## Understanding the Ethical Frameworks of Internet Measurement Studies

**Eric Pauley**, Patrick McDaniel University of Wisconsin-Madison



#### Applications of Internet Measurement

#### Botnets (e.g., Mirai)



#### User Behavior (e.g., CrUX)



Ruth et al. '22 [2]

#### **Novel Vulnerabilities**

Cloud Squatting	: Vulnerability at S	Scale 🕅 PennState
Cloud Services	Third-Party Services	DNS
<ul> <li>&gt;5M messages</li> </ul>	<ul> <li>&gt;3M messages</li> </ul>	<ul> <li>5400 Websites</li> </ul>
<ul> <li>4 cloud services</li> </ul>	<ul> <li>Numerous Services</li> </ul>	• 23 top-1000
⑦ ∅ ∅ ∅	\$\$ <b>1</b>	
Example Sens	sitive Data Received	intel 📲
Financial Personal Location	Remote Code Passwords Images	
		T II IY 🤜 redhat.
Measuring and Mitigating the Risk of IP Reuse on Public Clouds		

Pauley et al. '22 [3]

[1] Understanding the Mirai Botnet

[2] A world wide view of browsing the world wide web

[3] Measuring and Mitigating the Risk of IP Reuse on Public Clouds



#### Challenge: Ethical measurement

- Increased focus on ethical security research
- Required ethics considerations at conferences
- But: what does it mean for research to be *ethical*?
  - Is it legal?
  - Is it IRB-approved (read: *exempt*)?
  - Are reviewers convinced it's ethical?
- Goal: develop a cohesive, normative framework (a *classifier?*) for ethical Internet measurement

Conference	Ethics in CFP since
ACM IMC	2009 [6]
<b>USENIX</b> Security	2013 [7]
NDSS	2015 [8]
ACM CCS	2017 [10]
ACM ASIACCS	2017 [9]
IEEE S&P	2017 [11]
IEEE EuroS&P	2017 [5]
ACM SIGMETRICS	2018 [12]
ACSAC	2021 [2]



#### An existing classifier: Institutional Review Boards (IRB)

- Required (In US) for federally-funded research
- In reality: required by Universities (and conference CFPs)
- False Accept (doing unethical research):

Failing to identify human subjects
 Incomplete/missing anonymization
 Unforeseen harms

False Reject (rejecting ethical research):

Reasonable expectation of measurement Statistically improbable impacts





Towards a framework of IM ethical considerations

**Study Goal: understand considerations and** emergent consensus on ethical measurement

Broad expectations from venues

- **Ethical risks in papers**
- Considerations by authors

**Ultimately: develop a cohesive normative** framework for ethical Internet measurement



**Building an Ethical Classifier?** 



Venue	Vantage Point		
ASIACCS '18	Campus Net		
IMC '19	DNS Resolver		
IMC '19	CDN IPs		
CCS '21	Cloud IPs		
SEC '21	Cloud IPs		
EuroS&PW '22	Campus Net		
SEC '22	Container Registries		
S&P '22	Cloud IPs		
IMC '22	Web Browser		
IMC '22	Darknet		

Data points: accepted conference papers

Venues with ethics in CFP

• Features: presence and mitigation of possible ethical concerns



Venue	Vantage Point	Data Collected		
ASIACCS '18	Campus Net	Transport-Layer		
IMC '19	DNS Resolver	DNS Queries		
IMC '19	CDN IPs	Transport Layer		
CCS '21	Cloud IPs	DDoS Traffic		
SEC '21	Cloud IPs	Application Layer		
EuroS&PW '22	Campus Net	Application Layer		
SEC '22	Container Registries	Download counts		
S&P '22	Cloud IPs	Application Layer		
IMC '22	Web Browser	Aggregate Browsing Behavior		
IMC '22	Darknet	Passive IP + DNS		

Works tend to properly scope ethical considerations to data collected.





Venue	Vantage Point	Data Collected	Target Parties	Incidental Parties	
ASIACCS '18	Campus Net	Transport-Layer	Scanners	End-Users	
IMC '19	<b>DNS</b> Resolver	DNS Queries	<b>Recursive Resolvers</b>	End-Users	
IMC '19	CDN IPs	Transport Layer	Scanners		
CCS '21	Cloud IPs	DDoS Traffic	Scanners	End-Users	
SEC '21	Cloud IPs	Application Layer	Scanners	End-Users	
EuroS&PW '22	roS&PW '22 Campus Net Application Layer		Scanners		
SEC '22 Container Registries Download counts		End-Users			
S&P '22	S&P '22 Cloud IPs Application Layer		Scanners, End-Users		
IMC '22	Web Browser	Aggregate Browsing Behavior	End-Users		
IMC '22	Darknet	Passive IP + DNS	Scanners, DNS Servers		

# Measurement papers often miss risk of incidental end-user data collection.



Venue	Vantage Point	Data Collected	Target Parties	Incidental Parties	Ethics Sec.	Anon. <sup>1</sup>
ASIACCS '18	Campus Net	Transport-Layer	Scanners	End-Users	0	D
IMC '19	DNS Resolver	DNS Queries	<b>Recursive Resolvers</b>	End-Users		
INIC 19	CDN IPS	Transport Layer	Scanners		0	D
CCS '21	Cloud IPs	DDoS Traffic	Scanners	End-Users		D
SEC '21	Cloud IPs	Application Layer	Scanners	End-Users		D
EuroS&PW '22	Campus Net	Application Layer	Scanners		0	D
SEC '22	Container Registries	Download counts	End-Users			
S&P '22	Cloud IPs	Application Layer	Scanners, End-Users			D
IMC '22	Web Browser	Aggregate Browsing Behavior	End-Users			
INIC 22	Darknet	Passive IP + DNS	Scanners, DNS Servers		•	D

#### Anonymization can be a *technical* contribution



Venue	Vantage Point	Data Collected	Target Parties	Incidental Parties	Ethics Sec.	Anon. <sup>1</sup>	Impact <sup>2</sup>
ASIACCS '18	Campus Net	Transport-Layer	Scanners	End-Users	0	D	0
IMC '19	DNS Resolver	DNS Queries	<b>Recursive Resolvers</b>	End-Users	$\bullet$		O
IMC '19	CDN IPs	Transport Layer	Scanners		0	O	0
CCS '21	Cloud IPs	DDoS Traffic	Scanners	End-Users	$\bullet$	D	
SEC '21	Cloud IPs	Application Layer	Scanners	End-Users	$\bullet$	D	D
EuroS&PW '22	Campus Net	Application Layer	Scanners		0	D	D
SEC '22	Container Registries	Download counts	End-Users		$\bullet$		O
S&P '22	Cloud IPs	Application Layer	Scanners, End-Users		$\bullet$	O	•3
IMC '22	Web Browser	Aggregate Browsing Behavior	End-Users		$\bullet$		0
IMC '22	Darknet	Passive IP + DNS	Scanners, DNS Servers		$\bullet$	O	•3



# Studies sufficiently mitigated harms to users due to interactivity

#### Examining venue expectations

Conference	Ethics in CFP since	Latest CFP <sup>8</sup>	IRB <sup>1</sup>	Impact <sup>3</sup>	Disclosure <sup>4</sup>	Legal <sup>5</sup>	REC <sup>6</sup>	Framework <sup>7</sup>
ACM IMC	2009 [6]	2022	•	•	0	0	0	Belmont [28] (B/C)
<b>USENIX</b> Security	2013 [7]	2023	•2	$\bullet$	$\bullet$	0	$\bullet$	Menlo [22] (B)
NDSS	2015 [8]	2023	$\bullet$	0	$\bullet$	$\bullet$	0	
ACM CCS	2017 [10]	2022	$\bullet$	0	•	0	0	
ACM ASIACCS	2017 [9]	2023	0	0	$\bullet$	$\bullet$	0	
IEEE S&P	2017 [11]	2023	●2	0	0	$\bullet$	$\bullet$	
IEEE EuroS&P	2017 [5]	2023	●2	0	$\bullet$	0	0	Menlo [22] (B)
ACM SIGMETRICS	2018 [12]	2023	●2	0	0	0	0	Menlo [22] (B/C)
ACSAC	2021 [2]	2022	$\bullet$	$\bullet$	•	0	0	

#### **Recommendations:**

- Apply learnings from other venues
- Emphasize technical merit in ethical considerations



# Technical vs. Ethical "Innovation"? SOTA **Technical Innovation** No concrete framework **Reviewer discretion** Ethical "Innovation" No negative examples Х MADS&P

Future work towards cohesive ethical norms

## Soliciting structured feedback from reviewers

- Community survey with hypothetical ethical concerns
- Aggregated feedback on acceptable norms and ethical risks

# Result: criteria with exemplars to clarify expectations at major venues



Future work towards cohesive ethical norms

## Analysis of negative ethical examples

- Paper retractions (rare, low signal)
- Rejected papers (requires PC collaboration)
- **Recommendation**: anonymized ethical post-mortems

## **Result: practical negative ethical examples**











pauley.me/EthiCS23

