



## King's Research Portal

### *Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

### *Citation for published version (APA):*

Abrishamchi, B., Farhoudi, M. H., Mihailovic, A., & Aghvami, A-H. (2018). A Novel Minimum Set Cover Routing Plane Construction Approach for Random Wireless IP Access Networks. *Wireless Communications and Networking Conference (WCNC) 2018*.

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# A Novel Minimum Set Cover Routing Plane Construction Approach for Random Wireless IP Access Networks

Benyamin Abrishamchi, Mohammad Farhoudi, Andrej Mihailovic, and A.H. Aghvami

Centre for Telecommunications Research, Faculty of Natural and Mathematical Sciences, King's College London, UK

Email: {benyamin.b.abrishamchi, mohammad.farhoudi, andrej.mihailovic, hamid.aghvami}@kcl.ac.uk

**Abstract**—The expected variable nature of future access network structures calls for network providers to adapt accordingly in order to facilitate loss free networks supported with the developing IP infrastructure. To this end, a routing optimization mechanism applicable to randomly shaped all-IP access networks must be developed. In this paper, Multi-Plane Routing (MPR) that consolidates various aspects in all-IP infrastructure as a whole is redesigned and reformulated to provide a comprehensive solution in consideration of the randomness of future access networks. We prove the offline Routing Plane (RP) construction problem as being the generalization of the Minimum Set Cover (MSC) problem which is NP and also NP-complete. To this end, we propose a novel MSC-based paths-diverse offline TE algorithm which is comprehensively applicable to random wireless access network structures. Our simulation results demonstrate the constructed RPs for complex random networks of different sizes and sparseness (i.e. meshing). Hence, the comprehensive applicability of our novel approach is verified. Moreover, we propose a new optimization framework along with a dynamic cost function (considering capacity and correlation of paths) that formally describe our novel offline TE mechanism for future randomly shaped wireless IP access networks.

## I. INTRODUCTION

Internet was invented as an open global platform that would enable everyone, everywhere to share information, access wide ranging opportunities, and cooperate across geographic and cultural boundaries. The exponential growth of Internet has turned it into a multi-faceted collaborative environment connecting a wide range of users. The impact of this growth is shaping the way users on different scales interact with each other as reflected in our daily lives. Accordingly, the emergence of exciting new devices along with new highly demanding applications have put even more burden on the Internet in the recent years. Hence, it is becoming essential for cellular network operators and IP Network Providers (INPs) to adopt Traffic Engineering (TE) [1]. Moreover, the surge of highly demanding devices and applications at the edges of the network has extended the focus from core to access networks. Hereafter, we refer to this design space as all-IP wireless access networks. Considering that such networks are positioned at the edges of the Internet routing fabric, we focus on intra-domain TE accordingly.

Intra-domain TE is classified into the Multi-Protocol Label Switching (MPLS) - and IP-based TE. MPLS is extensively deployed in access networks where encapsulated IP packets are delivered over Labelled Switched Paths (LSPs). MPLS enables

explicit routing and arbitrary splitting of traffic with the scalability and robustness being issues due to the complexity and overhead associated with building and maintaining LSPs. IP-based TE is achieved through the manipulation of link weights in case of Interior Gateway Protocols (IGPs) such as OSPF (an intra-domain dynamic link-state IP IGP). In contrast with MPLS TE, IP TE does not intrinsically enable arbitrary splitting of traffic and explicit routing. Equal-Cost Multi-Path (ECMP) is an extension of OSPF based on which traffic is split equally between several paths of equivalent cost through hop-by-hop forwarding. ECMP cannot be configured in complex large-scale topologies as the quality of OSPF TE can become arbitrarily poor compared to optimal TE due to the computational intractability to derive optimal link weights for large-scale networks [2]. There have been notable studies on improving the optimality of IP-based intra-domain TE mechanisms based on ECMP. These studies aim for immediate load balancing using path diversity by capitalising on the flexible use of network topology [3], [4]. Nonetheless, the dependence of the aforementioned ECMP-based schemes on link weight setting, make them subject to performance degradation, slow convergence times and deviation from optimal TE.

Multi-Plane Routing (MPR) which was proposed in [5] and investigated in more detail in [6] for IP access networks aims to address the deficiencies associated with MPLS and ECMP. MPR is based on Multi-Topology OSPF (MT-OSPF) principle and allows for path diversity with minimal protocol overhead. MT-OSPF was originally proposed for fast re-route in case of node/link failure whereas MPR employs MT-OSPF for comprehensive network-wide load balancing. MPR applies an offline TE method in order to build logical Routing Planes (RPs) (i.e. instances of OSPF) that render a set of shortest paths ahead of the traffic flow in the network that is controlled by an online TE approach. The path diversity potential benefits in access networks was investigated in [7]. This study undermined the necessity for next-generation access networks' evolution to more meshed topologies to exploit path diversity (as facilitated by multi-path routing). MPR achieves explicit routing and arbitrary splitting of traffic (as achieved under MPLS TE approach) by applying a purely IP-based TE approach. Therefore, the complexity and overhead associated with the MPLS TE approach can be averted. MPR is already proven to achieve performance gains against MPLS and OSPF considering the whole network topology and a significant

number of key performance criteria in extensive simulations for limited pre-defined tree-like topologies [6].

In consideration of the uncertainties regarding the future wireless all-IP access networks' structure and given the disparately shaped existing access structures, a routing optimization method applicable to randomly shaped wireless IP access networks must be developed. Such an approach will represent a comprehensive solution for future wireless access networks. MPR's legacy offline TE approach [8] that aims to achieve full path diversity suffers from deficiencies and limitations in terms of applicability to randomly shaped access networks and it is limited to meshed tree topologies that follow essential configuration of networks. This is because of the algorithm being an arbitrary RP construction method with a strict constraint of every link being used at most  $N - 1$  times which can prove inapplicable in case of random graphs that are likely to occur in practical scenarios of network installations and configurations (e.g. HetNets [9]). In this work, we focus on the offline TE aspect (network planning phase) of MPR which has the physical topology with associated randomly allocated link capacities as the input (depending on link's level). Under the new scenario, the topology independent RP construction is redesigned and reformulated to suit the randomly shaped access networks with the aim of achieving maximum path diversity and use of entire routing resources.

#### A. Contributions

In this paper, a novel paths-diverse routing optimization approach for future all-IP wireless access networks with random topological structure is proposed. This protocol expands upon MPR which is not suited to random graphs. This study considers a comprehensive routing scheme applicable to randomly shaped access networks with different levels of meshing. To the best of our knowledge, the study of such a comprehensive TE mechanism which is applicable to randomly shaped access networks is absent in literature. We propose a new optimization framework (considering capacity and correlation of paths) that formally describes our novel offline TE mechanism. Furthermore, we revisit the multi-topology (i.e. RPs) construction problem in intra-domain networks. To this end, we prove the offline RP construction problem as being the generalization of the Minimum Set Cover (MSC) problem which is NP-complete. Hence, the adoption of heuristics is justified. Correspondingly, a dynamic MSC-based cost function is formulated to construct a diverse set of routing planes in a random access network. Consequently, a novel comprehensive dynamic MSC-based RP construction algorithm is proposed. We also introduce the possibility for existence of more than one gateway in a given access network.

## II. ANALYTICAL MODEL

### A. Problem Formulation

The main objective of MPR's offline algorithm is as follows: Given an underlying network topology, we aim to to extract a set of *disjoint* routing planes (RPs), such that each routing plane (RP) has a path between every source-destination pair in addition to every link in the network *appearing once* in a routing plane (ensuring every link is used). These constraints are aimed at achieving maximum path diversity across the

RP-set over the state-of-the art approaches. In addition to these constraints, it may also be desirable to impose the usual constraints commonly imposed in TE, such as least cost paths, maximum capacity constraints and so on. We can therefore precisely summarize our prime objectives as follows:

**Problem 1.** *Given a network represented by an underlying arbitrary graph  $G = (\mathcal{V}, \mathcal{E})$ , with the vertex and edge sets  $\mathcal{V}$  and  $\mathcal{E}$  respectively, retrieve a set  $\mathcal{R} = \{R_p\}_{p=1}^P$  of  $|\mathcal{R}| = P$  RPs, such that the following properties hold:*

- Each RP constructed contains a valid path for every source-destination pair in the underlying network  $G$ ; in other words for every  $p = 1, 2, \dots, P$ ,  $R_p \in \mathcal{R}$  should be a connected subgraph of  $G$  (in practical terms, an instance of OSPF routing installation).*
- Ideally every link in the underlying network topology should appear in one of the constructed routing planes. Denote the  $p$ -th RP as the subgraph  $R_p = (\mathcal{V}_p, \mathcal{E}_p)$  with vertex and edge sets  $\mathcal{V}_p$  and  $\mathcal{E}_p$  respectively, then if  $e \in \mathcal{E}$  then there exists a  $p$  such that  $e \in \mathcal{E}_p$ . This also implies that the union  $\bigcup_{p=1}^P R_p = G$ .*
- All RPs chosen are completely different from each other. This means that for each  $p, q = 1, 2, \dots, P$  then if  $e_p \in \mathcal{E}_p$  and  $e_q \in \mathcal{E}_q$ , it follows that  $\mathcal{E}_p \cap \mathcal{E}_q = \{\emptyset\}$  if  $p \neq q$  whilst  $\mathcal{E}_p \cap \mathcal{E}_q = \mathcal{E}_p \cup \mathcal{E}_q = \mathcal{E}_p$  for  $p = q$ .*
- Finally, the cost, as prescribed by some function  $f$  (of the capacity for example) is optimal between every source-destination pair for each plane.*

It is not difficult to see that the constraints, specifically (b) and (c) in Problem 1, are very restrictive and may even mean that no feasible solution would exist to the problem in some proverbial cases. In Figure 1 for example, a connected subgraph of  $G$  is  $R_1 = (\mathcal{V}, \{e_1, e_2, e_3\})$ . Meanwhile, it is no longer possible to construct another subgraph of  $G$  that simultaneously satisfies (b) and (c), in particular, the only unused links are  $e_4$  and  $e_5$ , but the subgraph  $R_2 = (\mathcal{V}, \{e_4, e_5\})$  is not connected and hence not a valid routing plane. To avoid such situations it makes sense to relax restriction (c) to allow for links to be reused. Despite this relaxation, we still aim for the chosen subgraphs to be as different as possible. Sticking with the same example, this means we may prefer to choose  $R_2 = (\mathcal{V}, \{e_1, e_4, e_5\})$  instead of  $R_2 = (\mathcal{V}, \{e_1, e_2, e_5\})$ , given the choice  $R_1$ , as the latter reuses two links instead of one. In light of this, we would need to define a suitable graph similarity metric in order to compare any set of graphs and determine precisely how similar they are. With this similarity metric (as defined following section), a penalty can be imposed for reusing links in our optimization problem.

### B. Measuring the similarity of graphs

**Definition 1** (Graph Correlation metric). *The graph correlation metric is a measure of similarity between a pair of graphs which we define herein to be the total number of different edges present between the graphs normalized by the number of edges in the graph containing the highest number of edges. Let  $R_p = (\mathcal{V}_p, \mathcal{E}_p)$  and  $R_q = (\mathcal{V}_q, \mathcal{E}_q)$  then it follows that:*

$$\rho(G_p, G_q) = \frac{|\mathcal{E}_p \cap \mathcal{E}_q|}{\max(|\mathcal{E}_p|, |\mathcal{E}_q|)} \quad (1)$$

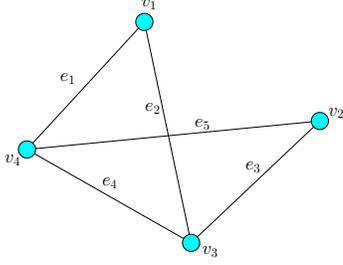


Fig. 1: An underlying network topology  $G$  with four vertices and five edges.

It is easy to see that  $\rho(G_p, G_q) \geq 0$ , in fact  $\rho(G_p, G_q) \in [0, 1]$  with  $\rho(G_p, G_q) = 1$  if and only if  $G_p = G_q$ ; whilst  $\rho(G_p, G_q) = 0$  if  $\mathcal{E}_p \cap \mathcal{E}_q = \{\emptyset\}$  (as a technicality, to avoid dividing by zero, we may assume that at least either  $\mathcal{E}_p$  or  $\mathcal{E}_q$  is nonempty). In addition we can also deduce that if  $\rho(G_p, G_q) = 0$ , then for any two nodes in  $G_p$  joined by a single edge, the same pair of nodes (*Ingress - Egress*) in  $G_q$  will be joined by a path with at least two different edges. Furthermore, from Definition 1, we can also deduce an expression for  $\rho(G_p, G_q)$  in terms of their corresponding adjacency matrices  $\mathbf{A}_p, \mathbf{A}_q \in \mathbb{R}^{N \times N}$ , where  $N = |\mathcal{V}|$  is the number of nodes in the network as:

$$\rho(G_p, G_q) = \frac{\mathbf{1}^\top (\mathbf{A}_p \odot \mathbf{A}_q) \mathbf{1}}{\max(\mathbf{1}^\top \mathbf{A}_p \mathbf{1}, \mathbf{1}^\top \mathbf{A}_q \mathbf{1})}, \quad (2)$$

where  $\odot$  represents the element-wise (Hadamard) product between two matrices and  $\mathbf{1} \in \mathbb{R}^N$  is a vector with all its entries equal to unity.

### C. Enforcing diversity by minimizing link presence

As it has been allowed for links to be reused across multiple RPs as a result of relaxation, we may want to further penalize for this re-usage to minimize the chances of a link being used in all routing planes. Hence, the following definitions:

**Definition 2** (Link Presence). *The link presence is a function with a binary output indicating whether or not a link  $e \in \mathcal{E}$  is used in the RP  $R_p$ . Denoting it as  $LP : \mathcal{E} \times \mathcal{R} \rightarrow \{0, 1\}$ :*

$$LP(e, R_p) = \begin{cases} 1, & \text{if } e \in R_p \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

**Definition 3** (Full Link Presence). *The full link presence indicates the appearance of an edge in all RPs:*

$$FLP(e) = \begin{cases} 1, & \text{if } \sum_{p=1}^P LP(e, R_p) = P \\ 0, & \text{otherwise} \end{cases}. \quad (4)$$

### D. The optimization problem

The objective in our problem is to maximize the projected capacity for each routing plane that is to be constructed. Let  $\mathbf{C}$  be the capacity matrix, which we assume to be normalized so that its entries only take on values between 0 and 1, also denote by  $C_e \in [0, 1]$  the capacity of the edge  $e$  and by  $r_m^{i,j}$  the  $m$ -th path (which is simply an ordered sequence of edges) between the source-destination pair  $i$  and  $j$ , so that the collection of all possible routes between  $i$  and  $j$  is  $\{r_m^{i,j}\}_{m=1}^{M_{i,j}}$ . With these definitions, we can state our main objectives:

- 1) In the path linking the nodes  $(v_i, v_j)$ , the link with the smallest capacity is the source of the main bottleneck, hence maximizing the minimum capacity in this path is the objective.
- 2) We aim to maximize the minimum projected capacity for all pairs of nodes, i.e. for all  $i, j = 1, 2, \dots, N$  and  $i \neq j$ ; across all the  $P$  RPs.

Therefore to the MPR as described in Problem 1 can be stated as the following optimization problem below:

$$\begin{aligned} & \underset{\mathcal{R}=\{R_p\}_{p=1}^P}{\text{maximize}} && \sum_{p=1}^P \sum_{k=1}^N \sum_{\substack{l=1 \\ l \neq k}}^N \min_{e \in r^{l,k}(R_p)} F(C_e) \\ & \text{subject to} && r^{l,k}(R_p) \neq \{\emptyset\}, \forall k, l = 1, \dots, N, l \neq k \\ & && \text{and } p = 1, \dots, P. \\ & && \rho(R_p, R_q) = 0, \forall p, q = 1, \dots, P \text{ and } p \neq q. \\ & && FLP(e) = 0, \forall e \in \mathcal{R}. \\ & && P = \bar{P}. \end{aligned} \quad (5)$$

where  $r^{l,k}(R_p)$  is used to denote the path between the pair of nodes  $(l, k)$  in the subgraph  $R_p$ ,  $F(x)$  is some well defined monotonically increasing function of  $x$  and  $\bar{P}$  is the desired number of routing planes. The objective is to utilize the path with the highest available capacity between every *source-destination* pair when formulating each plane. Furthermore, the constraints imposed ensures the following: first, each  $R_p$  is a connected subgraph of  $G$ ; second, It is guaranteed that any pair of planes are disjoint with respect to the edges; third, no link is used in all planes and finally that we have more than  $\bar{P}$  RPs. We have seen that there can be situations where this problem is infeasible, and so we would need to relax the third constraint, as described in Problem 1. This relaxation is such that we are instead interested in minimizing the overlap between the chosen planes and a link can appear across multiple planes; in so doing, we can move the constraints into the objective to obtain the new problem below: (maximize)

$$\begin{aligned} & \alpha \sum_{p=1}^P \sum_{k=1}^N \sum_{\substack{l=1 \\ l \neq k}}^N \min_{e \in r^{l,k}(R_p)} F(C_e) + \beta \sum_{p=1}^P \sum_{\substack{q=1 \\ q \neq p}}^P \rho(R_p, R_q) \\ & \text{subject to: } && r^{l,k}(R_p) \neq \{\emptyset\}, \forall k, l = 1, \dots, N, l \neq k \\ & && \text{and } p = 1, \dots, P. \\ & && P = \bar{P}. \end{aligned} \quad (6)$$

where  $\alpha \in (0, 1)$  and  $\beta = 1 - \alpha$  are arbitrary tuning parameters, chosen *a priori*. They can be interpreted as a way

of assigning more importance to the terms in the objectives; specifically  $\alpha \gg \beta$  means more importance is assigned to finding planes with high capacity but they will not necessarily be diverse. This new relaxed problem as we will show in the following section is NP-complete, thus we will have to resort to finding heuristic algorithms for computing useful solutions in Section III. In particular, for some network  $G$ , the approach we will adopt is to: firstly, construct a highly redundant collection of subgraphs of the network and from these subgraphs, find a set of connected subgraphs within this new collection of connected subgraphs. To this end, we present an approach in Section III-A; followed by Section IV constructing the RPs.

### E. A Proof of NP-Completeness

In this section we demonstrate that the problem of interest, given the requirements stated in Section II-A, is NP-complete. This is achieved herein by demonstrating that a well-known NP-complete problem is reducible, in polynomial time, to our problem. We commence this proof by summarizing the minimum set cover problem [10]:

**Definition 4 (Minimum Set Cover Problem).** *Given a finite collection  $\mathcal{S} = \{S_i\}_{i=1}^I$  of subsets of a universe  $U$ , a set cover  $\mathcal{C} \subseteq \mathcal{S}$  is a subcollection of the sets whose union is  $U$ , i.e.  $\bigcup_{S \in \mathcal{C}} S = U$ . Moreover, each  $S \in \mathcal{C}$  has an associated non-negative cost  $c_S$ . The MSC problem is to compute such a subcollection  $\mathcal{N} \subseteq \mathcal{C}$  such that it is a set cover for  $U$  and its associated cost  $\sum_{S \in \mathcal{N}} c_S$  is simultaneously minimized.*

Note that the unweighted set cover and unweighted  $k$ -set cover problems are special cases of the weighted set cover and weighted  $k$ -set cover problems, respectively. Furthermore, it is known that the MSC problem is NP-complete. To show that our problem is NP-complete it suffices to show that it is in NP and that the weighted MSC problem is reducible to our problem in polynomial time. Let the network topology being considered be defined by its connectivity graph  $G = (\mathcal{V}, \mathcal{E})$  and its associated link weight function  $w : \mathcal{E} \mapsto \mathbb{R}_+$ , that assigns a non-negative weight  $w(e)$  to each edge  $e \in \mathcal{E}$ . It is understood that  $\mathcal{V}$  and  $\mathcal{E}$  are the set of all vertices/nodes and edges/links respectively of the underlying network. Define  $\mathcal{S}$  to be a collection of distinct subsets of  $\mathcal{E}$ , that is to say:  $\mathcal{S} = \{S_i\}_{i=1}^I$  where  $S_i \subset \mathcal{E}$  for each  $i = 1, 2, \dots, I$  so that for any  $i \neq j$ ,  $S_i \neq S_j$ . Such a collection of subsets can be obtained, for example by constructing all trees of the underlying graph  $G$ , e.g. [11]. Each spanning tree  $S_i$  is simply a subset with  $|\mathcal{V}| - 1$  elements (edges ‘ $e$ ’) chosen from  $\mathcal{E}$ ; with an associated cost  $c_{S_i} = \sum_{e \in S_i} w(e)$ . It is easy to see that  $\bigcup_{\forall i} S_i = \mathcal{E}$ . Now given  $\mathcal{S}$ , each element  $S_i$  of it has a path between all possible source-destination pairs and so it is a valid RP. However our task here is to find the subcollection  $\mathcal{N}$  of RPs of minimal cost that utilizes every link in the network. In other words, we desire an  $\mathcal{N}$  such that  $\bigcup_{S_i \in \mathcal{N}} S_i = \mathcal{E}$  and  $\sum_{S_i \in \mathcal{N}} c_{S_i}$  is minimized. This is clearly the MSC problem, by definition. It is now clear that our MPR problem is a generalization of this MSC problem, therefore it is in NP and also NP-complete.

## III. ALGORITHMS FOR CONSTRUCTING ROUTING PLANES

In what follows we develop an algorithm for constructing multiple valid routing planes. To achieve this, we adopt the following approach: 1) Initially, given the underlying network topology  $G$ , we construct a set  $\mathcal{S}$  whose elements  $S_i$  are subsets of the set of all edges  $\mathcal{E}$  of the underlying network. We impose a further constraint that each set  $S_i$  produces a connected network – i.e. we can reach any destination from any source. Clearly each  $S_i$  is a candidate plane. 2) Next with access to these candidate planes  $S_i$ , each candidate plane having an associated cost  $c_{S_i} = \sum_{e \in S_i} w(e)$ , we then aim to find the minimum set cover  $\mathcal{N}$  using any state-of-the-art algorithm such as [10].

### A. Constructing Multiple Candidate Planes II: All Spanning Trees

We now investigate how to construct the candidate planes  $S_i$  and hence a valid collection  $\mathcal{S}$  which includes a cover for the set  $\mathcal{E}$ . Every subset  $S_i$  of edges must contain a link between every vertex in  $G$ , in other words each subgraph  $G_i \stackrel{\text{def}}{=} (\mathcal{V}, S_i)$  of  $G$  must be connected. One possible solution to this is to construct distinct spanning trees of  $G$ , in fact we introduce a result in graph theory that recovers all spanning trees of  $G$ . For some given  $G$ , the problem of enumerating all its spanning trees has a long history (see for example [11], [12] and references therein). We adopt the method proposed in [12] which enumerates spanning trees by swapping edges in a fundamental cycle; in fact, their method can find and list all trees for the unweighted and undirected graph. The spanning trees constructed using the algorithm in [12] form the elements of the collection  $\mathcal{S}$ . Given  $\mathcal{S}$  we propose the use of a greedy algorithm to solve the NP-complete MSC problem.

## IV. FINDING MULTIPLE MINIMUM COST RPs: THE MINIMUM SET COVER

Out of the candidate RPs in  $\mathcal{S}$ —found using the approaches outlined in Section III for example—we wish to select several planes from  $\mathcal{S} = \{S_i\}_i$  so that their union coincides with the underlying network topology, and also produces a minimum cost (wrt certain cost function associated with each  $S_i$ ). As discussed earlier this can be recast as the MSC problem; indeed several approaches have been proposed in recent years on how to solve the MSC problem, a useful survey of current approximate algorithms can be found in [13]. A well known and straightforward greedy algorithm for solving this problem exists [10], [14]. Starting with an empty collection of subsets in the solution and no covered item, the following procedure is iterated. Define  $N_{S_i}$  to be the number of uncovered items in  $S_i$  and define the *current* ratio  $r_{S_i} = \frac{f(S_i)}{N_{S_i}}$ , let  $S^*$  be a set with minimal  $r_{S^*}$ . The algorithm then adds  $S^*$  to the collection of subsets of the solution and defines the items in  $S^*$  as covered and assigns a cost  $r_{S^*}$  to all items that are now covered by were uncovered before this iteration. In fact a modification, that improves on the complexity, of this algorithm called greedy algorithm with withdrawals was proposed by Hassin and Levin in [10]. This modification can be used which we propose to use in order to construct the minimum cover from  $\mathcal{S}$ . To solve our problem, we propose a

modification of the simple greedy algorithm, studied in [14], wherein the cost  $f(S_i)$  associated with the unused subsets  $S_i$  changes with each greedy iteration. This new modification aims to dynamically capture the requirement that the selected topologies are as diverse as possible from each other (with respect to the links used). Therefore we use our similarity metric (1) to adjust the cost associated with selecting a plane at each greedy iteration  $t \in \mathbb{N}_0$ . We denote this new dynamic cost function by  $f_t(S_i)$ , where the index  $t$  has been included to emphasize the fact that the cost changes with  $t$ . In what follows we provide an explicit mathematical expression for our dynamic cost function and consequently outline the steps of the proposed algorithm in Sections IV-A and IV-B respectively.

#### A. The cost function

The associated cost function with which to perform the MSC is explicitly defined as:

$$f_t(S_i) = \begin{cases} \alpha \sum_{k=1}^N \sum_{\substack{l=1 \\ l \neq k}}^N \max_{e \in r^{l,k}(S_i)} F(C_e), & \text{for } t = 0 \\ \alpha \sum_{k=1}^N \sum_{\substack{l=1 \\ l \neq k}}^N \max_{e \in r^{l,k}(S_i)} F(C_e) + \frac{\beta}{|\tilde{\mathcal{R}}_t|} \sum_{S_j \in \tilde{\mathcal{R}}_t} \rho(S_i, S_j), & \\ \text{for } t > 0. \end{cases} \quad (7)$$

where  $F(x) = \frac{1}{x}$  and  $\tilde{\mathcal{R}}_t$  is a set of those planes that have already been selected as RPs. Note that at the start of the MSC algorithm the set  $\tilde{\mathcal{R}}_t$  will be empty, i.e.  $\tilde{\mathcal{R}}_t = \emptyset$  with cardinality zero, hence we choose the best RP in terms of capacity in the first iteration. The first term of the cost function  $f_t$  will be large when  $C_e$  is small and therefore penalizes planes containing paths with very low capacity. Consequently, this sum ensures that planes having the best capacity between all source destination pairs are preferred. In addition, the second term in the cost function introduces a measure of the correlation between the current plane  $S_i$  and every other RP already chosen during the implementation of the MSC algorithm. Therefore this sum will be small when  $S_i$  is most different from the already chosen planes.

#### B. The MSC-based RP construction algorithm

The following iterative procedure finds the near-optimal cover for the edge set of  $G$  given our RP construction constraints (that set our termination point).

- (S.1) Set  $\tilde{\mathcal{R}}_t = \emptyset$  and  $t = 0$ .
- (S.2) If  $G = \bigcup_{R_p \in \tilde{\mathcal{R}}_t} G$  then stop and return  $\mathcal{R} = \tilde{\mathcal{R}}_t$ , since  $\tilde{\mathcal{R}}_t$  is a cover. Otherwise, compute the ratio:

$$r_{S_i} = \frac{|S_i|}{f_t(S_i)}, \quad (8)$$

and select the plane  $S_j$  that maximizes this ratio.

- (S.3) Add  $S_j$  to  $\tilde{\mathcal{R}}_t$  as follows: set  $R_{t+1} = S_j$  and then replace  $\tilde{\mathcal{R}}_{t+1}$  with  $\tilde{\mathcal{R}}_t \cup R_{t+1}$ .
- (S.4) Set  $S_j = \emptyset$  and return to step (S.2).

### V. PERFORMANCE EVALUATION

We evaluate our proposed framework using extensive random simulations. The algorithm used to generate the graph

can best be described as a random walk. Every one of  $V$  nodes is added one by one and is connected to a previous node, within its constraints. This algorithm results in a random connected graph that includes all the nodes. Missing edges are then randomly added until the desired random sparseness is achieved. Randomly generated sets of topologies with different randomly applied meshings/sparseness (i.e. average node degree) levels, varying possible number of gateways and destinations (i.e. ARs) are presented in TABLE I (the remaining nodes are transit (i.e. not source nor sink)). The capacity between any two nodes is randomly generated based on its distance from a gateway, with higher capacity corresponding to lower distance from the gateway (reflecting heterogeneity and practical deployments).

Topology	Nodes	ARs	Avg. Node degree	Gateways	RPs
T1S1	7	3	2.35	1	2
T1S2	7	3	3.71	1	3
T1S3	7	3	4.26	1	3
T1S4	7	3	4.65	1	3
T2S1	18	6	3.23	2	4
T2S2	18	6	3.67	2	5
T2S3	18	6	4.16	2	7
T2S4	18	6	4.91	2	8
T3S1	33	14	3.18	3	8
T3S2	33	14	3.62	3	9
T3S3	33	14	4.45	3	11
T3S4	33	14	4.88	3	15
T4S1	60	25	3.45	5	13
T4S2	60	25	3.87	5	17
T4S3	60	25	4.69	5	22
T4S4	60	25	5.87	5	27

TABLE I: Setup of the Randomly generated topologies

In general, the process of finding all trees as an input set into a function is extremely computationally expensive. For complete graphs, the number of such trees is on the order of  $O(v^{v-2})$ . As such, it is useful to be able to approximate a solution by an iterative approach. To this end, a greedy heuristic is developed for each step of MSC RPs' heuristic approach to find a near optimal tree at each iteration (based on our cost function) that concatenate to represent the near optimal MSC RP set solution (as generally set out in equation (8) with our dynamic cost function). Given an input graph, our heuristic approach applies the hill climbing algorithm in order to obtain a set of near-optimal RPs. Initially, the minimum tree on the input graph is obtained (first RP) followed by changing every edge (i.e. hill climbing) on the concerning tree in the attempt to find the near-optimum diverse RP set. Therefore, after obtaining the initial RP that we use as reference, in case a more suitable candidate RP (tree) is obtained in our heuristic search, the concerning candidate would replace its predecessor in each iteration. The algorithm will stop once the best near optimum minimum set cover relative to our objective function's criteria (namely capacity and correlation) is obtained. The number of diverse MSC RPs obtained for the different randomly generated graphs with random sparseness levels using the aforementioned approach is presented in TABLE I. It is important to note that legacy

MPR's offline algorithm failed to generate any diverse set of RPs in our studied random scenarios. The applicability of our approach in constructing a paths diverse set of RPs for such wide ranging varying random scenarios is indicative of the comprehensiveness of our solution. We note that density in topologies could be very high and the network becomes highly meshed and complex. The results for networks of such scale are indicative of the great scalability of our approach. To this end, the effectiveness of our paths diverse approach for large-scale random topologies is demonstrated which distinguishes our approach from the legacy MPR's RP construction method.

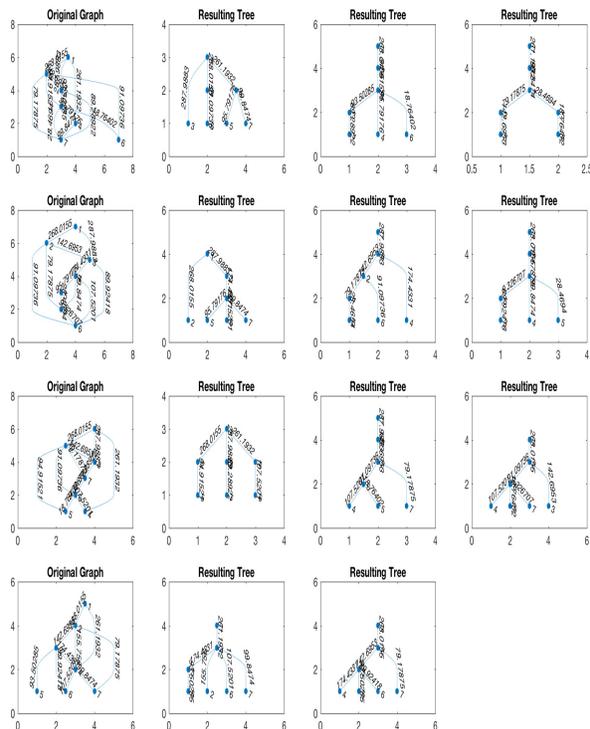


Fig. 2: Random network topologies (T1Sy) with the resulting Trees (i.e. RPs)

Fig. 2 illustrates the *T1Sy* topologies along with the resulting trees for four different random levels of sparseness (as presented in TABLE I). The links' capacities are randomly assigned based on their distance from the gateway (higher if closer - reflecting on heterogeneity [9]). We have chosen not to illustrate the other topologies due to the high link and node density as exemplified in Fig. 3.

## VI. CONCLUSION

In this paper, IP TE-based MPR has been extensively redesigned and reformulated to suit the future random wireless all-IP access network structures. The variable and disparate nature of access network structures as reflected in the evolutionary access networks, along with the rise of IP-based real-time applications demand a new consistent routing optimization paradigm. To this end, a comprehensive TE mechanism

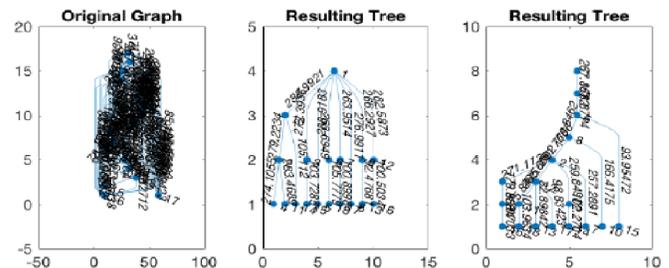


Fig. 3: 18 Node Randomly Generated graph (i.e. T2S3) with two exemplary Resulting Trees (i.e. RPs)

applicable to random access networks is proposed. This mechanism allows the network to maintain a set of diverse independent logical topologies which can be used to balance the traffic load in the network. Our approach adopts a dynamic minimum set-cover based cost function in order to construct a diverse set of RPs while also considering capacity. The results demonstrate the effectiveness of our approach by creating a diverse set of RPs for a wide-ranging set of randomly generated graphs of different sparseness levels where the legacy MPR's TE mechanism is redundant. For future work, SDN will be integrated with our dynamic approach in order to achieve a near-optimal network performance.

## REFERENCES

- [1] N. Wang, K. Ho, G. Pavlou, and M. Howarth, "An overview of routing optimization for internet traffic engineering," *Communications Surveys Tutorials, IEEE*, vol. 10, no. 1, pp. 36–56, Jan. 2008.
- [2] B. Fortz and M. T. At, "Increasing internet capacity using local search," *Computational Optimization and Applications*, vol. 29, pp. 13–48, 2004.
- [3] A. Sridharan, R. Guerin, and C. Diot, "Achieving near-optimal traffic engineering solutions for current OSPF/IS-IS networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 234–247, April 2005.
- [4] K. Nemeth, A. Korosi, and G. Retvari, "Optimal ospf traffic engineering using legacy equal cost multipath load balancing," in *IFIP Networking Conference, 2013*, May 2013, pp. 1–9.
- [5] A. Jaron, A. Mihailovic, and A. Aghvami, "Introducing multi-plane routing for next-generation all-IP wireless access networks," in *IEEE International Conference on Communications*, June 2012.
- [6] M. Farhoudi, B. Abrishamchi, A. Mihailovic, and A. H. Aghvami, "Analysis of practical aspects of multi-plane routing-based load balancing approach for future link-state convergent all-IP access networks," *IEEE Transactions on Mobile Computing*, vol. PP, no. 99, 2017.
- [7] L. Muscariello, D. Perino, and D. Rossi, "Do next generation networks need path diversity?" in *IEEE International Conference on Communications*, June 2009, pp. 1–6.
- [8] M. Farhoudi, A. Jaron, A. Mihailovic, and H. Aghvami, "A new approach for routing plane construction in future multi-plane routing based wireless IP access networks," in *2016 IEEE Wireless Communications and Networking Conference*, April 2016, pp. 1–6.
- [9] M. Mirahsan, R. Schoenen, H. Yanikomeroglu, G. Senarath, and N. Dung-Dao, "User-in-the-loop for hetnets with backhaul capacity constraints," *IEEE Wireless Communications*, vol. 22, no. 5, 2015.
- [10] R. Hassin and A. Levin, "A better-than-greedy approximation algorithm for the minimum set cover problem," *SIAM Journal on Computing*, vol. 35, no. 1, pp. 189–200, 2005.
- [11] H. N. Gabow and E. W. Myers, "Finding all spanning trees of directed and undirected graphs," *SIAM Journal on Computing*, vol. 7, no. 3.
- [12] S. Kapoor and H. Ramesh, "Algorithms for enumerating all spanning trees of undirected and weighted graphs," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 247–265, 1995.
- [13] V. T. Paschos, "A survey of approximately optimal solutions to some covering and packing problems," *ACM Comput. Surv.*, vol. 29, no. 2, pp. 171–209, Jun. 1997.
- [14] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of Operations Research*, vol. 4, no. 3, pp. 233–235, 1979.