ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization *Vol.3, Special Issue 4, April 2015*

National Conference On Emerging Trends in Information, Digital & Embedded Systems (NC'e-TIDES -15)

Organized by

Dept. of ECE, Annamacharya Institute Of Technology & Sciences, Rajampet, Andhra Pradesh-516126, India held on 28th February 2015

Optimization Process for Cost Reduction In the Overlay Network

Thammisetti Pushpalatha

Student, Department of CSE, New Boyanapalli, Rajampet, Kadapa(dist), Andhra Pradesh, India

ABSTRACT: Overlay routing is a very attractive scheme that allows improving certain properties of the routing without the need to change the standards of the current underlying routing. However, deploying overlay routing requires the placement and maintenance of overlay infrastructure. This gives rise to the following optimization problem: Find a minimal set of overlay nodes such that the required routing properties are satisfied. It rigorously study this optimization problem. It shows that it is NP-hard and derive a nontrivial approximation algorithm for it, where the approximation ratio depends on specific properties of the problem at hand. It examine the practical aspects of the scheme by evaluating the gain one can get over several real scenarios.

KEYWORDS: Overlay network, resource allocation,TCP throughput.

I. INTRODUCTION

The Internet is organized as independently operating autonomous systems (AS's) that peer together. In this architecture, detailed routing information is maintained only within a single AS and its constituent networks, usually operated by some network service provider. The information shared with other providers and AS's is heavily filtered and summarized using the Border Gateway Protocol (BGP-4) running at the border routers between AS's,which allows the Internet to scale to millions of networks.

Overlay routing has been proposed in recent years as an effective way to achieve certain routing properties, without going into the long and tedious process of standardization and global deployment of a new routing protocol. overlay routing was used to improve TCP performance over the Internet, where the main idea is to break the end-to-end feedback loop into smaller loops. This requires that nodes capable of performing TCP Piping would be present along the route at relatively small distances. Other examples for the use of overlay routing are projects like RON and Detour, where overlay routing is used to improve reliability.Yet another example is the concept of the "Global-ISP" paradigm , where an overlay node is used to reduce latency in BGP routing.

In order to deploy overlay routing over the actual physical infrastructure, one needs to deploy and manage overlay nodes that will have the new extra functionality. This comes with a non negligible cost both in terms of capital and operating costs. Thus, it is important to study the benefit one gets from improving the routing metric against this cost.

It concentrate on this point and study the minimum number of infrastructure nodes that need to be added in order to maintain a specific property in the overlay routing. In the shortest-path routing over the Internet BGP-based routing example, the question is mapped to: What is the minimum number of relay nodes that are needed in order to make the routing between a group of autonomous systems (ASs) use the underlying shortest path between them? In the TCP performance example, this may translate to: What is the minimal number of relay nodes needed in order to make sure that for each TCP connection, there is a path between the connection endpoints for which every predefined round-trip time (RTT), there is an overlay node capable of TCP Piping.

Regardless of the specific implication in mind, It define a general optimization problem called the Overlay Routing Resource Allocation (ORRA) problem and study its complexity .It turns out that the problem is NP-hard, and It present a nontrivial approximation algorithm for it.

Note that they are only interested in improving routing properties between a single source node and a single destination ,then the problem is not complicated, and finding the optimal number of nodes becomes trivial since the potential candidate for overlay placement is small, and in general any assignment would be good. However, when it consider one-to-many or many-to-many scenarios, then a single overlay node may affect the path property of many paths, and thus choosing the best locations becomes much less trivial.

By testing this algorithm in three specific such cases, where It have a large set of source–destination pairs and the goal is to find a minimal set of locations, such that using overlay nodes in these locations allows to create routes (routes are either underlay routes or routes that use these new relay nodes) such that a certain routing property is satisfied.

II. RELATED WORK

Using overlay routing to improve network performance is motivated by many works that studied the inefficiency of varieties of networking architectures and applications. Analyzing a large set of data, explore the question: How "good" is Internet routing from a user's perspective considering round-trip time, packet loss rate, and bandwidth? They showed that in 30%–80% of the cases, there is an alternate routing path with better quality compared to the default routing path.The authors show that TCP performance is strictly affected by the RTT. Thus, breaking a TCP connection into low-latency sub connections improves the overall connection performance. The authors show that in many cases, routing paths in the Internet are inflated and the actual length (in hops) of routing paths between clients is longer than the minimum hop distance between them.

Using overlay routing to improve routing and network performance has been studied before in several works. In the authors studied the routing inefficiency in the Internet and used an overlay routing in order to evaluate and study experimental techniques improving the network over the real environment. While the concept of using overlay routing to improve routing scheme was presented in this work, it did not deal with the deployment aspects and the optimization aspect of such infrastructure .A resilient overlay network (RON), which is an architecture for application-layer overlay routing to be used on top of the existing Internet routing infrastructure, has been presented. Similar to our work, the main goal of this architecture is to replace the existing routing scheme, if necessary, using the overlay infrastructure. This work mainly focuses on the overlay infrastructure (monitoring and detecting routing problems, and maintaining the overlay system), and it does not consider the cost associated with the deployment of such system.

Here mainly focuses on relay placement problem, in which relay nodes should be placed in an intra domain network. An overlay path, in this case, is a path that consists of two shortest paths, one from the source to a relay node and the other from the relay node to the destination. The objective function in this work is to find, for each source–destination pair, an overlay path that is maximally disjoint from the default shortest path.This problem is motivated by the request to increase the robustness of the network in case of router failures.They introduce a routing strategy, which replaces the shortest-path routing, that routes traffic to a destination via predetermined intermediate nodes in order to avoid network congestion under high traffic variability.

The first to actually study the cost associated with the deployment of overlay routing infrastructure. Considering two main cases, resilient routing, and TCP performance,they formulate the intermediate node placement as an optimization problem, where the objective is to place a given number intermediate nodes in order to optimize the overlay routing and suggested several heuristic algorithms for each application. Following this line of work,the resource allocation problem in this paper as a general framework that is not tied to a specific application, but can be used by any overlay scheme. Moreover, unlike heuristic algorithms, the approximation placement algorithm presented in our work, capturing any overlay scheme, ensures that the deployment cost is bounded within the algorithm approximation ratio.

Node placement problems have been studied before in different contexts in many works, considering web cache and web server placement .overlay node placement is fundamentally different from these placement problems where the objective is to improve the routing using a different routing scheme rather than pushing the content close to the clients.

III. PROPOSED ALGORITHM

Given a graph G=(V,E) describing a network, let P_U be the set of routing paths that is derived from the underlying routing scheme and let P_0 be the set of routing paths that is derived from the overlaying routing scheme.Note that both P_U and P_0 can be defined explicitly as a set of paths or implicitly. Given a pair of vertices s, t €V denoted by $P_0^{s,t}$ the set of overlay paths between s and t, namely $P_0^{s,t} \leq P_0$, and the endpoints of p are s and t.

Definition 3.1:Given a graph G=(V,E),a pair of vertices(s,t),a set of underlay paths $P_{U,}$ a set of overlay paths $P_0^{s,t}$,and a set of vertices U≤V. We say that U covers(s,t) if there exists p belongs to $P_0^{s,t}$ such that p is a concatenation of one or more underlying paths ,and the endpoints of each one of these underlay paths are in U.

Definition 3.2: Given a graph G=(V,E), a set of source-destination pairs Q= $\{(s_1,t_1),(s_2,t_2),\ldots.(s_n,t_n)\}$, a set of underlay paths P_U and a set of overlay paths $P_{0,}$find a subset of vertices U≤V such that,U covers(s_i,t_i).

Definition 3.3:Given an instance of the ORRA problem and a nonnegative weight function W: V ->R over the vertices,one needs to find a set U_{opt}<V such that:1) U_{opt} is feasible; and 2) the cost of U_{opt} is minimal among all feasible sets.

IV. PSEUDO CODE

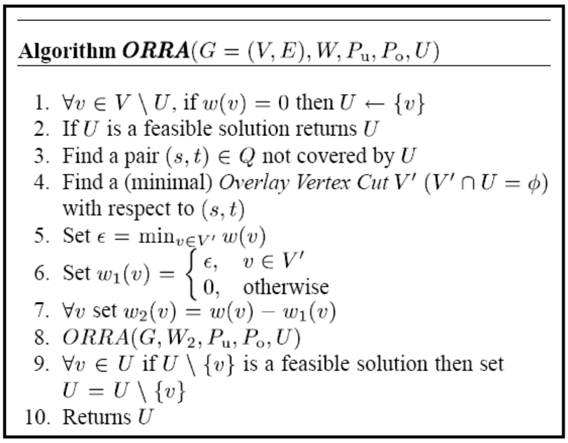
In this section, mainly focuses on the complexity of the ORRA problem. In particular, It show that the -ORRA problem is NP-hard and it cannot be approximated within a factor of O(logn).

While the reduction and the hardness result hold even for the simple case where all nodes have an equal cost (i.e., the cost associated with a relay node deployment on each node is equal), the approximation algorithm can be applied for an arbitrary weight function, capturing the fact that the cost of deploying a relay node may be different from one node to another.

Theorem 1:
1) The -ORRA problem is NP-hard.
2) The MIN-ORRA problem cannot be approximated within a factor of (1-€).In(n) for any €>0 unless NP≤DTIME($n^{\log\log n}$).

Proof: It prove the theorem using a reduction from the SC problem. First, It show that the reduction holds and then It show that it is an approximation preserving reduction. Thus, the ORRA problem is hard as the minimization version of Set Cover and the latter cannot be approximated within a factor of(1-€).In(n) for any €>0. In the SC problem, given a finite set,a positive number k, and a set of subsets C of S, one should test if there is a set of subsets such that every element in S appears at least in one set and(without loss of generality, It consider an instance of the Set Cover problem, such that for each element in S, there is at least one subset in C that contains this element).

Given an instance of the SC problem, i.e.,a finite set S, a positive number k and a set of subsets C of S, It construct an instance of the k-*ORRA* problem as follows. For each element si in the set S , It match a vertex V_{si} and for each subset Cj €C, It match a vertex V_{cj}.



```
Algorithm ORRA($G = (V, E), W, P_u, P_o, U$)

$\forall v \in V \setminus U$, if $w(v) = 0$ then $U \leftarrow \{v\}$
If $U$ is a feasible solution returns $U$
Find a pair $(s, t) \in Q$ not covered by $U$
Find a (minimal) Overlay Vertex Cut $V'$ $(V' \cap U = \phi)$
    with respect to $(s, t)$
Set $\epsilon = \min_{v \in V'} w(v)$
Set $w_1(v) = \begin{cases} \epsilon, & v \in V' \\ 0, & \text{otherwise} \end{cases}$
$\forall v$ set $w_2(v) = w(v) - w_1(v)$
$ORRA(G, W_2, P_u, P_o, U)$
$\forall v \in U$ if $U \setminus \{v\}$ is a feasible solution then set
    $U = U \setminus \{v\}$
Returns $U$
```

The recursive algorithm ORRA(G=(V,E),W,P_U,P_0,U) receives an instance of the *ORRA* problem (a graph , a nonnegative weight function W over the vertices, a set of underlay and overlay paths $P_{U,}P_0$, respectively) and a set of relay nodes U and returns a feasible solution to the problem. The set of relay nodes in the first call is empty (i.e.,U=ø).

At each iteration, the algorithm picks vertices with weight that is equal to zero until a feasible set is obtained (steps 1 and 2 of the algorithm). Thus, since at each iteration at least one vertex gets a weight that is equal to zero with respect to W_2 (steps 5–7), then in the worst case the algorithm stops after V iterations and returns a feasible set. In Step 9, unnecessary vertices are removed from the solution, in order to reduce its cost. While this step may improve the actual performance of the algorithm, it is not required in the approximation analysis below and may be omitted in the implementation.

V. SIMULATION RESULTS

A. BGP Routing Scheme

BGP is a policy-based inter domain routing protocol that is used to determine the routing paths between autonomous systems in the Internet. In practice, each AS is an independent business entity, and the BGP routing policy reflects the commercial relationships between connected ASs. A *customer–provider* relationship between ASs means that one AS (the customer) pays another AS (the provider) for Internet connectivity, a *peer–peer* relationship between ASs means that they have mutual agreement to serve their customers,while a *sibling–sibling* relationship means that they have mutual-transit agreement (i.e., serving both their customers and providers). These business relationships between ASs induce a BGP export policy in which an AS usually does not export its providers and peers routes to other providers and peers. The authors showed that this route export policy indicates that routing paths do not contain so-called *valleys* nor *steps*. In other words, after traversing a *provider–customer* or a *peer–peer* link, a path cannot traverse a *customer–provider* or a *peer-peer* link. This routing policy may cause, among other things, that data packets will not be routed along the shortest path.While routing policy is a fundamental and important feature of BGP, some application may require to route data using the shortest physical paths.3 In this case, using overlay routing, one can perform routing via shortest paths despite the policy. In this case, relay nodes should be deployed on servers located in certain carefully chosen ASs.Considering such a scenario, the corresponding *ORRA* instance consists of the AS topology graph , the set of valid routing paths.

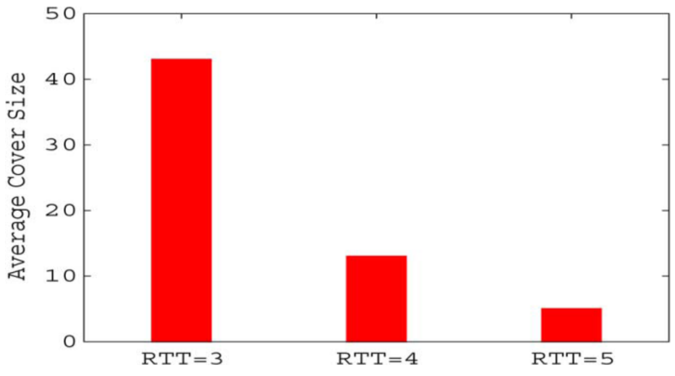


Fig:Algorithm coverage for different RTT values.

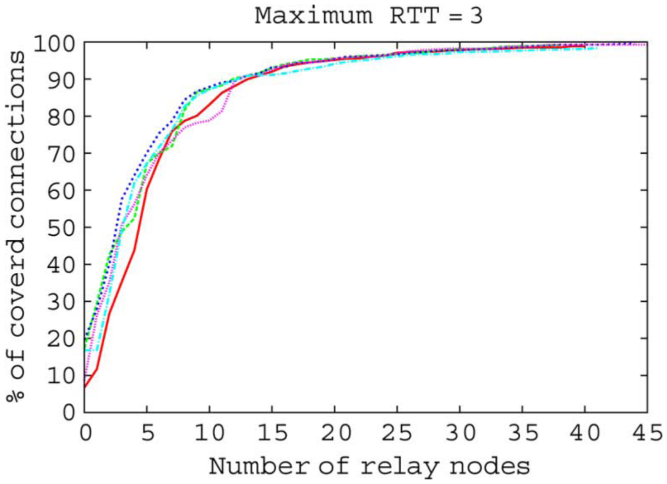


Fig. . Covered connections versus number of relay nodes.

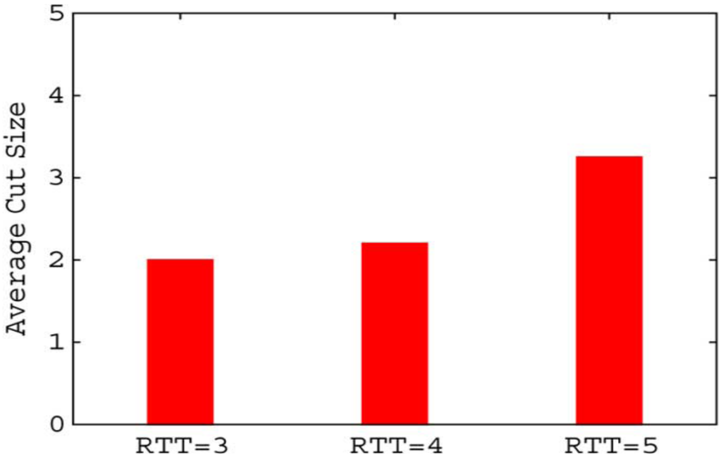


Fig. Algorithm average cut size for different RTT values.

VI. CONCLUSION AND FUTURE WORK

While using overlay routing to improve network performance was studied in the past by many works both practical and theoretical, very few of them consider the cost associated with the deployment of overlay infrastructure. In this paper, It addressed fundamental problem developing an approximation algorithm to the problem. Rather than considering a customized algorithm for a specific application or scenario, it suggestes a general framework that fits a large set of overlay applications. Considering three different practical scenarios, It evaluates the performance of the algorithm, showing that in practice the algorithm provides close-to-optimal results. Many issues are left for further research. One interesting direction is an analytical study of the vertex cut used in the algorithm. It would be interesting to find properties of the underlay and overlay routing that assure a bound on the size of the cut. It would be also interesting to study the performance of our framework for other routing scenarios and to study issues related to actual implementation of the scheme .In particular, the connection between the cost in terms of establishing overlay nodes and the benefit in terms of performance gain achieved due to the improved routing is not trivial and it is interesting to investigate it. The business relationship between the different players in the various use cases is complex and thus it is important to study the economical aspects of the scheme .For example, the one-to-many BGP routing scheme can be used by a large content provider in order to improve the user experience of its customers. The VoIP scheme can be used by VoIP services (such as Skype) to improve call quality of their customers. In both these cases, the exact translation of the service performance gain into actual revenue is not clear and can benefit from further research.

REFERENCES

[1] Rami Cohen and Danny Raz :"cost-effective resource allocation of overlay routing relay nodes".
[2] h. Pucha and y. C. Hu, "overlay tcp: multi-hop overlay transport for high throughput transfers in the internet," purdue university, itst lafayette, in, usa, tech. Rep., 2005.
[3] d. Andersen, h. Balakrishnan, f. Kaashoek, and r. Morris, "resilient overlay networks," in *proc. 18th acm sosp*, 2001, pp. 131–145.
[4] s. Savage, t. A. A. Aggarawl, t. Anderson, a. Aggarwal, d. Becker, n. Carditll, a. Collins, e. Hoffman, j. Snell, a. Vahdat, g. Voelker, and j. Zahorjan, "detour: a case for informed internet routing and transport," *ieee micro*, vol. 19, no. 1, pp. 50–59, jan.–feb. 1999.
[5] r. Cohen and a. Shochot, "the "global-isp" paradigm," *comput.netw.*, vol. 51, no. 8, pp. 1908–1921, 2007.
[6] l. Gao and f. Wang, "the extent of as path inflation by routing policies,"in *proc. Ieee globecom*, 2002, vol. 3, pp. 2180–2184.
[7] s. Savage, a. Collins, e. Hoffman, j. Snell, and t. Anderson, "the end-to-end effects of internet path selection," in *proc. Acm sigcomm*,1999, pp. 289–299.
[8] r. Cohen and s. Ramanathan, "using proxies to enhance tcp performance over hybrid fiber coaxial networks," *comput. Commun.*, vol. 20,no. 16, pp. 1502–1518, jan. 1998.
[9] n. Spring, r. Mahajan, and t. Anderson, "the causes of path inflation,"in *proc. Acm sigcomm*, 2003, pp. 113–124.646 ieee/acm transactions on networking, vol. 22, no. 2, april 2014.
[10] h. Tangmunarunkit, r. Govindan, s. Shenker, and d. Estrin, "the impact of routing policy on internet paths," in *proc. Ieee infocom*, 2001, pp. 736–742.
[11] m. Cha, s. Moon, c.-d. Park, and a. Shaikh, "placing relay nodes for intra-domain path diversity," in *proc. Ieee infocom*, apr. 2006, pp.1–12.

BIOGRAPHY

Thammisetti Pushpalatha is a post graduate in the Computer Science Department, College of Annamacharya Institute of Engineering & Technology,Rajampet,Kadapa district, Andhra pradesh. She received bachelor degree in 2013 from Rishi M.S.institute of Engineering &Technology for Women,Hyderabad, India. Her research interests are Computer Networks (wireless Networks) and big data.