

**INSTITUCIÓN UNIVERSITARIA POLITÉCNICO GRANCOLOMBIANO**  
**FACULTAD DE INGENIERÍA Y CIENCIAS BÁSICAS**  
**MAESTRÍA EN INGENIERÍA DE SISTEMAS**  
**GRUPO DE INVESTIGACIÓN FICB-PG**  
**LÍNEA DE PROFUNDIZACIÓN: MINERÍA DE DATOS**

**“ANÁLISIS DE RIESGO DE SER VÍCTIMA DE SUPLANTACIÓN DE IDENTIDAD  
CON BASE A INFORMACIÓN PERSONAL EXPUESTA EN INTERNET: UN  
CASO DE ESTUDIO SOBRE PERSONAS NATURALES EN COLOMBIA”**

**PRESENTA:**

**DIEGO ORLANDO ROA MALAGÓN**  
**CÓDIGO 1420020248**

**ASESOR TEMÁTICO:**

**JEFFREY STEVE BORBÓN SANABRIA MSc.**

**Enero de 2018**

## **RESUMEN**

Gracias a la rápida expansión de Internet, en la actualidad las personas cuentan con una enorme cantidad de información disponible en la red. Dichas fuentes de información se alimentan de manera constante con datos que cada individuo carga, publica, almacena en la red de manera consciente o inconsciente. Es aquí donde personas mal intencionadas toman partida y usan esos datos para obtener algún beneficio.

Este trabajo busca estimar el grado de exposición de las personas naturales en Colombia de ser víctimas de suplantación de identidad, con base a la ejecución de un análisis de riesgo a partir de datos recolectados en Internet; utilizando como fuente números de cédula de ciudadanía. También se quiere exponer aquellas fuentes de información que facilitan en mayor medida la obtención de datos personales que eventualmente podrían ser utilizados para la ejecución de suplantación de identidad.

### **PALABRAS CLAVE**

Suplantación de identidad, análisis de riesgo, minería de datos.

## **ABSTRACT**

Because of the fast expansion of the Internet, people have a huge amount of information available. These sources of information feed on a constant basis with data that everybody upload to the network. It is here where criminals take advantage and use that information to commit fraud.

This work seeks to estimate the degree of exposure of Colombian citizens to being victims of identity fraud based on the execution of a risk analysis from data collected on the Internet using as a source ID numbers. It also wants to expose those sources of information that facilitate the collection of personal data that could eventually be used for the commission of identity fraud.

### **KEY WORDS**

Identity fraud, risk analysis, data mining.