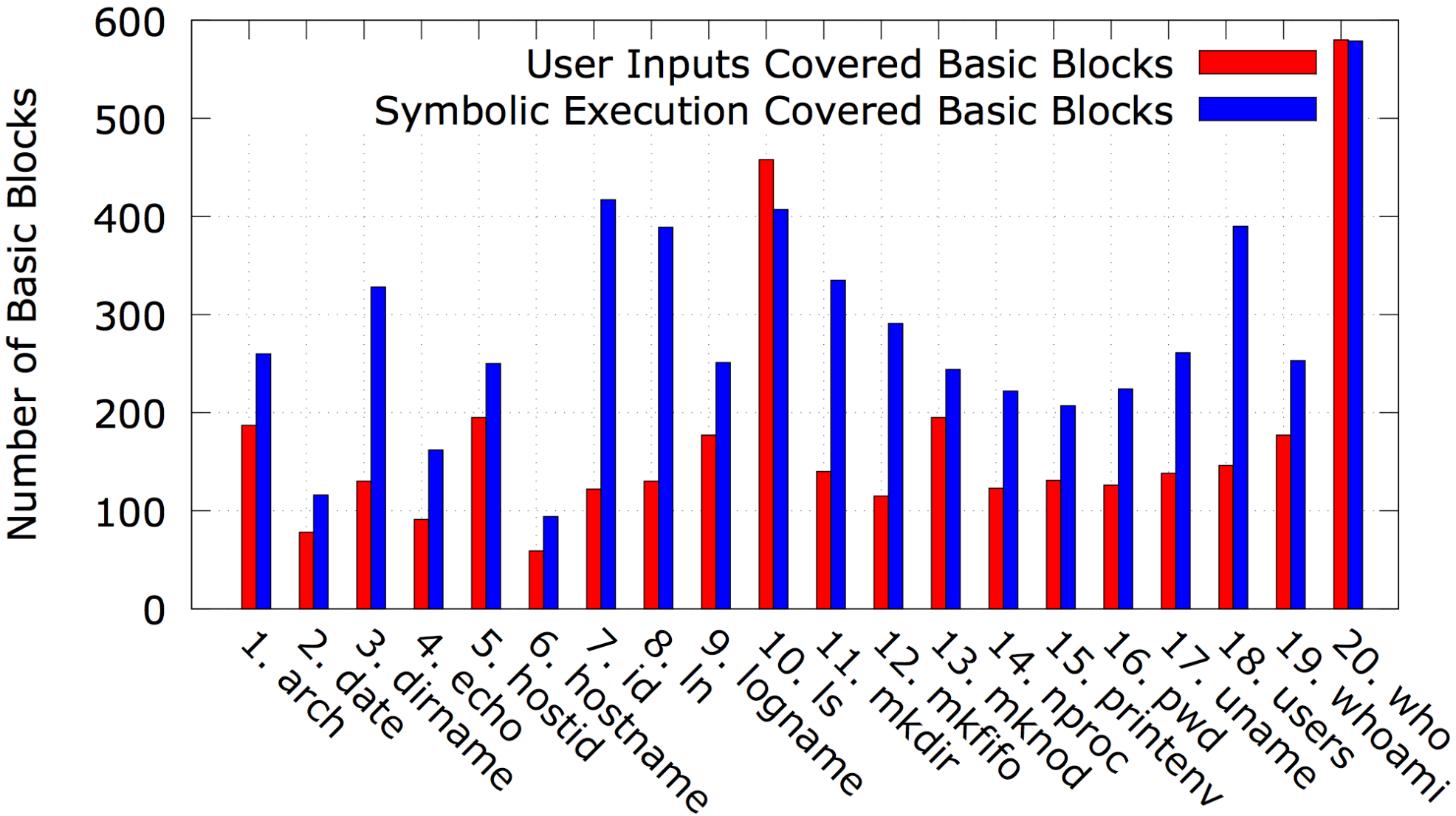
  Dirlist

  Filetype

  Userinfo

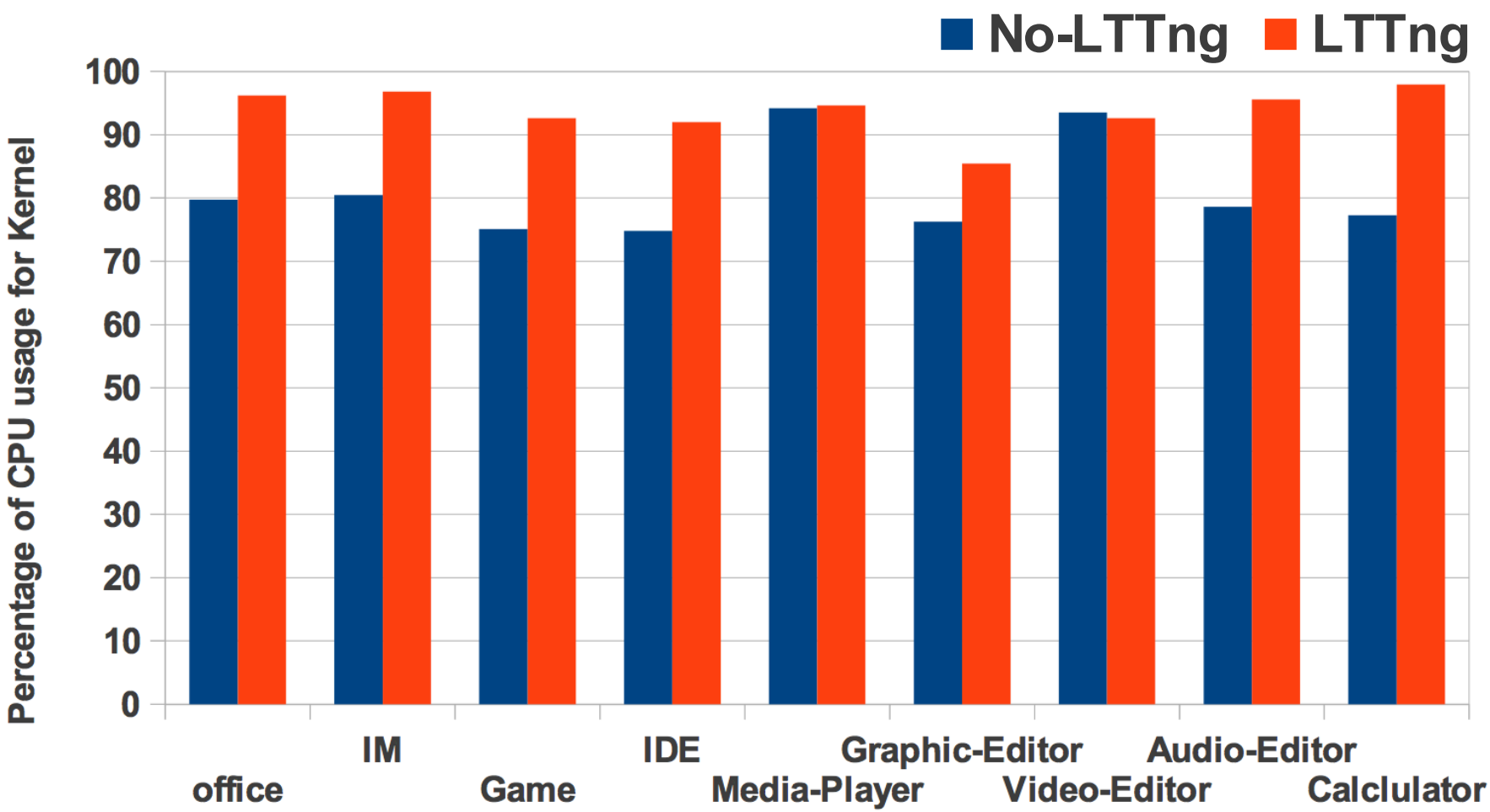
  Systeminfo
- Features collected from symbolic execution give us 52% precision
- Features collected from user execution give us 76% precision

User Execution VS Symbolic Execution



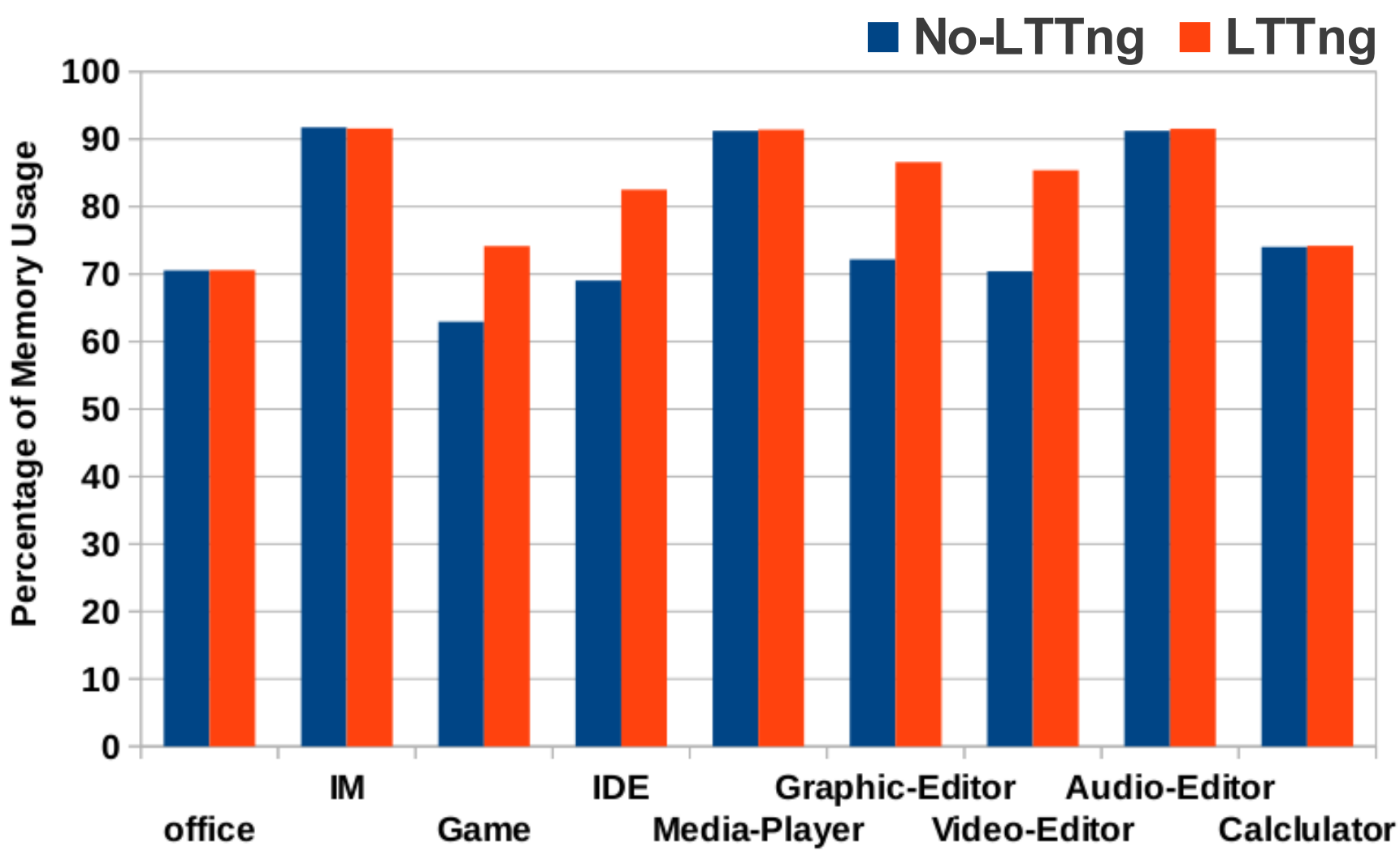
Performance Overhead

- CPU usage:



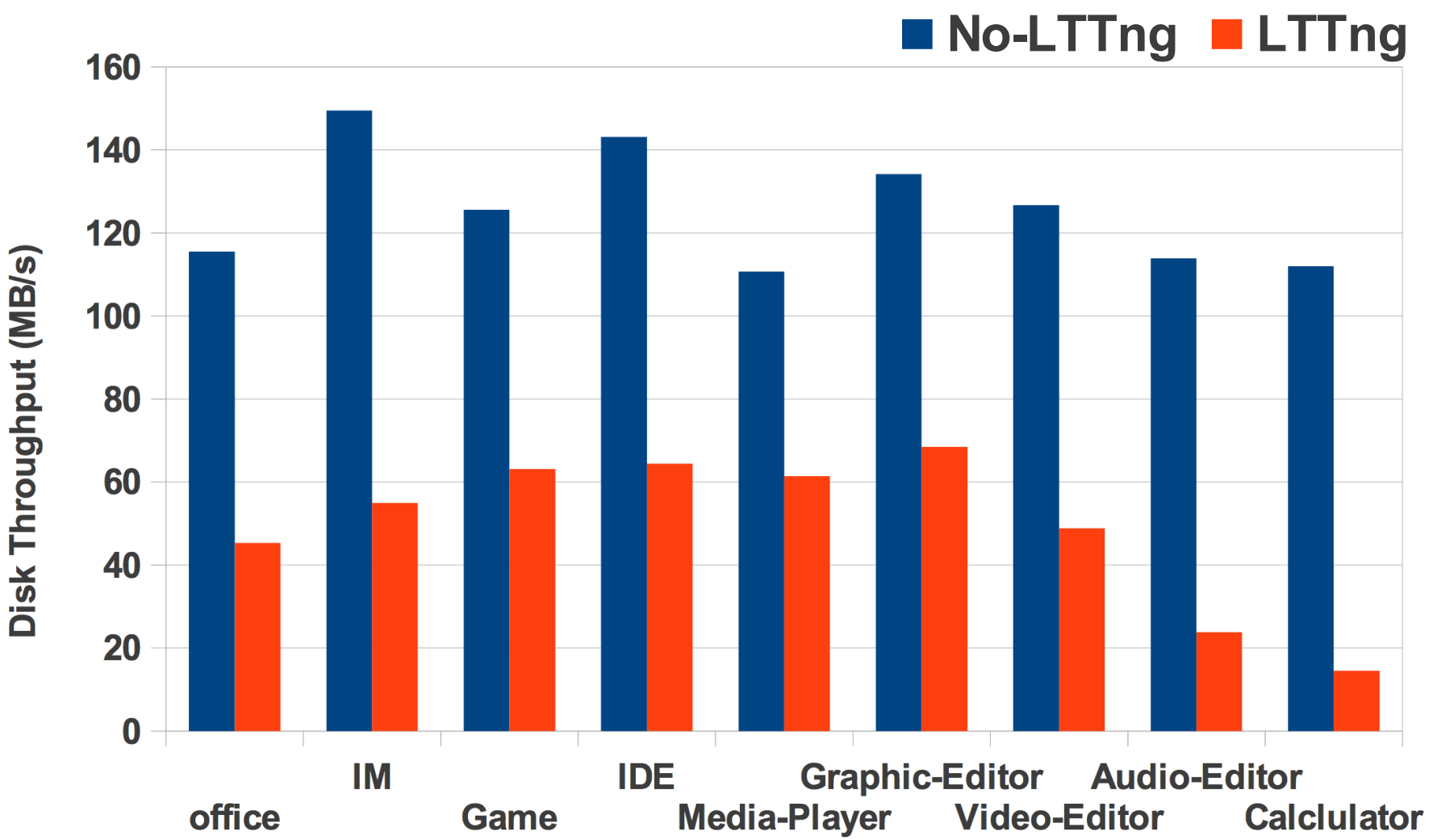
Performance Overhead

- Memory usage:



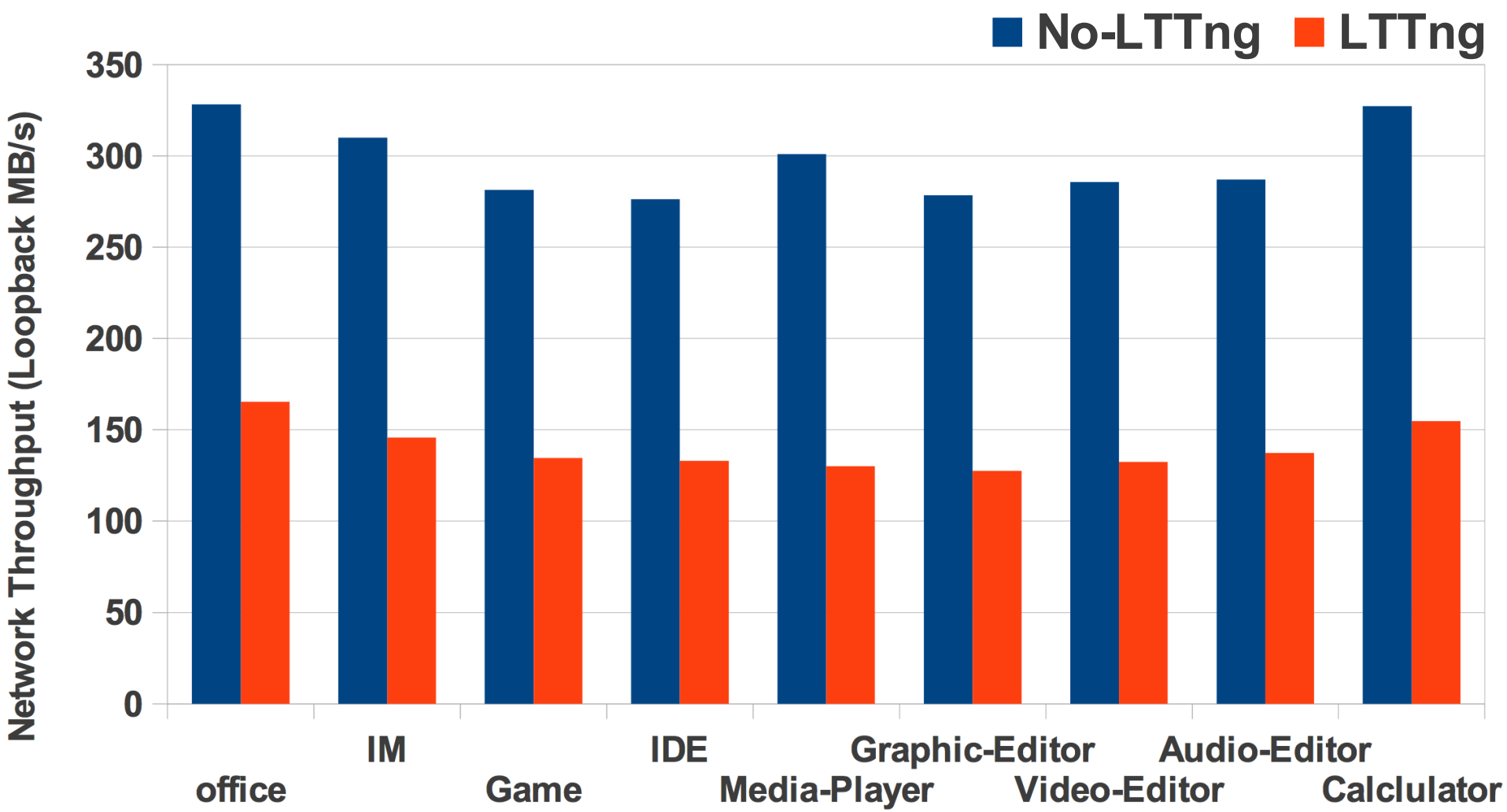
Performance Overhead

- Disk throughput:



Performance Overhead

- Network throughput:



Conclusions

- TROGUARD detects Trojans based on the premise that applications with **similar functionalities** expose **similar system-level behaviors**
- TROGUARD can detect Trojan application download by bridging the gap between the user perceived functions and genuine software functions

Thank you!

Questions?

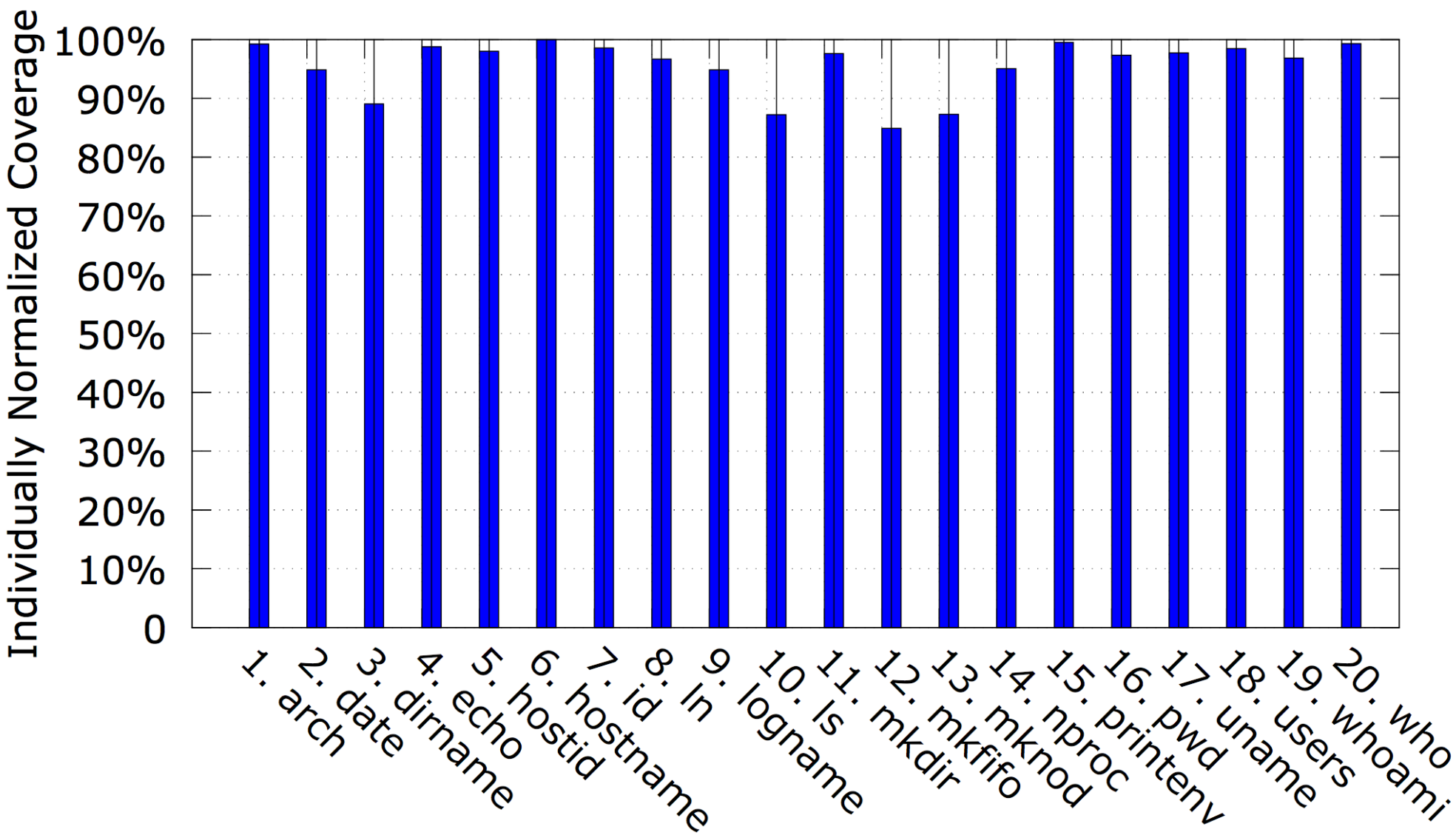
Rui Han
r.han@umiami.edu

Conclusions

- TROGUARD detects Trojans based on the premise that applications with **similar functionalities** expose **similar system-level behaviors**
- TROGUARD can detect Trojan application download by bridging the gap between the user perceived functions and genuine software functions

Questions?

Symbolic Execution Code Coverage



SE Code Coverage Evolution