Slowing Routing Table Growth by Filtering Based on Address Allocation Policies

Steve Bellovin, Randy Bush, Timothy G. Griffin, and Jennifer Rexford
{smb,randy,griffin,jrex}@research.att.com

Abstract— **BGP routing tables have been growing at an alarming rate in recent years. In this paper we investigate how BGP table size can be reduced and the rate of growth slowed by applying filters that enforce the allocations boundaries documented by the numbering authorities. In addition, this appears to be possible while losing reachability to only a small percent of addresses.**

I. Introduction

The rapid growth of the Internet during the past few years has led to increased concerns about the scalability of the underlying routing infrastructure. The Internet consists of thousands of autonomous systems (ASes) that interact to coordinate the delivery of IP traffic. Neighboring ASes use the Border Gateway Protocol (BGP) to exchange routing information [1], [2], [3]. Each BGP route advertisement concerns a particular block of IP addresses (a *prefix*) and includes a list of the ASes in the path, along with other attributes. A router stores its best and alternate routes for each prefix in a BGP routing table and uses this information to construct a forwarding table that controls the forwarding of each incoming packet to the next hop in its journey. The number of prefixes in BGP routing tables has important implications on storage requirements, computational load, forwarding performance, and protocol overheads for Internet routers. In this paper, we evaluate the effectiveness of filtering techniques that network operators can apply to reduce the number of prefixes and the rate of growth of their views of the routing table.

An IP address consists of a network portion (or prefix) and a host portion. Routing through the Internet is based only on the network portion of the address. Initially, IP addresses were allocated in three main block sizes, or classes, based on the number of octets devoted to the network and host portions. A Class A address starts with a 0 in the first bit and uses the first octet for the network address, a Class B address starts with a 10 in the first two bits and uses the first two octets for the network address, and a Class C address starts with 110 and uses the first three octets for the network address. The remaining addresses are reserved for multicast groups (Class D) and future use (Class E). The restriction of fixed block sizes was lifted with the introduction of Classless InterDomain Routing (CIDR) [4], [5]. CIDR permits an arbitrary division between the network and host portions of the address. A mask length identifies the number of bits devoted to the network part of the address. For example, the prefix 204.70.2.0/23 has a 23-bit mask, leaving nine bits for the host portion of the address.

CIDR allows network providers to allocate small blocks of IP addresses to different customers while advertising a large, aggregated block to the rest of the Internet [6]. The deployment of CIDR and the CIDR-compatible BGP-4 slowed the rate of routing table growth. However, the growth of the Internet in the late 1990s led to a new surge in the size of BGP tables [7], [8]. The growth stems, in part, from the allocation of new blocks of IP addresses by the Regional Internet Registries (RIRs). In addition, some ASes advertise small address blocks (i.e., prefixes with large mask lengths) to balance the traffic load over multiple paths through the network. In other cases, an AS may advertise a small block of addresses on behalf of a customer that connects to two or more upstream providers. Otherwise, the customer's prefix would not be reachable through each of the providers. Finally, in some cases, an AS may advertise a small address block due to misconfiguration. Each of these factors may contribute to some extent

to the increasing size of BGP routing tables.

Network operators configure their routers to apply filters to incoming BGP route advertisements. These filters prevent the router from accepting inappropriate advertisements, such as routes to private addresses. In this paper, we investigate the potential for route filtering to help control the growth of BGP routing tables. We focus on three types of filtering rules. First, operators typically filter routes for so-called "martian" addresses that should not appear in the global routing tables. Second, in practice some operators filter prefixes with a mask length that is longer than 24; in the Class B range of addresses, operators sometimes filter prefixes with a mask length that exceeds 16. Third, the RIRs publish allocation rules that dictate the maximum mask length for prefixes in certain regions of the address space; operators could reduce the size of BGP routing tables by applying filters that enforce these allocation rules. We discuss these filtering policies in more detail in Section II. We also describe the collection of routing table data that we use to evaluate these filtering policies over the past few years and from multiple vantage points in the Internet.

Then, in Section III we analyze the growth of the BGP routing table subject to the various routing filter policies. We show that the number of prefixes that have masks lengths that are longer than the RIR allocation policies is growing at a faster rate than other prefixes. We also show that, by applying the entire set of filters, it is possible to reduce the size of BGP tables from 90,000–110,000 prefixes to just over 70,000 prefixes and divide the growth rate roughly in half. Larger BGP tables reduce by a more significant amount since a larger number of the prefixes that exceed the RIR policies. In some cases, a filtered prefix may be covered by a larger block of addresses in the routing table (e.g., 204.70.2.0/23 would be covered by 204.70.0.0/16). However, aggressive filtering may make some parts of the Internet address space unreachable. We attempt to quantify the potential loss of reachability. We show that the most aggressive filtering policy leaves about 0.3% of the address space uncovered by any remaining prefix in the routing table. We analyze the main contributors to the prefixes that exceed the RIR policies and discuss ways to prevent reachability problems. The paper concludes in Section IV with a discussion of future directions.

II. Routing Data

The size of a routing table depends on the vantage point of the router and the filtering policy applied by the network operator. In this section, we discuss the filtering policies and routing table data that we analyze in the remainder of the paper.

A. Policies for Prefix-Based Filtering

Network operators can configure their routers to filter certain routes based on the region of the address space and the mask length of the prefix in the advertisement. Operators are advised to filter martian addresses. Some operators filter prefixes with mask lengths longer than 24, or longer than 16 in the Class B portion of the IP address space. Operators could also filter prefixes that have larger mask lengths than the address allocation guidelines published by the Regional Internet Registries (RIRs).

Martians: The IPv4 address space includes several "special use" prefixes [9] that have been reserved by the Internet Assigned Numbers Authority (IANA). Network operators should not accept or send advertisements for these martian addresses, as summarized in Table I. The Class A block 0.0.0.0/8 includes the address 0.0.0.0 which is commonly used for default routes. The 127.0.0.0/8 prefix is reserved for loopback addresses used by a host or router to identify itself. Three prefixes are reserved for private networks that use the IP protocols, as discussed in RFC 1918 [10]. The 224.0.0.0/3 block is devoted to Class D (multicast) and Class E (reserved) addresses. The 169.254.0.0/16 block is dedicated for auto-configuration of hosts when no DHCP (Dynamic Host Configuration Protocol) server is available. The prefix 192.0.2.0/24 is

Category	Prefix(es)
Default/broadcast	0.0.0.0/8
Loopback addresses	127.0.0.0/8
Private addresses	10.0.0.0/8 172.16.0.0/12 192.168.0.0/16
Class D/E	224.0.0.0/3
Auto-configuration	169.254.0.0/16
Test network	192.0.2.0/24
Exchange points	192.41.177.0/24 192.157.69.0/24 198.32.0.0/16 206.220.243.0/24
IANA reserved	128.0.0.0/16

TABLE I
MARTIAN ADDRESS BLOCKS

used for example IP addresses in documentation and code fragments. Several prefixes are allocated to the infrastructure at the public Internet exchange points. In addition, prefix 128.0.0.0/16 is reserved by IANA.

Operator policies: To limit routing table size, some network operators configure their routers to filter all prefixes with a mask length larger than 24. This protects the routers in the AS from storing and processing routes for a large number of small address blocks. In practice, an AS might apply this filter only to routes learned from a peer or upstream provider while still accepting prefixes with larger mask lengths from customers. That is, an AS may be willing to carry these prefixes on behalf of a paying customer but not for other ASes in the Internet. In addition to removing routes for small address blocks, some operators filter advertisements in the Class B space that have a prefix with a mask longer than 16. The RIRs do not allocate small address blocks within the Class B space to other institutions. Rather, the RIRs allocate address blocks in the Class A and Class C regions of the IP address space, as discussed below.

RIR allocation policies: The remainder of the IPv4 address space is allocated by IANA to the three RIRs—APNIC (Asia-Pacific Network Information Centre), ARIN (American Registry for Internet Numbers), and RIPE (Réseaux IP Européens)—with new registries proposed for Africa and Latin America. Each regional registry allocates address blocks to Local Internet Registries (LIRs) and other organizations within their regions. To reduce the impact on the size of the routing tables, the three regional registries limit the allocation sizes in different parts of the Class A and Class C portions of the address space [11], [12], [13], as summarized in Appendix A. For example, ARIN does not make allocations in the 63.0.0.0/8 space with a mask length longer than 19; similarly, APNIC does not make allocations in the 211.0.0.0/8 space with a mask length longer than 23. These allocation policies are publicized to aid the ISP community in filtering and other policy decisions. Users announcing smaller blocks (with longer masks) are warned that network operators throughout the Internet may choose to filter prefixes that exceed the address allocation guidelines.

The various filtering rules can be represented in a simple, common format consisting of three fields—a 32-bit address and an integer mask to represent a block of IP addresses, along with another integer for the maximum mask length permitted for prefixes in this address range. For example, the APNIC rule for addresses in 211.0.0.0/8 is represented as $(211.0.0.0, 8, 23)$, whereas the martian 127.0.0.0/8 is represented as $(127.0.0.0, 8, 0)$ since no prefixes are permitted in this block. The Class B rule is represented as $(128.0.0.0, 2, 16)$, whereas the rule that filters all prefixes with a mask length greater than 24 is represented as $(0.0.0.0, 0, 24)$.

B. BGP Routing Tables

Our evaluation of filtering policies draws on publicly-available BGP routing tables collected from several locations with the permission of the owners of the data, as summarized in Table II. The University of Oregon Route-Views project [14] and the RIPE Route Information Service project [15] provide an archive of BGP routing table data from multiple vantage points in the Internet. For

Table	Dates	Format
RouteViews [16]	12/01/97–12/01/00	IOS
RIPE NCC [17]	01/01/00–06/10/01	MRT
Telstra [18]	03/16/01	IOS
Verio	03/16/01	IOS

TABLE II
COLLECTION OF ROUTING TABLES

example, the RouteViews router has multi-hop BGP sessions with several dozen ASes that agree to advertise their routes. Figure 1 shows a fragment of a sample dump of a RouteViews table, available from a public Web site provided by NLANR [16]. The table was originally generated by Telneting to the RouteViews router and running "show ip bgp" at the command line. The ASCII output includes a list of routes for each prefix (e.g., 3.0.0.0/8 and 9.2.0.0/16), where each route has a next hop (the IP address of the BGP speaker), an AS path, and various other attributes. Focusing on a particular next-hop IP address enables us to construct a view of the routing table from the vantage point of a particular router in the corresponding next-hop AS. For example the next-hop IP address 193.140.0.1 corresponds to AS 8517 (i.e., UlakNet in Turkey).

In contrast to the RouteViews data, the RIPE NCC site [17] has an archive of BGP routing tables in the binary MRT (Multi-threaded Routing Toolkit) format. The RouteViews and RIPE data provide a unique opportunity to study routing tables from multiple vantage points over a period of time. However, the RouteViews and RIPE routers do not have control over whether a participating ASes advertises all of the prefixes in its routing tables. In practice, most participating ASes treat these multi-view routers as a customer. However, a router inside a participating AS may have a larger table that includes additional routes that do not appear in the RouteViews and RIPE tables. For example, an ISP may have prefixes that correspond to parts of their infrastructure (e.g., individual routers and interfaces) and fine-grain prefixes used to balance load over a set of links to the same customer. An ISP may choose not to apply strict filtering policies to these internal prefixes. To study filtering policies from the vantage point of a router in a particular AS, we also analyze routing tables collected from inside Telstra [18] and from a Verio customer.

Our software for analyzing the BGP tables accepts data in IOS or MRT format in gzipped or uncompressed form. Each routing table is converted into a fixed-format, bar-separated stream of records with one line for each route. The software can apply a list of filtering rules to the stream of records, with the option of focusing on routes advertised by a particular next-hop IP address. In the next section, we summarize our results in applying the filtering rules to the collection of BGP routing tables.

III. PERFORMANCE EVALUATION

Applying the filtering policies reduces the number of prefixes in the BGP tables and the growth over time. In this section, we evaluate the influence of the filtering policies on routing table size and examine the possibility that the IP addresses in a filtered prefix become unreachable.

A. Routing Table Size

To evaluate the influence of filtering policies on routing table size, we applied a collection of filtering rules to two routing tables from May 16, 2001. We focus on the Telstra and Verio data to provide a view of routing tables from two ASes in different parts of the world. These two routers hear different views of the routing tables from their neighboring ASes and apply different filtering policies. Both of these factors contribute to differences in the number and type of prefixes seen in the routing tables. The Telstra and Verio tables start with 108,400 and 86,997 unique prefixes, respectively, as summarized in Figure 2. Starting with the initial routing table, we first removed any prefixes that fall within the martian address blocks. Next, we simulated the common practice of removing prefixes with mask lengths longer than 24. Then, we applied the filtering rules that relate to the Class A, B, and C portions of the address space.

```
BGP table version is 2108687, local router ID is 198.32.162.100
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network      Next Hop         Metric LocPrf Weight Path
*  3.0.0.0      167.142.3.6                           0 5056 701 80 i
*               4.0.0.2            2095               0 1 701 80 i
*               204.42.253.253                        0 267 2914 701 80 i
*               212.4.193.253                         0 8918 701 80 i
*               205.215.45.50                         0 4006 701 80 i
*               193.140.0.1                           0 8517 9000 2548 701 80 i
*               165.87.32.5                           0 2685 701 80 e
*               203.62.252.21                         0 1221 16779 1 701 80 i
...
*  9.2.0.0/16   167.142.3.6                           0 5056 701 i
*               4.0.0.2            2095               0 1 701 i
*               204.42.253.253                        0 267 2914 701 i
*               212.4.193.253                         0 8918 701 i
*               205.215.45.50                         0 4006 701 i
*               193.140.0.1                           0 8517 9000 2548 701 i
*               203.62.252.21                         0 1221 5727 701 i
```

Fig. 1. Example fragment of BGP table from the RouteViews server

Since the address classes are non-overlapping, the three sets of rules could be applied in any order without affecting the number of prefixes removed in each stage. We removed the prefixes with mask lengths that exceed the RIR allocation policies for the Class A portion of the address space, prefixes in Class B space that have a mask length larger than 16, and prefixes that exceed the RIR allocation policies for the Class C portion of the address space.

Each routing table includes 33 martian prefixes. Upon further inspection, each of these prefixes falls in the set of addresses allocated to public exchange points. These prefixes appear in the routing tables to enable these two service providers to communicate with other routers at the exchange points. Filtering prefixes with a mask length larger than 24 removes 5984 prefixes from the Telstra table; 169 of these prefixes originate directly from Telstra. This filter removes 80 prefixes from the Verio table; 13 of these prefixes originate directly from Verio. The Class A, B, and C rules remove additional prefixes. The most significant reduction comes from removing the prefixes that exceed the RIRs' allocation policies in the Class C portion of the address space. Enforcing these allocation rules would remove 16,398 prefixes from the Telstra table and 15,971 prefixes from the Verio table. Applying the entire set of filters would reduce the number of prefixes to 70,892 and 69,198, respectively. Experiments with the RouteViews and RIPE data from the vantage point of other ASes reveal similar trends, with a final table size of about 70,000 prefixes after applying the entire collection of filtering rules.

In addition to reducing the number of prefixes, applying the filtering policies reduces the growth rate in the routing tables over time. Figure 3(a) plots the number of prefixes for Cable and Wireless (AS 3561) over a three-year period. The top curve shows that the number of prefixes advertised to the RouteViews server grew steadily from 46,343 in December 1997 to 88,700 in November 2000—an increase of 42,357 prefixes. Filtering martians and prefixes with mask lengths larger than 24 does not make an appreciable difference in the number of prefixes. Applying these two rules to the November 2000 data only reduces the table size by

	prefixes removed	prefixes left
initial		108400
martians	33	108367
24	5984	102383
A	8118	94265
B	6975	87290
C	16398	70892

(a) Telstra

	prefixes removed	prefixes left
initial		86997
martians	33	86964
24	80	86884
A	708	86176
B	1007	85169
C	15971	69198

(b) Verio

Fig. 2. Routing table size for May 16, 2001

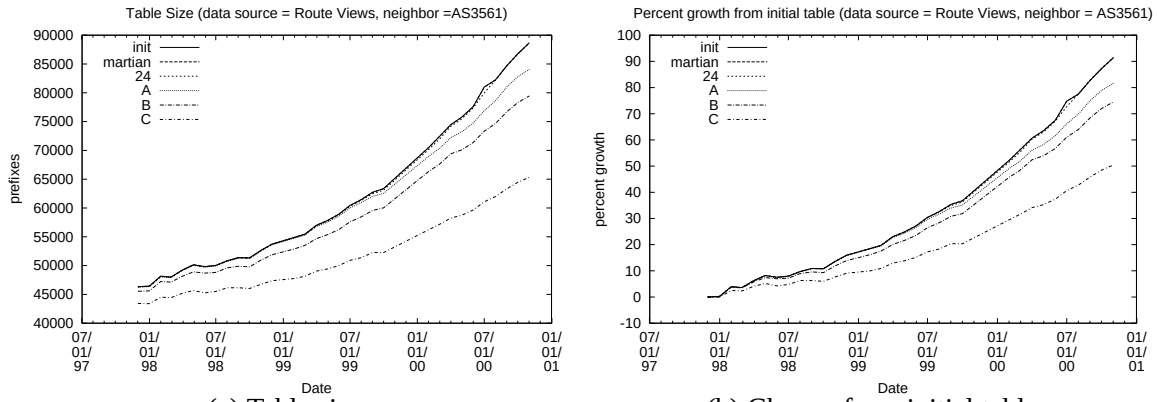
a total of 27 prefixes—from 88,700 to 88,673. The Class A, B, and C filtering policies have a more significant influence on table size, with the bulk of the benefit coming from the Class C rules. The bottom curve in Figure 3(a) shows that applying the full set of filtering rules would have reduced the size of the November 2000 table from 88,700 prefixes to 65,320 prefixes—a decrease of 23,380 prefixes. The final size of the table is comparable to the results in Figure 2.

Perhaps more importantly, the route filters slow the *rate* of growth of the table, as demonstrated in Figure 3(b). Each curve plots the percentage change in the number of prefixes relative to the first data point for December 1997. For example, the top curve shows that the number of prefixes advertised by AS 3561 increased by more than 90% from December 1997 to November 2000. The bottom curve shows that, if all of the filtering rules had been applied, the routing tables would have only grown by 50% over the same time period (from 43,424 to 65,320 prefixes). Prefixes exceeding the RIR guidelines are responsible for a significant portion of the growth in the routing table. Applying the Class A and Class C rules to the December 1997 data removed 17 and 2111 prefixes, respectively. For the November 2000 data, these rules removed 4611 and 14,144 prefixes, respectively. Experiments with other ASes and with the RIPE routing tables show the same trends, as shown in the graphs in Appendix B.

B. Unreachable Addresses

Prefixes that exceed the RIR allocation policies are responsible for about half of the growth in the BGP routing tables we studied. Filtering prefixes based on these allocation policies could help stem the growth of the tables. However, these filtering policies could conceivably cause certain parts of the Internet to become unreachable. In the worst case, any IP address that falls within a filtered prefix may not be reachable from a router that applies the filtering policy. In practice, the router might still be able to direct traffic toward many of these destination IP addresses based on another *covering prefix* with a smaller mask length. For example, the Class A filter removes prefix 24.240.209.0/24 because allocations in the 24.0.0.0/8 region of the address space should not have a mask length of more than 20. The November 2000 routing table from Cable and Wireless contains a route for the prefix 24.240.209.0/24 as well as a covering prefix 24.240.208.0/20 that would not be filtered by this rule. Both routes are originated by the same origin AS. In fact, all of the routes sent to RouteViews for these two prefixes were originated by the same origin AS.

Addresses in the 24.240.209.0/24 prefix could still be reached by forwarding traffic toward the origin AS based on the route to 24.240.208.0/20. The routers in the origin AS would know how to direct the traffic toward the appropriate end-point, based on the more specific 24.240.209.0/24 prefix. Even if the addresses in 24.240.209.0/24 remain reachable, filtering the route might disrupt some other goals that the origin AS (or some other party) is trying to achieve. For example, the origin AS may advertise a separate route for 24.240.209.0/24 to have more control over the flow of traffic. In other cases, a filtered prefix may be covered by a prefix originated by a



(a) Table size (b) Change from initial table

Fig. 3. Table size and growth for Cable & Wireless (AS 3561) from the RouteViews tables. Rules were applied in the order martian, 24, A, B, and C. Results are cumulative, so the curve for B represents the results of applying rule B after martian, 24, and A have been applied.

	prefixes uncovered	addresses uncovered	percent uncovered
A	74	61184	0.005
B	551	1517824	0.136
C	3058	2018304	0.180

(a) Telstra

	prefixes uncovered	addresses uncovered	percent uncovered
A	39	101120	0.009
B	91	1337856	0.121
C	3038	2012160	0.181

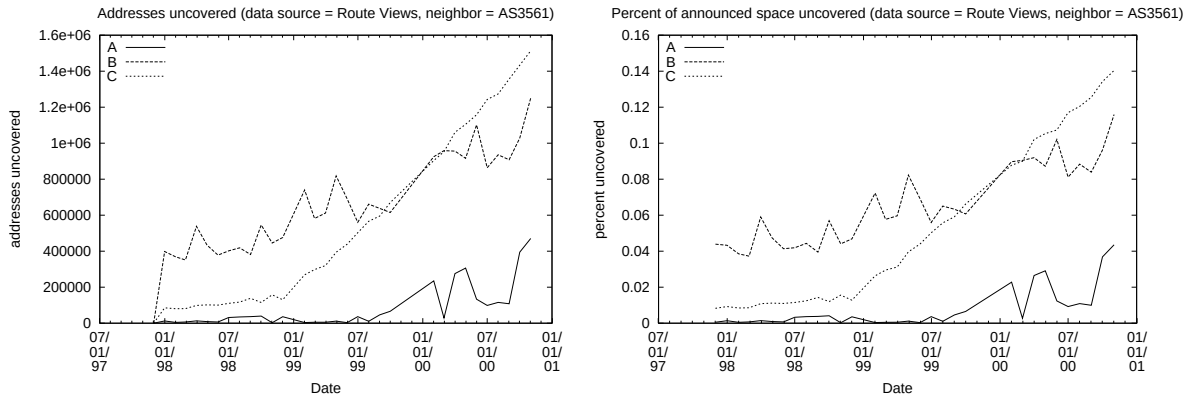
(b) Verio

Fig. 4. Uncovered prefixes for May 16, 2001

different AS. This AS may be the provider of the AS that originates the filtered prefix and, in fact, may have allocated the smaller address block in the first place. Although the provider may be able to reach the customer's prefix, the customer might also want to receive traffic via other upstream providers for load balancing and fault tolerance. Filtering the prefix may disrupt these arrangements.

In the worst case, a filtered prefix may not be covered by *any* prefix. Table 4 summarizes the results for the Telstra and Verio tables from May 16, 2001. Although the Class A filtering rules remove 8118 prefixes from the Telstra table (see Figure 2(a)), all but 74 of them are covered by some other prefix that remains in the routing table (see Figure 4(a)). These uncovered prefixes span a set of 61,184 IP addresses that would not match any entry in the routing table if the Class A filtering rules were applied. Fortunately, these addresses come from a very small proportion of the filtered prefixes and account for only 0.005% of the reachable address space represented in the initial table. Similar results hold for the Verio tables, as shown in Figure 4(b). In total, applying all of the filtering rules leaves about 0.3% of the address space uncovered. Similar results hold for the RouteViews and RIPE data. Figure 5 plots these results for the Cable and Wireless routing data from the RouteViews table over a period of three years. Figure 5(a) shows that the number of uncovered addresses increases over time if the filtering rules were applied. However, the reachable address space in the Internet was also increasing over this time period. Figure 5(b) plots the percentage of address space left uncovered, relative to the reachable address space in the routing table at that that time. The prefixes left uncovered by the Class C rules account for an increasing proportion of the address space.

Attacking potential reachability problems requires a more detailed analysis of the parties responsible for the bulk of the filtered prefixes. Ultimately, more detailed



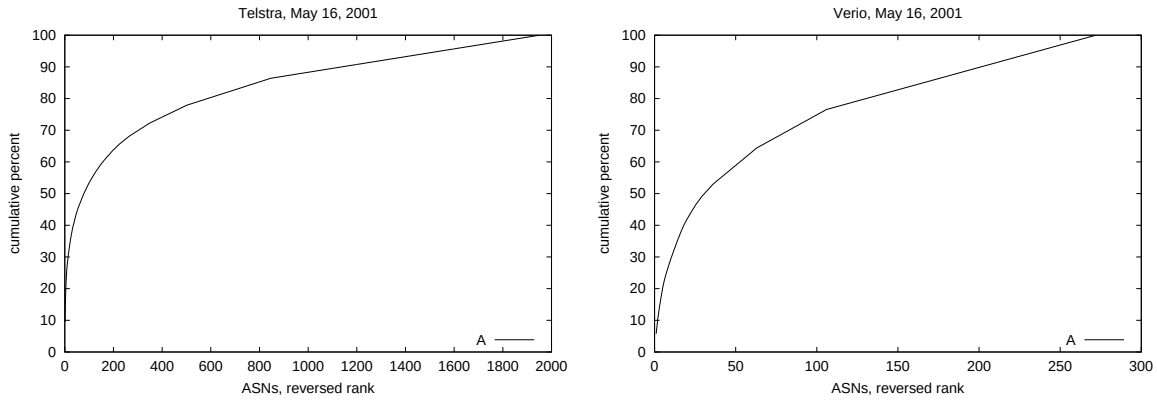
(a) # addresses uncovered (b) % addresses uncovered

Fig. 5. Prefixes uncovered by each rule, for Cable & Wireless (AS 3561) from RouteViews. The percentage in (b) was calculated each day with respect to the address space covered by the intial table of that day.

study is required to identify the parties who allocated these addresses and bear responsibility for ensuring that they are reachable from the rest of the Internet. In some cases, simple configuration changes may solve a potential reachability problems. For example, an AS that originates multiple contiguous prefixes that exceed the RIR allocation policies could advertise a larger block that adheres to the policies. In other cases, a customer and its provider(s) could arrange alternate ways to achieve the load-balancing goals without requiring other parts of the Internet to carry multiple routes to small address blocks. This approach is consistent with the documentation at the RIR Web sites that warn potential users that network operators may choose to filter prefixes that exceed the allocation guidelines. In particular, an AS may choose to carry these routes on behalf of paying customers but not for the customers of their peers.

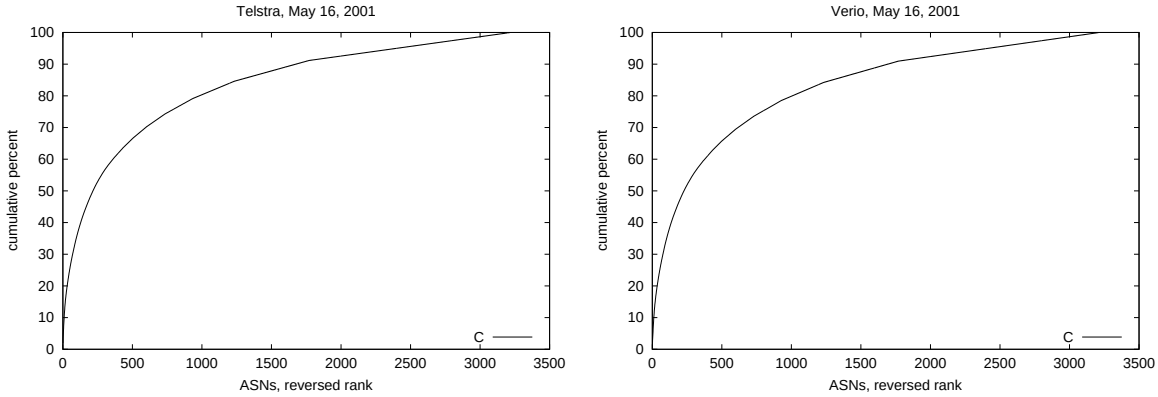
Characterizing the origin of the routes that exceed the RIR guidelines is an important first step. A fairly large number of ASes originate routes that would be filtered by the Class A and Class C rules, as shown in Figure 6 and Figure 7, respectively. In each plot, the x-axis ranks the origin ASes, starting with the AS that originated the largest number of prefixes that exceed the filtering rules. The y-axis plots the percentage of filtered prefixes that were originated by the top x originating ASes. For example, Figure 6(a) shows that the top 20 ASes originate one-third of the prefixes filtered by the Class A rules. However, 1949 origin ASes contribute to the entire set of filtered prefixes, and 425 origin ASes contribute 75% of these prefixes. For both the A and C rules, a relatively large number of origin ASes contribute to the set of filtered prefixes. This suggests that focusing on origin ASes may not be the best way to attack the problem.

Instead, identifying the parties that allocated the small address blocks may be a better approach. A relatively small number of the Class A and Class C rules are responsible for the bulk of the filtered prefixes. The breakdown of prefixes by filtering rule is presented in Figure 8 in Appendix A. For example, there are 22 address blocks in the Class C filtering policy (11 for ARIN, 6 for RIPE, and 5 for APNIC). Three of these address blocks (208.0.0.0/8, 209.0.0.0/8, and 216.0.0.0/8) are responsible for 13,222 of the 16,398 prefixes filtered from the Telstra routing table by the Class C rules. Similar results hold for the Verio data. In fact, the number of prefixes that would be filtered by each of the Class C rules is very similar between the two routing tables. These results suggest that a potentially small number of Local Internet Registries (mostly large Internet Service Providers), and their downstream customers, may be responsible for most of the prefixes that do not adhere to these rules. A more detailed analysis of our data would shed light on the ASes that could influence



(a) Telstra (b) Verio

Fig. 6. ASes responsible for prefixes filtered by the Class A rules



(a) Telstra (b) Verio

Fig. 7. ASes responsible for prefixes filtered by the Class C rules

the advertisement of these prefixes and ensure that their customers do not lose reachability if other parts of the Internet starting applying the filtering policies.

IV. Conclusion

We have shown how BGP table size and the rate of growth of such tables can be reduced by aggressively filtering prefixes to enforce the allocation boundaries documented by the numbering authorities. Rather than advocating one set of rules over another, we have simply demonstrated the effectiveness of filtering based on allocation rules. For this type of filtering to become widespread, the community will have to clarify these allocation rules and arrive at some consensus about how and where they should be applied.

Two main issues require further investigation. First, we need to improve our understanding of whether a filtered prefix becomes "unreachable" or not. One method would be to check for a less-specific covering prefix that has same origin AS. However, this approach may be too conservative. Another approach would be to check if the routes for the filtered prefix and the covering prefix both have the same next-hop AS number. This would guarantee that the next-hop AS has a more-specific route, unless that AS also applies the filtering policies.

Second, some ASes may advertise small blocks of IP addresses for legitimate reasons related to traffic engineering and multi-homing. As suggested in recent NANOG discussions, we need to find ways to allow fine-grain prefixes to be announced in a scoped way. That is, it should

be possible to announce fine-grained routes only to your direct and indirect customers and providers, avoiding leakage to peers along the way. One possibility is to establish a reserved and well known BGP communities to label these routes. Routers could be configured to filter these routes on BGP sessions between peers.

ACKNOWLEDGMENTS

The authors would like to thank the people and institutions who have made their BGP routing tables publicly available. We would also like to thank the people at ARIN, RIPE, and APNIC for making their address allocation policies publicly available on the Web and for answering our questions.

REFERENCES

[1] Y. Rekhter and T. Li, "A Border Gateway Protocol," RFC 1771, IETF, March 1995. http://www.rfc-editor.org/rfc/rfc1771.txt.

[2] B. Halabi, *Internet Routing Architectures*. Cisco Press, 1997.

[3] J. W. Stewart, *BGP4: Inter-Domain Routing in the Internet*. Addison-Wesley, 1998.

[4] Y. Rekhter and T. Li, "An Architecture for IP Address Allocation with CIDR," RFC 1518, IETF, September 1993. http://www.rfc-editor.org/rfc/rfc1518.txt.

[5] V. Fuller, T. Li, J. Y. Yu, and K. Varadhan, "Classless Inter-Domain Routing (CIDR): An Address Assignment and Aggregation Strategy," RFC 1519, IETF, September 1993. http://www.rfc-editor.org/rfc/rfc1519.txt.

[6] E. Chen and J. Stewart, "A Framework for Inter-Domain Route Aggregation," RFC 2519, IETF, February 1999. http://www.rfc-editor.org/rfc/rfc2519.txt.

[7] A. Ahuja and R. Bush, "Effects of Aggregation and Filtering on Routing Table Growth." Internet Draft draft-ptomaine-taxonomy-00.txt, March 2001. http://www.merit.edu/~ahuja/draft-ptomaine-taxonomy-00.txt.

[8] G. Huston, "Analyzing the Internet BGP routing table," *Internet Protocol Journal*, March 2001. http://www.cisco.com/warp/public/759/ipj_4-1/ipj_4-1_bgp.html.

[9] B. Manning, "Documenting Special Use IPv4 Address Blocks," January 2001. http://www.isi.edu/~bmanning/dsua.html.

[10] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear, "Address Allocation for Private Internets," RFC 1918, IETF, February 1996. http://www.rfc-editor.org/rfc/rfc1918.txt.

[11] http://www.apnic.net/db/min-alloc.html.

[12] http://www.arin.net/regserv/IPStats.html#cidr.

[13] http://www.ripe.net/ripe/docs/ripe-211.html.

[14] "University of Oregon RouteViews project." http://www.routeviews.org/.

[15] "BGP tables from RIPE NCC." http://abcoude.ripe.net/ris/.

[16] BGP tables from the University of Oregon RouteViews Project. http://moat.nlanr.net/AS/Data/.

[17] "RIPE Routing Information Service." http://www.ripe.net/ripencc/pub-services/np/ris-index.html.

[18] "Daily BGP table from Telstra." http://www.telstra.net/ops/bgp/bgptable.txt.

APPENDIX

I. GUIDELINES

The guidelines for ARIN are documented in [12], the guidelines for RIPE are documented in [13], and the guidelines for APNIC are documented in [11]. Figure 8 presents these guidelines along with the number of prefixes exceeding them counted in the Verio and Telstra data from May 16, 2001.

II. PLOTS

This appendix presents the results of applying filters to several other neighbors of the RouteViews router. In addition, it shows AT&T data derived from MRT table dumps taken from RIPE NCC.

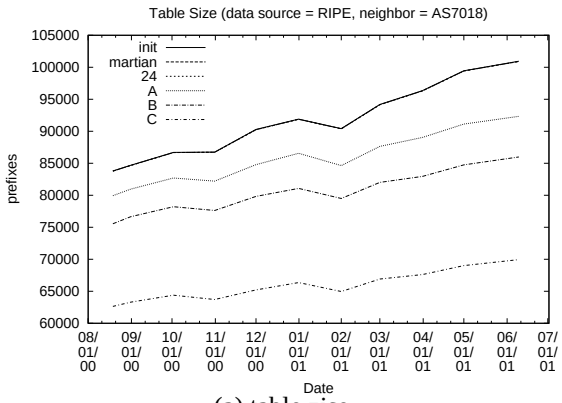
A guidelines

RIR	supernet	mask limit	Telstra filtered	Verio filtered
ARIN	24.0.0.0/8	20	918	19
ARIN	63.0.0.0/8	19	2894	310
ARIN	64.0.0.0/8	20	2543	318
ARIN	65.0.0.0/8	20	863	7
ARIN	66.0.0.0/8	20	645	7
RIPE	62.0.0.0/8	19	224	37
APNIC	61.0.0.0/8	22	31	10

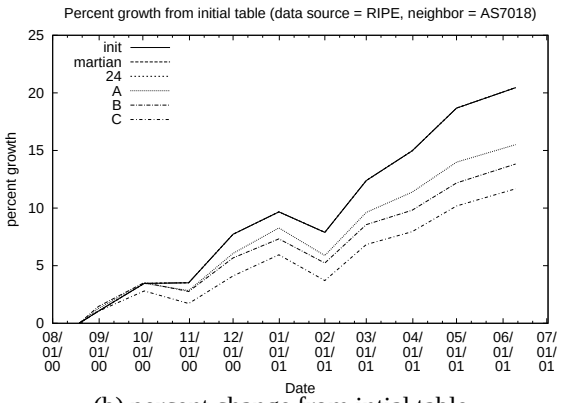
C guidelines

RIR	supernet	mask limit	Telstra filtered	Verio filtered
ARIN	196.0.0.0/8	24	0	0
ARIN	198.0.0.0/8	24	0	0
ARIN	199.0.0.0/8	24	0	0
ARIN	200.0.0.0/8	24	0	0
ARIN	204.0.0.0/8	24	0	0
ARIN	205.0.0.0/8	24	0	0
ARIN	206.0.0.0/8	24	0	0
ARIN	207.0.0.0/8	24	0	0
ARIN	208.0.0.0/8	20	4549	4535
ARIN	209.0.0.0/8	20	4132	4047
ARIN	216.0.0.0/8	20	4541	4526
RIPE	193.0.0.0/8	29	0	0
RIPE	194.0.0.0/8	29	0	0
RIPE	195.0.0.0/8	29	0	0
RIPE	212.0.0.0/8	19	1238	1238
RIPE	213.0.0.0/8	19	671	637
RIPE	217.0.0.0/8	20	309	293
APNIC	202.0.0.0/8	24	0	0
APNIC	203.0.0.0/8	24	0	0
APNIC	210.0.0.0/8	22	797	535
APNIC	211.0.0.0/8	23	161	160
APNIC	218.0.0.0/8	20	0	0

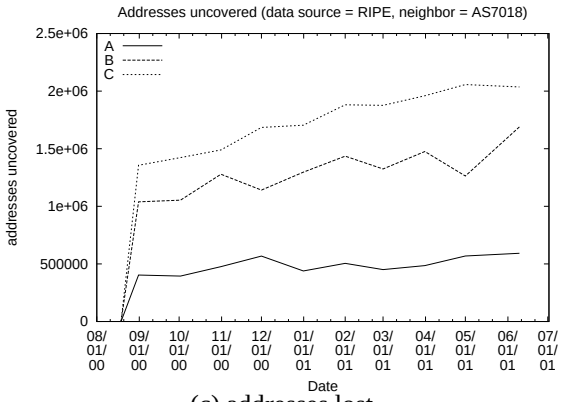
Fig. 8. RIR A and C guidelines and corresponding counts of filtered prefixes for Telstra and Verio, May 16 2001



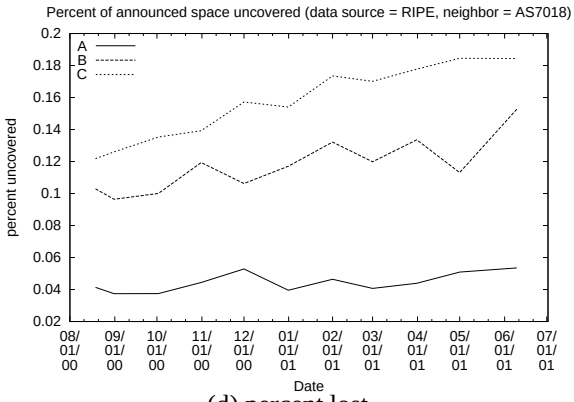
(a) table zise



(b) percent change from intial table

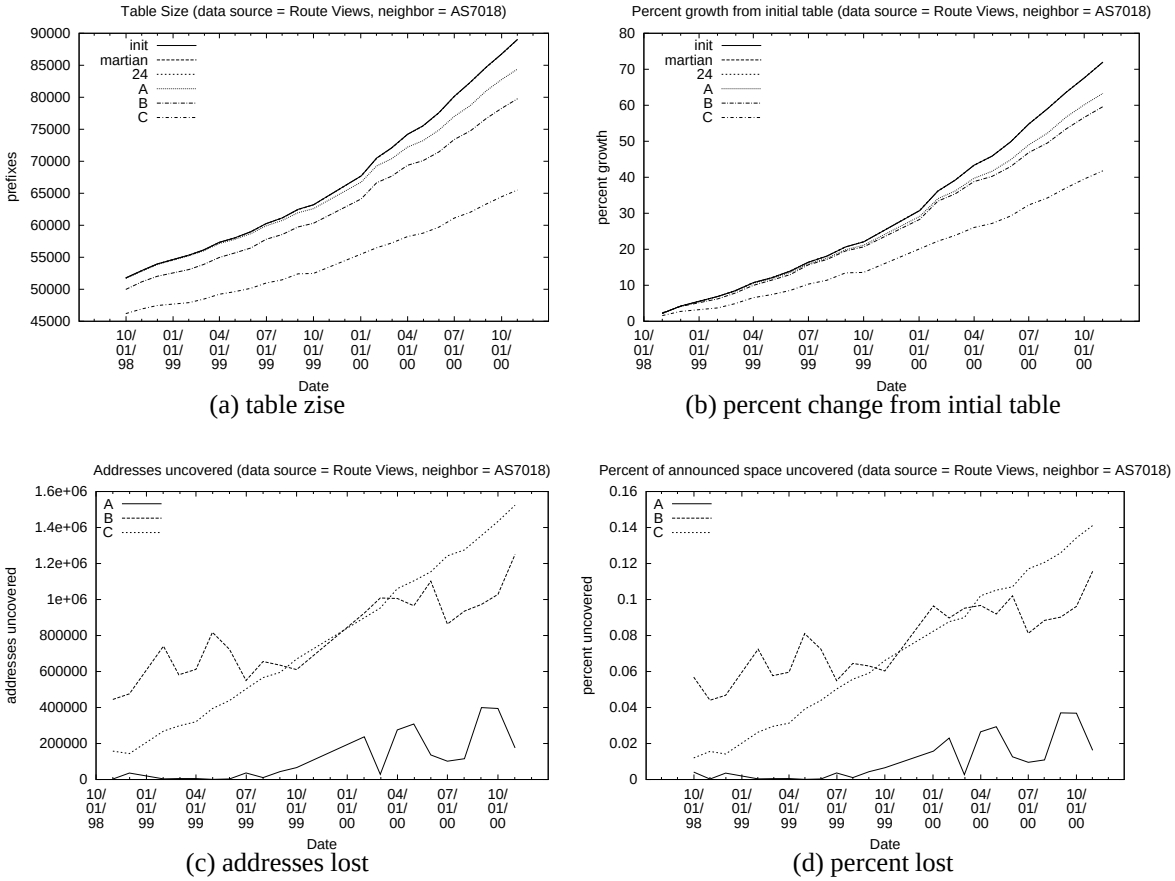


(c) addresses lost



(d) percent lost

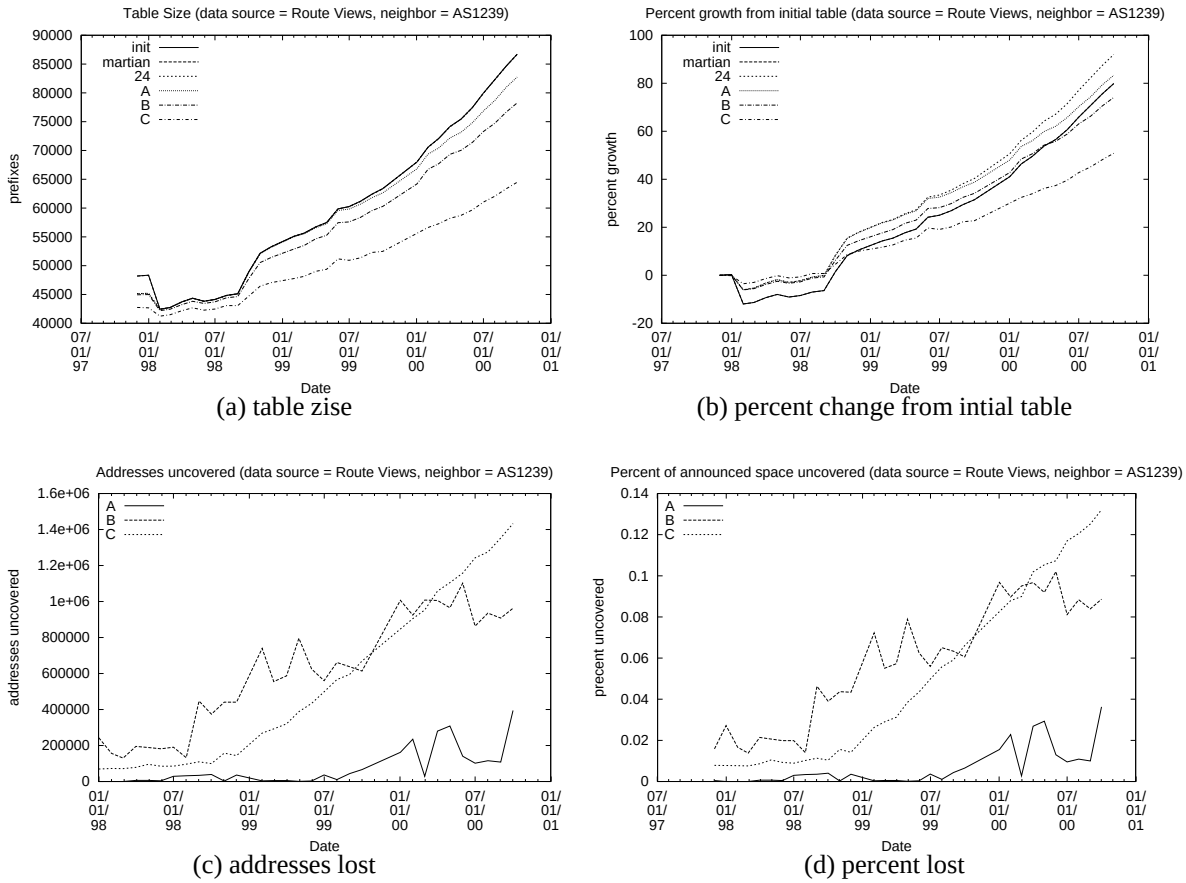
Fig. 9. AT&T (AS 7018) from RIPE NCC



(a) table zise (b) percent change from intial table

(c) addresses lost (d) percent lost

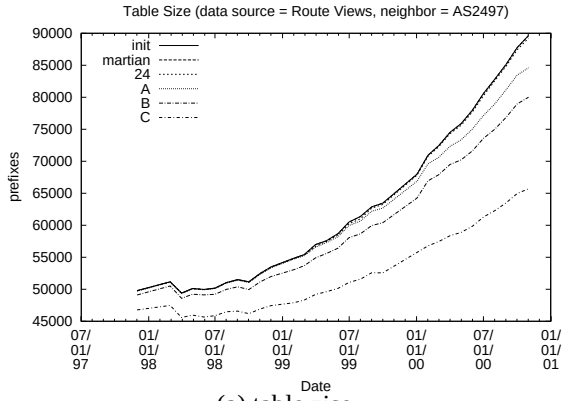
Fig. 10. AT&T (AS 7018) from RouteViews



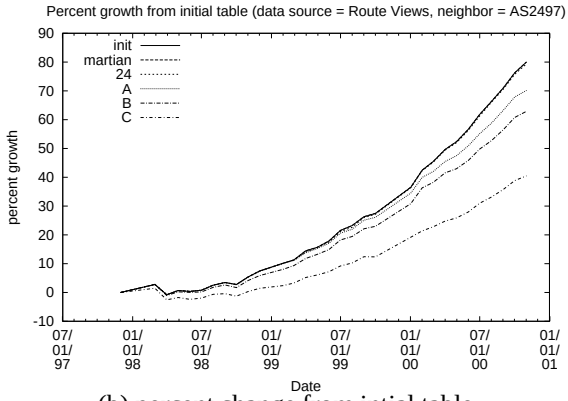
(a) table zise (b) percent change from intial table

(c) addresses lost (d) percent lost

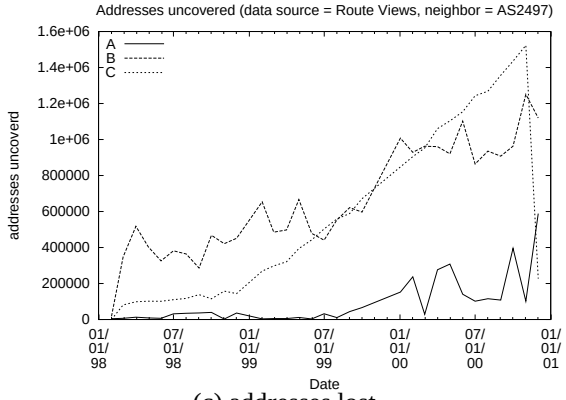
Fig. 11. Sprint (AS 1239) from RouteViews



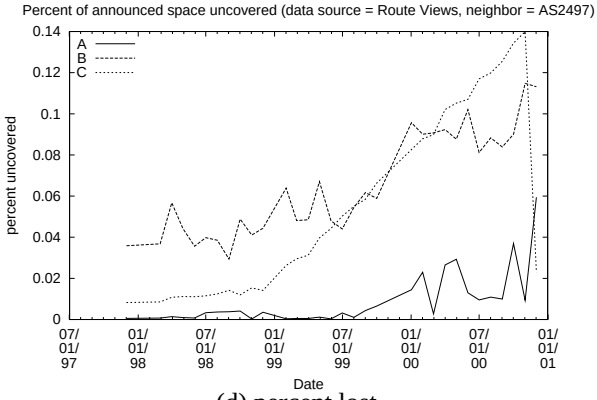
(a) table zise



(b) percent change from intial table



(c) addresses lost



(d) percent lost

Fig. 12. IIJ (AS 2497) from RouteViews