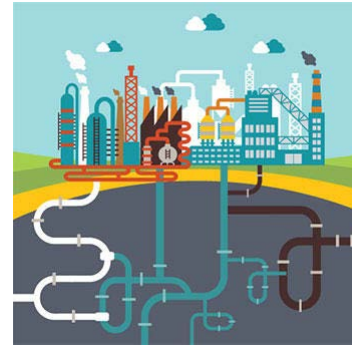


# Our Research on New AI and Blockchain Techniques for Network Security



**Raj Jain**

Washington University in Saint Louis  
Saint Louis, MO 63130

[Jain@wustl.edu](mailto:Jain@wustl.edu)

A talk in “CSE 591: Introduction to Graduate Studies in CSE”  
September 17, 2021

These slides and a video recording of this talk are at:  
<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>



Scan this to  
download slides



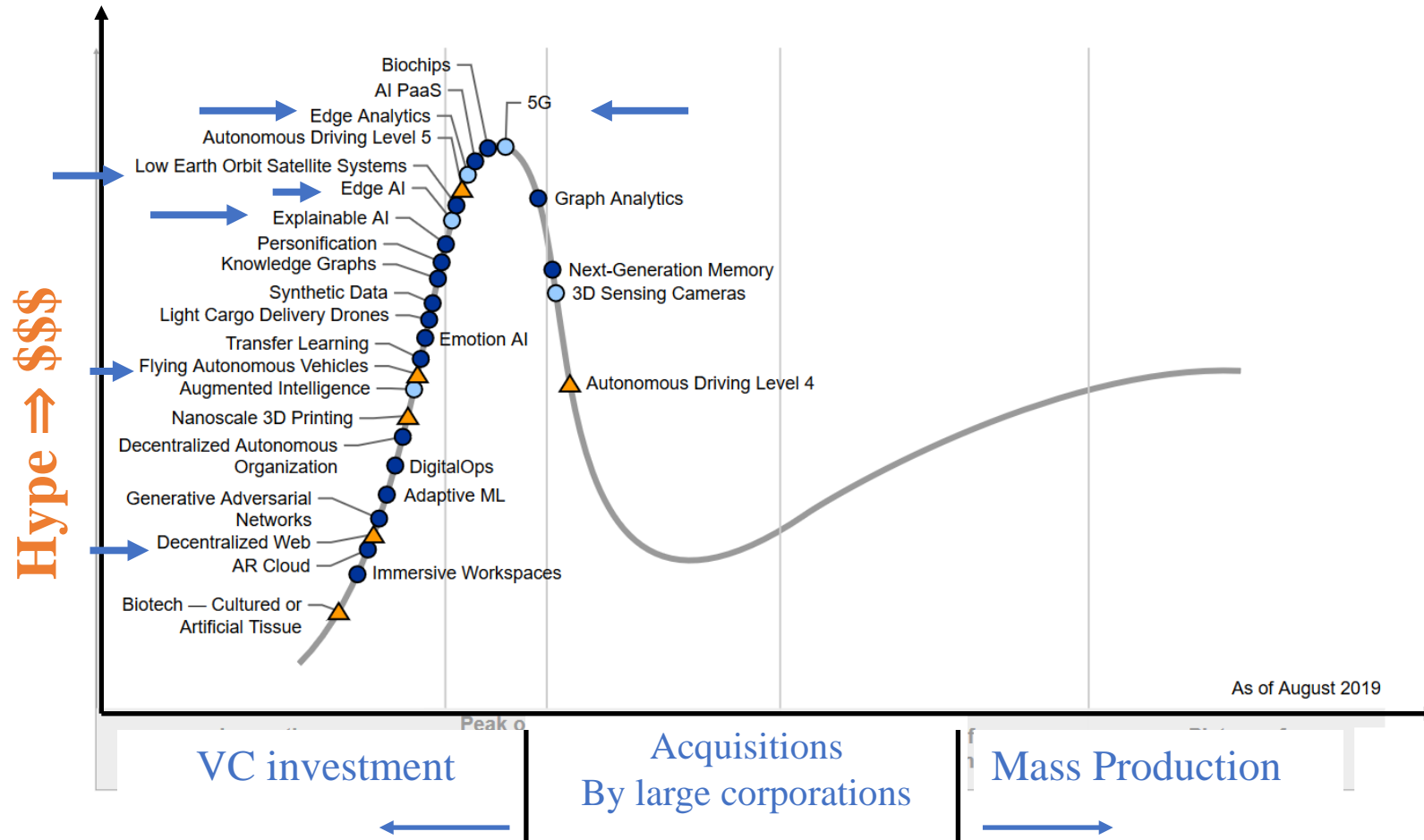
## 1. Trends in:

- AI
- Blockchain
- Cybersecurity

## 2. Our Research Projects in these areas

## 3. Key distinction of our research

# Gartner Hype Cycle of Emerging Tech 2019

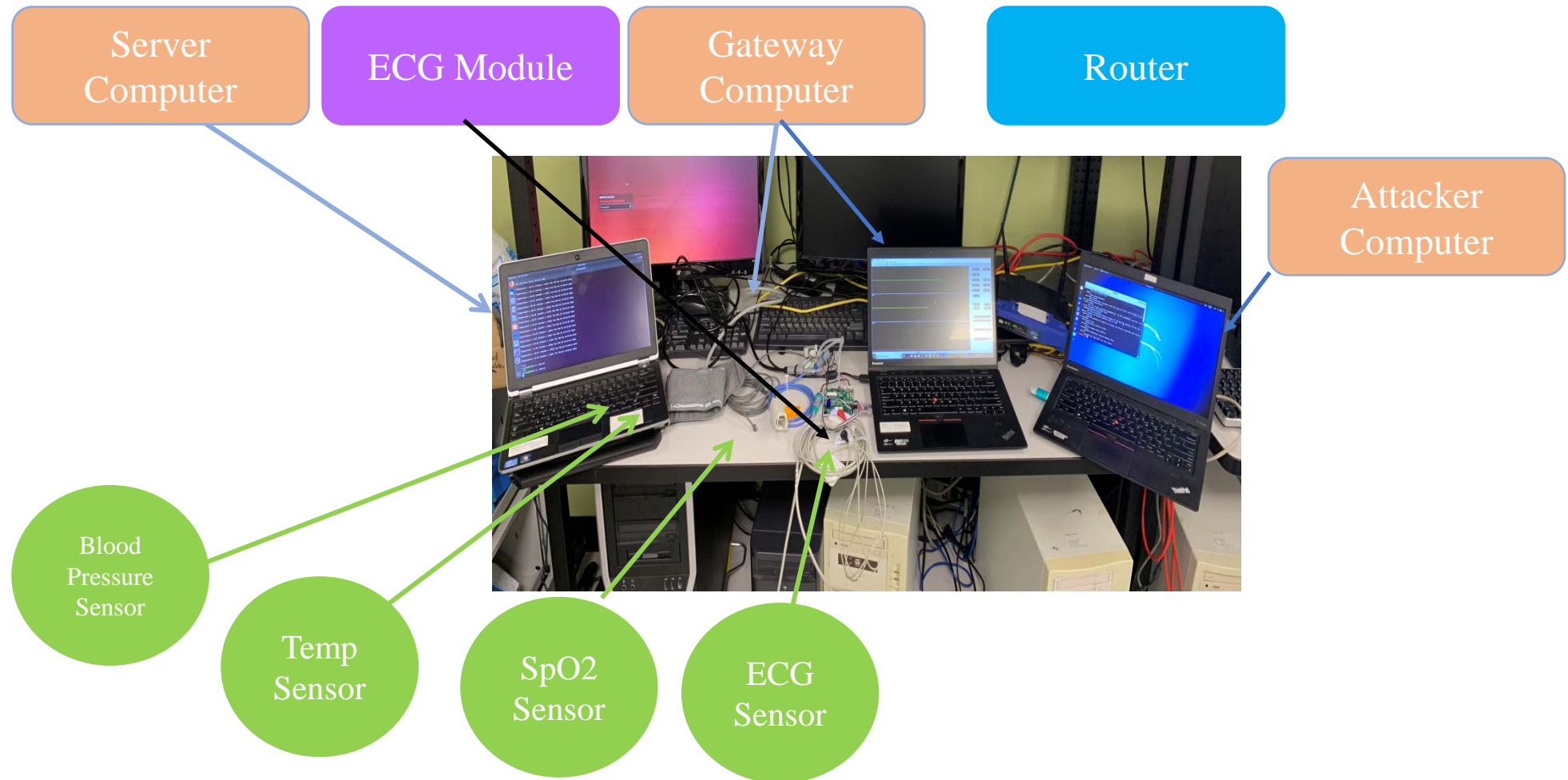


Ref: B. Burke, D. Smith, "Hype Cycle for Emerging Technologies, 2019," Gartner Report G00370466, 6 Aug. 2019, 68 pp.

Washington University in St. Louis

<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>

# Internet of Medical Things Security Using AI

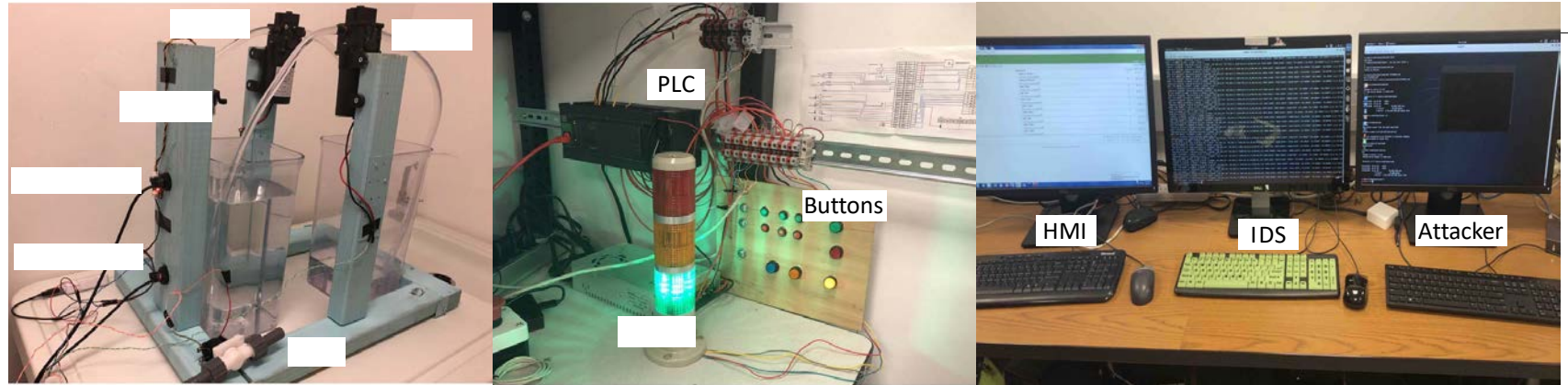
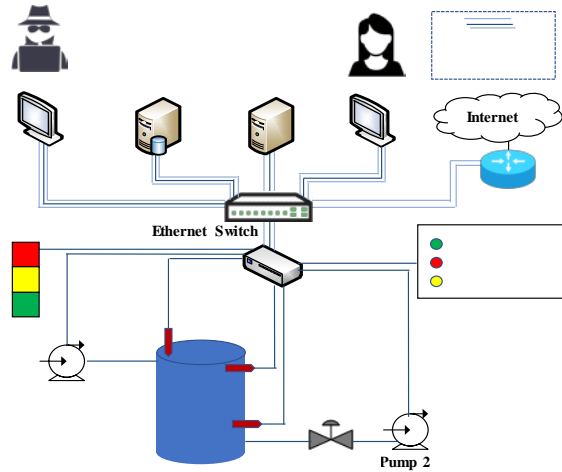


Ref: Anar A. Hady, Ali Ghubaish, Tara Salman, Devrim Unal, and Raj Jain, "Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study," IEEE Access, June 2020, <http://www.cse.wustl.edu/~jain/papers/hms.htm>

Washington University in St. Louis

<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>

# Industrial Control Systems Security Using AI

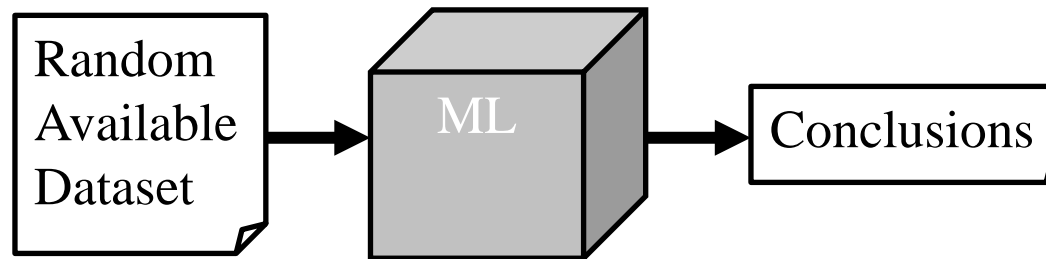


Ref: Maede Zolanvari, Marcio A. Teixeira, Lav Gupta, Khaled Khan, Raj Jain, "Machine Learning Based Network Vulnerability Analysis of Industrial Internet of Things," IEEE Internet of Things Journal, Vol. 6, Issue 4, Aug. 2019, pp. 6822-6834, <http://www.cse.wustl.edu/~jain/papers/vulnerab.htm>

Marcio Andrey Teixeira, Tara Salman, Maede Zolanvari, Raj Jain, Nader Meskin, and Mohammed Samaka, "SCADA System Testbed for Cybersecurity Research Using Machine Learning Approach," Future Internet 2018, 10(8), 76, [http://www.cse.wustl.edu/~jain/papers/ics\\_ml.htm](http://www.cse.wustl.edu/~jain/papers/ics_ml.htm)

# Machine Learning Challenges

- Machine learning is currently a black box
- ML algorithms are developed/used without domain expertise
- Data cleanliness, labeling, feature extractions, all require domain knowledge, e.g., What is the distance between Port 80, Port 81, and Port 8080?
- Synthetic data is used  $\Rightarrow$  Garbage-In, Garbage-Out
- Results are stated without model validation.



Ref: Lav Gupta, Tara Salman, Maede Zolanvari, Aiman Erbad, Raj Jain, "**Fault And Performance Management In Multi-Cloud Virtual Network Services Using AI: A Tutorial And A Case Study**," Computer Networks, Pre-Proof published on 14 Oct 2019,

[http://www.cse.wustl.edu/~jain/papers/fp\\_comst.htm](http://www.cse.wustl.edu/~jain/papers/fp_comst.htm)

Washington University in St. Louis

<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>



# AI for Security Issue 1: Imbalance

- AI started with image analysis but needs to be extended for security
- Security data is very different from image data
  - Most security datasets are not representative of real world.
  - In most papers, 10-15% of the packets are attack packets
- In real-world, 1 in a billion packets is an attack packet
  - Mis-classify the attack packet  $\Rightarrow$  99.9999% accuracy
  - Current metrics and methods not suitable for highly imbalanced data
- **Data imbalance** is a key issue in AI for security



1% attack

Ref: Maede Zolanvari, Marcio A. Teixeira, Raj Jain, "Effect of Imbalanced Datasets on Security of Industrial IoT Using Machine Learning," 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), Miami FL, Nov. 9 - 11, 2018, 6 pp.,

[http://www.cse.wustl.edu/~jain/papers/imb\\_isi.htm](http://www.cse.wustl.edu/~jain/papers/imb_isi.htm)

# AI for Security Issue 2: Wrong Metrics

- In Image analysis:  
Cost of predicting 0 when it is 1 = Cost of predicting 1 when it is 0  
⇒ Cost of errors is symmetric
- In Cyber Security:
  - Cost of missing an attack =  $10^6 \times$  Cost of false attack prediction
- Accuracy is not a good metric for security
- Need new metric to find the best algorithm  
⇒ Use Safety Score

Ref: Tara Salman, Ali Ghubaish, Devrim Unal, Raj Jain, "Safety Score as an Evaluation Metric for Machine Learning Models of Security Applications," IEEE Networking Letters, Vol. 2, Issue 4, December 2020, pp. 207-211, <http://www.cse.wustl.edu/~jain/papers/safety.htm>



# Trustable AI

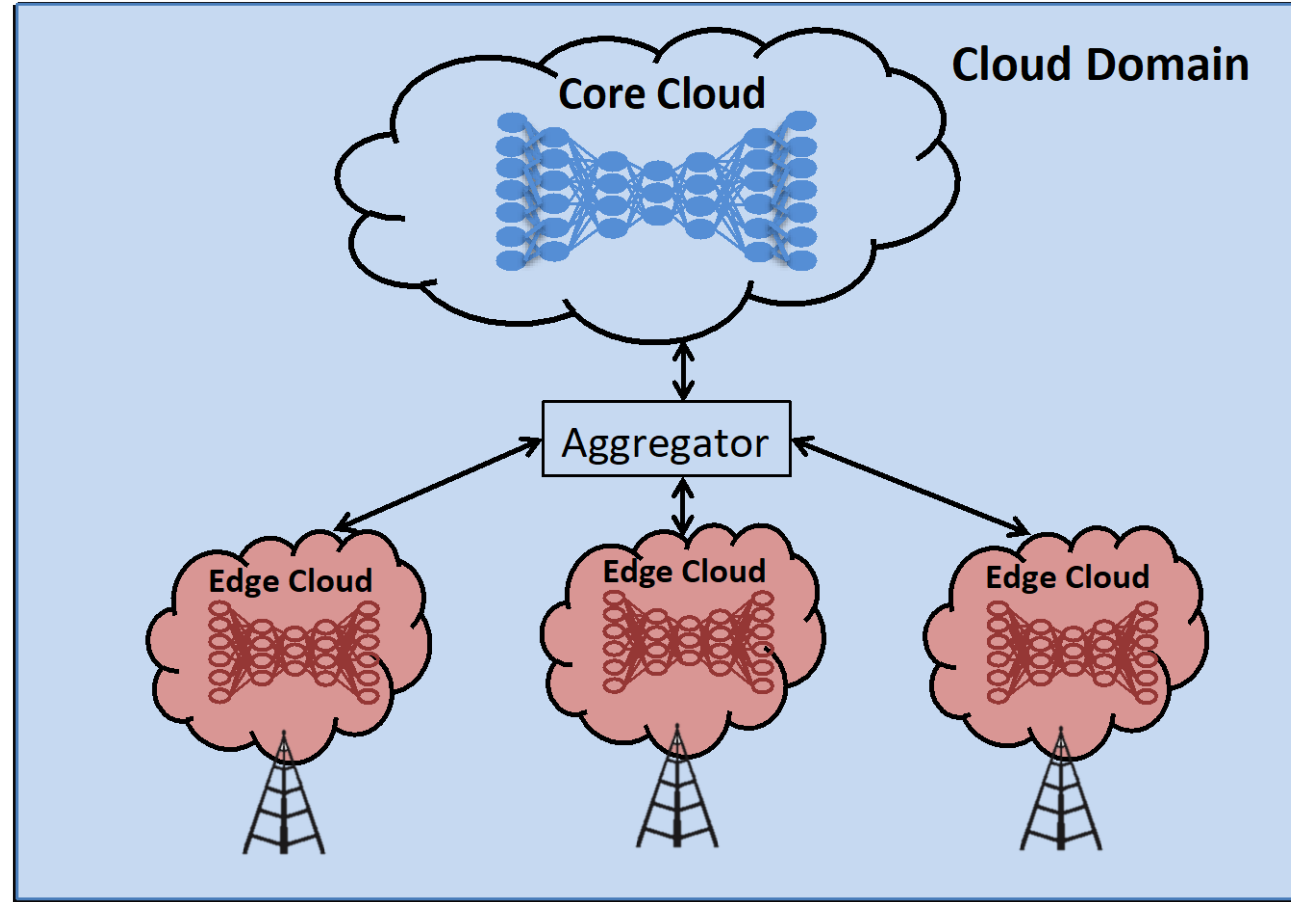
- Would you trust AI to diagnose your disease?
- Would you trust AI to National Cybersecurity?
- No, because you have no idea of why the results are what they are  
⇒ Can't discover bugs in ML model implementations
- Trustable AI = Explainable AI  
⇒ AI that can be understood by humans



*Machine Learning is what only machines can do,  
but human cannot do and cannot explain*

Ref: Maede Zolanvari, Zebo Yang, Khaled Khan, Raj Jain, and Nader Meskin, "TRUST XAI: A Novel Model for Explainable AI with An Example Using IIoT Security," IEEE IoT Journal, preliminary acceptance, September 2021.

# Distributed/Federated AI



Ref: Lav Gupta, Tara Salman, Ali Ghubaish, Devrim Unal, Abdulla Khalid Al-Ali, Raj Jain, "Cybersecurity of Multi-Cloud Healthcare Systems: A Hierarchical Deep Learning Approach" Under Review.

Washington University in St. Louis

<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>

# Blockchains



1. Satoshi Nakamoto invented Bitcoin
2. He used blockchains to make it decentralized
3. Since then blockchains have found numerous other applications
4. Blockchains are distributed, decentralized (no single point of control), trustless (no need to trust other parties), secure (Elliptic Curve Cryptography) with non-repudiation guarantee (all transactions are signed)



# Trend: Centralized → Decentralized

- **Banks:** Allow money transfer between two accounts
- **City Records:** Wedding registers, Property ownership
- **Networks:** Certificate Authorities, DNS
- In all cases:
  - There is a central third party to be trusted
  - Central party maintains a large database ⇒ Attracts Hackers
  - Central party may be hacked ⇒ Affects millions
  - Central party is a single point of failure. Can malfunction or be bribed

# Our Goal

- Moving the chain from deterministic to **probabilistic**
  - Moving the chain from storage to **computation**
  - Moving the chain from data to **knowledge**
  - Moving the chain from information to **decision making**
- ⇒ A blockchain that provides knowledge ⇒ A **knowledge chain** would be more useful

Ref: Tara Salman, Raj Jain, Lav Gupta, "A Reputation Management Framework for Knowledge-Based and Probabilistic Blockchains," IEEE 1st International Workshop on Advances in Artificial Intelligence for Blockchain (AIChain 2019), held in conjunction with the 2019 IEEE International Conference on Blockchain, Atlanta, July 14, 2019, <http://www.cse.wustl.edu/~jain/papers/rpmcewa.htm>

# Key Distinction of Our Research

- Goal: Impact to the real-world  
DECbit congestion indication in almost all networking architectures since its invention
- Funded by industry partners:  
Intel, Cisco, Broadcom, Boeing, ...
- Impact real-world by participating in standards organizations and industry forums:  
ATM Forum, IEEE Standards, American National Standards Institute (ANSI), Internet Engineering Task Force (IETF), WiMAX Forum
- Work on long term as well as short term research

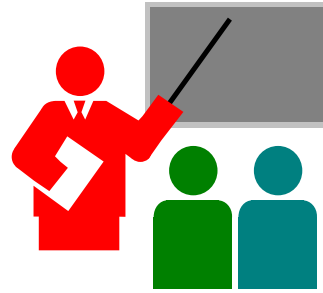




# Requirements

- Have 3 students working on 3 funded projects + 1 approved
- Need 1 more new Ph.D. student
- Requirements:
  - Background and interest in networking: CSE 473
  - Flexibility to work on the latest issues
  - Good communication skills
  - Machine learning (optional)
  - Preferably with a masters degree

# Summary



1. AI → Explainable AI → Federated AI
2. Blockchains  
→ Blockchains with AI = Knowledge Chains
3. IoT Intelligent IoT Secure and Intelligent IoT
  - Industrial systems
  - Medical systems
  - Agriculture Systems
4. Research for Impact

# References: Class Recordings

- Recordings of all of my classes and talks are available on YouTube and on my website:
  1. CSE 473: Introduction to Computer Networks, <http://www.cse.wustl.edu/~jain/cse473-21/index.html>
  2. CSE 571S: Network Security, <http://www.cse.wustl.edu/~jain/cse571-17/index.html>
  3. CSE 574S: Wireless Networks, <http://www.cse.wustl.edu/~jain/cse574-20/index.html>
  4. CSE 567: Computer Systems Analysis  
<http://www.cse.wustl.edu/~jain/cse567-17/index.html>
  5. CSE 570: Recent Advances in Networking  
<http://www.cse.wustl.edu/~jain/cse570-19/index.html>

# Our Publications on AI

- Tara Salman, Ali Ghubaish, Devrim Unal, Raj Jain, "**Safety Score as an Evaluation Metric for Machine Learning Models of Security Applications**," IEEE Networking Letters, Vol. 2, Issue 4, December 2020, pp. 207-211, <http://www.cse.wustl.edu/~jain/papers/safety.htm>
- Maede Zolanvari, Marcio A. Teixeira, Lav Gupta, Khaled Khan, Raj Jain, "**Machine Learning Based Network Vulnerability Analysis of Industrial Internet of Things**," IEEE Internet of Things Journal, Vol. 6, Issue 4, Aug 2019, <http://www.cse.wustl.edu/~jain/papers/vulnerab.htm>
- Marcio Andrey Teixeira, Tara Salman, Maede Zolanvari, Raj Jain, Nader Meskin, and Mohammed Samaka, "**SCADA System Testbed for Cybersecurity Research Using Machine Learning Approach**," Future Internet 2018, 10(8), 76, [http://www.cse.wustl.edu/~jain/papers/ics\\_ml.htm](http://www.cse.wustl.edu/~jain/papers/ics_ml.htm)
- Lav Gupta, Tara Salman, Ria Das, Aiman Erbad, Raj Jain, Mohammed Samaka, "**HYPER-VINES: A HYbrid Learning Fault and Performance Issues ERadicator for Virtual Network Services over Multi-cloud**," International Workshop on Computing, Networking and Communications (CNC19) at the International Conference on Computing, Networking and Communications (ICNC 2019), Honolulu, Hawaii, Feb. 18-21, 2019, 7 pp., <http://www.cse.wustl.edu/~jain/papers/hypervin.htm>

# Our Publications on AI (Cont)

- Maede Zolanvari, Marcio A. Teixeira, Raj Jain, "**Effect of Imbalanced Datasets on Security of Industrial IoT Using Machine Learning**," 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), Miami FL, Nov. 9 - 11, 2018, 6 pp., [http://www.cse.wustl.edu/~jain/papers/imb\\_isi.htm](http://www.cse.wustl.edu/~jain/papers/imb_isi.htm)
- Tara Salman, Deval Bhamare, Aiman Erbad, Raj Jain, Mohammed Samaka, "**Machine Learning for Anomaly Detection and Categorization in Multi-cloud Environments**," The 4th IEEE International Conference on Cyber Security and Cloud Computing (IEEE CSCloud 2017), New York, June 26-28, 2017, <http://www.cse.wustl.edu/~jain/papers/cscloud.htm>
- Deval Bhamare, Tara Salman, Mohammed Samaka, Aiman Erbad, Raj Jain, "**Feasibility of Supervised Machine Learning for Cloud Security**," 3rd International Conference on Information Science and Security (ICISS2016), December 19th - 22nd, 2016, Pattaya, Thailand, <http://www.cse.wustl.edu/~jain/papers/iciss16.htm>

# Our Publications on ICS Security

1. M. Elnour, N. Meskin, K. Khan, R. Jain, "A Dual-Isolation-Forests-Based Attack Detection Framework for Industrial Control Systems," IEEE Access, Vol. 8, 19 February 2020, pp. 36639 - 36651, <http://www.cse.wustl.edu/~jain/papers/dif.htm>
2. D. Bhamare, M. Zolanvari, A. Erbad, R. Jain, K. Khan, N. Meskin, "Cybersecurity for Industrial Control Systems: A Survey," Computers and Security, Elsevier, Volume 89, February 2020, Article 101677, [http://www.cse.wustl.edu/~jain/papers/ics\\_survey.htm](http://www.cse.wustl.edu/~jain/papers/ics_survey.htm)
3. M. Zolanvari, M. Teixeira, L. Gupta, K. Khan, R. Jain, "Machine Learning Based Network Vulnerability Analysis of Industrial Internet of Things," IEEE Internet of Things Journal, Vol. 6, Issue 4, Aug 2019, <http://www.cse.wustl.edu/~jain/papers/vulnerab.htm>
4. M. Zolanvari, M. Teixeira, R. Jain, "Effect of Imbalanced Datasets on Security of Industrial IoT Using Machine Learning," 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), Miami FL, Nov. 9 - 11, 2018, 6 pp., [http://www.cse.wustl.edu/~jain/papers/imb\\_isi.htm](http://www.cse.wustl.edu/~jain/papers/imb_isi.htm)



# Our Publications on Healthcare Security

1. Ali Ghubaish, Tara Salman, Maede Zolanvari, Devrim Unal, Abdulla Khalid Al-Ali, Raj Jain, "**Recent Advances in the Internet of Medical Things (IoMT) Systems Security**," IEEE Internet of Things Journal, Accepted December 2020, [http://www.cse.wustl.edu/~jain/iomt\\_iot.htm](http://www.cse.wustl.edu/~jain/iomt_iot.htm)
2. Tara Salman, Ali Ghubaish, Devrim Unal, Raj Jain, "**Safety Score as an Evaluation Metric for Machine Learning Models of Security Applications**," IEEE Networking Letters, Vol. 2, Issue 4, December 2020, pp. 207-211, <http://www.cse.wustl.edu/~jain/papers/safety.htm>
3. Anar A. Hady, Ali Ghubaish, T. Salman, Devrim Unal, and R. Jain, "**Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study**," IEEE Access, June 2020, <http://www.cse.wustl.edu/~jain/papers/hms.htm>

# Our Publications on Blockchains

1. T. Salman, M. Zolanvari, A. Erbad, R. Jain, and M. Samaka, "**Security Services Using Blockchains: A State of the Art Survey**" IEEE Communications Surveys and Tutorials, First Quarter 2019, Volume 21, Issue 1, 858-880 pp., <http://www.cse.wustl.edu/~jain/papers/bcs.htm>
2. T. Salman, R. Jain, L. Gupta, "**A Reputation Management Framework for Knowledge-Based and Probabilistic Blockchains**," IEEE 1st International Workshop on Advances in Artificial Intelligence for Blockchain (AICChain 2019), held in conjunction with the 2019 IEEE International Conference on Blockchain, Atlanta, July 14, 2019, <http://www.cse.wustl.edu/~jain/papers/rpmcewa.htm>
3. T. Salman, R. Jain, and L. Gupta, "**Probabilistic Blockchains: A Blockchain Paradigm for Collaborative Decision-Making**," 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON 2018), New York, NY, November 8-10, 2018, 9 pp., [http://www.cse.wustl.edu/~jain/papers/psc\\_uem.htm](http://www.cse.wustl.edu/~jain/papers/psc_uem.htm)

# Acronyms

- 3GPP Third Generation Partnership Project
- AI Artificial Intelligence
- ANSI American National Standards Institute
- AT&T American Telephone and Telegraph
- BSS Business Support Services
- CA California
- CGNAT Carrier Grade Network Address Translator
- CSE Computer Science and Engineering
- DECbit Digital Equipment Corporation Bit
- IEEE Institution of Electrical and Electronic Engineering
- IoT Internet of Things
- ML Machine Learning
- MO Missouri
- MS Master of Science
- NFV Network Function Virtualization
- NTT Nippon Telephone and Telegraph

# Acronyms (Cont)

- OpenADN      Open Application Delivery Networking
- OSS            Operations Support Services
- SON            Self-Organizing Networks
- TV             Television
- UK             United Kingdom
- US             United States
- VC             Venture Capital
- WAN            Wide Area Network
- WiMAX        Worldwide Interoperability for Microwave Access
- WUSTL        Washington University in St. Louis

**Scan This to Download These Slides**



Raj Jain  
[Rajjain.com](http://Rajjain.com)

<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>

Washington University in St. Louis

<http://www.cse.wustl.edu/~jain/talks/cs59121.htm>

# Our Courses on YouTube



CSE567M: Computer Systems Analysis (Spring 2013),  
[https://www.youtube.com/playlist?list=PLjGG94etKypJEKjNAa1n\\_1X0bWWNyZcof](https://www.youtube.com/playlist?list=PLjGG94etKypJEKjNAa1n_1X0bWWNyZcof)

CSE473S: Introduction to Computer Networks (Fall 2011),  
[https://www.youtube.com/playlist?list=PLjGG94etKypJWOSPMh8Azcg5e\\_10TiDw](https://www.youtube.com/playlist?list=PLjGG94etKypJWOSPMh8Azcg5e_10TiDw)



CSE 570: Recent Advances in Networking (Spring 2013)  
<https://www.youtube.com/playlist?list=PLjGG94etKypLHyBN8mOgwJLHD2FFIMGq5>

CSE571S: Network Security (Fall 2011),  
<https://www.youtube.com/playlist?list=PLjGG94etKypKvzfVtutHcPFJXumyyg93u>



Video Podcasts of Prof. Raj Jain's Lectures,  
<https://www.youtube.com/channel/UCN4-5wzNP9-ruOzQMs-8NUw>