**Big Data Technical Challenges in Security and Privacy**

**Paper 1 Summary (Challenges and Opportunities with Big Data)**

Big Data is crucial for obtaining actionable knowledge. There are some technical challenges to be overlooked. Companies are not pointing towards the volume but also variety which means representation and heterogeneity of data types and velocity means the time data arrives and the time in which it must be acted upon. These three are important but it fails to include some important requirements like privacy and usability. Big data Analysis involves multiple phases which introduce technical challenges. This paper consists of technical challenges as different phases of data analysis pipeline.

**Paper 2 Summary (Security Problems and Challenges in a Machine Learning-Based Hybrid Big Data Processing Network Systems)**

To extract, analyze and visualize data Machine Learning techniques are used. For huge amount of data, which is, produced from data sources with high volume and high velocity with different data types, which creates big, data cause problems and challenges to Machine Learning techniques. To overcome this, a hybrid network model is used which consists of some components like Hadoop distributed File System (HDFS), cloud storage, security model and Machine Learning Unit. Processing Big Data in networking level with ML techniques require user interaction and storage. There will be a gap between data transmission and security vulnerability in public cloud storage. We need a key-based encryption algorithm to reduce delay-time and strengthen the security. A model is proposed for using public cloud trust levels to encryption types data storage in Big Data Analysis network Topology.

**Paper 3 Summary (Big Data – Security and Privacy)**

Industry uses web applications in different ways to quickly process and analyze large volumes of data extracted from datasets which is critical information for security related tasks. The use of data in privacy is a major security related task. This paper is about security with privacy. In order to achieve security we must give up privacy. This paper tells about few Privacy-Enhancing techniques that focus on efficiently reconciling security with privacy.

**Paper 4 Summary (Emerging Trends around Big Data Analytics and Security)**

This paper discuss on big data security analytics. Internet based application for business, finance and banking need huge amount of data, which is collected from big data repositories and provides meaningful insight from it. Security and operational problems are the areas where big data analytics is expected to play a crucial role. To overcome this challenges a group of people from industry and government will provide a insight into big data security analytics.

**Paper 5 Summary (Attack Tolerant Architecture for Big Data File Systems)**

Big data analytics involves collecting, storing and processing large volumes of data and security such as confidentiality, integrity and availability has become a major problem. The valuable information extracted from big data set needs to be protected by cyber attackers. This paper tells about different efforts which can reduce the attack tolerance. GFS and Hadoop can defense the availability. Shadoop is used to prevent attacks by strong authentication and SCALA programming for Map Reduce query Processing.

**Bibliography**

1. Divyakant Agrawal, UC Santa Barbara, Philip Bernstein, Microsoft, Elisa Bertino, Purdue Univ., Susan Davidson, Univ. of Pennsylvania, Umeshwar Dayal, HP, Michael Franklin, UC Berkeley, Johannes Gehrke, Cornell Univ, Laura Haas, IBM, Alon Halevy, Google, Jiawei Han, UIUC, H. V. Jagadish, Univ. of Michigan (Coordinator), Alexandros Labrinidis, Pittsburgh Univ., Sam Madden, MIT, Yannis Papakonstantinou, UC San Diego, Jignesh M. Patel, Univ. of Wisconsin, Raghu Ramakrishnan, Yahoo!, Kenneth Ross, Columbia Univ., Cyrus Shahabi, Univ. of Southern California, Dan Suciu, Univ. of Washington, Shiv Vaithyanathan, IBM, Jennifer Widom, Stanford Univ “*Challenges and Opportunities with Big Data*” Purdue Nov. 2011 to Feb. 2012 < <https://www.purdue.edu/discoverypark/cyber/assets/pdfs/BigDataWhitePaper.pdf>>
2. Jeff Whitworth, Shan Suthaharan, “*Security Problems and Challenges in a Machine Learning-Based Hybrid Big Data Processing Network Systems”* in April 2014
3. [E Bertino](http://scholar.google.com/citations?user=EfuT9RoAAAAJ&hl=en&oi=sra), [M Kantarcioglu](http://scholar.google.com/citations?user=qXb4xQMAAAAJ&hl=en&oi=sra) , “*Big Data – Security with Privacy*” in September 2014 proceedings at university of Texas Dallas
4. [Rafae Bhatti](http://dl.acm.org/author_page.cfm?id=81100545165&coll=DL&dl=ACM&CFID=448937646&CFTOKEN=45410205), [Ryan LaSalle](http://dl.acm.org/author_page.cfm?id=81502800412&coll=DL&dl=ACM&CFID=448937646&CFTOKEN=45410205), [Rob Bird](http://dl.acm.org/author_page.cfm?id=81502676297&coll=DL&dl=ACM&CFID=448937646&CFTOKEN=45410205), [Tim Grance](http://dl.acm.org/author_page.cfm?id=81100311943&coll=DL&dl=ACM&CFID=448937646&CFTOKEN=45410205), [Elisa Bertino](http://dl.acm.org/author_page.cfm?id=81409594658&coll=DL&dl=ACM&CFID=448937646&CFTOKEN=45410205), “*Emerging Trends around Big Data Analytics and Security*” in June 2012 at **SACMAT '12:**Proceedings of the 17th ACM symposium on Access Control Models and Technologies.
5. Manoj Banik, Bharat B. Madan, “*Attack Tolerant Architecture for Big Data File Systems”* in March 2014 proceeding at [SIGMETRICS](http://dl.acm.org/sig.cfm?id=SP937&CFID=450597202&CFTOKEN=15879571) ACM Special Interest Group on Measurement and Evaluation at New York USA.