

"Cross-Device Consumer Identification"

Submitted by [grigby1](#) on Thu, 02/23/2017 - 11:58am

Title "Cross-Device Consumer Identification"
Publication Type Conference Paper
Year of Publication 2015
Authors [G. Kejela](#), [C. Rong](#)
Conference Name 2015 IEEE International Conference on Data Mining Workshop (ICDMW)
Date Published Nov
Publisher IEEE
ISBN Number 978-1-4673-8493-3
Accession Number 15774351

Keywords [advertising companies](#), [advertising data processing](#), [Computational modeling](#), [Computers](#), [consumer behaviour](#), [consumer identity](#), [cross-device consumer identification](#), [Data models](#), [Decision trees](#), [Deep Learning](#), [Ensemble](#), [GBDT](#), [GBM](#), [gradient boosting decision trees](#), [ICDM2015 contest](#), [Internet](#), [IP networks](#), [learning \(artificial intelligence\)](#), [performance evaluation](#), [personal information](#), [Predictive models](#), [pubcrawl170105](#), [Random Forest](#), [random forest algorithm](#), [Training](#), [Xgboost](#)

Abstract Nowadays, a typical household owns multiple digital devices that can be connected to the Internet. Advertising companies always want to seamlessly reach consumers behind devices instead of the device itself. However, the identity of consumers becomes fragmented as they switch from one device to another. A naive attempt is to use deterministic features such as user name, telephone number and email address. However consumers might refrain from giving away their personal information because of privacy and security reasons. The challenge in ICDM2015 contest is to develop an accurate probabilistic model for predicting cross-device consumer identity without using the deterministic user information. In this paper we present an accurate and scalable cross-device solution using an ensemble of Gradient Boosting Decision Trees (GBDT) and Random Forest. Our final solution ranks 9th both on the public and private LB with F0.5 score of 0.855.

URL <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7395888&isnumber=7395635>
DOI [10.1109/ICDMW.2015.241](https://doi.org/10.1109/ICDMW.2015.241)

Citation
Key

7395888



[advertising companies](#) [advertising data processing](#) [Computational modeling](#) [Computers](#) [consumer behaviour](#) [consumer identity](#) [cross-device consumer identification](#) [Data models](#) [Decision trees](#) [deep learning](#) [Ensemble](#) [GBDT](#) [GBM](#) [gradient boosting](#) [decision trees](#) [ICDM2015 contest](#) [internet IP networks](#) [learning \(artificial intelligence\)](#) [performance evaluation](#) [personal information](#) [Predictive models](#) [pubcrawl170105](#) [Random Forest](#) [random forest algorithm](#) [Training](#) [Xgboost](#)
