



## **Seyed Hossein Khasteh**

K. N. Toosi University of Technology

**Seyed Hossein Khasteh** is an Assistant Professor at K. N. Toosi University of Technology. He received his Bachelor's Degree in Electrical Engineering, MSc and PhD in Computer Engineering from Sharif University of Technology in 2004, 2006 and 2012, respectively. He has conducted numerous projects in the areas of artificial intelligence, social network analysis, telecommunications network management and hardware / software hybrid projects. Dr. Khasteh has been teaching Big-data analytics course since 2015 at K. N. Toosi University of Technology. He is a member of the board of directors of the Computer Society of Iran, the National Elite Foundation, the Case Study Group of Knowledge Base Companies, and the Monitoring Committee on how to fund innovation and prosperity fund loans. He has (co-/) authored more than 50 research and tutorial educational papers in journals, conferences and magazines. He has served as the secretary of the Student Committee of Electrical Engineering Conference in 2017, the Secretary of the Computer Engineering Conference Publication Committee of 2018 and as the referee of many domestic and foreign conferences and journals.

### **Big data, concepts and applications**

What is Big Data and what are its features? Perhaps one answer is: Big-data is a term used to describe a large volume of complex and variable data generated at high velocity. Advanced methods and technologies are needed to capture, store, distribute, analyze, and manage such data.

Big data is a reality in today's world, and it's a problem that real systems have to deal with. Modern information societies are defined by vast repositories of data; they can be public or private. Thus, with the increase in the datasets of interest, all practical applications should be able to scale.

In this lecture, key concepts in the field of big data such as: 3V, 4V and 5V characteristics of big data, key frameworks for managing and working with big data such as: Hadoop, Hadoop ecosystem and software in this ecosystem, map-reduce framework, New methods of stream-data processing, Spark and Storm tools, and some of the new algorithms used to process data streams and the important points of these algorithms are discussed.

There are also some practical examples of Big-data-driven operating systems, how to filter spam emails, analyze big-store sales data (such as Amazon), and analyze Telegram Persian channels information, some algorithms and technologies employed for this purpose are discussed.