Advances in QoS Path(s) Selection Problem

Krishnaiyan Thulasiramn, Ying Xiao
School of Computer Science
University of Oklahoma
Norman OK, USA
{thulasi, ying_xiao}@ou.edu

Guoliang Xue
Department of Computer Science and Engineering
Arizona State University
Tempe AZ, USA
xue@asu.edu

Abstract— In this paper we provide a survey of three classes of combinatorial optimization problems that arise in communication network analysis. These problems, referred to as QoS (Quality of Service) route selection problems, have assumed great importance in view of the stringent demands imposed on communication protocols to provide certain quality of service guarantees. We first give a broad overview of advances in these areas and then provide a discussion of algorithms that are of fundamental importance and significance to the networking community as well as operations research and computer science and engineering communities.

I. INTRODUCTION

Recently, there has been considerable emphasis on designing communication protocols which deliver certain end-to-end performance guarantees. This is the result of an explosive growth in applications such as digital video and audio which require stringent quality of service (QoS) guarantees. In this context three classes of combinatorial optimization problems arise. These problems referred to as QoS route selection problems are defined below.

Consider a directed network $G(V, E)$ where V is the set of nodes and E is the set of links of the network. Each link $(u, v) \in E$ is associated with two integer weights $c_{uv} > 0$ (cost) and $d_{uv} > 0$ (delay). For any path p define

$$c(p) = \sum_{(u, v) \in p} c_{uv} \text{ and } d(p) = \sum_{(u, v) \in p} d_{uv} .$$

Given two nodes s, t and integer $\Delta > 0$, a directed s-t path p is said to be feasible if $d(p) \leq \Delta$.

CSP (Constrained Shortest Path) problem: Find an s-t path $p_{opt} = \arg \min\{c(p)| \ p \text{ is a feasible } s\text{-}t \text{ path}\}$. This is illustrated with the example in Fig. 1.

CSDP(k) (Constrained k Shortest Disjoint-Paths) problem: Here the objective is to find a set of k link-disjoint paths between a source node and a destination node with minimum total cost and with the total delay satisfying certain pre-specified bound. This problem arises in the context of providing alternate QoS paths to achieve protection against link failures.

MP-DCP Problem (Most Probable Delay Constrained Path Selection under Inaccurate Information): Objective here is to identify a path that has the highest probability of satisfying a delay bound. The delay of each link is a random variable. This problem is of great importance since accurate state of a network (parameter information) is not often available.
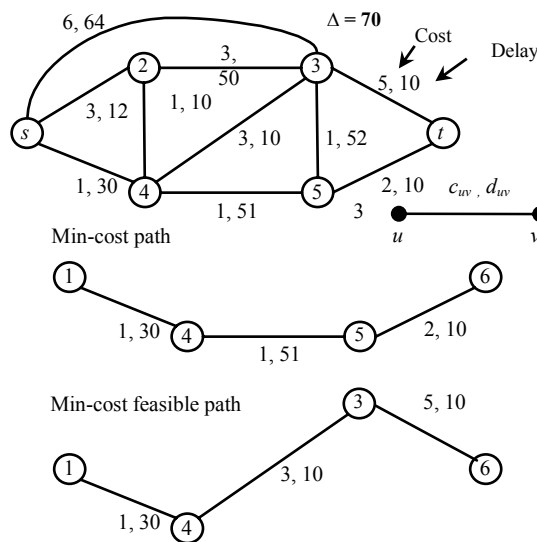

Figure 1.   An example of CSP problem

The rest of the paper is organized as follows. We first provide in section II a broad overview of current literature on advances in this area. We then provide in sections III-V a discussion of algorithms that are of fundamental importance and significance for the three problems defined above.

II. OVERVIEW OF LITERATURE

The CSP problem is NP-complete [1]. So, in the literature, heuristic approaches and approximate algorithms have been proposed. Heuristics, in general, do not provide performance guarantees on the quality of solution produced, though they are usually fast in practice. On the other hand, ε-

approximation algorithms deliver solutions within arbitrarily specified precision requirement. To the best of our knowledge, the first ε-approximation algorithm for the CSP problem was developed by Warburton [2] on acyclic networks. His work was improved later by Hassin [3] and Lorenz and Raz [4]. The best achieved time complexity due to Lorenz and Raz is $O(m\,n\,(log\,log\,n + \varepsilon^{-1}))$, where m and n are the number of links and number of nodes of the network [4]. Given a specified tolerance on the delay of a path, in [5] the authors presented an approximation algorithm that produces a minimum cost path satisfying this tolerance. A simple and fast heuristic algorithm LHWHM based on Dijkstra's shortest path algorithm and its variations based on Bellman-Ford-Moore algorithm are discussed in [6]. In perhaps the most recent work the CSP problem under multiple constrains and related concepts are discussed in [7].

The CSP problem can be formulated as an integer linear programming (ILP) problem that is NP-hard due to the integrality constraints. Relaxing the integrality constraints is a natural strategy to design heuristics. The relaxed ILP problem can be solved using different tools available in linear programming literature. To the best of our knowledge, all these algorithms are based on solving the dual of the relaxed ILP formulation of the CSP problem. These algorithms are based on convex hull [8] or Lagrangian relaxation [9], [10] and the linear programming problem is reduced to a series of shortest path problems. We call these algorithms as LARAC proposed in [10]. Some variations of LARAC can be found in [11]. Recently, a parametric search based strong polynomial algorithm was proposed for the relaxed CSP problem [12]. This algorithm has the best time complexity to date. In a most recent work [13] the CSP problem was studied using the primal simplex method of linear programming and the algorithm presented has performance comparable to LARAC. The relaxation of the integrality constraints results in a gap between the actual optimal solution to the original problem and the optimal solution to the relaxed problem. Ziegelmann [14] pointed that the gap may be arbitrarily large for specially constructed networks. Closing the gap involves a second phase for applications demanding stringent quality of the solution. In [9], [12] and [14], the techniques to close the gap are discussed.

The literature on the CDSP(k) and MP-DCP problems will be presented in sections IV - VI.

III.   LP BASED ALGORITHMS FOR THE CSP PRROBLEM

The CSP problem can be presented as a linear programming problem as in equations (1) - (4).

$$\textbf{CSP}: \text{Min} \sum_{(u,\,v)\,\in\,E} c_{uv}\,x_{uv} \qquad (1)$$

Subject to

$$\sum_{\{v|(u.v)\in E\}} x_{uv} - \sum_{\{v|(v.u)\in E\}} x_{vu} = \begin{cases} 1 & if \quad u = s \\ -1 & if \quad u = t \\ 0 & otherwise \end{cases} \qquad (2)$$

$$\sum_{(u,\,v)\,\in\,E} - d_{uv}\,x_{uv} - w = -\Delta,\, w \geq 0 \qquad (3)$$

$$x_{uv} \in \{0,\,1\}, \text{ for all } (u,\,v) \in E. \qquad (4)$$

In (3), w is the slack variable for the delay constraint. The CSP problem can be seen as a minimum cost flow problem with a side constraint (3).

The main difficulty with the CSP problem lies with the integrality condition that requires that the variables x_{uv} be 0 or 1. Removing or relaxing this requirement from the above integer linear program leads to RELAX-CSP, the relaxed CSP problem. It is often convenient to solve the dual of the relaxed form of the CSP problem which we present below.

A.   Dual Based Approach: LARAC Algorithm

The dual involves s-t paths and a variable $\lambda \geq 0$. For each link $(u,\,v)$, let the aggregated cost c_λ be defined as $c_{uv} + \lambda\,d_{uv}$. For a given λ, let $c_\lambda(p) = c(p) + \lambda\,d(p)$ denote the aggregated cost of the path p. Finally define $L(\lambda)$ as:

$$L(\lambda) = \text{min}\{c_\lambda(p): p \in P(s,\,t)\} - \lambda\,\Delta. \qquad (5)$$

Note that in the above, $\text{min}\{c_\lambda(p): p \in P(s,\,t)\}$ is the same as the minimum aggregated cost of an s-t path with respect to a given value of λ. This can be easily obtained by applying Dijkstra's algorithm using aggregated link costs. Let the s-t path which has minimum aggregated cost with respect to a given λ be denoted as p_λ. Then $L(\lambda) = c_\lambda(p_\lambda) - \lambda\,\Delta$ and the dual of the RELAX-CSP can be presented as follows.

DUAL- RELAX-CSP: Max $L(\lambda)$ for $\lambda \geq 0$.

We note that the problem of maximizing $L(\lambda)$ as above is also called the Lagrangian dual problem. The value of λ that achieves the maximum $L(\lambda)$ in DUAL-RELAX-CSP will be denoted by λ^*. Note that L^*, the optimum value of DUAL-RELAX-CSP is a lower bound on the optimum cost of the path solving the corresponding CSP problem. The key issue in solving DUAL-RELAX-CSP is how to search for the optimal λ. The LARAC algorithm of [10] presented in Fig. 2 is one such efficient search procedure.

Procedure LARAC ($s,\,t,\,\Delta$)
  $p_c := Dijkstra(s,\,t,\,c)$ /*Compute minimum cost s-t path*/
  if $d(p_c) \leq \Delta$ then return p_c
  $p_d := Dijkstra(s,\,t,\,d)$ /*Compute minimum delay s-t path*/
  if $d(p_d) > \Delta$ then return "no solution"
  repeat
     $\lambda = (c(p_c) - c(p_d)) / (d(p_d) - d(p_c))$
    /*Compute minimum aggregate cost s-t path*/
     $\gamma = Dijkstra(s,\,t,\,c_\lambda)$
     if $c_\lambda(\gamma) = c_\lambda(p_c)$ then return p_d
     else if $d(r) \leq \Delta$ then $p_d := \gamma$ else $p_c := \gamma$
  end repeat
end procedure

Figure 2.   LARAC algorithm

B. Primal Simplex Based Approach

In [13] an efficient algorithm based on the revised primal simplex method, denoted as NBS algorithm, was proposed to solve the RELAX-CSP problem. To solve the CSP problem efficiently, the problem is transformed to an equivalent one on a transformed network defined as follows. The transformed network is the same as the original network except that each link delay in the transformed network is twice the corresponding link delay in the original network. Also the path delay bound is set equal to $2\Delta + 1$.

The transformed problem is denoted as the TCSP problem. It can be shown that an s-t path p^* is a feasible solution (an optimal solution) to the CSP problem iff it is a feasible solution (an optimal solution) to the TCSP problem.

Efficiency is achieved by adopting the following strategies: (1) Systems of linear equations are solved using the special graph structure of basic solutions. (2) Effective pivot rules are developed to avoid cycling and speedup the convergence of the simplex algorithm. (3) A method is proposed to avoid any degenerate pivots.

NBS algorithm has be proven to terminate within $2n^3D^2C$ pivots, where n is the number of nodes and D (*resp*. C) is the maximal link delay (*resp*. cost). Extensive numerical simulations show that NBS algorithm outperforms currently available algorithms.

Basically, the LARAC and NBS algorithms cannot provide the performance guarantee beforehand because the gap between the exact optimal solution and the solution got by LARAC or NBS can be arbitrarily large (see [13] and [14]). However, from the solution of the NBS algorithm, we can evaluate the quality of the solution as given in the following theorem. Note that a basic solution is associated with a spanning tree of G or a spanning tree with a cycle.

Theorem 1: If the optimal basic solution is a spanning tree T, the s-t path in T is the optimal solution to the CSP problem. Otherwise, the flow of the optimal basic solution to RELAX-TCSP problem can be decomposed exactly onto two paths p_1 and p_2 with $d(p_1) \leq \Delta < d(p_2)$, $c(p_1) \geq c(p_{opt}) \geq c(p_2)$, $c(p_1) / c(p_{opt}) \leq 1 + (1 - c(p_2) / c(p_1))(1 - \lambda) / \lambda$ and $d(p_2) / \Delta \leq 1 + (1 - d(p_1) / \Delta) \lambda / (1 - \lambda)$, where λ (*resp*. $1 - \lambda$) is the flow on path p_1 (*resp*. p_2) at termination.

IV. QoS-Disjoint Paths Selection Problem

In the QoS-disjoint paths selection problem, we need to search for a set of disjoint (link disjoint or node disjoint) s-t paths with each of them satisfying certain delay bounds and minimizing the total cost. This problem is computationally prohibitive. Even to find two disjoint paths satisfying the delay bounds is NP-hard and is not $2 - \varepsilon$ approximable for any $\varepsilon > 0$ [15]. In the literature, only heuristics are available for the QoS-disjoint paths selection problem. All these heuristics only take care of the total delay and total cost of disjoint paths which is actually the CSDP(k) problem defined in section I.

A. Primal Simplex Approach

The NBS algorithm [13] can also be generalized to the disjoint paths selection problem. Using the same transformation as in Sec. III and relaxing the integrality requirement, the CSDP(k) problem can be formulated as follows [16]:

RELAX-TCSDP(k): Min $\sum_{(u, v) \in E} c_{uv} x_{uv}$ (6)

Subject to

$$\sum_{\{v|(u,v)\in E\}} x_{uv} - \sum_{\{v|(v,u)\in E\}} x_{vu} = \begin{cases} k & if \quad u = s \\ -k & if \quad u = t \\ 0 & otherwise \end{cases}$$ (7)

$$\sum_{(u, v) \in E} - d_{uv} x_{uv} - w = -\Delta, \; w \geq 0$$ (8)

$$0 \leq x_{uv} \leq 1, \text{ for all } (u, v) \in E \; .$$ (9)

B. Generalized LARAC algorithm

A generalized version of LARAC algorithm G-LARAC applicable to a more general class of optimization problems is presented in [17]. By replacing the shortest path algorithm in Fig. 2 with Suurballe's disjoint paths algorithm [18] to compute k disjoint paths with minimum total cost will give an approximation algorithm for the CSDP(k) problem.

V. ε- Approximation Algorithms

A. ε-Approximation Algorithms for the CSP Problem

An approximation algorithm for a minimization problem obtains a solution whose cost is within a specified multiple of the optimum cost. This idea is formally stated as follows [19]. An approximation scheme for a problem P is an algorithm that, given an instance I and a desired degree of accuracy $\varepsilon > 0$, constructs a problem solution with value $F(I)$, such that, if $F^*(I) > 0$ is the value of an optimal solution to I, then

$$| F^*(I) - F(I)| / F^*(I) \leq \varepsilon \; .$$

A fully polynomial time approximation scheme (FPAS) for a graph optimization problem is an approximation scheme whose computing time is a polynomial function of the problem size and $1/\varepsilon$. A strongly polynomial time approximation scheme for a graph/network optimization problem is an approximation scheme whose computing time is a polynomial function of m, n and $1/\varepsilon$. In the literature, there has been an extensive discussion of approximation algorithms for the CSP problem. Hassin [3] presented a fully polynomial time ε-approximation algorithm and Lorenz and Raz [4] presented a strongly polynomial time approximation algorithm.

B. Orda's Approximation Algorithm

Orda *el al*. studied the CSDP(2) problem in [20]. They proposed (α, β)-approximation algorithms defined next. Given an instance of the CSDP(2) problem, an (α, β)-

approximate solution $P = (p_1, p_2)$ to CSDP(2) is a solution for which

$$d(p_1) + d(p_2) \leq \alpha \Delta \text{ and } c(p_1) + c(p_2) \leq \beta\ OPT, \text{ where}$$

OPT is the total cost of the optimal solution.

Four approximation algorithms with different trade off between the cost (the value of α) and delay (the value of β) were proposed in [20] (see Table 1). These algorithms are extensions of the minimum cost flow augmentation path algorithms. In Table 1, the parameters ε and w capture the trade-off between the violation of the delay constraint, the cost of the approximate solution and the computational complexity of the algorithms. For a positive integer w, γ is a small value bounded by $2\ (\log w + 1)\ /\ w$.

TABLE I.   APPROXIMATION ALGORITHMS FOR CSDP(2) PROBLEM.

Alg.	Approx. Ratio	Complexity
DP-1	$(1.5, 1.5\ (1 + \varepsilon))$	$O(mn(1\ /\ \varepsilon + \log \log n))$
DP-2	$(1 + 1\ /\ w, w\ (1 + \gamma))$	$O(mn\ OPT \log w \log(CD))$
DP-3	$(1+1/w, w\ (1+ \gamma)(1 + \varepsilon))$	$O(\log\ (CD)\ mn^2/\varepsilon)$
DP-4	$(1 + 1/w, w\ (1 + \gamma)(1+\varepsilon))$	$O(\log\ (CD)\ (mn^2w^2\log w\ )/\varepsilon)$

VI.   MP-DCP PRBLEM: DELAY CONSTRAINTED PATH SELECTION UNDER INACCURATE STATE INFORMATION

The above results for the CSP problem have been developed assuming that the exact state of the network is known. However, in practice this is not the case. For several reasons [21], full knowledge of the network state is not available. This has led researchers to study the routing problem with uncertain parameters [21], [22], [23], and [24]. The objective in these papers is to identify a path that is most likely to satisfy the delay requirement. This problem is referred to as the MP-DCP problem. In their pioneering works [20] and [21], Orda and Sprintson studied several aspects of this problem and related computational issues. Unlike the CSP problem that involves two deterministic metrics, namely, link cost and link delay, only one link metric, say delay, is considered in the MP-DCP problem. In [22], a heuristic similar to LARAC algorithm was proposed based on the assumption that the path delay is normally distributed. This assumption is fully justified by the central limit theorem. In [24], Xiao *et al*. proposed an exact algorithm, a FPAS algorithm and a strongly polynomial algorithm for the MP-DCP problem.

VII.   SUMMARY

We have presented an extensive survey of advances in three classes of QoS path(s) selection problems. The results presented constitute significant and fundamental contributions of interest to researchers in networking, computer science, and operations research.

REFERENCES

[1]   Z. Wang and J. Crowcroft, "Quality-of-Service routing for supporting multimedia applications," *IEEE JSACs*, vol.14, no.7, pp.1228-1234, Sept. 1996.

[2]   A. Warburton, Approximation of pareto optima in multiple-objective shortest path problems, *Operations Research*, 35:70-79,1987

[3]   R. Hassin, "Approximation schemes for the restricted shortest path problem," *Math. of Oper. Res.*, 17(1), 1992, pp.36-42.

[4]   D. Lorenz and D. Raz, "A simple efficient approximation scheme for the restricted shortest paths problem," *Oper. Res. Letters*, vol. 28, pp. 213-219.

[5]   A. Goel, K. G. Ramakrishnan, D. Kataria and D. Logothetis, "Efficient computation of delay-sensitive routes from one source to all destinations," IEEE INFOCOM, 2001.

[6]   R. Ravindran, K. Thulasiraman, A. Das, K. Huang, G. Luo and G. Xue, "Quality of services routing: heuristics and approximation schemes with a comparative evaluation," *ISCAS*, 2002.

[7]   P. V. Mieghem and F. A. Kuipers, "Concepts of exact QoS routing algorithms," IEEE/ACM Trans. on Networking, vol. 12, pp. 851-864, Oct. 2004.

[8]   K. Mehlhorn and M. Ziegelmann, "Resource constrained shortest path," *Proc. 8th European Symposium on Algorithms* (*ESA*2000)

[9]   G. Handler and I. Zang, "A dual algorithm for the constrained shortest path problem," *Networks* 10, 293-310, 1980

[10] A. Jüttner, B. Szviatovszki, I. Mécs and Z. Rajkó, "Lagrange relaxation based method for the QoS routing problem," *IEEE INFOCOM*-2001, pp. 859-868.

[11] G. Xue, "Minimum-cost QoS multicast and unicast routing in communication networks," IEEE Trans. On Comm., vol. 51, no.5, May 2003, pp.817-824.

[12] Y. Xiao, K. Thulasiraman, G. Xue and A. Jüttner, "The constrained shortest path problem: algorithmic approaches and an algebraic study with generalization," submitted for publication.

[13] Y. Xiao, K. Thulasiraman and G. Xue, "The primal approach to the QoS routing problem," in *proc. First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks*, 2004.

[14] Mark Ziegelmann, Constrained Shortest Paths and Related Problems, PhD thesis, Max-Planck Institute für Informatik, Germany, 2001

[15] C.-L. Li, S. T. McCormick, and D. Simchi-Levi, "The complexity of finding two disjoint paths with min-max objective function," Discrete Applied Math. 26, 1990, pp. 105-115.

[16] Y. Xiao, K. Thulasiraman and G. Xue, "Disjoint QoS paths selection for protection against failures: a network programming based approach," Allerton Conference on Control, Communication and Computing, University of Illinois, Urbana-Champaign, Oct. 2004.

[17] D. Blokh and G. Gutin, "An approximation algorithm for combinatorial optimization problems with two parameters," *Australasian Journal of Combinatorics*, vol. 14, 1996, pp.157-164.

[18] J. W. Suurballe, "Disjoint paths in a network", *Networks*, 4, 1974, pp. 125-145.

[19] Sartaj Sahni, General Techniques for Combinatorial Approximation, *Oper. Res.* 25, 920-936, 1977

[20] A. Orda and A. Sprintson, "Efficient algorithms for computing disjoint QoS paths, "in *Proc. IEEE INFOCOM*, 2004.

[21] R. Guerin and A. Orda, "QoS routing in networks with inaccurate information: theory and algorithms," IEEE/ACM Trans. on Networking, vol. 7, pp. 350-364, June 1999.

[22] T. Korkmaz and M. Krunz, "Bandwidth-delay constrained path selection under inaccurate state information," IEEE/ACM Trans. on Networking, June 2003

[23] D. H. Lorenz and A. Orda, "QoS routing in networks with uncertain parameters," IEEE/ACM Trans. Networking, vol. 6, pp. 768-778, Dec. 1998

[24] Y. Xiao, K. Thulasiraman and G. Xue, "Approximation and Heuristic Algorithms for Delay Constrained Path Selection under Inaccurate State Information," in *proc. First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks*, 2004.