An Evolutionary Optimization Approach for Categorical Data Protection

**Title**: An Evolutionary Optimization Approach for Categorical Data Protection

**Publication Type**: Conference Paper

**Year of Publication**: 2012

**Authors**: Marés J [1], Torra V [2]

**Conference Name**: Privacy and Anonymity in the Information Society 2012

**Conference Location**: Berlin

**Date Published**: 30/03/2012

**Keywords**: Categorical Data [3], data mining [4], Data Privacy [5], Disclosure risk [6], genetic algorithms [7], Information loss [8]

**Abstract**: The continuous growing amount of public sensible data has increased the risk of breaking the privacy of people or institutions in those datasets. Many protection methods have been developed to solve this problem by either distorting or generalizing data but taking into account the difficult trade-off between data utility (information loss) and protection against disclosure (disclosure risk). In this paper we present an optimization approach for data protection based on an evolutionary algorithm which is guided by a combination of information loss and disclosure risk measures. In this way, state-of-the-art protection methods are combined to obtain new data protections with a better trade-off between these two measures. The paper presents several experimental results that assess the performance of our approach.