GRACE HOPPER CELEBRATION

/ ANITA B.ORG

Shagufta Mehnaz

Ph.D. candidate in the Computer Science department at Purdue University. The field of my research is information security/privacy.

Intrusion Detection

Privacy-preserving Machine Learning

#GHC19

Shagufta Mehnaz

I grew up in Dhaka, the capital of Bangladesh and received my BSc. in Computer Science and Engineering from Bangladesh University of Engineering and Technology (BUET).

#GHC19

$162 MILLION

TARGET

$148 MILLION

UBER

Everyone's
~~Private Driver~~
Privacy Breached!

600,000 drivers' license numbers,
57 MILLION user information

#GHC19

SALTED HASH- TOP SECURITY NEWS

By Steve Ragan, Senior Staff Writer, CSO | FEB 4, 2015 9:23 PM PST

About

Fundamental security insight to help you minimize risk and protect your organization

NEWS

Anthem confirms data breach, but full extent remains unknown

Breach could impact millions

MENU CNBC

Insider threats may be the biggest cyberthreats an organization faces

Harriet Taylor

Wednesday, 5 Oc

CNBC

Employees i
cybersecuri
defending a
public and p

The Washington Post

Democracy Dies in Darkness

Business

NSA case highlights growing concerns over insider threats

DEC 11, 2014 @ 01:13 PM 26,956 VIEWS

Massive Security Breach At Sony What You Need To Know

Joseph Steinberg, CONTRIBUTOR

I cover cybersecurity and entrepreneurship. FULL BIO

Opinions expressed by Forbes Contributors are their own.

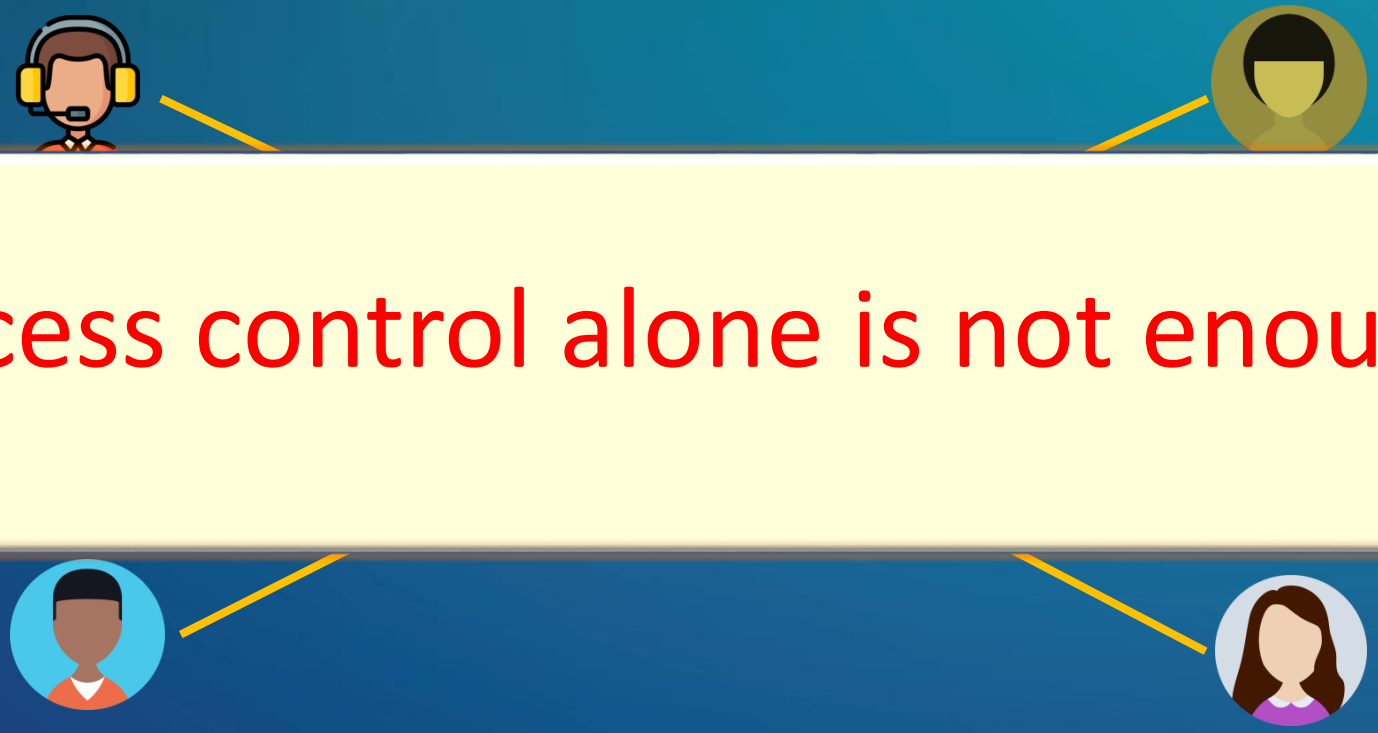
Privacy & Security

Insider threat health data breaches doubled in February, Protenus says

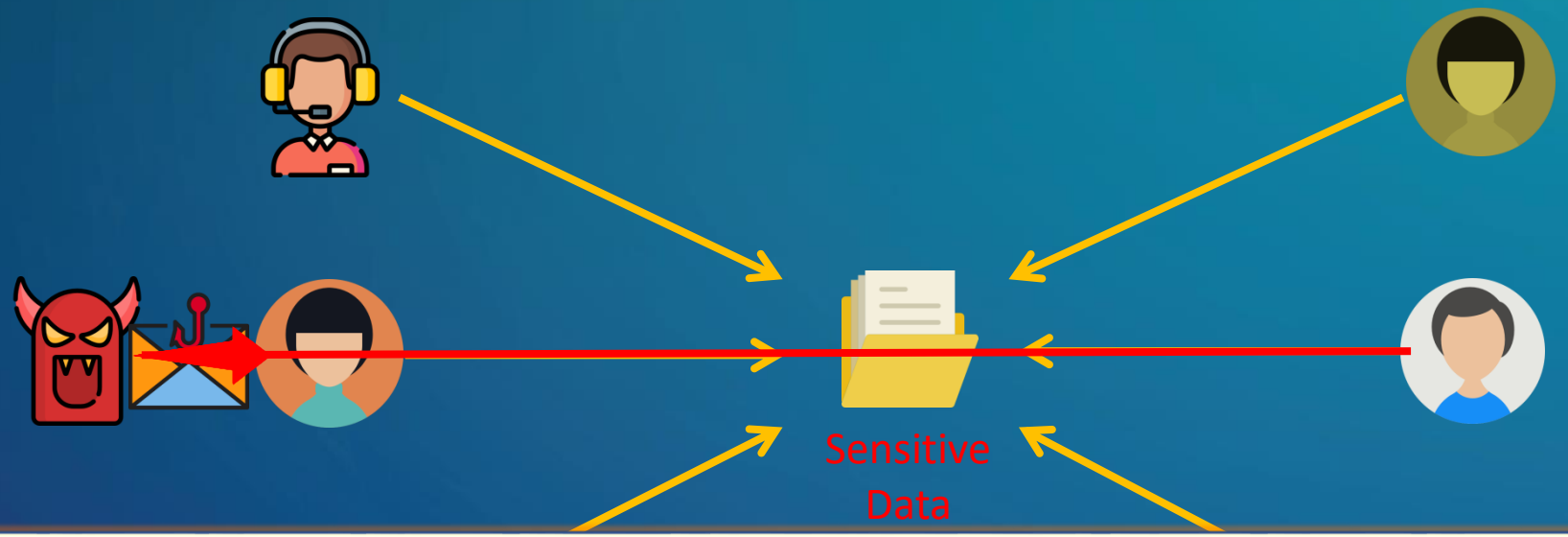
The good news? Hacking was down and only comprised 12 percent of reported incidents during the last month.

By Jessica Davis | March 20, 2017 | 04:45 PM

#GHC19



- Access control alone is not enough!



- Anomaly detection: a mechanism to differentiate between normal and anomalous usage of data.

CODASPY 2017

Ghostbuster: A Fine-grained Approach for Anomaly Detection in File System Accesses

#GHC19

Existing techniques

ANOMALY DETECTION AT FILE LEVEL

?

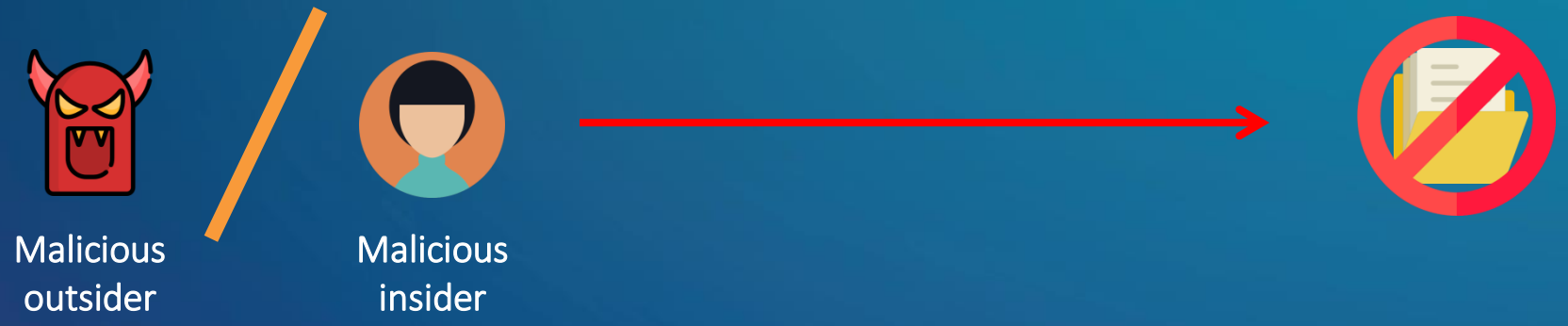
Our solution

FINE-GRAINED ANOMALY DETECTION (BLOCK LEVEL)

?

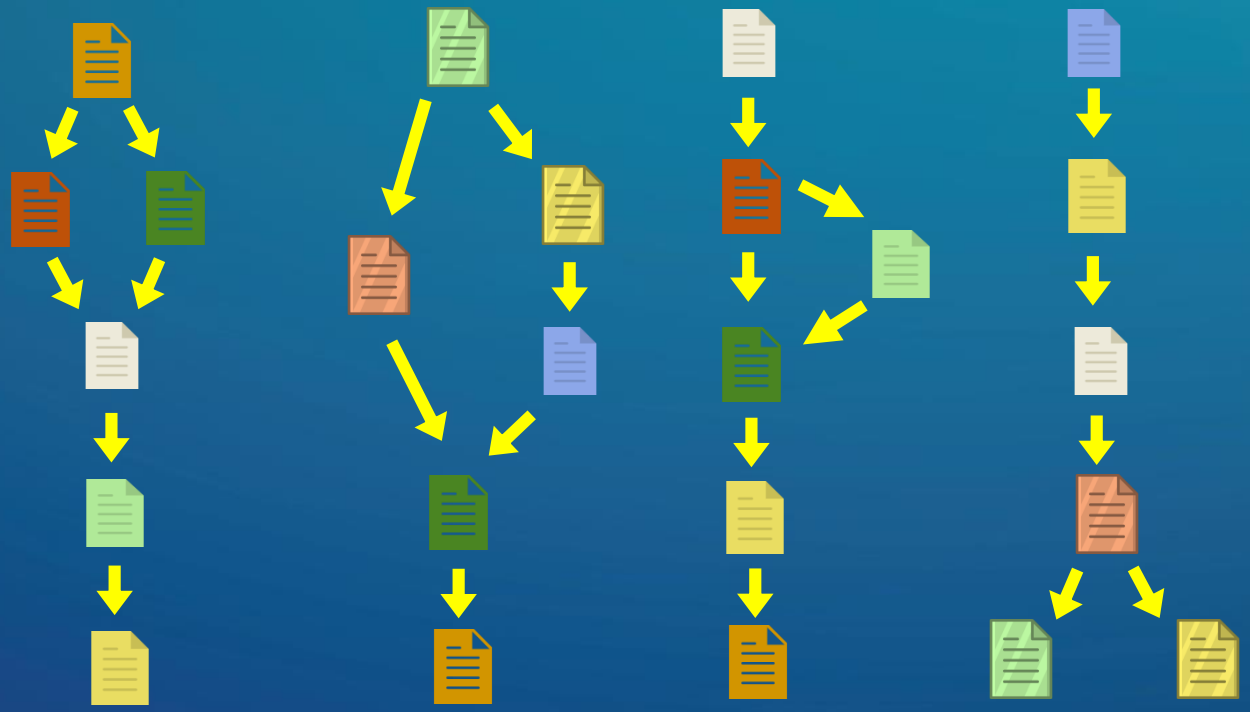
Anomalous file access without permission

Anomalous file access without permission



	Existing techniques	Our solution
	ANOMALY DETECTION AT FILE LEVEL	FINE-GRAINED ANOMALY DETECTION (BLOCK LEVEL)
Anomalous file access without permission	✓	✓
Anomalous access out of context	?	?

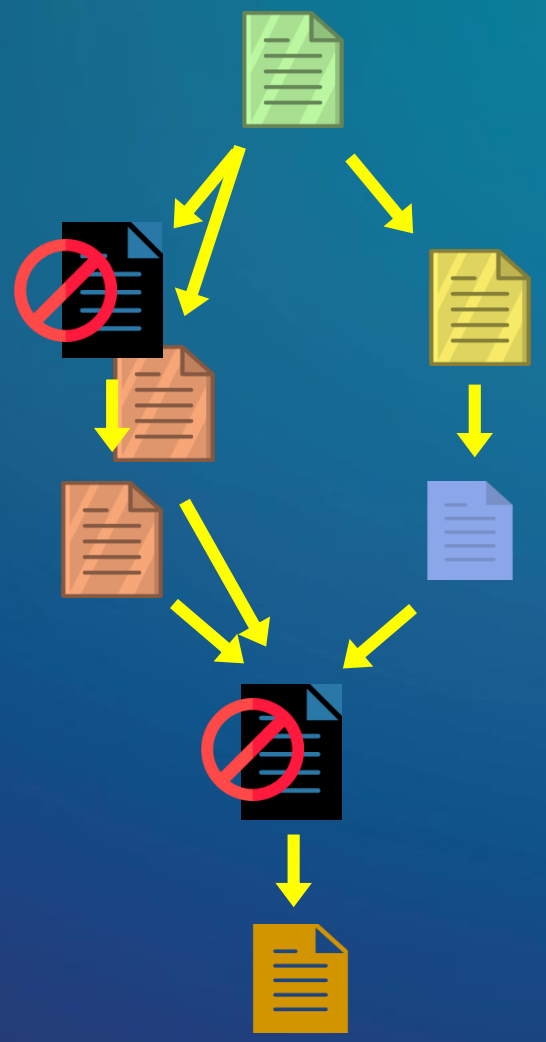
Anomalous access out of context



Frequent Episode Mining

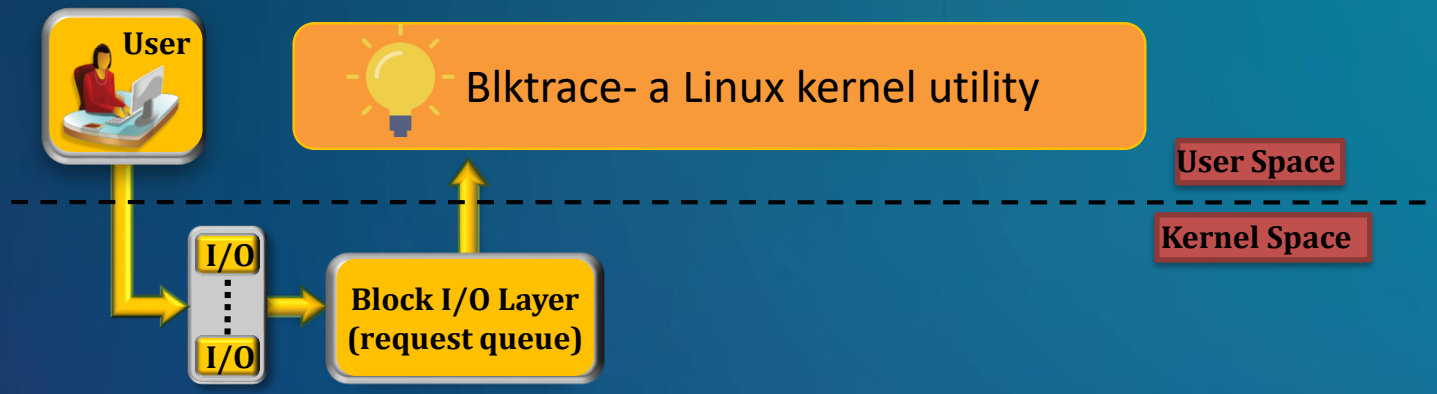
Finite State Machine

Anomalous access out of context



	Existing techniques	Our solution
	ANOMALY DETECTION AT FILE LEVEL	FINE-GRAINED ANOMALY DETECTION (BLOCK LEVEL)
Anomalous file access without permission	✓	✓
Anomalous access out of context	✗	✓
Anomalous size/segment of file access	?	?

Anomalous size/segment of file access

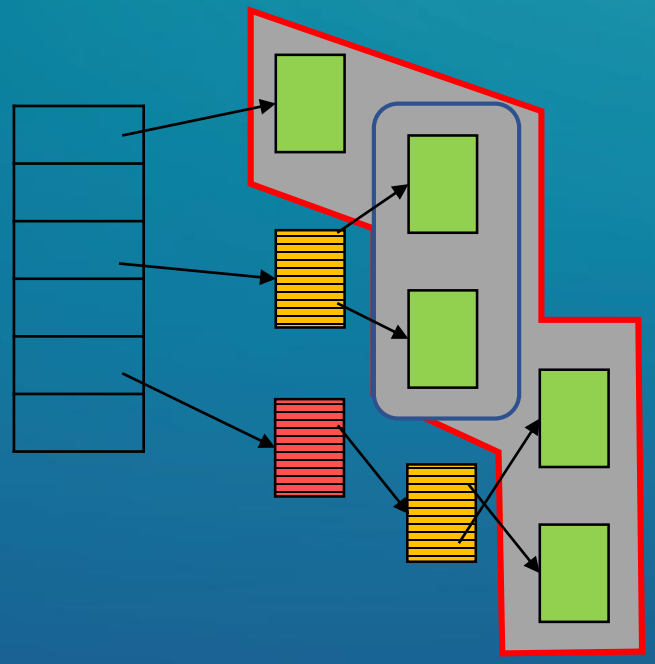


8,0	0	427	8.06743	57768	I	R	1729336	+32	[gedit]
8,0	0	428	8.52346	57768	I	R	1729368	+32	[gedit]
8,0	0	429	8.97686	57768	I	R	1729400	+32	[gedit]
:	:	:	:	:	:	:			:
:	:	:	:	:	:	:			:
:	:	:	:	:	:	:			:

Timestamp

Starting sector number

Number of sectors read



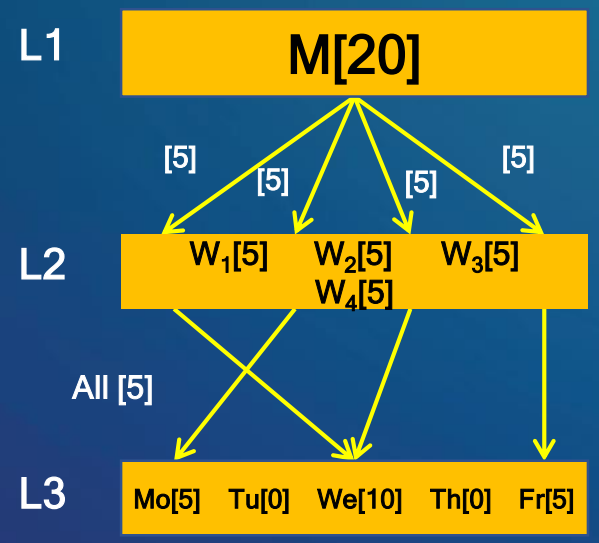
	Existing techniques ANOMALY DETECTION AT FILE LEVEL	Our solution FINE-GRAINED ANOMALY DETECTION (BLOCK LEVEL)
Anomalous file access without permission	✓	✓
Anomalous access out of context	✗	✓
Anomalous size/segment of file access	✗ ?	✓ ?
Anomalous frequency of file access	?	?

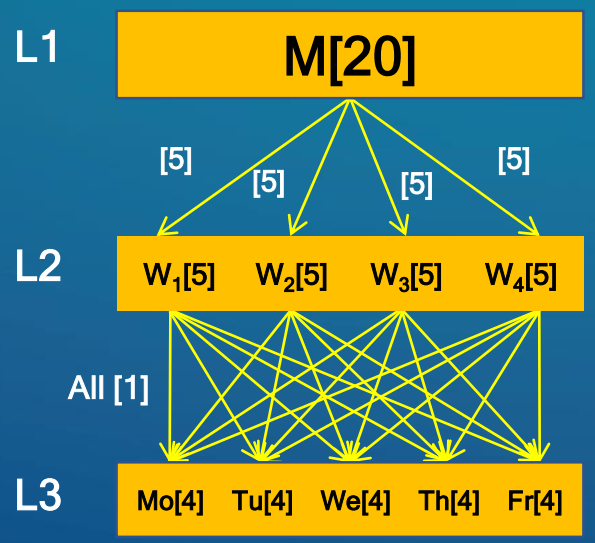
Anomalous frequency of file access

Case A File is accessed on an arbitrary weekday in a week, with a frequency of 5

Case B File is accessed on each weekday in a week, with a frequency of 1

Multi-level Time Granularity Approach





Case A ∀ M[20], ∀ W[5], ∃ D[5]

Case B ∀ M[20], ∀ W[5], ∀ D[1]

#GHC19

	Existing techniques ANOMALY DETECTION AT FILE LEVEL	Our solution FINE-GRAINED ANOMALY DETECTION (BLOCK LEVEL)
Anomalous file access without permission	✓	✓
Anomalous access out of context	✗	✓
Anomalous size/segment of file access	✗	✓
Anomalous frequency of file access	✗ ?	✓ ?

#GHC19

Experiments

- File repository
  - 560 files, 77 users
- Accesses for a duration of 2 months
  - training dataset: 4 weeks, test dataset: 4 weeks
- Four test sets (TS)

TS-I: data remains unchanged

TS-II: 3% of data anomalous

TS-III: 25% of data anomalous

TS-IV: all data anomalous

Experiment Results

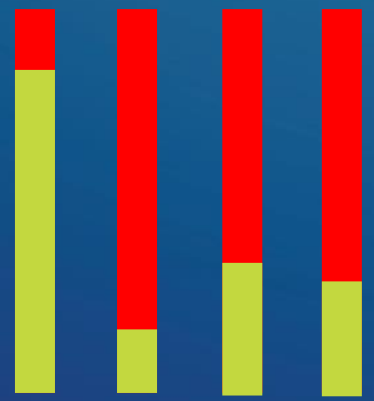
Existing techniques

ACCESS CONTROL

ANOMALY DETECTION
AT FILE LEVEL

Our solution

FINE-GRAINED
(AD at BLOCK LEVEL)

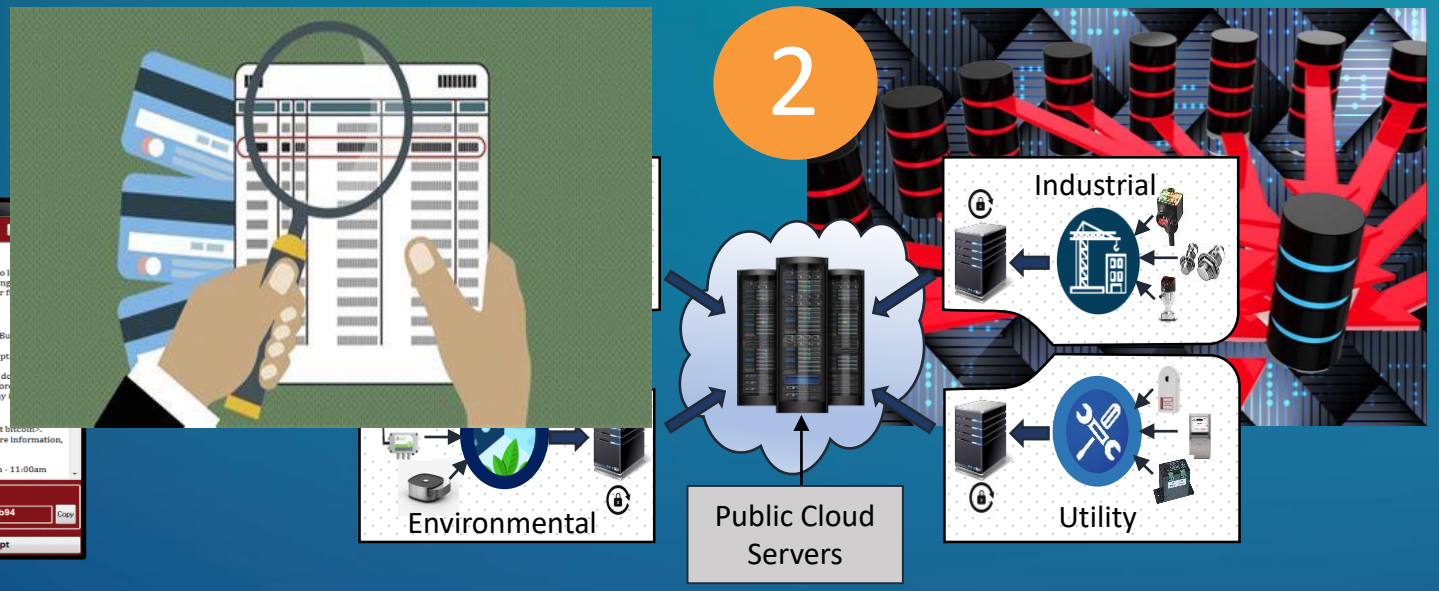


Conclusion

- Fine-grained approach for anomaly detection in file system accesses
- Comparison between the existing techniques and our solution
- Validated the performance of our solution on a real dataset

Follow-up work and future directions

A Real-time Detection System
Against Cryptographic Ransomware
[RAID, 2018]



Privacy-preserving Real-time Anomaly Detection Using Edge Computing
[ICDE, 2020]

Please remember to complete the session survey in the mobile app.

THANK YOU

YOU CAN *FOLLOW ME* @

https://sites.google.com/view/shaguftamehnaz/

https://www.linkedin.com/in/shagufta-mehnaz-539a20a3/

https://twitter.com/MehnazShagufta

https://www.facebook.com/shagufta.mehnaz

GRACE HOPPER CELEBRATION

ANITA B.ORG

#GHC19