

Accepted Manuscript

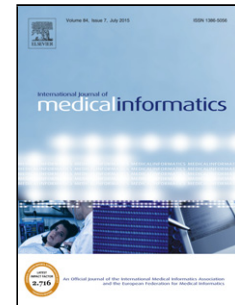
Title: Anticipating Anonymity in Screening Program Databases

Author: Rafael Caballero Sagar Sen Jan F Nygård

PII: S1386-5056(17)30080-1
DOI: <http://dx.doi.org/doi:10.1016/j.ijmedinf.2017.04.003>
Reference: IJB 3494

To appear in: *International Journal of Medical Informatics*

Received date: 18-1-2017
Revised date: 22-3-2017
Accepted date: 1-4-2017



Please cite this article as: Rafael Caballero, Sagar Sen, Jan F Nygård, Anticipating Anonymity in Screening Program Databases, *International Journal of Medical Informatics* (2017), <http://dx.doi.org/10.1016/j.ijmedinf.2017.04.003>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Anticipating Anonymity in Screening Program Databases

Rafael Caballero^{a,1,**}, Sagar Sen^{b,2,*}, Jan F Nygård^{c,3,*}

^a*University Complutense of Madrid*

^b*Certus SFI*

^c*Cancer Registry of Norway (Kreftregisteret)*

Abstract

In this paper, we propose a technique for improving anonymity in screening program databases to increase the privacy for the participants in these programs. The data generated by the invitation process (screening centre, appointment date) is often made available to researchers for medical research and for evaluation and improvement of the screening program. This information, combined with other personal quasi-identifiers such as the ZIP code, gender or age, can pose a risk of disclosing the identity of the individuals participating in the program, and eventually their test results. We present two algorithms that produce a set of screening appointments that aim to increase anonymity of the resulting dataset. The first one, based on the constraint programming paradigm, defines the optimal appointments, while the second one is a suboptimal heuristic algorithm that can be used with real size datasets. The level of anonymity is measured using the new concept of generalized k-anonymity, which allows us to show the utility of the proposal by means of experiments, both using random data and data based on screening invitations from the Norwegian Cancer Registry.

Keywords: Anonymity, screening programs, constraint programming

*Corresponding author

**Principal corresponding author

¹Rafael Caballero partially supported by the Spanish MINECO project CAVI-ART (TIN2013-44742-C4-3-R), Madrid regional project N-GREENS Software-CM (S2013/ICE-2731), UCM grant GR3/14-910502, and the NILS grant

²Funded by the Norwegian Research Council

³Funded by The Cancer Registry of Norway

Download English Version:

<https://daneshyari.com/en/article/4966586>

Download Persian Version:

<https://daneshyari.com/article/4966586>

[Daneshyari.com](https://daneshyari.com)