ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

On the resilience of canonical reducible permutation graphs☆

Lucila M.S. Bento[a,b], Davidson R. Boccardo[b], Raphael C.S. Machado[b], Vinícius G. Pereira de Sá[a,*], Jayme Luiz Szwarcfiter[a,b,c]

[a] *Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*

[b] *Instituto Nacional de Metrologia, Qualidade e Tecnologia, Rio de Janeiro, Brazil*

[c] *COPPE Sistemas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*

ARTICLE INFO

Article history:
Received 11 July 2015
Accepted 28 September 2016
Available online xxxx

This paper is dedicated to Marty Golumbic on his 65th birthday

Keywords:
Reducible permutation graphs
Graph-based watermarking
Linear-time algorithms
Software security

ABSTRACT

An ingenious graph-based watermarking scheme recently proposed by Chroni and Nikolopoulos encodes integers as a special type of reducible permutation graphs. It was claimed without proof that those graphs can withstand attacks in the form of a single edge removal. We introduce a linear-time algorithm which restores the original graph after removals of $k \leq 2$ edges, therefore proving an even stronger result. Furthermore, we prove that $k \leq 5$ general edge modifications (removals/insertions) can always be detected in polynomial time. Both bounds are tight. Our results reinforce the interest in regarding Chroni and Nikolopoulos's scheme as a possible software watermarking solution for numerous applications.

1. Introduction

Among the modern methods to fight the illegal reproduction of software, the embedding of digital watermarks deserves attention. Roughly speaking, software watermarks hide encoded identification data into a program. They allow for the timely retrieval of authorship and/or ownership information, therefore discouraging piracy.

Soon after the creation of the first software watermark in 1996 by Davidson and Myhrvold [8], many interesting ideas have followed, including encoding a binary – the *identifier* – as a special digraph embedded into the software's control-flow graph, an idea which was patented by Venkatesan and Vazirani in 2006 [12]. *Graph-based* watermarking schemes have received a lot of attention ever since, and due emphasis must be given to the contributions of Collberg et al. in a series of papers [6,7,5]. More recently, Chroni and Nikolopoulos presented an ingenious such scheme [3,4], where the generated watermark graphs constitute a subclass of reducible flow graphs [9–11]. Such subclass possesses desirable features, among which its ease of implementation and its linear-time running time. A third feature would be its alleged resilience to attacks. However, though its ability to withstand single edge removals has been conjectured in [4], proving or disproving it was still an open problem.

☆ A preliminary version containing partial results of this paper was presented as an extended abstract entitled "Towards a provably resilient scheme for graph-based watermarking" at the 39th International Workshop on Graph Theoretic Concepts in Computer Science, WG 2013, and appeared in *Lecture Notes in Computer Science* **8165** (2013), 50–63.

* Corresponding author.

E-mail addresses: lucilabento@ppgi.ufrj.br (L.M.S. Bento), drboccardo@inmetro.gov.br (D.R. Boccardo), rcmachado@inmetro.gov.br (R.C.S. Machado), vigusmao@dcc.ufrj.br (V.G. Pereira de Sá), jayme@nce.ufrj.br (J.L. Szwarcfiter).

http://dx.doi.org/10.1016/j.dam.2016.09.038
0166-218X/

In [1,2], a formal characterization of the class of graphs produced by Chroni and Nikolopoulos's encoding function was given. They were referred to as *canonical reducible permutation graphs*. We have also formulated a robust polynomial-time algorithm that, given a watermark with an arbitrary number $k \geq 0$ of deleted edges, either retrieves the encoded identifier or proves that to be an impossible task. In the present paper, we disclose the actual resilience of Chroni and Nikolopoulos's watermark by proposing a linear-time procedure which *always* succeeds in reconstituting a watermark from which $k \leq 2$ edges were removed, a bound which is the best possible. Moreover, our results imply that $k \leq 5$ edge deletions and/or insertions can always be *detected* in polynomial time, a bound that is also tight.

Even though the resilience against two edge removals may look modest, notice that, from the attacker's standpoint, the fact that the watermark can withstand even a single edge removal may already be hard to surmount. Indeed, because the location of the watermark in the software binary is unknown, one cannot do much better than the trial-and-error approach, hoping to spoil the watermark by removing as few as possible arbitrarily chosen edges, so not to spoil the very functionality of the software. If the watermarking scheme is resilient to some number $k \geq 1$ of edge removals, though, then the attacker should remove at least $k+1$ arbitrarily chosen edges, and the probability that the software functionality is affected grows with k. Moreover, the total number of brute-force trials – the $\binom{m}{k+1}$ subsets of $k+1$ edges – also grows fast with k when k is very small compared to the total number m of edges.

This paper is organized as follows. In Section 2, we recall the watermark from Chroni and Nikolopoulos. In Section 3, we revisit some necessary definitions and previous results. In Section 4, we formulate linear-time algorithms to reconstruct the original graph and recover the encoded data even if two edges are missing. The proof of one of the central results in that section, namely Theorem 11, is somewhat involved, and we dedicate to it the whole Section 5. Section 6 concludes the paper with our final remarks.

Throughout the text, we let $V(G)$ and $E(G)$ respectively denote, as usual, the vertex set and edge set of a given graph G. Also, we let $N_G^+(v)$ and $N_G^-(v)$ be the sets of out-neighbors and in-neighbors of vertex v in G, with $d_G^+(v)$ and $d_G^-(v)$ their respective sizes. If J is a subset of either $V(G)$ or $E(G)$, then $G-J$ corresponds to the graph obtained from G by the removal of J.

2. The watermark from Chroni and Nikolopoulos

We recall the encoding algorithm described in [4]. The index of the first element in all considered sequences is 1.

Let ω be a positive integer identifier, and n the size of the binary representation B of ω. Let also n_0 and n_1 be the number of 0's and 1's, respectively, in B, and let f_0 be the index of the leftmost 0 in B. The extended binary B^* is obtained by concatenating n digits 1, followed by the one's complement of B and by a single digit 0. We let $n^* = 2n+1$ denote the size of B^*, and we define $Z_0 = (z_i^0)$, $i = 1, \ldots, n_1+1$, as the ascending sequence of indexes of 0's in B^*, and $Z_1 = (z_i^1)$, $i = 1, \ldots, n+n_0$, as the ascending sequence of indexes of 1's in B^*.

Let S be a sequence of integers. We denote by S^R the sequence formed by the elements of S in backward order. If $S = (s_i)$, for $i = 1, \ldots, t$, and there is an integer $k \leq t$ such that the subsequence consisting of the elements of S with indexes less than or equal to k is ascending, and the subsequence consisting of the elements of S with indexes greater than or equal to k is descending, then we say S is *bitonic*. If all t elements of a sequence S are distinct and belong to $\{1, \ldots, t\}$, then S is a *permutation*. If S is a permutation of size t, and, for all $1 \leq i \leq t$, the equality $i = s_{s_i}$ holds, then we say S is *self-inverting*. In this case, the unordered pair (i, s_i) is called a *2-cycle* of S, if $i \neq s_i$, and a *1-cycle* of S, if $i = s_i$. If S_1, S_2 are sequences (respectively, paths in a graph), we denote by $S_1 \parallel S_2$ the sequence (respectively, path) formed by the elements of S_1 followed by the elements of S_2.

Back to Chroni and Nikolopoulos's algorithm, we define $P_b = (b_i)$, with $i = 1, \ldots, n^*$, as the bitonic permutation $Z_0 \parallel Z_1^R$. Finally, the self-inverting permutation $P_s = (s_i)$ is obtained from P_b as follows: for $i = 1, \ldots, n^*$, element s_{b_i} is assigned value b_{n^*-i+1}, and element $s_{b_{n^*-i+1}}$ is assigned value b_i. In other words, the 2-cycles of P_s correspond to the n unordered pairs of distinct elements of P_b that share the same minimum distance to one of the extremes of P_b, that is, the pairs $(p, q) = (b_i, b_{n^*-i+1})$, for $i = 1, \ldots, n$. Since the central index $i = n+1$ of P_b is the solution of equation $n^* - i + 1 = i$, element b_{n+1} – and no other – will constitute a 1-cycle in P_s. We refer to such element of P_s as its *fixed element*, and we let f denote it.

The watermark generated by Chroni and Nikolopoulos's encoding algorithm [4] is a directed graph G whose vertex set is $\{0, 1, \ldots, 2n+2\}$, and whose edge set contains $4n+3$ edges, to wit: a *path edge* $(u, u-1)$ for $u = 1, \ldots, 2n+2$, constituting a Hamiltonian path that will be unique in G, and a *tree edge* from u to $q(u)$, for $u = 1, \ldots, n^*$, where $q(u)$ is defined as the vertex $v > u$ with the greatest index in P_s to the left of u, if such v exists, or $2n+2$ otherwise. The rationale behind the name *tree edge* is the fact that such edges induce a spanning tree of $G \setminus \{0\}$.

Let us glance at an example. For $\omega = 349$, we have $B = 101011101$, $n = 9$, $n_0 = 3$, $n_1 = 6$, $f_0 = 2$, $B^* = 111111110101000100$, $n^* = 19$, $Z_0 = (10, 12, 14, 15, 16, 18, 19)$, $Z_1 = (1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 17)$, $P_b = (10, 12, 14, 15, 16, 18, 19, 17, 13, 11, 9, 8, 7, 6, 5, 4, 3, 2, 1)$, $P_s = (10, 12, 14, 15, 16, 18, 19, 17, 13, 1, 11, 2, 9, 3, 4, 5, 8, 6, 7)$ and $f = 11$. The watermark associated to ω presents, besides the path edges in the Hamiltonian path $20, 19, \ldots, 0$, the tree edges $(1, 13)$, $(2, 11)$, $(3, 9)$, $(4, 9)$, $(5, 9)$, $(6, 8)$, $(7, 8)$, $(8, 9)$, $(9, 11)$, $(10, 20)$, $(11, 13)$, $(12, 20)$, $(13, 17)$, $(14, 20)$, $(15, 20)$, $(16, 20)$, $(17, 19)$, $(18, 20)$ and $(19, 20)$.

Please cite this article in press as: L.M.S. Bento, et al., On the resilience of canonical reducible permutation graphs, Discrete Applied Mathematics (2016), http://dx.doi.org/10.1016/j.dam.2016.09.038