Optimizing IBGP Route Reflection Network

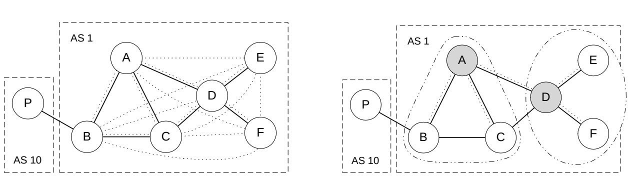
Li Xiao, Jun Wang and Klara Nahrstedt
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, IL 61801

***Abstract*— In the conventional Internal Border Gateway Protocol (IBGP), the BGP sessions between all BGP speakers in a single Autonomous System (AS) form a full mesh. For scalability reasons, route reflection is proposed as an alternative to the full mesh inside an AS. The selection of route reflectors and their clients determines the paths used by IBGP route advertising. Thus, the design of the route reflection graph is an important issue for improving IBGP operating efficiency and reliability.**

This paper focuses on the topology optimization for the route reflection graph, i.e., the selection of the reflectors and the interconnections between reflectors and clients. We propose the Optimum Reflection Graph (ORG) problem to find the best topology for IBGP reflection according to the efficiency or reliability metrics. We give the solvability conditions for the problem and present a solution based on the Integer Programming model. Our approach is also highly flexible. Human decisions or constraints can be easily incorporated to find a topology which satisfies AS administrators' manual configurations.

I. Introduction

Border Gateway Protocol (BGP) [1] is the de-facto routing protocol for exchanging network reachability information at the inter-domain level. BGP is running on millions of routers, called *BGP speakers* in the Internet. BGP speakers communicate with each other via TCP connections, and the communication exchange is called a *BGP session*. According to the relation between two BGP speakers which are connected by a BGP session, BGP can be divided into two parts: External BGP (EBGP) and Internal BGP (IBGP). An EBGP session connects two speakers which reside in different ASes; An IBGP session links two speakers which belong to the same AS. The speaker which runs EBGP sessions is a border node. In Fig. 1(a), for example, there are two ASes, $AS1$ and $AS10$. The nodes represent BGP speakers. P and B are connected by an EBGP session. A and B are connected by an IBGP session. P and B are border nodes.



(a) IBGP with full mesh (b) IBGP with route reflection

Fig. 1. Two IBGP implementations: full mesh and route reflection.

This work was supported by NSF under contract number NSF ANI 00-73802. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

In the conventional design of IBGP, all IBGP sessions in one AS form a full mesh over the BGP speakers. This is shown in Fig. 1(a). The solid lines represent physical links. The dotted lines represent logical links, i.e., IBGP sessions. The border node B informs all its peer routers in $AS1$ about the known network addresses outside the domain.

The full mesh IBGP design is not scalable [2][3]. For example, if there are n BGP speakers in a domain, the total number of IBGP sessions (TCP connections) is $(n^2-n)/2$ and each speaker has to handle $n-1$ IBGP sessions concurrently. There are two practical techniques to solve this scalability problem: route reflection[4] and confederation[5].

The basic idea of route reflection is to divide the BGP speakers into multiple clusters. In each cluster, there are one or more Route Reflectors (RR), and other routers are Route Reflection Clients (RRC). Clients establish IBGP sessions only with the reflectors of the same cluster. All reflectors in the AS establish a full mesh via IBGP sessions. A reflector is responsible for reflecting the routing information to the peer reflectors and its clients. A client only communicates with its reflectors. For example, in Fig. 1(b), the six routers are divided into two clusters. A and D, represented by the shade nodes, are route reflectors and the dotted lines represent IBGP sessions. B advertises the route about P to its reflector A. A summarizes all the routing information and only reflects the best result to the peer reflector D and the other client C. D further informs its clients, E and F, of its routing result.

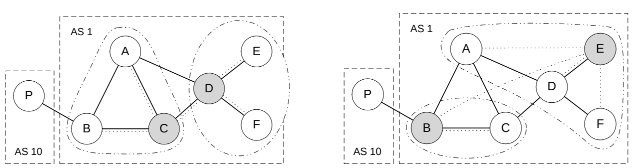
Route reflection has two main advantages. First, it decreases the maximum number of IBGP sessions which a router has to handle concurrently. Second, the implementation of route reflection is easier than confederation which breaks an AS into multiple sub-ASes. In route reflection, only the reflectors need to be modified, and the clients are the same as conventional IBGP speakers. While, in confederation, every BGP router has to be changed.

There is some research on the IBGP configuration with route reflection recently. Without considering Multi-Exit Discriminator (MED), Griffin [6] presents sufficient conditions to guarantee deterministic and unique IBGP routing and to avoid forwarding deflections. Basu [7] expands the advertised routes in IBGP to prevent route oscillation.

However, the design of route reflection topology has not been well studied so far. The guideline for setting up reflection topology is usually to follow the physical topology [8]. In a Point of Presence (PoP), the core routers are selected as the reflectors. But, in general, where to put reflectors and how to cluster routers have not been discussed in the literature. The route reflection topology determines directly which IGP (Interior Gateway Protocol) paths are used to setup IBGP connections. Therefore, the topology can influence the efficiency and reliability of IBGP, and it is thereby inspiring our further investigation.

Fig. 2(a) shows that the reflection topology can affect the reliability of IBGP. It is the same network as Fig. 1, but the IBGP configuration is changed. C and D are reflectors, A and B are the clients of C, and E and F are clients of D. If IGP indicates that link $A \rightarrow D$ is less reliable than other links, then this setup is better than the one previously shown in Fig 1(b). The reason is that if $A \rightarrow D$ is used in the reflection network, the failure of link $A \rightarrow D$ will make A and D close their IBGP connections and invalidate all the routes learned from each other. Rerouting and route oscillation thus happen. However, if link $C \rightarrow D$, which is more reliable, is used for IBGP sessions, then the IBGP reflection network is more reliable and has higher probability to survive.

Fig. 2(b) shows the impact of the reflection topology on the IBGP efficiency. The IBGP topology in the figure is not a good choice, because it is not as efficient as previous designs. The IBGP session between reflectors B and E has at least 3 hops ($B \rightarrow C \rightarrow D \rightarrow E$). The sessions between reflectors and clients are $B \rightarrow C$, $E \rightarrow D$, $E \rightarrow D \rightarrow A$ and $E \rightarrow D \rightarrow F$. The total length of all the IBGP sessions is 9 in terms of hop count, while it is 5 in the other two figures.



(a) An alternative design in case that link $A \rightarrow D$ is unreliable.

(b) A bad design.

Fig. 2. Two other options for reflection topology design.

The route reflection topology design is a constrained optimization problem. There are two kinds of constraints. First, the maximum number of IBGP sessions, which belong to one BGP router, is constrained to a limited value. The star topology, where all clients connect to one reflector, is not scalable. Second, some human configuration decisions restrict the solution space. For example, administrators may want to set some routers as the reflectors manually, or assign a router to a reflector as the client. We use the graph reliability and hop count as two metrics for optimization. The target is to find the most reliable or the most efficient reflection graph. We will show later that these two metrics can be generalized into one type of metrics, so that a unified solution is possible.

There are two main contributions in this paper. First, we propose the Optimum Reflection Graph (ORG) problem to find the best reflection topology and prove the solvability condition for the problem. Second, we present an optimization solution based on Integer Programming, which is highly flexible. Human decisions in topology configuration can be easily incorporated.

The rest of the paper is organized as follows: In Section II, we give the system model and the route reflection optimization problem is defined formally. In Section III, we formulate the problem into integer programming model. Section IV shows the implementation and the result analysis. Section V concludes the paper.

II. System Model

A. Network Models

In one AS, the BGP speakers communicate via IBGP sessions which are based on TCP connections and the IGP routing. If we view each IBGP session as one link, IBGP actually runs on an overlay network, which consists of BGP speakers and the IBGP sessions, above the physical network. Thus, we define the physical graph and logical graph as follows.

1) Physical Graph: A typical network in an AS is represented as graph $G(V, E)$. Node set V represents all the routers. E is the set of physical links, i.e. $E = \{e = (u, v) | u, v \in V, u$ and v are connected by a physical link.$\}$. For each physical link e, define $\gamma(e)$ to be the reliability of link e, which means the probability of successfully transferring a packet over e, or the percentage of time when link e works properly. $0 \leq \gamma(e) \leq 1$.

Given any two nodes s and t, IGP routing (e.g., OSPF) provides a path from s to t, written P_{st}. $P_{st} = (s, v_1, v_2, \ldots, v_n, t)$, where (s, v_1), (v_i, v_{i+1}) and (v_n, t) belong to E. Denote the hop count of P_{st} as $Hop(P_{st})$. The reliability of path P_{st} is $\gamma(P_{st}) = \prod_{e \in P_{st}} \gamma(e)$.

2) Logical Graph: The logical graph, also called *reflection graph*, is based on the physical graph presented above. We denote the logical graph as $G_l(V_l, E_l)$. V_l is the set of nodes running IBGP. $V_l \subseteq V$, and some nodes in V may not have BGP deployed. Define $n = |V_l|$ as the number of IBGP routers. E_l represents the set of IBGP sessions, i.e., $E_l = \{l = (u, v) | u, v \in V_l$, u and v share an IBGP session and l corresponds to IGP path P_{uv} $\}$. Denote the hop count of l as $Hop(l)$, and $Hop(l) = Hop(P_{uv})$. Denote the reliability of l as $\gamma(l)$, and $\gamma(l) = \gamma(P_{uv})$. Define the number of IBGP sessions associated with node v as the degree of v. Define the largest degree of all nodes in $G_l(V_l, E_l)$ as $max_D(G_l(V_l, E_l))$.

In the conventional IBGP, the logical graph is a full mesh, i.e., there is an IBGP session between any pair of nodes in V_l. If route reflection is installed, the nodes in V_l are divided into M clusters: $\mathcal{P}_1$, $\mathcal{P}_2$, $\ldots$, $\mathcal{P}_M$, where $\bigcup_{1 \leq i \leq M} \mathcal{P}_i = V_l$, and $\mathcal{P}_i \cap \mathcal{P}_j = \emptyset, \forall i \neq j$. Denote the set of route reflectors in cluster $\mathcal{P}_i$ as $\mathcal{R}_i$, and $\mathcal{R}_i \subseteq \mathcal{P}_i$. Denote $\mathcal{R} = \bigcup_{1 \leq i \leq M} \mathcal{R}_i$, and $\mathcal{C}_i = \mathcal{P}_i \setminus \mathcal{R}_i$. The link set E_l can be divided into three categories, $E_l = E_l^m \cup E_l^r \cup E_l^c$. E_l^m is the full mesh among all nodes in $\mathcal{R}$, i.e., $E_l^m = \{(u, v) | \forall u, \forall v, u \neq v, \text{ and } u, v \in \mathcal{R}\}$. E_l^r represents the links between clients and the correspondent reflectors, i.e., $E_l^r = \bigcup_{1 \leq i \leq M} \{(u, v) | \forall u, \forall v, u \in \mathcal{C}_i, \text{ and } v \in \mathcal{R}_i\}$. E_l^c represents links between some clients in one cluster, which is optional in route reflection design. An example of route reflection graph is in Fig. 3.

The hop count of the logical graph is the sum of the hop count of all links, i.e., $Hop(G_l) = \sum_{e \in E_l} Hop(e) = \sum_{e \in E_l} Hop(P_e)$. The hop count of the logical graph measures the size of the reflection graph. We use this value to represent IBGP efficiency. If the value is smaller, the traffic of routing

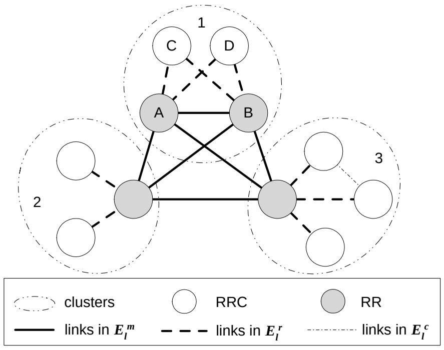


Fig. 3. An example of route reflection logical graph

message is less and it takes less time to propagate route information to all routers.

The reliability of the logical graph is the product of the reliability of all links, i.e., $\gamma(G_l) = \prod_{e \in E_l} \gamma(e) = \prod_{e \in E_l} \gamma(P_e)$. If a IBGP session breaks, all routes associated with this session will be marked as invalid, and this causes a large amount of route oscillation and rerouting. Thus, finding a reliable logical graph is important.

Note: $\gamma(G_l)$ and $Hop(G_l)$ are measurements for the reflection graph, which show, to some extent, the reliability and efficiency of operating IBGP. Although, IBGP reliability and efficiency are influenced by other factors too. For example, IBGP can be made more reliable by using multiple reflectors in a cluster. In Fig. 3, cluster 1 has two reflectors. C can get reflected routes from either A or B. In this paper, we will not discuss the reliability, affected by the redundant design. We only focus on how to find the most reliable or efficient reflection graph which does not have redundancy.

3) Assumptions: The reflection graph, defined in previous section, only shows the two-level reflection. Theoretically, there can be multiple levels of reflection, i.e., a route reflector could be the client of a reflector at a higher level. However, in practice, the two-level reflection is most often used. In this paper, we only focus on the two-level reflection. We further assume that the reflection graph does not have redundancy, i.e., there is only one reflector in each cluster and there is no E_l^c link. Based on these two assumptions, let $G_{l0}(V_l, E_l)$ denote the reflection graph $G_l(V_l, E_l)$, which uses two-level reflection and does not have redundancy.

B. Problem Formulation

Reflection graph optimization problem is to find the most reliable or efficient topology for IBGP. For convenience, $G_{l0}(V_l, E_l, \alpha)$ is defined as the reflection graph $G_{l0}(V_l, E_l)$, whose largest node degree is equal or less than α, i.e., $max_D(G_l(V_l, E_l)) \leq \alpha$.

Definition 1 (The Most Efficient Reflection Graph Problem): Let us assume that V_l, the set of IBGP routers, and α, the upper bound of the node degree, are given. Set $\{Hop(P_{ij}) | \forall i, j \in V_l, i \neq j\}$ is known from IGP routing. The most efficient reflection graph problem means to find a logical graph $G^*_{l0}(V_l, E^*_l, \alpha)$ which satisfies $Hop(G^*_{l0}(V_l, E^*_l, \alpha)) \leq Hop(G_{l0}(V_l, E_l, \alpha))$, for any $G_{l0}(V_l, E_l, \alpha)$.

Definition 2 (The Most Reliable Reflection Graph Problem): Let us assume that V_l, the set of IBGP routers, and α, the upper bound of the node degree, are given. Set $\{\gamma(P_{ij}) | \forall i, j \in V_l, i \neq j\}$ is known from IGP routing. The most reliable reflection graph problem means to find a logical graph $G^*_{l0}(V_l, E^*_l, \alpha)$ which satisfies $\gamma(G^*_{l0}(V_l, E^*_l, \alpha)) \geq \gamma(G_{l0}(V_l, E_l, \alpha))$, for any $G_{l0}(V_l, E_l, \alpha)$.

Although the two problems use additive and multiplicative metrics respectively, they can be unified into one problem as follows. The solution for the Most Reliable Reflection Graph Problem is $\arg\max_{E_l} \gamma(G_{l0})$. Because

$$\begin{aligned} \arg\max_{E_l} \gamma(G_{l0}) &= \arg\max_{E_l} \prod_{e \in E_l} \gamma(e) = \arg\max_{E_l} \sum_{e \in E_l} \log(\gamma(e)) \\ &= \arg\min_{E_l} \sum_{e \in E_l} (-1)\log(\gamma(e)) \end{aligned}$$

and the solution for the Most Efficient Reflection Graph Problem is $\arg\min_{E_l} \sum_{e \in E_l} Hop(e)$, we can unify these two problems into the same linear optimization problem. We define general weight w_e as $Hop(e)$ or $(-1)\log(\gamma(e))$, and $w_{G_{l0}}$ as $\sum_{e \in E_l} Hop(e)$ or $\sum_{e \in E_l} (-1)\log(\gamma(e))$. The unified optimization problem is presented as follows.

Definition 3 (Optimal Reflection Graph Problem (ORG)): Given the set of IBGP routers V_l, upper bound of node degree α, and the weight between any pair of nodes w_{ij}, ORG problem means to find $\arg\min_{E_l} w_{G_{l0}}$, i.e., $\arg\min_{E_l} \sum_{e \in E_l} w_e$.

The ORG problem has very large solution space, which is affected by the number of clusters, the clustering variations, and the placement of reflectors. If the α constraint is ignored, there are $\sum_{M=1}^{N} \sum_{i=0}^{M} (-1)^i \frac{(M-i)^N}{i!(M-i)!}$ variations to cluster N nodes.

The ORG problem generalizes the formulations of the Most Efficient and the Most Reliable Reflection Graph Problems. However, in order to find a reflection graph which is both acceptably reliable and efficient, it is also meaningful to discuss the multi-metric optimization problem. A simplified approach is to aggregate these two metrics together as a new metric: $w_{G_{l0}} = \sum_{e \in E_l} [(-1)\theta \log(\gamma(e)) + (1-\theta)Hop(e)]$, where $\theta \in [0, 1]$. In practice, the reliability metric is more important than the efficiency. Thus, we choose a θ that is close to 1.

C. The Solvability Condition for ORG Problem

α gives the largest number of IBGP sessions that a node can have. If the value of α is too small, it may be impossible to construct a reflection graph, i.e., $G_{l0}(V_l, E_l, \alpha)$ does not exist, which means the ORG problem is not solvable. The theorems below give conditions, for the ORG problem to have feasible solutions.

Theorem 1: The sufficient and necessary condition for ORG problem to be solvable is $\alpha \geq \lceil \frac{n}{\lceil \sqrt{n} \rceil} \rceil + \lceil \sqrt{n}\ \rceil - 2$, where n is the number of IBGP routers.

Proof: Suppose $k = |\mathcal{R}|$. Define $G_{l0}(k)$ as the reflection graph $G_{l0}(V_l, E_l, \alpha)$ which has k reflectors. First, we will find the minimum largest node degree a reflection graph can have.

Define $f(k) = \min max_D(G_{l0}(k))$ as the minimum largest node degree in $G_{l0}(k)$. One of the reflectors should have the largest node degree, which includes $k-1$ links to other reflectors and at least $\lceil (n-k)/k \rceil$ links to its clients. Thus,

$$f(k) = k - 1 + \lceil \frac{n-k}{k} \rceil = \lceil \frac{n}{k} + k - 2 \rceil \quad (1)$$

Because $k-2+n/k$ reaches its minimum value when $k = \sqrt{n}$, we have

$$\begin{aligned} \min_k f(k) &= \min(\lceil \frac{n}{\lfloor \sqrt{n} \rfloor} \rceil + \lfloor \sqrt{n} \rfloor - 2, \lceil \frac{n}{\lceil \sqrt{n} \rceil} \rceil + \lceil \sqrt{n} \rceil - 2) \\ &= \lceil \frac{n}{\lceil \sqrt{n} \rceil} \rceil + \lceil \sqrt{n} \rceil - 2 \quad (2) \end{aligned}$$

We get equation 2, because it can be proved $\lceil \frac{n}{\lfloor \sqrt{n} \rfloor} \rceil + \lfloor \sqrt{n} \rfloor = \lceil \frac{n}{\lceil \sqrt{n} \rceil} \rceil + \lceil \sqrt{n} \rceil$. When $\alpha \geq \min_k f(k)$, ORG problem is solvable, and vice versa. Therefore, the theorem is proved. ■

Corollary 1: Let n be the number of IBGP routers. If $\alpha \geq 2\lceil \sqrt{n} \rceil - 2$, reflection graph $G_{10}(V_l, E_l, \alpha)$ exists and the ORG problem is solvable.

Proof: It is easy to show that $\alpha \geq 2\lceil \sqrt{n} \rceil - 2 \geq \lceil \frac{n}{\lceil \sqrt{n} \rceil} \rceil + \lceil \sqrt{n} \rceil - 2$. Therefore, the ORG problem is solvable. ■

On the other hand, if α is given, the number of reflectors $|\mathcal{R}|$ is also constrained, which is shown in the following theorem.

Theorem 2: Given α and n, the condition for the existence of G_{l0} is $\frac{\alpha+2-\sqrt{\alpha^2+4\alpha-4n+4}}{2} \leq |\mathcal{R}| \leq \frac{\alpha+2+\sqrt{\alpha^2+4\alpha-4n+4}}{2}$.

Proof: Suppose $|\mathcal{R}| = k$. From equation (1), the constraint is $\alpha \geq \lceil n/k + k - 2 \rceil$. Thus,

$$\alpha - k + 2 \geq \lceil n/k \rceil \quad (3)$$

Denote all the integers k which satisfy formula (3) as S_k. Define S'_k as the set of the solutions for $\alpha - k + 2 \geq n/k$. Because $\lceil n/k \rceil \geq n/k$, we have $S_k \subseteq S'_k$. $\forall k' \in S'_k$, if $k'|n$, obviously $k' \in S_k$. Otherwise, since k' is integer, $\alpha - k' + 2 \geq \lfloor n/k \rfloor + 1 = \lceil n/k \rceil$. Thus, $k' \in S_k$. Therefore, we have $S_k = S'_k$. We can further prove the theorem by obtaining S'_k easily. ■

III. Integer Programming Optimization

In this section, we give the solution to the ORG problem from Integer Programming approach.

A. Integer Programming Formulation

Let us define $\mathcal{I}$ as all integers in interval $[1, n]$, where n is the number of IBGP routers. We define following binary variables, $i, j \in \mathcal{I}$:

$$x_i = \begin{cases} 1 & : \text{ node } i \text{ is a route reflector.} \\ 0 & : \text{ node } i \text{ is a client.} \end{cases} \quad (4)$$

$$r_{ij} = \begin{cases} 1 & : \text{ node } i \text{ is the route reflector of node } j. \\ 0 & : \text{ otherwise.} \end{cases} \quad (5)$$

$$s_{ij} = \begin{cases} 1 & : \text{ node } i \text{ and } j \text{ are connected by an IBGP session.} \\ 0 & : \text{ otherwise.} \end{cases} \quad (6)$$

$$w_{ij} \quad : \quad \text{the weight of the IGP path from } i \text{ to } j \quad (7)$$

x_i controls if a node is a reflector. r_{ij} controls the reflection relationship. Each client node connects to one and only one reflector. s_{ij} controls if two nodes are IBGP peers. w_{ij} is the general weight, which represents $Hop(P_{ij})$, $(-1)\log(\gamma(P_{ij}))$ or the weighted aggregation of the two. w_{G_l} is denoted as W.

$$W = \sum_{i,j \in \mathcal{I}, i \neq j} s_{ij} w_{ij} \quad (8)$$

The optimization objective is to minimize W.

$$\min_{s_{ij}} W \quad (9)$$

The optimization problem is subjected to the following constraints.

$$\sum_{\substack{i \in \mathcal{I} \\ i \neq j}} r_{ij} = 1 - x_j \qquad \forall j \in \mathcal{I} \quad (10)$$

$$\sum_{\substack{j \in \mathcal{I} \\ i \neq j}} r_{ij} \leq N x_i \qquad \forall i \in \mathcal{I} \quad (11)$$

$$s_{ij} \geq r_{ij} + r_{ji} \quad \forall i, j \in \mathcal{I}, i \neq j \quad (12)$$

$$s_{ij} \geq x_i + x_j - 1 \quad \forall i, j \in \mathcal{I}, i \neq j \quad (13)$$

$$\sum_{\substack{j \in \mathcal{I} \\ i \neq j}} s_{ij} \leq (\alpha - 1) x_i + 1 \qquad \forall i \in \mathcal{I} \quad (14)$$

$$s_{ij} = s_{ji} \quad \forall i, j \in \mathcal{I}, i \neq j \quad (15)$$

Equation (10) guarantees that a client node has exactly one reflector, and a reflector can not be the client of other nodes. Equation (11) ensures that a client can not be the reflector for any other nodes. Equation (12) guarantees that if one of the two nodes is the reflector of the other, they are IBGP peers. Equation (13) ensures that if two nodes are both reflectors, they are IBGP peers (link in E_l^m). Equation (14) ensures that the largest number of IBGP sessions that a node establishes can not exceed α. Equation (15) guarantees that IBGP session is a bidirectional connection.

B. Manual Reflector and Client Configuration

Manual configuration is one of the major determinants in the reflection graph design. For example, some nodes are preset to be clients for special reasons. In our optimization model, human configuration can be incorporated easily. If node i is configured as the route reflector and node j is the client, $x_i = 1$ and $x_j = 0$ are added into the constraint set. In general, if at least n_r nodes in $\mathcal{W}_r$ are configured as route reflectors, and at least n_c nodes in $\mathcal{W}_c$ are configured as clients, the following conditions are added to the constraint set:

$$\sum_{i \in \mathcal{W}_r} x_i \geq n_r \quad (16)$$

$$\sum_{i \in \mathcal{W}_c} x_i \leq |\mathcal{W}_c| - n_c \quad (17)$$

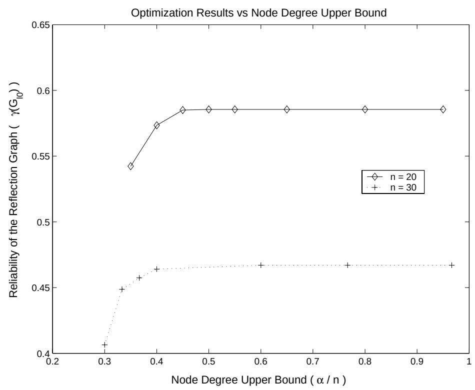


Fig. 4. The Relation between Optimization Results and α

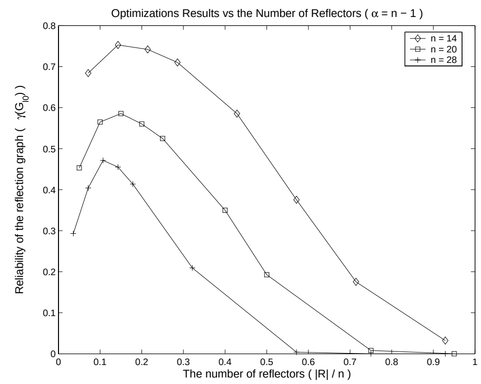


Fig. 5. The Impact of $|\mathcal{R}|$ on Optimization Results.

IV. Optimization Result Analysis

In order to evaluate the Integer Programming model introduced in previous sections for the design of reflection graphs, the optimization model is implemented, and the reliability metric is used as an example.

The proposed Integer Programming model is solved by ILOG CPLEX, a mixed Integer Programming solver[9]. The reflection graph can be constructed based on the optimized results $\{x_i\}$ and $\{r_{ij}\}$.

The network topology is generated by BRITE generator[10]. The Waxman model is used and nodes are placed according to the heavy-tail distribution. The reliability of each physical link is generated randomly from $[0.95, 1.0]$. We assume that the minimum hop-count routing is used in IGP. Based on the IGP routing result, we can calculate $\{w_{ij}\}$ for any pair of nodes. The target is to find the most reliable reflection graph. We use the reliability of the reflection graph $\gamma(G_{l0})$ to evaluate the optimization results in different scenarios.

We present the impact of the node degree upper bound and the impact of the number of reflectors on the optimization results.

A. Impact of α

By solving the Integer Programming model, we can find the optimum reflection graph given a node degree constraint α. Intuitively, if α is smaller, we have less feasible solutions and thereby the optimized graph reliability is lower.

Fig. 4 shows the impact of α on the optimization results. The vertical axis represents the reflection graph reliability, and the horizontal axis is the ratio α/n, where n is the number of IBGP routers. When α is large enough, the construction of the reflection graph is not constrained by α, and the most reliable graph can be found. When α decreases below certain value, the reliability of the reflection graph decreases. When α is smaller than $\lceil \frac{n}{\lceil\sqrt{n}\rceil} \rceil + \lceil \sqrt{n}\ \rceil - 2$, according to Theorem 1, no feasible solution exists.

B. Impact of $|\mathcal{R}|$

Fig. 5 shows the impact of the number of reflectors on the optimization results. $\sum_{i\in\mathcal{I}} x_i = |\mathcal{R}|$ is added into the constraint set to calculate the optimal reflection graph in scenarios of different numbers of reflectors. α is set to $n-1$ to remove the impact of the node degree upper bound.

The horizontal axis is the ratio $|\mathcal{R}|/n$, and the vertical axis shows the reliability. When $|\mathcal{R}|$ is a large number, the reflection graph is similar to a full mesh, and too many E_l^m links make the graph reliability low. On the other hand, if $|\mathcal{R}|$ is too small, clients have to use long paths to connect to reflectors, and E_l^r is the main determinant of the graph reliability. When $|\mathcal{R}|$ is some number in the middle, e.g., $[0.1n, 0.2n]$ in the figure, the optimal reflection graph can be found.

V. Conclusion

In this paper, we propose the Optimized Reflection Graph (ORG) problem for the IBGP reflection graph design. We present the conditions under which the problem is solvable and give a solution based on the Integer Programming to find the best reflection graph. Our model is highly flexible to incorporate manual configurations.

References

[1] Y. Rekhter and T. Li, *A Border Gateway Protocol 4 (BGP-4). RFC 1771.* Network Working Group, March 1995.

[2] R. Dude, "A comparison of scaling techniques for BGP," *ACM Computer Communication Review*, vol. 29, Jul. 1999.

[3] J. G. Scudder and R. Dude, "BGP scaling techniques revisited," *ACM Computer Communication Review*, vol. 29, Oct. 1999.

[4] T. Bates, R. Chandra, and E. Chen, *BGP Route Reflection - An Alternative to Full Mesh IBGP. RFC 2796*. Network Working Group, April 2000.

[5] P. Traina, D. McPherson, and J. Scudder, *Autonomous System Confederations for BGP. RFC 3065*. Network Working Group, February 2001.

[6] T. G. Griffin and G. Wilfong, "On the correctness of IBGP configuration," in *Proceedings of ACM SIGCOMM*, 2002.

[7] A. Basu, C.-H. L. Ong, A. Rasala, F. B. Shepherd, and G. Wilfong, "Route oscillations in I-BGP with route reflection," in *Proceedings of ACM SIGCOMM*, 2002.

[8] S. Halabi and D. McPherson, *Internet Routing Architectures*. Cisco Press, 2000.

[9] "ILOG CPLEX," in *http://www.ilog.com/products/cplex/*.

[10] "Boston university representative internet topology generator," in *http://cs-www.bu.edu/brite/*.