Operations Research Letters 41 (2013) 664–669

Contents lists available at ScienceDirect

Operations Research Letters

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

Routing by ranking: A link analysis method for the constrained dial-a-ride problem

Lauri Häme*, Harri Hakula

Department of Mathematics and Systems Analysis, Aalto University School of Science, P.O. Box 11100, FI-00076 Aalto, Finland

ARTICLE INFO

Article history:
Received 16 March 2012
Received in revised form
13 September 2013
Accepted 24 September 2013
Available online 27 September 2013

Keywords:
Dial-a-ride problem
Link analysis
HITS
Topological ordering

ABSTRACT

The dial-a-ride problem involves the dispatching of a fleet of vehicles in order to transport a set of customers from specific pick-up nodes to specific drop-off nodes. Using a modified version of hyperlink-induced topic search (HITS), we characterize hubs as nodes with many out-links to other hubs and calculate a hub score for each pick-up and drop-off node. Ranking the nodes by hub score gives guidance to a backtracking algorithm for efficiently finding feasible solutions to the dial-a-ride problem.

0. Introduction

Several web information retrieval (IR) methods have been developed for finding the most appropriate web pages corresponding to queries given to search engines. The most sophisticated methods, such as HITS [21], PageRank [7] and SALSA [23] in use today make use of the hyperlinked structure of the web, since the goodness of a web page and the position of the page with respect to other web pages seem to have a certain connection. For example, a web page may be considered good if there are many other web pages linking to that page. In other words, web pages are ranked by search engines not only by means of the content of the page, but also by exploiting information regarding the hyperlink-induced relationships between pages.

The bringing of hyperlinks to bear on the ordering of web pages has given rise to a mathematical analysis related to hyperlink-induced web IR methods, such as in [14,22,25,1], in which the behavior of several IR methods is studied from the computational point of view.

In this work, we focus on the HITS (Hyperlink-Induced Topic Search) algorithm, which defines *authorities* (web pages with several in-links) and *hubs* (several out-links). The HITS thesis is that *good hubs point to good authorities and good authorities are pointed to by good hubs*. Based on this thesis, HITS assigns both a hub score and authority score to each web page [22]. In this paper, we use the term "hub" similarly as in the above context and not in the context of travel dispatch centers.

We present an application of HITS on the dial-a-ride problem (DARP) [27,20,24,29,9,12,30,8,28,17,2–4,15], in which the goal is to construct a set of vehicle routes, serving a set of customers, satisfying the given time, capacity and precedence constraints. The DARP is examined as a *constraint satisfaction problem*, in which the goal is to find a set of m feasible vehicle routes that serve all customers, where m is the number of vehicles, as in [4]. This reference suggests that an algorithm for checking the feasibility of a DARP instance has two main applications: (1) Determining the feasibility can be the first phase in an optimization algorithm in a *static* setting, where all trip requests are known, for example, one day in advance. (2) In *dynamic* services, a constraint satisfaction algorithm can be used for deciding whether to accept or reject incoming user requests.

In the context of the dial-a-ride problem, we define links between nodes as feasible transitions with respect to the constraints of the problem: If node j can be visited after i, a link from i to j is formed. Thus, a good hub score of a pick-up or drop-off node i means that many nodes can be reached in time from i. Thus, in order to efficiently find feasible solutions to the dial-a-ride problem, we suggest that nodes with large hub scores should be visited first since there are many nodes that can be visited after such nodes.

We emphasize that the algorithm presented in this paper is designed for problems with tight time windows and that the goal is to find a feasible solution instead of an optimal one. Thus, in contrast to most works related to vehicle routing and dial-a-ride problems, no specific cost function is considered. For approaches with cost functions, we refer to [5,6,16,9,3].

This work is partly motivated by a flexible transport service currently operating in Helsinki. The service, run by Helsinki Region Transport, The Finnish Transport Agency and Aalto University, was

* Corresponding author. Tel.: +358 40 576 3585; fax: +358 9 470 23016.
E-mail addresses: lauri.hame@gmail.com, Lauri.Hame@aalto.fi (L. Häme), Harri.Hakula@aalto.fi (H. Hakula).

0167-6377/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.orl.2013.09.007

deployed at the end of 2012. The service is designed to operate on a demand-responsive basis, that is, each trip is booked in advance and the vehicle routes are modified according to the trips. The main difference from previous services is that no pre-order times for trips are required and the trips can be booked "on the fly" by means of an interactive user interface.

1. The ranking method

In the following, we introduce a ranking method based on the HITS algorithm (Sections 1.1 and 1.2) and show how it can be used to solve the dial-a-ride problem (Section 1.3).

1.1. The HITS algorithm [21]

Given a web graph $G = (V, A)$ consisting of pages V and links A between pages, the authority and hub scores a_i and h_i are computed for each page i as follows. Letting (i, j) represent a link from page i to page j, given that each page has been assigned an initial authority score $a_i(0)$ and hub score $h_i(0)$, HITS successively refines these scores by computing

$$a_i(k) = \sum_{(j,i)\in A} h_j(k-1), \qquad h_i(k) = \sum_{(i,j)\in A} a_j(k-1)$$

for $k \in \{1, 2, \ldots\}$. By using matrix notation, these equations can be written in the form $a(k) = L^T h(k-1)$ and $h(k) = La(k-1)$, where $a(k)$ is the authority vector containing the authority scores of each page at step k, $h(k)$ is the corresponding hub vector and L is the adjacency matrix of the graph with elements $L_{i,j} = 1$ if $(i, j) \in A$ and $L = 0$ otherwise [22].

It has been shown in [14] that the authority and hub vectors describing the authority and hub scores of nodes of a given graph in the limit are given by the dominant eigenvectors of the matrices $L^T L$ and LL^T (or equivalently, dominant singular vectors of L), where L is the adjacency matrix of the graph.

1.2. Modified HITS

For constrained routing problems, we present a modified version of the HITS algorithm, in which only the hub scores are considered. Our thesis is that *good hubs point to good hubs*. This formulation is motivated by Theorem 1, which states that for a specific class of graphs, the hub score of a node i corresponds to the number of self-avoiding paths [26] from i to a given destination node. When constructing a path that visits all nodes, or maximizing the number of visited nodes, the modified HITS idea induces an intuitive policy: Good hubs are visited first, since many nodes can be reached from good hubs.

The hub scores are calculated as follows. Let L denote the adjacency matrix of a directed graph G. Similarly as in the HITS algorithm, the hub vector containing the *hub scores* of nodes is first initialized, $h(0) = (1, 1, \ldots, 1)$ and the hub vector is successively updated by means of the power method

$$h(k) = Lh(k-1) \quad \text{for } k \in \{1, 2, \ldots\}. \tag{1}$$

Similarly as in the original HITS algorithm, the hub vector converges to a dominant eigenvector of L.

Theorem 1 characterizes the hub scores produced by the modified HITS algorithm for *sink graphs* defined as follows.

Definition. Let $G = (V, A)$ be a directed acyclic graph and let $s \in V$ be a node such that $(s, i) \notin A$ for all $i \in V$. The graph $G_s = (V, A \cup (s, s))$ is called a *sink graph*.

In other words, a sink graph is a directed acyclic graph (V, A) with the exception that one node $s \in V$ with zero out-degree is associated with a loop (s, s).

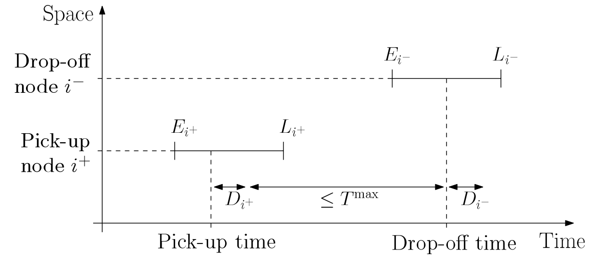

Fig. 1. Pick-up and drop-off time windows. The pick-up point of customer i is denoted by i^+ and the drop-off point is denoted by i^-. The customer should be picked up at i^+ within the time window $[E_{i^+}, L_{i^+}]$ and the customer should be dropped off at i^- within the time window $[E_{i^-}, L_{i^-}]$. The service times needed for the customer to get on the vehicle and get off the vehicle are denoted by D_{i^+} and D_{i^-}. The time between the drop-off and the pick-up (excluding D_{i^+}) should not exceed the maximum ride time $T^{\max}$.

Theorem 1. *Let L denote the adjacency matrix of a sink graph $G_s = (V, A)$, where $V = \{1, \ldots, |V|\}$, let h_i denote the number of self-avoiding paths from i to s for $i \in V \setminus \{s\}$ and let $h_s = 1$. Then, $h = (h_1, \ldots, h_{|V|})^T$ is a unique dominant eigenvector of L.*

Proof. Since the adjacency matrix of a directed acyclic graph is an upper triangular matrix with zeros on the diagonal, the adjacency matrix of a sink graph is an upper triangular matrix with diagonal elements $L_{i,i} = 0$ except for the sink node s, for which we have $L_{s,s} = 1$. The eigenvalues of an upper triangular matrix are equal to the diagonal elements and thus there exists a unique dominant eigenvalue 1. Let us show that $h = (h_1, \ldots, h_{|V|})^T$ is the corresponding eigenvector of L.

Since all paths in a directed acyclic graph are self-avoiding and by definition we have $h_s = 1$, the number h_i of self-avoiding paths from node i to the sink node s in the sink graph G_s satisfies $h_i = \sum_{j\in V} L_{i,j} h_j$ for $i \in V \setminus \{s\}$ and h_s satisfies $h_s = 1 = L_{s,s} h_s = \sum_{j\in V} L_{s,j} h_j$. In matrix form, we have $h = Lh$ and thus h is the unique eigenvector corresponding to eigenvalue 1. □

Note that since $h_i \geq h_j$ for all $i, j \in V$ for which $L_{i,j} = 1$, the vector h defines a *topological ordering* [10] of the nodes for which $h_i > 0$. Although Theorem 1 considers a special class of graphs, the result gives us an idea of the behavior of the modified HITS method: There are many paths beginning from nodes with high hub scores.

In the following, we show how the hub scores are used to guide a backtracking algorithm for the dial-a-ride problem.

1.3. The dial-a-ride problem

The dial-a-ride problem is defined as follows [4]. Let $G = (V, A)$ be a complete and directed graph with node set $V = \{0\} \cup P$, where node 0 represents the depot, and P represents the set of pick-up and drop-off nodes, where $(|P| = 2n)$. The set P is partitioned into sets P^+ (pick-up nodes) and P^- (drop-off nodes). Each arc $(i, j) \in A$ has a non-negative travel time T_{ij}. The travel times satisfy the triangular inequality, that is, $T_{ij} + T_{jk} \geq T_{ik}$ for all $i, j, k \in V$. With each node $i \in V$ associate a time window $[E_i, L_i]$, a service duration D_i and a load q_i, where $D_0 = 0$ and $q_0 = 0$. Let $H = \{1, \ldots, n\}$ be the set of customers and let $T^{\max}$ be the maximum ride time for any customer [19]. The maximum ride time constraints are taken into account by updating the upper bounds L_i of the time windows during run-time (one method for doing this is described in [18]). With each customer i is associated a pickup node $i^+ \in P^+$, a delivery node $i^- \in P^-$ and a load $q_{i^+} = q_{i^-}$. The time parameters are illustrated in Fig. 1.

Let $\{1, \ldots, m\}$ be the set of available vehicles, each with capacity Q. A *route* is a directed circuit over a set of nodes in P, starting and finishing at node 0. The goal is to construct m vehicle routes