Applied Mathematical Sciences, Vol. 7, 2013, no. 47, 2335 - 2339
HIKARI Ltd, www.m-hikari.com

Restricted Common Superstrings

Anna Gorbenko

Department of Intelligent Systems and Robotics
Ural Federal University
620083 Ekaterinburg, Russia
gorbenko.ann@gmail.com

Vladimir Popov

Department of Intelligent Systems and Robotics
Ural Federal University
620083 Ekaterinburg, Russia
Vladimir.Popov@usu.ru

Abstract

In this paper we consider an approach to solve the restricted common superstring problem. This approach is based on an explicit reduction from the problem to the satisfiability problem.

Keywords: restricted common superstring, **NP**-complete, satisfiability

Computational complexity of different problems of finding regularities and efficient algorithms for these problems are thoroughly studied in theoretical computer science (see e.g. [1] – [6]). In particular, complexity of the restricted common superstring problem and some approximation algorithms for the problem was considered in [7, 8].

Let $\Sigma = \{a_1, \ldots, a_m\}$ be a finite alphabet. Let $\mathcal{S} = \{S_1, \ldots, S_n\}$ be a set of strings over Σ. We assume that $S[i]$ is the ith letter in string S. Also, we assume that $S[i, j]$ is the substring of S consisting of the ith letter through the jth letter. The length of a string S is the number of letters in it. We assume that $|S|$ is the length of S. We use $\#occ(X, Y)$ to denote $|\{i \mid X = Y[i, j]\}|$.

The decision version of the restricted common superstring problem can be formulated as following.

THE RESTRICTED COMMON SUPERSTRING PROBLEM (RCSSTR):

INSTANCE: *A set $\mathcal{S}$ of strings over Σ, a string T, and a positive integer k.*

QUESTION: *Is there a string S such that $|S| = |T|$, $S[\pi(i)] = T[i]$, for all $1 \leq i \leq |T|$, and $|\{i \mid \#occ(S_i, S) \geq 1\}| \geq k$?*

The problem RCSSTR is **NP**-complete [7]. Note that encoding different hard problems as instances of SAT and solving them with efficient satisfiability algorithms has caused considerable interest (see e.g. [9] – [20]). In this paper, we consider an approach to solve the RCSSTR problem. Our approach is based on an explicit reduction from the problem to the satisfiability problem.

Let

$$\varphi[1] = \wedge_{1 \leq i \leq |T|} \vee_{1 \leq j \leq |T|} w[i, j],$$

$$\varphi[2] = \wedge_{1 \leq i \leq |T|} \wedge_{1 \leq j[1] < j[2] \leq |T|} (\neg w[i, j[1]] \vee \neg w[i, j[2]]),$$

$$\varphi[3] = \wedge_{1 \leq i \leq |T|} \vee_{1 \leq j \leq m} x[i, j],$$

$$\varphi[4] = \wedge_{1 \leq i \leq |T|} \wedge_{1 \leq j[1] < j[2] \leq m} (\neg x[i, j[1]] \vee \neg x[i, j[2]]),$$

$$\varphi[5] = \wedge_{1 \leq i \leq |T|} \wedge_{1 \leq j \leq |T|} \wedge_{1 \leq p \leq m, T[j] \neq a_p} (\neg w[i, j] \vee \neg x[i, p]),$$

$$\varphi[6] = \wedge_{1 \leq i \leq k} \vee_{1 \leq j \leq n} y[i, j],$$

$$\varphi[7] = \wedge_{1 \leq i \leq k} \wedge_{1 \leq j[1] < j[2] \leq n} (\neg y[i, j[1]] \vee \neg y[i, j[2]]),$$

$$\varphi[8] = \wedge_{\substack{1 \leq i \leq k, \\ 1 \leq j \leq n}} (\neg y[i, j] \vee (\vee_{1 \leq p \leq |T| - |S_j| + 1} z[j, p])),$$

$$\varphi[9] = \wedge_{\substack{1 \leq i \leq k, \\ 1 \leq j \leq n}} \wedge_{1 \leq p[1] < p[2] \leq |T| - |S_j| + 1} (\neg y[i, j] \vee \neg z[j, p[1]] \vee z[j, p[2]]),$$

$$\varphi[10] = \wedge_{\substack{1 \leq i \leq k, \\ 1 \leq j \leq n, \\ 1 \leq p \leq |T| - |S_j| + 1, \\ p \leq q \leq p + |S_j| - 1}} \wedge_{1 \leq r \leq m, S_j[q - p + 1] \neq a_r} (\neg y[i, j] \vee \neg z[j, p] \vee \neg x[q, r]),$$

$$\xi = \wedge_{i=1}^{10} \varphi[i].$$

It is easy to check that there is a string S such that $|S| = |T|$, $S[\pi(i)] = T[i]$, for all $1 \leq i \leq |T|$, and $|\{i \mid \#occ(S_i, S) \geq 1\}| \geq k$ if and only if ξ is satisfiable. It is clear that ξ is a CNF. So, ξ gives us an explicit reduction from RCSSTR to SAT. Now, using standard transformations (see e.g. [21]) we can obtain an explicit transformation ξ into ζ such that $\xi \Leftrightarrow \zeta$ and ζ is a 3-CNF. Clearly, ζ gives us an explicit reduction from RCSSTR to 3SAT.

We have designed generators of natural instances for RCSSTR. We consider our genetic algorithms OA[1] (see [22]), OA[2] (see [23]), OA[3] (see [24]), and OA[4] (see [25]) for SAT. We used heterogeneous cluster. Each test was runned

time	average	max	best
OA[1]	33.7 min	2.32 h	7.11 min
OA[2]	1.21 h	2.94 h	3.44 min
OA[3]	1.43 h	2.27 h	9.52 min
OA[4]	48.2 min	2.66 h	4.83 min

Table 1: Experimental results for RCSSTR.

on a cluster of at least 100 nodes. Selected experimental results are given in Table 1.

ACKNOWLEDGEMENTS. The work was partially supported by Analytical Departmental Program "Developing the scientific potential of high school" 8.1616.2011.

References

[1] V. Yu. Popov, Computational complexity of problems related to DNA sequencing by hybridization, *Doklady Mathematics*, 72 (2005), 642-644.

[2] V. Popov, The approximate period problem for DNA alphabet, *Theoretical Computer Science*, 304 (2003), 443-447.

[3] V. Popov, The Approximate Period Problem, *IAENG International Journal of Computer Science*, 36 (2009), 268-274.

[4] V. Popov, Approximate Periods of Strings for Absolute Distances, *Applied Mathematical Sciences*, 6 (2012), 6713-6717.

[5] V. Popov, Multiple genome rearrangement by swaps and by element duplications, *Theoretical Computer Science*, 385 (2007), 115-126.

[6] V. Popov, Sorting by prefix reversals, *IAENG International Journal of Applied Mathematics*, 40 (2010), 247-250.

[7] R. Clifford, Z. Gotthilf, M. Lewenstein and A. Popa, Restricted Common Superstring and Restricted Common Supersequence, *Lecture Notes in Computer Science*, 6661 (2011), 467-478.

[8] P. Bonizzoni, R. Dondi, G. Mauri and I. Zoppis, Restricted and Swap Common Superstring: A Parameterized View, *Lecture Notes in Computer Science*, 7535 (2012), 49-60.

[9] A. Gorbenko and V. Popov, The Minimum Test Collection Problem, *Applied Mathematical Sciences*, 7 (2013), 1191-1193.

[10] A. Gorbenko and V. Popov, The Farthest Substring Problem, *Applied Mathematical Sciences*, 7 (2013), 1209-1212.

[11] A. Gorbenko and V. Popov, Computational Experiments for the Problem of Hamiltonian Path with Fixed Number of Color Repetitions, *Advanced Studies in Theoretical Physics*, 7 (2013), 121-126.

[12] A. Gorbenko and V. Popov, On Hamilton Paths in Grid Graphs, *Advanced Studies in Theoretical Physics*, 7 (2013), 127-130.

[13] A. Gorbenko and V. Popov, Computational Experiments for the Problem of Sensor-Mission Assignment in Wireless Sensor Networks, *Advanced Studies in Theoretical Physics*, 7 (2013), 135-139.

[14] A. Gorbenko and V. Popov, The Swap Common Superstring Problem, *Applied Mathematical Sciences*, 7 (2013), 609-614.

[15] A. Gorbenko and V. Popov, The String Barcoding Problem, *Applied Mathematical Sciences*, 7 (2013), 615-622.

[16] A. Gorbenko and V. Popov, On Multiple Occurrences Shortest Common Superstring Problem, *Applied Mathematical Sciences*, 7 (2013), 641-644.

[17] A. Gorbenko and V. Popov, Longest Common Parameterized Subsequences with Fixed Common Substring, *Applied Mathematical Sciences*, 7 (2013), 645-650.

[18] A. Gorbenko and V. Popov, The Hamiltonian Alternating Path Problem, *IAENG International Journal of Applied Mathematics*, 42 (2012), 204-213.

[19] A. Gorbenko and V. Popov, The Problem of Finding Two Edge-Disjoint Hamiltonian Cycles, *Applied Mathematical Sciences*, 6 (2012), 6563-6566.

[20] A. Gorbenko and V. Popov, The set of parameterized k-covers problem, *Theoretical Computer Science*, 423 (2012), 19-24.

[21] A. Gorbenko and V. Popov, The c-Fragment Longest Arc-Preserving Common Subsequence Problem, *IAENG International Journal of Computer Science*, 39 (2012), 231-238.

[22] A. Gorbenko and V. Popov, On the Problem of Placement of Visual Landmarks, *Applied Mathematical Sciences*, 6 (2012), 689-696.

[23] A. Gorbenko and V. Popov, Computational Experiments for the Problem of Selection of a Minimal Set of Visual Landmarks, *Applied Mathematical Sciences*, 6 (2012), 5775-5780.

[24] A. Gorbenko and V. Popov, Task-resource Scheduling Problem, *International Journal of Automation and Computing*, 9 (2012), 429-441.

[25] A. Gorbenko and V. Popov, SAT Solvers for the Problem of Sensor Placement, *Advanced Studies in Theoretical Physics*, 6 (2012), 1235-1238.

Received: February 12, 2013