2017 June class:
Individual and Group Anonymity. Introduction

Lecturer: Oleg CHERTOV,
Professor, Sc.D. (Doctor of Technical Sciences),
Head of the Applied Mathematics Department,
National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic
Institute”, Kyiv, Ukraine

Duration: 2 hours

Objectives: The objective of this class is to introduce students to some key ideas for
guaranteeing of individual and group data anonymity. Major challenge in Privacy-preserving
data publishing is to simultaneously preserve both the privacy and information utility in the
anonymous data. Privacy is a double-edged sword – there should be enough privacy to ensure
that sensitive information about the individuals is not disclosed by the views and at the same
time, there should be enough data to perform the data analysis.

Agenda.

Day 1 (June 8), 1 hour

I. What is “Privacy-preserving data publishing”?
   1. There are no much data
   2. A typical scenario of data collection and publishing
   3. Three negative cases (Massachusetts voters, AOL Search Data, Netflix Prize) &
      one positive case (U.S. Census)

II. Individual Anonymity
   1. Individual data anonymity as a property of information about an individual to be
      unidentifiable within a dataset
   2. Microdata and microfile. Explicit Identifiers, Quasi Identifiers, Sensitive Attributes,
      Non-Sensitive Attributes
   3. Anonymization techniques
   4. Information loss

Day 2 (June 9), 1 hour

III. Group Anonymity
   1. Sensitivity to social impact — “Group harm”
   2. Vital and Parameter Attributes of microfile
   3. Group anonymity scheme. Different ways to construct a goal representation
      (Quantity signal, Concentration signal, Concentration difference signal)
   4. Minimizing microfile distortion.