DISTINGUISHED WEBINAR SERIES IN ARTIFICIAL INTELLIGENCE AND CYBER SECURITY

Chernoff Information as a Privacy Constraint for Adversarial Classification and Membership Advantage

Featuring Dr. Ayşe Ünsal

Research fellow

Digital Security Department, EURECOM

Abstract:

This work inspects a privacy metric based on Chernoff information, namely Chernoff differential privacy, due to its significance in characterization of the optimal classifier's performance and adversary's membership advantage. Adversarial classification, as any other classification problem is built around minimization of the (average or correct detection) probability of error in deciding on either of the classes in the case of binary classification. In this work, we focus on the relationship between the best error exponent of the average error probability and $\varepsilon-$ differential privacy. Accordingly, we re-derive Chernoff differential privacy in terms of $\varepsilon-$differential privacy via the Radon-Nikodym derivative and show that it satisfies the composition property. Subsequently, we present numerical evaluation results, which demonstrates that Chernoff information outperforms Kullback-Leibler divergence as a function of the privacy budget, the impact of the adversary's attack and global sensitivity for adversarial classification in Laplace mechanisms. Lastly, we introduce a novel upper bound on the adversary's advantage in membership inference attacks and compare its performance against existing ones.

Biography:

Dr. Ayşe ÜNSAL is working at the Digital Security Department of EURECOM as a research fellow since September 2020. Her current position is funded by ANR Frame XG project AIRSEC- Air Interface Security for 6G and European project TRUMAN-Trustworthy Human-Centric AI. Previously, she worked as a post-doctoral researcher on coded caching at the Communication Systems Dept. of EURECOM (September 2017- February 2019) where she had also received her Ph.D. degree from Telecom ParisTECH with a specialization in Electronics and Communication in November 2014. Her Ph.D. was funded by an FP7 European project (248993-LOLA) which focused on information-theoretic characterizations of transmission strategies to lower latency in wireless communications and explain the technical basis behind it. After having obtained her degree, she worked as a post-doctoral researcher and lecturer at the Paderborn University, Germany and CITI Lab of INRIA/INSA Lyon.

DATE:
Thursday, July 17th, 2025

TIME:
11:00-11:50 a.m. CST

LOCATION:
Virtual

Webinar LINK:
Join Directly

The **Distinguished Speaker Webinar Series** is aimed to advance the state-of-the-art concepts and methods in artificial intelligence and cyber security areas. The series is jointly hosted by the Center for Cyber Security Research (C2SR), the Artificial Intelligence Research (AIR) Initiative, and the School of Electrical Engineering and Computer Science (SEECS) at the University of North Dakota College of Engineering & Mines with support from University of Minnesota, North Dakota State University, University of Miami, Texas A&M Kingsville, University of Connecticut and West Virginia University.

For inquires please contact UND.C2SR@und.edu

Center for Cyber Security Research
College of Engineering & Mines
University of North Dakota

Artificial Intelligence Research Initiative
College of Engineering & Mines
University of North Dakota

West Virginia University
POWERGRID CENTER OF EXCELLENCE IN CYBER SECURITY
NDSU NORTH DAKOTA STATE UNIVERSITY
University of Connecticut

University of Minnesota
Texas A&M University Kingsville
University of Miami