



## **Integrating Big Data, Data Science and Cyber Security with Applications in Internet of Transportation and Infrastructures**

**Dr Bhavani Thuraisingham**

**The University of Texas at Dallas**

The collection, storage, manipulation, analysis and retention of massive amounts of data have resulted in new technologies including big data analytics and data science. It is now possible to analyze massive amounts of data and extract useful nuggets. However, the collection and manipulation of this data has also resulted in serious security and privacy considerations. Various regulations are being proposed to handle big data so that the privacy of the individuals is not violated. Furthermore, the massive

amounts of data being stored may also be vulnerable to cyber attacks.

Big Data, Data Science and Security are being integrated to solve many of the security and privacy challenges. For example, machine learning techniques are being applied to solve security problems such as malware analysis and insider threat detection. However, there is also a major concern that the machine learning techniques themselves could be attacked. Therefore, the machine learning techniques are being adapted to handle adversarial attacks. This area is known as adversarial machine learning. In addition, privacy of the individuals is also being violated through these machine learning techniques as it is now possible to gather and analyze vast amounts of data and therefore privacy enhanced data science techniques are being developed.

With the advent of the web, computing systems are now being used in every aspect of our lives from mobile phones to autonomous vehicles. It is now possible to collect, store, manage, and analyze vast amounts of sensor data emanating from numerous devices and sensors including from various transportation systems. Such systems collectively are known as the Internet of Transportation, which is essentially the Internet of Things for Transportation, where multiple autonomous transportation systems are connected through the web and coordinate their activities. However, security and privacy for the Internet of Transportation and the infrastructures that support it is a challenge. Due to the large volumes of heterogenous data being collected from numerous devices, the traditional cyber security techniques such as encryption are not efficient to secure the Internet of Transportation. Some Physics-based solutions being developed are showing promise. More recently, the developments in Data Science are also being examined for securing the Internet of Transportation and its supporting infrastructures.

To assess the developments on the integration of Big Data, Data Science and Security over the past decade and apply them to the Internet of Transportation, the presentation will focus on three aspects. First it will examine the developments on applying Data Science techniques for detecting cyber security problems such as insider threat detection as well as the advances in adversarial machine learning. Some developments on privacy aware and policy-based data management frameworks will also be discussed. Second it will discuss the developments on securing the Internet of Transportation and its supporting infrastructures and examine the privacy implications. Finally, it will describe ways in which Big Data, Data Science and Security could be incorporated into the Internet of Transportation and Infrastructures.

### **Biography of Dr Bhavani Thuraisingham**

Dr. Bhavani Thuraisingham is the Founders Chair Professor of Computer Science and the Executive Director of the Cyber Security Research and Education Institute at the University of Texas at Dallas (UTD). She is also a visiting Senior Research Fellow at Kings College, University of London and an elected Fellow of the ACM, IEEE, the AAAS, the NAI and the BCS. She was a Cyber Security Policy Fellow at the New America Foundation for 2017-

2018 and focused on engaging rural America in cyber security. Her research interests are on integrating cyber security and artificial intelligence/data science for the past 35 years (where it used to be computer security and data management/mining). She has received several awards including the IEEE CS 1997 Technical Achievement Award, ACM SIGSAC 2010 Outstanding Contributions Award, the IEEE Comsoc Communications and Information Security 2019 Technical Recognition Award, the IEEE CS Services Computing 2017 Research Innovation Award, the ACM CODASPY 2017 Lasting Research Award, the IEEE ISI 2010 Research Leadership Award, the 2017 Dallas Business Journal Women in Technology Award, and the ACM SACMAT 10 Year Test of Time Awards for 2018 and 2019 (for papers published in 2008 and 2009). She was named one of 500 most influential business leaders in North Texas for 2021 by the D Magazine's D CEO Magazine. She co-chaired the Women in Cyber Security Conference (WiCyS) in 2016 and delivered the featured address at the 2018 Women in Data Science (WiDS) at Stanford University serves as the Co-Director of both the Women in Cyber Security and Women in Data Science Centers at UTD. She has spent around 20 years promoting Diversity, Equity and Inclusion in Cyber Security and Data Science and has chaired multiple panels including her recent panel at IEEE ISI 2020 (Intelligence and Security Information). Her 40-year career includes industry (Honeywell), federal research laboratory (MITRE), US government (NSF) and US Academia. Her work has resulted in 130+ journal articles, 300+ conference papers, 170+ keynote and featured addresses, seven US patents, fifteen books as well as technology transfer of the research to commercial products and operational systems. She received her PhD from the University of Wales, Swansea, UK, and the prestigious earned higher doctorate (D. Eng) from the University of Bristol, UK.