Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2011-038 August 6, 2011

Tragedy of the routing table: An analysis of collective action amongst Internet network operators

Stephen Robert Woodrow

massachusetts institute of technology, cambridge, ma 02139 usa — www.csail.mit.edu

Tragedy of the routing table: An analysis of collective action amongst Internet network operators

by

Stephen Robert Woodrow

B.Sc. Computer Engineering, University of Manitoba (2007)

Submitted to the Engineering Systems Division

and

Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degrees of

Master of Science in Technology and Policy

and

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

Author . . . .

Engineering Systems Division
Department of Electrical Engineering and Computer Science
June 10, 2011

Certified by . . . .

Karen R. Sollins
Principal Research Scientist, Computer Science and Artificial Intelligence Laboratory
Thesis Supervisor

Accepted by . . . .

Leslie A. Kolodziejski
Professor of Electrical Engineering
Chair, Department Committee on Graduate Students

Accepted by . . . .

Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program

Tragedy of the routing table: An analysis of collective action amongst Internet network operators

by

Stephen Robert Woodrow

Submitted to the Engineering Systems Division
and
Department of Electrical Engineering and Computer Science
on June 10, 2011, in partial fulfillment of the
requirements for the degrees of
Master of Science in Technology and Policy
and
Master of Science in Electrical Engineering and Computer Science

Abstract

This thesis analyzes and discusses the effectiveness of social efforts to achieve collective action amongst Internet network operators in order to manage the growth of the Internet routing table. The size and rate of growth of the Internet routing table is an acknowledged challenge impeding the scalability of our BGP interdomain routing architecture. While most of the work towards a solution to this problem has focused on architectural improvements, an effort launched in the 1990s called the CIDR Report attempts to incentivize route aggregation using social forces and norms in the Internet operator community. This thesis analyzes the behavior of Internet network operators in response to the CIDR Report from 1997 to 2011 to determine whether the Report was effective in achieving this goal.

While it is difficult to causally attribute aggregation behavior to appearance on the CIDR report, there is a trend for networks to improve their prefix aggregation following an appearance on the CIDR Report compared to untreated networks. This suggests that the CIDR Report did affect network aggregation behavior, although the routing table continued to grow. This aggregation improvement is most prevalent early in the study period and becomes less apparent as time goes on. Potential causes of the apparent change in efficacy of the Report are discussed and examined using Ostrom's Common Pool Resource framework. The thesis then concludes with a discussion of options for mitigating routing table growth, including the continued use of community forces to better manage the Internet routing table.

Thesis Supervisor: Karen R. Sollins
Title: Principal Research Scientist, Computer Science and Artificial Intelligence Laboratory

Acknowledgments

This thesis marks the culmination of my journey at MIT—arguably not the smoothest (or most direct) of sails. While often challenging, my time at MIT has also been stimulating and broadening, and has caused me to grow in ways that I am not sure I would have needed to or been able to elsewhere. I am grateful for the opportunity I have had to deeply study and grow even more excited about the Internet—a space where the interplay between technology, policy, and business is incredibly rich and interesting. Participating in network operator communities, talking with experts directly, reading papers, writing code, and analyzing data have all contributed to this interest and excitement, and to this thesis. I am also grateful for this opportunity to obtain the CS education that I "missed" during my undergraduate computer engineering career, while simultaneously learning about other ways to think about and address the world, its institutions, and its challenges through TPP.

While probably contributing to what I perceived to be somewhat uncertain seas of my graduate school career, I must ultimately thank my adivsors Karen and Dave for allowing me the flexibility and freedom to spend my time fully on what would eventually and somewhat unexpectedly become my thesis research. You might say I was given enough rope to hang myself, but thankfully, with their help, I managed to avoid such a fate. With this long leash and the problems that inevitably occur when doing things for the first time, I gained an appreciation for what it really means to do research for a living, including the most important first steps of picking a good question to answer and being excited about the work. I also learned how important it is to talk to people about your work, for feedback, criticism, new ideas, or in the case of Internet operators, an understanding of how the Internet actually works, rather than how we might learn about it at the university.

Finally (last but not least), the obligatory round of thanks to and acknowledgment of those who helped and influenced me during this effort:

My first and greatest thanks are to Karen Sollins, my adivsor. She first helped me by hiring me when I was searching for an RA position related to my interests in the Internet in November of 2008, and has been a constant source of encouragement, understanding, and good advice on both my career as a graduate student and beyond since then. Karen also took efforts to involve me in a number of interesting projects and has offered much feedback and helpful perspectives on this work, for which I am greatly appreciative.

I would also like to thank David Clark for his feedback and advice as he became involved in my project as I changed to focus on what would ultimately become this thesis. His experience with and insights about Internet policy and economics together with his technical knowledge was very helpful, as was his suggestion to use Elinor Ostrom's CPR framework to think about the Internet routing table.

Steven Bauer and I first met when I was searching for an RA, but we ultimately became better friends when he returned to MIT to work on another project. I would argue that Steve has helped keep me sane, as he has been a great sounding board, a source of strong feedback, and an advocate, as well as someone to "geek out" with on topics such as Python, TCP, and startups. Thanks for all these things, Steve.

Arthur Berger has been invaluable in helping me plan my investigation and analysis and offering suggestions and assistance in interpreting the resulting data, for which I am immensely grateful. Arthur and I also had several discussions about my professional aspirations that were helpful in my deciding what to do next, and I appreciate his thoughts and advice to this end.

I would like to thank a number of members of the Internet community for volunteering to talk with me about some of the various questions I have had about the Internet. In particular, these individuals' insights and opinions were extremely valuable in understanding how the Internet works *in practice*, which is often quite different from theory and academic viewpoints. Thank you Martin Hannigan, Geoff Huston, Tony Li, Bruce Davie, and Jeff Schiller.

Thanks to Susan Perez, the ANA administrative assistant, for always being incredibly helpful and cheerful. I must acknowledge my officemates in 32-G806 for always providing good company and discussion—especially Jesse Sowell and Chintan Vaishnav—who have offered research advice and discussion when asked, and passionate conversation always.

Thanks are due to my family, and in particular my mother, for their support and encouragement through this process. It started with the drive down to Massachusetts in 2008 and has continued to the present with the Calvin and Hobbes comic strip that has graced my inbox daily since April.

Finally, I acknowledge the financial support of the Intel Corporation, the MIT Communications Futures Program, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Office of Naval Research under contract N00014-08-1-0898.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The autonomous and distributed nature of Internet networks, an intentional consequence of the Internet architecture, can pose difficulty in coordinating networks and their operators to achieve collective action. This observation is visible in the slow progress being made on a number of the challenges that face the Internet today: taking action against spam and malicious behavior, transitioning to a larger Internet address space (IPv6), or adopting more efficient network protocols. These are all cases where the costs and benefits of individual action towards solutions are generally not commensurate, and so no action takes place.

At the same time, the Internet has a relatively strong social network and sense of community amongst its network operators—arguably a historical artifact [MC10]—that can be used to both promote and stigmatize behaviors through social forces such as peer pressure and adherence to norms. This has arguably been successful in achieving economically non-rational behavior on the Internet[1]. While these social forces are probably less strong than they once were when the Internet was more homogeneous, and must now compete with commercial forces stemming from providing Internet service as a business, they are still present as any participant in a network operator community can attest to.

The purpose of this thesis is to investigate the effectiveness of one case where social

[1]Cases of note include Stanford University returning their /8 address block to IANA ("[A]s members of the network community, we need to think about this issue and do the right thing.... It's important for people that have large address space like ours to be good network neighbors." [Mar00]), and more recently, Interop returning their /8 address block to ARIN [ARI10], as well as organizing and participating in mutual aid efforts such as the Conficker Working Group, DDoS mitigation, etc.

forces in the Internet operations community may have played a role in mitigating an Internet collective action problem: the potentially unsustainable growth of the Internet routing table. The objectives of this investigation are twofold. The first is to understand whether the monitoring mechanism that ostensibly spurred the social forces—the CIDR Report—was effective in mitigating this problem. The second is to draw conclusions about approaches to managing the Internet routing table (and possibly inform work on other collective action problems affecting Internet operations and infrastructure) based on lessons learned from analyzing the CIDR Report's efficacy.

1.1 Context and motivation

While invisible to most Internet users, interdomain routing is one of the fundamental architectural elements of the Internet. Essentially the defining characteristic of the Internet as a network of networks, interdomain routing enables every network participating in the Internet to locate and exchange traffic with every other Internet-connected network.

In the face of continued growth, the interdomain routing system used by Internet networks faces scalability limits stemming from its design and operation that can limit the ability of network service providers to build and operate networks efficiently. The size and growth rate of the Internet routing table, and specifically the number of prefixes in the table, is one of the major sources of potential scaling issues [MZF07]. There are a number of causes for this growth, including the addition of new customers and networks in response to increased Internet demand and use, as well as network engineering and the expression of routing policy for existing networks. As shown in Figure 1-1, the Internet routing table has grown at a super-linear, sometimes exponential, rate over time [Hus01].

While the size of the routing table was once a serious concern of network operators [Li11b], engineering efforts have allowed the capabilities of modern Internet routers to scale more quickly than the growth of the routing table for the most part, allowing the Internet to continue to grow organically without concern. Between evolutionary architectural improvements in succeeding generations of Internet routers [McK06] and the near-guaranteed capacity and performance increases in semiconductors (i.e. Moore's Law), most

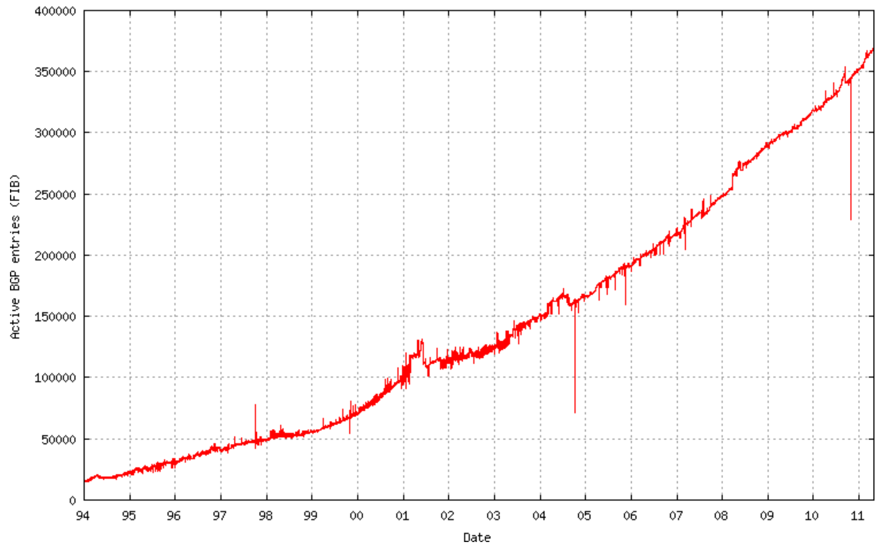


Figure 1-1: Growth of the Internet (DFZ) routing table, 1994-present [Husa].

routers today have capacities that exceed the needs of the Internet routing table. However, these engineering successes have not altered the underlying scaling properties of BGP, the current interdomain routing protocol. While some in the Internet engineering community claim that this is not a pressing concern [Hus11, Hus09], others [Li11b] claim that the lack of any mechanism to control or disincentivize routing table growth means that there is no guarantee that routing table growth will not outpace engineering developments in the future.

At present, table growth causes router vendors to make engineering trade-offs [Li11b, FIRG09] and requires planning and investment consideration by network operators [ZPS10]. While the actual size and growth rate of the routing table has not exceeded the advances provided by Moore's Law, it is exceeding Li's estimate of the constant cost sustainability as shown in Figure 1-2. Further, the related challenges of IPv4 exhaustion, which will likely result in advertisement of more, smaller address blocks due to address repurposing and transfers, and IPv6 adoption, which will drive growth of the IPv6 table[2] on dual-stacked routers, may potentially cause routing table growth to accelerate towards a point where routing becomes more expensive or even unsustainable (i.e. outpacing Moore's Law). In

[2]The IPv6 routing table is still small (approximately 6000 prefixes), but currently appears to be growing exponentially: `http://bgp.potaroo.net/v6/as6447/`

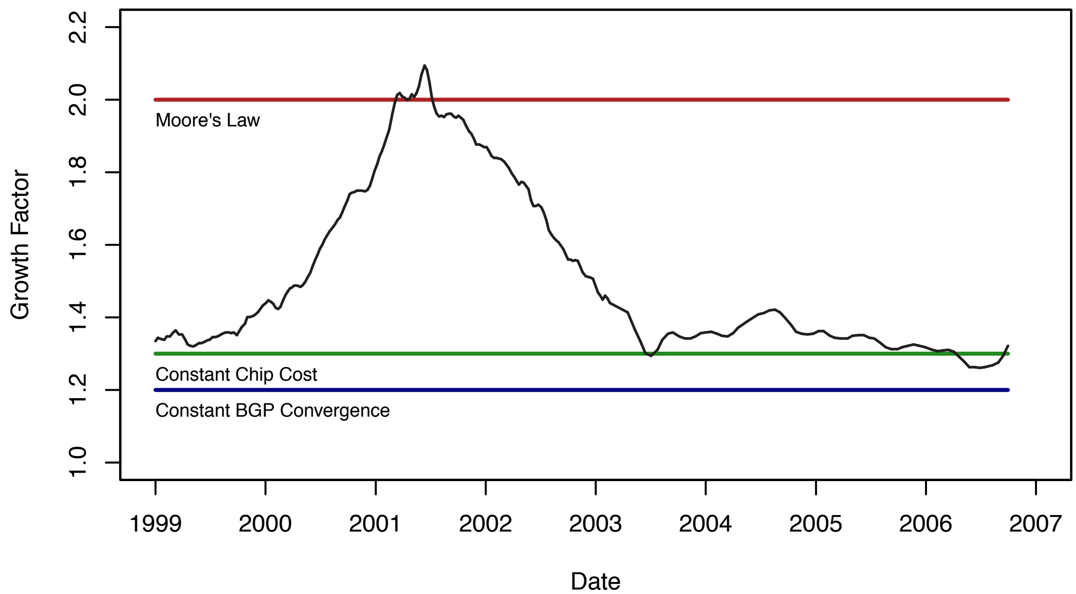


Figure 1-2: Normalized growth factor (smoothed over a 2 year window) of the Internet routing table relative to Moore's Law and other desirable engineering objectives, semiconductor chip cost and routing protocol convergence time, from 1999-2007 [Li06].

the long term, unchecked growth could potentially curtail the decentralized, laissez-faire growth and operation that has been a hallmark of the Internet up to this point, or at the very least cause Internet routing, and thus Internet service, to become more expensive than it presently is.

The specter of routing table scaling problems has been encountered once before in the history of the Internet. In the early 1990s, as the Internet was becoming more commercial and moving from a single hierarchical backbone to multiple backbone providers, the Internet routing table began to grow at a rate that would exceed the capabilities of routers available at the time [Hus01] and in some cases actually did affect the operational behavior of the routers [Li11b]. The solution to this problem was fundamentally a technical one: updating the addressing and routing architecture of the Internet to allow networks to be aggregated, or advertised in larger blocks, thus consuming fewer entries in the routing table. However, the adoption of these technical protocol changes and improvements in operational behavior required to enjoy benefit from these changes was promoted at least in part by social forces within the operator community.

The mechanism that was used by the Internet operations community to promote more efficient route advertisements was called the CIDR Report. Transmitted weekly since the mid-1990s to mailing lists associated with major network operator communities, the report contained a section called the "aggregation report": an ordered list of the thirty networks that could most reduce the number of entries in the Internet routing table by improving their route announcement behavior. This email report was initially seen by network operators as moderately successful in affecting network aggregation behavior, but at some later point it came to be viewed as ineffective[3]. It is interesting that this social mechanism, one that superficially seems to provide only very weak incentives and disincentives, was effective (at least as claimed by a number of network operators and others in the community) in improving on this collective action problem related to technology adoption and efficient route announcement.

1.2 The case of the CIDR Report

This thesis asks the question of whether the social forces in the Internet operations community are capable of inducing collective action. This question is asked in the context of one specific case: the CIDR Report and its ability to control the growth of the Internet routing table. The working hypothesis for this question is that appearing on the CIDR Report did not significantly affect operator behavior. This position is based on discussions with network operators and observation of the continually-increasing number of prefixes in the routing table as shown in Figure 1-1.

The case of the CIDR Report and routing table growth is of interest and particularly well suited for this analysis for a number of reasons. First, this case runs over a long period of time, starting in 1997, so there is a good opportunity to observe community and individual operator behavior. This long duration is backed by a large amount of publicly available data to support the analysis of the report, including the CIDR Report emails themselves as well as archived views of the Internet routing table. The CIDR Report is naturally suited to allow

[3]Operator views on the effectiveness of the CIDR Report come from correspondence with Martin Hannigan, Patrick Gilmore, Tony Li [Li11b], and Geoff Huston [Hus11], as well as [Ste10].

for a quasi-experiment in that only part of the population is “treated” by the CIDR Report, while the rest of the population is available for use as a control. Finally, the CIDR Report was and is a well-known and well-publicized phenomenon within the operator community, and was created with the intent of educating and also socially pressuring operators, and so should be a suitable case for assessing the effects of social forces within the Internet operator community.

Other potential cases involving the social forces within the operator community, such as the back-channel communications mentioned in [MC10], are difficult to study as there is typically a lack of data, a lack of publicity of the events that occur that motivate social pressure, and the events of potential interest for analysis are somewhat ad-hoc and randomly distributed (e.g. the hijacking of YouTube by Pakistan Telecom [BUZ08]).

As with any study of a single case, it is generally not possible to make generalizations about the broader question based the results of the study. Thus, while this thesis is motivated by the potential use of social forces to solve collective action problems of Internet operations, conclusions drawn from my study of the CIDR Report may not be useful in providing insights for other cases. However, any interesting insights about this case may be starting points for further study and exploration in other cases.

1.3 The routing table as a Common Pool Resource

The unconstrained growth of the Internet routing table is often considered a commons problem: individuals derive private benefit from adding entries to the table, but each entry incurs a public cost—a negative externality—for all others that participate in the Internet routing system. The public cost is not necessarily trivial either, with one network operator roughly estimating the marginal cost of a BGP prefix at $6000-$8000 per year [Her08] and a researcher estimating the same figure at $77,000 over the lifetime of the router [Cla10]. This cost is not borne by any one network, but is the estimated cost of the fraction of router resources consumed by one route across all BGP-speaking routers with a full (DFZ) routing table worldwide.

Discussions of the problem [Hus01, Cla10, BBGR01] often invoke Hardin’s [Har68] notion of the tragedy of the commons—that individually rational actors will attempt to maximize consumption of a resource because they privately enjoy the full benefit of this consumption but share the cost of its reduced capacity, making all actors worse off. Common solutions advocated for such problems are the establishment of a central regulator or private property rights. However, the Internet routing table and interdomain routing lacks both such features and has not yet been reduced to tragedy. This is arguably a result, at least in part, of the "property rights" and governance regime of the interdomain routing system. By understanding these characteristics, it may be helpful in understanding how and why the CIDR Report had effect while also offering insights into other approaches to enable more effective management of the routing table.

In contrast to the broad notion of a "commons problem", Ostrom [Ost90] presents the more nuanced concept of common pool resources (CPR) and CPR management problems. Her work has mostly focused on the management of natural resources such as fisheries, forests, and aquifers, but the general elements of the framework are applicable to other cases as well. In Ostrom's model of common pool resources, the *resource system* (such as a fishery) is considered separately from the subtractable *resource units* (in this example, fish) that can be extracted from the resource. The resource system produces some number of units that can be extracted sustainably, and beyond that point extraction causes harm to the system itself. Actors that extract resource units are referred to as *appropriators* (fishers), and actors that take efforts to improve or sustain the resource (such as farming and stocking bodies of water with fish) are *producers*. If the resource system is not naturally occurring, then it must be created or organized by *providers*. In all cases, the essential defining qualities of a CPR are that it is difficult to exclude others from using the resources (there are no private property rights), and that the resource is rivalrous (the use of the resource by one actor precludes its use by another).

The CPR framework can also be mapped relatively cleanly to the case of the Internet routing table. The Internet's interdomain routing system is a CPR resource system provided collectively by every network that maintains a router with a full table of BGP routes. There is not a single global routing table—each provider maintains its own version of the global routing table. However, the value in the routing table is its consistency and uniqueness across all Internet participants in order to allow global reachability. It is technically feasible for one network to exclude others' routes from the routing table, but the utility of the interdomain routing system comes from global reachability, and so it is difficult to exclude routes without potentially reducing the utility of the network's connectivity to the Internet.

Routing table capacity ("slots" or routes) are the resource units of the routing table CPR. Routing table entries are not intrinsically valuable in that they cannot be "extracted" and used elsewhere. However, they are valuable in that they permit interconnection and reachability to the rest of the Internet. It is necessary for each network connected to the Internet to occupy at least one slot in order to make the network available to the Internet, and its often desirable to consume multiple slots for engineering or routing policy reasons. Thus, networks must appropriate routing table slots to participate in the Internet. These slots are rivalrous in that each router has a limited capacity and slots used by one network's route announcements cannot be used by another. As a public system with rival resource units, the interdomain routing system has the hallmark characteristics of a CPR.

Mueller [Mue10] supports a similar view of the routing table as a CPR, arguing that while IP addresses and address blocks could be handled as private property, they are managed as common pool goods to protect the routing table. He suggests that RIR address allocation policy is used to conserve routing table slots by enforcing route aggregation and preventing IP address block fragmentation that might result from reselling address blocks.

Under the CPR model, Ostrom's appropriators are rational individuals who make decisions based on four inputs: the costs of a particular strategy, the benefits of a particular strategy, their internal discount rate (the relative perceived value of future benefits versus present benefits), and their internal norms. What is notable about the appropriators in some CPR cases is that they have learned about the limits of their resource system through trial and error and, unlike the fictional rational grazers of Hardin's commons, communicate with other appropriators to establish institutions that govern the appropriation of units in an effort to ensure sustainability of the resource system. Such institutions work to affect the decision-making algorithm of appropriators by increasing the costs of a particular strategy through sanctions, or establishing norms that are then internalized by the appropriator, which in turn affect perceptions of costs and benefits.

CPR governance institutions—which usually include commitment to acceptable behavior as well as mutual monitoring and sanction mechanisms to limit opportunistic behavior—are not guaranteed to exist when a CPR is provided, or to be effective even when they do exist. However, there is evidence that some form of community-based governance institution has had an effect on the routing table, such as Huston's observation of decreases in routing table size following IETF meetings in the mid-1990s [Hus01, Cla10], or the view of operators that the CIDR Report was effective in its early days. In the case of the routing table, there are no explicitly obvious sanction mechanisms beyond shame, criticism, and peer pressure of the network operator community, which may affect the reputation of network providers and the individuals operating their networks. Internalization of community norms of cooperation and collegiality [Abb00] likely also contributed in the case of the routing table.

In this context, the CIDR Report can be considered a mutual monitoring mechanism that provided information about adherence to norms for the loosely-defined, norm-based governance institution. It was not created or mutually agreed upon by the community, but instead offered to the community by a few individuals as part of the CIDR deployment effort, though its conclusions could be verified by anyone with access to a router. The information provided by the CIDR Report was embraced by some network operators as a basis for invoking sanctions of shame and peer pressure that are sometimes visible on the NANOG mailing list [NANa], and it also affected operator behavior via internal norms[4]. However, as many operators have observed or conceded, the CIDR Report appears to be less effective than it once was.

If we accept the view of the interdomain routing system and Internet routing table as a CPR, and the CIDR Report as a monitoring mechanism for the loosely-defined norms-based governance institution for the routing table, it may be helpful to apply Ostrom's framework and analysis to the problem of managing the routing table. The framework may be instructive in seeking to understand the causes of variation in behavior change induced by the CIDR Report over time that have caused operators to perceive the report as less

[4] In [Li11b], Li notes that Cisco clients sought advice and help to improve their route aggregation behavior after appearing on the CIDR Report.

effective. Further, it may provide insights about other approaches to managing the Internet routing table.

1.4 Contributions of this thesis

This thesis makes a number of contributions to the space of Internet routing table analysis, including:

- A history of the CIDR Report, as presented in Chapter 2.
- A well-documented, open-source implementation of the CIDR Report aggregation report algorithm that utilizes multiple vantage points, as described in Chapter 3 and Appendix A.
- An analysis of the characteristics of the CIDR Report, including the distribution of appearances of networks on the report and the fraction of networks it treats, as presented in Chapter 4.
- An analysis of the effects of appearing on the CIDR Report on the route announcement behavior of individual networks, also presented in Chapter 4.
- Consideration of the Internet routing table as a CPR and the CIDR Report as a monitoring mechanism for the CPR governance institution.

1.5 Roadmap for remaining chapters

The remainder of this thesis proceeds as follows. Chapter 2 presents a broad overview of relevant background information in this space, both for the reader who may be unfamiliar with interdomain routing or the Internet, as well as those who wish to understand some of the finer points that motivate Internet operations and routing table growth. This chapter also contains an overview CIDR Report and a brief history of the events that motivated its creation and evolution over time.

Chapter 3 describes the analytical approach taken to determine whether the CIDR Report was effective. It begins by describing the data sources that were used to conduct the analysis and the preprocessing steps taken against that data. Next, the algorithms and methods used to generate the CIDR Report and its aggregation report are described. Finally, the approach taken to analyze AS behavior over time is presented and described.

Chapter 4 presents the results of this analysis: it first considers both the overall characteristics of the CIDR Report and the networks that appear on it. It then proceeds to a specific analysis of the route announcement behavior of networks following an appearance on the CIDR Report, to determine if the CIDR Report had an effect.

Following the presentation of these analytical results, Chapter 5 discusses the meaning and implications of the results, as well reasons for why the efficacy of the CIDR Report may have changed over time, by considering the routing table growth situation through the lens of Ostrom's common pool resource framework.

Chapter 6 presents a review of related work from the literature and other sources on measurement and analysis of the growth of the Internet routing table, solutions to the scalability challenges facing interdomain routing, and characteristics of the Internet operations community.

Finally, Chapter 7 offers a concluding discussion and recommendations regarding the management of the Internet routing table, as well as how these observations might be used to design institutions to solve other problems facing the Internet, and suggestions for future research directions in this space.

Chapter 2

Background

This chapter presents background information about interdomain routing, the CIDR Report, and Internet coordination and operations that will be useful in providing conceptual and historic context about a number of the aspects of this thesis. The first section, which contains a primer on interdomain routing, is provided as a guide to the reader in order to discuss and articulate a number of the technical features of Internet routing and operations that are crucial to understanding this thesis and which are relied upon extensively in the approach used to answer the thesis question. Advanced readers may wish to skip ahead to the second and third sections, focused on the CIDR Report and Internet coordination respectively, which contain much more specific background information relevant to the this thesis that may not be widely known.

2.1 Routing and Internet operations

Interdomain routing (IDR) is essential to both the concept and operation of the Internet as a "network of networks". IDR enables the multitude of privately owned and operated networks that exist around the world to connect and exchange reachability information, which in turn enables a host connected in one network to contact another host connected to any other network. In the context of today's Internet, IDR is used to establish the topological location of blocks of network-layer identifiers—IP addresses—as well as routes to these blocks. This information allows routers within each network to send traffic destined for other networks in the appropriate "direction" in order to reach that network. The set of possible routes available for traffic from a network will be constrained by connectivity to other networks, as well as the commercial interests of Internet service providers who desire to route traffic to earn revenue or reduce costs.

This section begins with a broad overview of some of the fundamental aspects of IDR and the provision of Internet service from both a technical and a commercial perspective. From this foundation, the specific details of the Internet routing table and the protocol, BGP, used to establish it are discussed. Finally, commonly-used Internet operations activities that utilize BGP and thus affect the routing table are discussed.

2.1.1 Interdomain routing

The interconnection of diverse networks was the fundamental design goal of the Internet [Cla88]. While this goal prompted a number of design decisions that have resulted in the Internet as we know it today, the Internet's system of interdomain routing is one of the key operational components for realizing this goal.

Routing generally refers to the process by which paths to specific destinations are determined, as well as the forwarding of incoming packets along the the previously-determined best path towards the ultimate destination. Each router need not know the entire path to every other destination on the network; the router simply passes the packet along to the next hop that has indicated that it knows the way to the destination. In this context, a route can be considered abstractly as a claim that the router announcing a route knows of a path that should deliver a message to a given destination, and is willing to deliver it to the next closest point in the path that it knows of.

Routing takes place at two scopes in networks today. First, *intradomain* routing takes place within networks, the boundaries of which are typically demarcated by infrastructure ownership or policy (e.g. a business or campus network). Routes are typically established using dynamic interior routing protocols such as OSPF or IS-IS. The goals of interior routing are mainly to establish network-wide reachability efficiently, with little concern about the path taken. The presence of a route simply means that a link is up and can carry traffic.

Operators may adjust link metrics to balance traffic flow across networks, but there are not many other concerns about which parts of the network know about or utilize which routes.

The second scope is between networks, such as the multitude of networks that interconnect to make the Internet. Unlike intradomain routing, *interdomain* routing is focused on exchanging reachability information between various networks that are owned and operated with different policies and objectives. While reachability is still an important concern, other concerns relate to routing policy—the expression of desired behavior for network traffic. Routing policy is typically motivated by business concerns ultimately related to the carriage of traffic. In the context of interdomain routing, the exchange of routing information with another domain indicates a willingness to carry traffic to the route's destination. Thus, control of route announcements is important for business reasons. Routing policy objectives also include selective route advertisement in order to control traffic flow, as well as finer-grained capabilities to govern where traffic enters and exits a given network.

On the Internet, interdomain routing information is exchanged between *autonomous systems* or ASes. These represent domains of consistent routing policy [HB96], and typically consist of a single organization's network, though complex network designs sometimes organize an organization's networks into multiple ASes for various reasons. Each AS is identified by a unique number, an AS Number, assigned to it by the Internet Assigned Numbers Authority (IANA) or the appropriate Regional Internet Registry (RIR). These identifiers are used as short-hand operational handles to refer to networks: MIT holds AS number 3, and AT&T's North American backbone is often referred to as "AS 7018".

With the notions of routes and autonomous systems defined, we can consider that the Internet abstractly as a directed graph with ASes as vertices and edges representing the willingness to carry traffic (according to the advertised routes) for adjacent ASes. Let us now consider what the incentives are for ASes to interconnect, and the types of relationships this forms between networks as a result.

The value that comes from interconnecting networks is a sort of network effect—like other forms of communication, the Internet's value increases as the number of hosts reachable via the Internet increases. It is a long standing tradition on the Internet for pairs of networks that receive mutual benefit from interconnecting to do so without charging the other for the traffic sent. This practice is known as peering, and the pair of connecting networks would be considered to have a peer relationship. Peering is effective when the interconnecting networks derive similar benefits from the interconnection and when they are directly adjacent, but this is not often the case in the Internet. Unlike the case where peers gain benefit from connecting and providing access to each others users/customers, peers do not benefit when they carry traffic between two other peer ASes connected to it—what is known as providing *transit* for those two networks. A different approach is required for networks that do not connect directly but that still wish to reach each other via the Internet.

Generally speaking, there are two ways for an autonomous system to receive routes that enable them to access the rest of the networks that compose the Internet. First, ASes may pay the other networks that they connect with to carry traffic to and from the rest of the Internet on the AS' behalf. Payment makes the carrying of traffic for others beneficial to interconnecting networks in cases where it normally was not in the previous discussion of peering. In these cases, the paying AS is a transit *customer* of their upstream networks. Alternately, an AS operator can build its own network to have sufficient geographic breadth and presence in order to directly connect to other networks. This is a capital-intensive operation and so these networks are usually operated as businesses—network service *providers*—to provide transit service to other networks seeking connectivity to the Internet.

Each autonomous system is free to interconnect and exchange routing information with other networks as they choose, and many have their own unique policies in this regard. However, the generally-accepted logic is that routing policy is a rational decision-making process with the goal of maximizing income and minimizing payment. This can be distilled into a simple set of ordered preferences for selecting routes within a given AS. To reach a given prefix from a given AS, routes are preferred in the following order of availability:

1. Prefer a customer's route, as this allows the AS to charge the customer for this traffic.

2. If a customer route is not available, prefer a peer's route as this allows the AS to send the traffic without charge.

3. If a customer or peer route is not available, use the upstream provider's route. This

is the route of last resort, as the AS must pay its provider for this traffic.

The upstream provider route is usually installed as a "default" route, which means it matches any prefix that the AS does not have specific routing information about, on the assumption that the upstream provider has better connectivity to other networks than the AS. This is a reasonable assumption, especially for smaller networks at the edge of the Internet, and necessarily true for those that obtain Internet connectivity from a single provider.

There is a small set of networks that are sufficiently large that they achieve reachability to all other Internet networks by using customer and peer routes. These networks are large providers that have many customers of their own, and would never be customers of other large networks. This results in peering relationships with their other large counterparts in order to reach the rest of the Internet. These networks do not have default routes, and are thus collectively known as the default-free zone (DFZ). The DFZ represents the "core" of the Internet, and its members have the most complete view of the Internet and play a significant role in Internet routing.

2.1.2 The Border Gateway Protocol

The Border Gateway Protocol (BGP) is the standard interdomain routing protocol that runs between networks to achieve Internet connectivity. The BGP specification defines a protocol for exchanging reachability information between autonomous systems, as well as an algorithm for selecting paths and means for expressing routing policy based on path advertisement and selection.

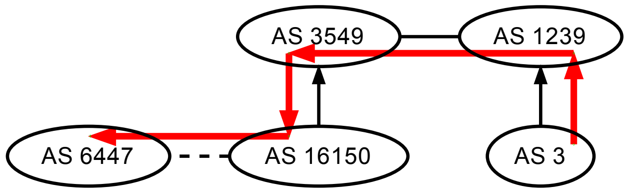
BGP operates as a bilateral session established between two BGP-speaking routers in adjacent autonomous systems. Once the session is established, network layer reachability information—IP address prefixes—that are known to each router is transmitted to the other AS' router, along with attributes corresponding to the prefixes, such as the origin, AS_PATH, or the IP address of the next hop router for packets destined for a prefix. The routes advertised to one's neighbors may either originate from within the AS (for customers and infrastructure) or be from other neighbors that one agrees to provide connectivity for.

As the paths to prefixes change (such as when a link comes up or goes down), or as new prefixes are advertised or become unreachable, update messages are sent advertising or withdrawing routes accordingly. As route updates are received, the router updates its routing information base and may select and advertise a new best path based on this change in information.

The AS_PATH feature of BGP is a distinguishing characteristic of the protocol and one of the most relevant characteristics of the protocol for the purposes of this thesis. It is constructed as announcements are received and re-announced by autonomous systems providing transit to each other. Each AS who announces a route appends their AS number to the AS_PATH of the prefix. This serves to provide loop prevention, as BGP speakers ignore routes with their own AS in their AS_PATH. The AS_PATH also serves as a source of topological information, identifying the AS that first announced the prefix—the origin AS—as well as the vantage point AS and the ASes in between.

An illustrative example of a BGP route announcement and corresponding topology that generates the AS_PATH is shown in Figure 2-1. Looking at the AS_PATH in the update message, we can see that the prefix was originated by MIT (AS 3), observed by Route Views peer Phonera (AS 16150) and transited by MIT's upstream provider Sprint (AS 1239) and intermediate provider Global Crossing (AS 3549). In this case Route Views (AS 6447) is the vantage point and so its AS number is not recorded in the AS_PATH. By observing the AS_PATH information contained in a diverse set of routing tables, such as is visible from Route Views [Rou], a great deal of information about connectivity and topology can be gathered.

Routing policy can be controlled and expressed in a few ways with BGP. First, route advertisements incoming from neighbors may be filtered to ensure that the routes accepted into one's own routing table agree with business and operational practices. Similarly, outgoing route advertisements to neighbors may be filtered to ensure that routing policy is maintained (such as not providing transit to peers). Finally, operators may tag their outbound route announcements with BGP Communities. These are (generally) provider-specific attributes that affect the way providers will treat route announcements they receive, such as prepending the route's AS_PATHs or announcing the route to only part of their net-



(a) An example BGP topology showing AS relationships (black edges) and BGP route announcement progression (red edges).

UPDATE
path attributes=   AS_PATH: 16150 3549 1239 3   NEXT_HOP: 217.75.96.60   ...
prefixes=   18.0.0.0/8   ...

(b) An example route advertisement (UPDATE) message received by AS 6447 reflecting the topology in 2-1(a).

Figure 2-1: A simplified BGP topology and corresponding route announcement.

work. There are a few well-known community attributes that are defined in the RFCs [CTL96], including the NO_EXPORT attribute, which should instruct a peer to not share (export) the prefix with any of its neighbors.

2.1.3 The Internet routing table

The Internet routing table is where global Internet reachability information is collected and maintained by each autonomous system. Each AS will construct their routing table based on connectivity with their neighbors and their own routing policy objectives. The Internet routing table can be constructed in a number of ways, including the manual configuration of routes by a router operator, but it is most commonly constructed by running BGP with one's neighbors to exchange reachability information.

While it's commonly referred to as *the* Internet routing table or the *global* routing table, there is no canonical definition of the routing table. Each AS will have its own unique view of the Internet through its routing table because of that AS' unique position in the topology of the Internet. The AS' own policy requirements, as well as the routing policies of its neighbors will also affect the set of routes that it observes relative to other vantage points on the Internet. However, most ASes will have a roughly similar view of the Internet.

The Internet routing table is implemented as two sets of related data in most routers and as used by most networks. The main table of Internet routing information in an interdomain router is more precisely defined in the BGP specification [RLH06] as the routing information base (RIB). Many routers also maintain a forwarding information base (FIB) as described below. All feasible routes received over all BGP sessions are stored in the router's RIB and remain there unless updated or withdrawn at a later time. The RIB may contain multiple possible routes for a given destination, and so cannot be used directly for forwarding packets. Instead, the BGP path selection algorithm is run over the data in the RIB to select the best path for each prefix. These best paths are then installed in the FIB where they are then used to forward packets to neighboring networks in order to reach their final destination. These best paths are also exported to any peers and customers of the AS, after filtering to enforce routing policy.

The FIB is typically a more compact version of the best paths in the RIB, and resembles a lookup table within each AS' border router that is used to determine the link along which incoming packets should be forwarded in order to reach the destination as specified in their IP packet header. An excerpt illustrating what a routing table might look like is shown in Figure 2-2.

PREFIX	NEXT HOP	INTERFACE
...	...	...
10.0.0.0/8	192.168.10.1	eth0
10.1.0.0/14	172.16.120.200	eth1
10.2.1.0/24	172.16.120.204	eth1
...	...	...

Figure 2-2: A partial example of a routing table showing next IP hop, and physical egress interface for each prefix.

For each incoming packet, the router searches the table for the most specific (longest) prefix that matches the destination address and forwards the packet to the router specified by the address in the "next hop" column of the table. In the example above, a packet destined to 10.1.2.3 would be forwarded to the router at 172.16.120.200 over the 'eth1' interface.

The resources within routers that are used to store the RIB and FIB and execute the BGP route selection algorithm are generally finite and fixed by the design of a given router. The RIB is stored in a router's route processor DRAM, and the route processor CPU executes the BGP path selection algorithm on receiving a BGP update message from a peer. These routes are then installed in the FIB, which is copied to high speed lookup memory (often content-addressable memory) associated with special packet forwarding processors on router linecards [ZPS10]. Given these limits, the scalability of the Internet is constrained to some degree by how its growth affects the size of the Internet routing table. The upper ceiling on router memory limits the absolute growth of the routing table in terms of the number of routes it can maintain. Each entry in a router's routing table memory is often colloquially referred to as a "slot".

While this memory limitation is less of a problem in modern routers with multi-gigabyte memories and multi-million route capacities, this was a significant problem in early routers as the Internet began growing quickly in the early 1990s, with routers failing or behaving unusually as the routing table consumed all available memory [Li11b]. In contrast, router CPU utilization poses less of an absolute limit on routing table growth, but does cause other challenges. As the size of the routing table grows, the amount of data required to be processed for a large routing update or BGP session startup or reset has become significant. It is apparently not unusual for this initial router startup process to take tens of minutes [Ste10] before routes are exchanged with peers and connectivity is fully established. More generally, as the time required for processing routing updates increases, the time for routes to converge after a BGP update will also increase, resulting in poor performance and potential partial loss of Internet connectivity during transient network failures. Concerns over the stability and scalability of Internet routing has made the growth of the routing table and BGP UPDATE dynamics a topic of interest for at least the past two decades.

2.1.4 Routing policy and BGP network operations

At its simplest and most efficient, the Internet routing table should contain one entry per autonomous system. Given that there are approximately 40,000 autonomous systems visible on the Internet today, this would yield a small and compact routing table, similar to that of the mid-late 1990s. However, the routing table cannot be made that simple for a number of reasons. The needs-based allocation policies used by RIRs [HKC+96] to allocate blocks of IP addresses to end-user organizations results in fragmentation of IP address blocks and allocation of non-contiguous blocks to organizations as they receive new addresses over time. Such blocks cannot be announced as a single prefix, and so multiple routing table slots must be consumed to advertise each disjoint block allocated to an organization. Beyond fragmentation, most cases of increased consumption of routing table slots are the result of implementing routing policy for traffic incoming to a particular AS using the mechanisms available in BGP. A number of these mechanisms are discussed in turn below.

Control of routing policy for traffic exiting an AS is relatively straightforward—the AS operator simply needs to express and adjust their preferences for which route of those that they receive from their peers, customers, and providers should be selected for carrying that traffic to the specified prefix. In contrast, managing routing policy for traffic entering an AS is much more difficult because selection of the best path for traffic originating from all other networks occurs independently in each remote network and is dependent on factors such as AS_PATH length that the AS advertising a route lacks control over.

Many of the operational requirements of network service providers today for control over inbound traffic are not directly facilitated by explicit BGP features. Indeed, BGP provides only one "knob" that an AS operator can adjust to influence its adjacent peers' path selection: the relatively limited multi-exit discriminator (MED) attribute [vB02]. To achieve desired routing policy in more advanced situations such as multihoming and traffic engineering, operators must implement routing policy using other aspects of the BGP protocol. By understanding how the BGP path selection algorithm works and assuming that most peers follow the standard algorithm, an operator can implement more effective and fine-grained routing policy.

It is important to note that in general, each of these behaviors allows an AS to obtain private benefit by imposing a public cost against all other participants in the global routing system. To announce a multihomed or traffic engineered network, which provides benefit to the operator initiating that behavior by allowing them to better manage their network, they necessarily add extra routes to the routing table of every network operator that collects a full routing table from their peers.

Multihoming Multihoming refers to the connection of a network to the Internet via more than one upstream provider. This is typically motivated by a desire for reliability, such that one will have network connectivity even in the event of a failure or misconfiguration on the part of an upstream provider. A basic approach to establishing connectivity for a multihomed AS would be to announce all of the AS' prefixes to all of the AS' upstream providers (sometimes referred to as anycasting). Depending on one's view of the Internet, a slot in RIB may be consumed for each of these announcements (i.e. for each of the upstream providers).

This will achieve the goal of multihoming, but may result in an imbalance of traffic between upstream links as distant ASes select the path to take based on factors that the origin AS likely does not have control over, such as the AS_PATH length. This can pose a problem if the AS' primary link is inexpensive and the backup link is expensive. The typical approach used to solve this problem is called AS_PATH prepending, where an AS will artificially lengthen their AS_PATH as observed by certain (e.g. high cost) providers in order to make specific links less attractive for incoming traffic. Other issues regarding balance over links may arise, and these are best handled by the general class of behavior called traffic engineering.

Traffic engineering Traffic engineering (TE) refers to balancing traffic across network links to avoid overloading particular links (resulting in high latency and potentially loss) and allowing headroom in a given link's utilization to allow for bursty traffic. TE is easy to perform for traffic leaving an AS with multiple upstream links, as the AS operator has control over their interior routing protocol metrics and their border routers themselves, and can thus direct traffic as they wish. Inbound TE is more difficult, again because of the lack of knobs that an operator has to select the ingress point for inbound traffic into their network.

While AS_PATH prepending and the MED can both be used effectively for inbound traffic engineering sometimes, the ultimate tool for fine-grained TE is the announcement of more-specific prefixes. Because the BGP path selection algorithm always prefers more specific prefixes, a network operator can distribute high-traffic destinations within their network across multiple prefixes or subnets within their prefix and announce these more-specific prefixes separately to their various peers and upstream providers in order to spread incoming traffic across multiple incoming links. Each of these more-specific prefixes generally cannot be aggregated with the less specific-covering prefix because the routing policy differs (the AS_PATHs will likely be different because of the selection of different upstream providers for different prefixes).

Prefix hijacking prevention Prefix hijacking is the announcement of an IP address block by a network other than the legitimate owner of an address block. This can sometimes result from configuration errors [BUZ08] but can also be an intentional attack against a network [PK08]. These attacks are particularly effective because routes with the longest matching prefix are selected, allowing the most specific prefix to "win" the traffic, even if its path is less optimal. To guard against the hijacking of address blocks, particularly for critical infrastructure such as authoritative name servers, some providers announce parts of their address blocks with the most specific prefix that they expect will be successfully pass global route filtering policy. This typically leads to the advertisement of /24 blocks for critical infrastructure[1].

2.2 CIDR-ization of the Internet and the CIDR Report

This section offers a mainly-historic background on an effort called classless interdomain routing (CIDR) to allow more efficient announcement of address blocks in the Internet's interdomain routing system. This section also provides background and history on the CIDR Report, a social mechanism used to encourage network operators to aggregate their route advertisements using CIDR.

[1] As a current example that is accurate as of the date of submission of this thesis, Google.com locates its authoritative DNS nameservers (e.g. ns1.google.com, which resolves to 216.239.32.10) in 216.239.32.0/24 and other /24 blocks, advertised separately in addition to the covering 216.239.32.0/19 block allocated by ARIN.

2.2.1 Classless interdomain routing

In contrast to individual end hosts connected to an IP network, which are fully identified by an IP address, there is also the need to refer to *networks*—collections of hosts connected together in such a way that they are all reached by the same path. In other words, networks are blocks of IP addresses that have the same routing policy and are thus reached in the same way. An IP address contains both the network address and the subnetwork address of that machine on the given network. The differentiation between the network and the subnetwork is given by partitioning the 32-bit IP address into a network part and subnet part, as shown in Figure 2-3. Networks are a useful abstraction in interdomain routing in that they allow routers to maintain a relatively small amount of state to reach a potentially large number of IP addresses, rather than maintaining routing information separately for each Internet-connected host.

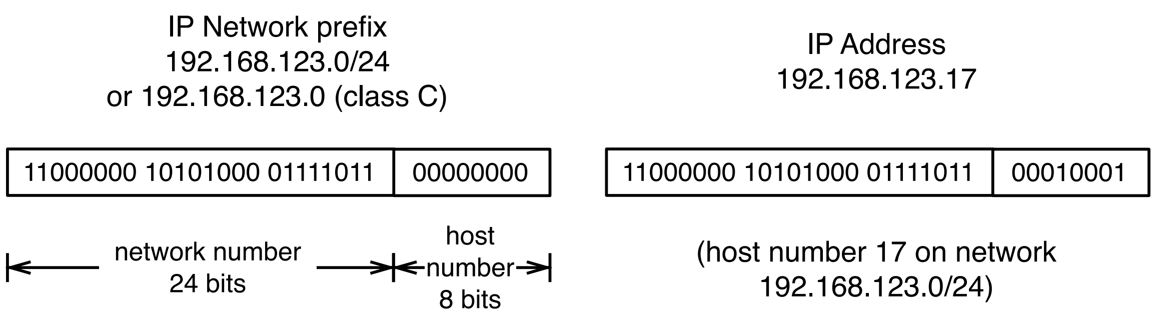


Figure 2-3: An example of an IP network prefix and full IP address, illustrating the distinction between the network number and host number components of the address. Note that the prefix length (e.g. /24) indicates the position of the partition between the network number and the host number, and thus the number of bits allocated to each.

The original specification of the Internet Protocol implicitly encoded the size of the network in the first octet of the network address itself. Three classes were available: class A, for large networks (2^{24} hosts), class B, for mid-sized networks (2^{16} hosts), and class C for small networks (2^{8} hosts). Networks with the first octets shown in Table 2.1 were assumed to be members of the class and thus be of that size. The distribution between class A, B, and C networks was specified arbitrarily by the protocol developers.

Assignment to permit use of these addresses by organizations was performed by the IANA, and addresses were allocated on a basis of justified need. Large organizations (e.g.

CLASS	ADDRESS RANGE	HOSTS/NETWORK	NETWORKS/CLASS
A	0.0.0.0–127.0.0.0	2^{24}	2^7 (128)
B	128.0.0.0–191.255.0.0	2^{16}	2^{14} (16384)
C	192.0.0.0–223.255.255.0	2^8	2^{21} (2097152)
D	224.0.0.0–239.255.255.255	multicast (224/4)	
E	240.0.0.0–255.255.255.255	experimental (240/4)	

Table 2.1: Classful IP addressing architecture [Pos81]

MIT, General Electric, etc.) typically planned to, or were assumed to have large networks and many hosts, and so were granted class A addresses. Organizations that were somewhere in the middle would be granted a class B, and organizations expecting to have less than 254 hosts would get a class C. As networks grew, they would often be assigned new blocks rather than trading up to a larger block size—this was particularly frequent in the class C space. This issue of granting new class C blocks was also exacerbated by the realization that the class B block was the most commonly needed network size, yet was under-allocated compared to class C networks. As ASes with class C blocks needed to grow, they were allocated multiple class C blocks to conserve class B blocks.

Like all address blocks, each block allocated to a given AS needed to be announced to the Internet via the interdomain routing protocol, EGP and later BGP, in order to exchange traffic with other networks. Thus, each block occupied a slot in the routing table. As time wore on, particularly with the commercialization of the Internet and the transition from a single backbone to multiple backbones, the routing table began to grow to the point that it was causing problems with DFZ provider routers, particularly with the RIB consuming all available DRAM in their routers, resulting in router failures and abnormal routing behavior [Li11b].

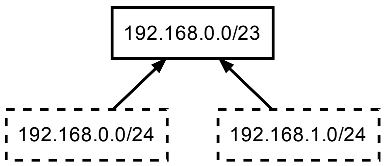
The solution that was ultimately proposed to deal with this was route aggregation—the announcement of a single route that only occupied one routing table slot but covered multiple network blocks. Originally called supernetting [FLYV92], this was implemented by explicitly specifying the size of the network that was to be announced, and relaxing the previously hard boundaries specified for class A, B, and C networks. The deprecation of network address classes resulted in the final name of this effort, Classless Interdomain Routing, or CIDR [FLYV93]. Under CIDR, networks of any size could be announced, and these networks were now specified as a prefix and an explicit prefix length, typically specified in the format 192.168.0.0/16, where 16 is the prefix length in this example. The prefix consists of the most significant bits of the network address that specify the network itself, and the prefix length indicates the position of the partition between the network number and the subnetwork, implying the size of the subnetwork.

CIDR enabled aggregation of prefixes by combining multiple adjacent prefixes into larger networks. Route aggregation is the announcement of routes in a way that provides the same reachability as before aggregation while requiring fewer route announcements to do so. Take for example the case of announcing the block of addresses from 192.168.0.0–192.168.255.255, the equivalent of a class B network. This could be announced with a single route (192.168.0.0/16), two routes (192.168.0.0/17, 192.168.128.0/17) or even 256 (class C) routes (192.168.0.0/24, 192.168.1.0/24, . . . , 192.168.255.0/24). If these multiple prefixes are all originated by the same AS and carried to the greater Internet by the same providers, then they provide the same connectivity and routing policy while consuming 1, 2, or 256 slots in the routing table. Thus, it is most efficient to announce address blocks as aggregated as possible.

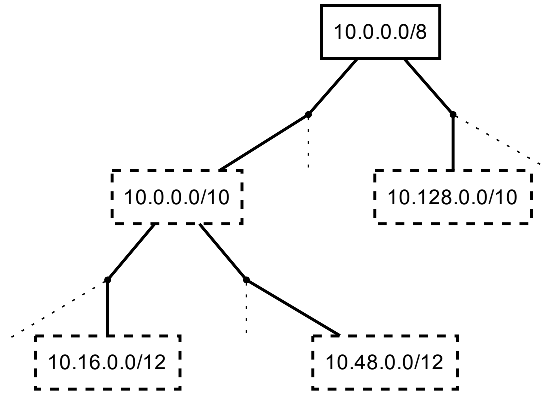
Determining how to announce blocks in aggregate can be done one of two ways. Adjacent, non-overlapping blocks issued by an RIR, such as the classful blocks allocated before CIDR, can be aggregated into a more specific covering prefix that is a synthetic announcement not corresponding to an RIR allocation. In this case, two blocks are considered adjacent if the network numbers of the two blocks are equal except for the least significant bit. The other approach is to announce a less specific covering prefix that corresponds to all, or a larger part of, the classless block of addresses allocated by the RIR. In both cases, the more specific prefixes can then be withdrawn, achieving a net reduction in prefixes announced. An example of each of these approaches is shown in Figure 2-4.

With the specification of classless network addresses codified in the CIDR standard, the Border Gateway Protocol was updated to support this new method of representing networks in the context of network reachability information. The new version of the protocol, version 4 (sometimes referred to as BGP4), was specified in 1994 [RL94].

The deployment of BGP4 required software and sometimes hardware upgrades to routers.



(a) Aggregation by synthesizing adjacent blocks into a less-specific covering block and withdrawing the more-specific prefixes (dashed boxes).



(b) Aggregation by withdrawing more-specific prefixes (dashed boxes) covered by a less-specific prefix (solid box).

Figure 2-4: Aggregation approaches under CIDR addressing

BGP4-speaking routers were not compatible with BGP3-speaking routers, as the previous version of the protocol did not support CIDR. This required some routers to speak both BGP4 and BGP3 to respective neighbors during the transition period to BGP4. Nevertheless, the transition was relatively rapid in terms of "Internet time", taking about 2 years [Li11b], and was strongly motivated by the relief in stress from routing table growth that many network operators had been facing. According to [Tra95]:

> BGP-4 was rushed into production use on the Internet because of the exponential growth of routing tables and the increase of memory and CPU utilization required by BGP. As such, migration issues that normally would have stalled deployment were cast aside in favor of pragmatic and intelligent deployment of BGP-4 by network operators.

Even with the deployment of BGP4-speaking routers, aggregation did not automatically occur by the agency of software running on the routers. Instead, network operators needed to determine that their assigned address blocks were aggregable and then announce these aggregates "by hand". This did not always occur, especially for operators and a community that was used to speaking in terms of classful addresses. Thus, even with the necessary

```
Aggregation Summary
The algorithm used in this report proposes aggregation only
when there is a precise match using the AS path, so as
to preserve traffic transit policies. Aggregation is also
proposed across non-advertised address space ('holes').

 --- 12Nov10 ---
ASnum     NetsNow NetsAggr  NetGain   % Gain   Description

Table      340755   208585   132170    38.8%   All ASes

AS6389       3751      407     3344    89.1%   BELLSOUTH-NET-BLK -
                                                BellSouth.net Inc.
AS4323       4556     1679     2877    63.1%   TWTC - tw telecom holdings,
                                                inc.
AS6503       2001      433     1568    78.4%   Axtel, S.A.B. de C.V.
AS19262      1780      316     1464    82.2%   VZGNI-TRANSIT - Verizon Online
                                                LLC
AS4766       1728      575     1153    66.7%   KIXS-AS-KR Korea Telecom
```

Figure 2-5: An excerpt from the 12 November 2010 CIDR Report

condition of BGP4 routers deployed, the potential savings through aggregation was not realized until network operators acted explicitly to aggregate their route advertisements and the reduce number of routing table slots they were consuming. It was this realization that motivated the creation of the CIDR Report in the mid-1990s.

2.2.2 The CIDR Report

The CIDR Report presents a summary of interesting routing table behavior related to routing table growth. The report contains a number of sections which have varied slightly over time, but it has always contained an overall summary of the size of the routing table over the past week, as well as the *aggregation report*, which is of greatest interest for this thesis. The CIDR Report is published via email every Friday to the major mailing lists that network operators participate in, including NANOG and similar lists in regions outside of North America. A recent CIDR Report is shown in Appendix B, with an excerpt of the aggregation report shown below in Figure 2-5.

With each week's CIDR Report, the aggregation report identifies the 30 ASes announcing the most aggregable routes and thus consuming the most slots in the global routing table that are unnecessary for the expression of routing policy. For the purposes of the aggregation report, an aggregable route is a route whose withdrawal would not cause a change in routing policy from the CIDR Report's BGP vantage point.

Elements of the CIDR Report, and in particular the aggregation report, date back to approximately 1994-1995 when it was implemented by Cisco Systems employee Tony Bates. Bates was a member of the IETF's CIDR Deployment (CIDRD) working group, and the report was conceived to educate and motivate networks about the the transition to from classful to CIDR-aggregated route announcements with the advent of BGP4. The earliest reports available on the Internet date back to 1994, when it was referred to as the "Top 10" report[2].

In mid-1996 the CIDRD group wound down[3], and Tony transitioned the report to the NANOG mailing list. The first report still accessible on the Internet, collected from the NANOG mailing list archives [NANa] dates from September 1996, when the report was briefly called the "Top-50" report before being renamed as the "CIDR Report". The report as originally conceived and implemented by Tony Bates was used until 23 August 2002. Geoff Huston then took responsibility for the report on 30 August 2002, producing a similarly-formatted report but with a new implementation that was developed without consulting Bates' source code [Hus11]. The CIDR Report's vantage points over time are in the table below[4]. The CIDR Report is still published every Friday (North American time zones) and has maintained a remarkably similar format over the 14 years where it can be observed on the NANOG mailing list.

AS number	AS Name	Date effective
AS 5413	Xara.net (at MAE-East)	17 September 1996
AS 5413	GX Networks	15 June 2001
AS 6447	Route Views	30 August 2002
AS 4637	REACH	11 October 2002
AS 2.0[5]	APNIC R&D	20 July 2007

Table 2.2: CIDR Report vantage point ASes over time

[2] `ftp://ftp.ietf.org/ietf-online-proceedings/94jul/area.and.wg.reports/ops/cidrd/cidrd.bates.slides.ps`

[3] `ftp://ftp.ietf.org/ietf/cidrd/cidrd-minutes-96jun.txt`

[4] The potential importance of vantage point selection is discussed further in Chapter 5.

[5] AS 2.0 is not visible in the routing table, but peers with AS 4777 and AS 4608 [Hus11].

2.2.3 The aggregation report

The purpose of the aggregation report is to identify ASes announcing routes that could be more efficiently announced via aggregated route announcements without altering expressed routing policy. In the specific context of the CIDR Report, routing policy is simplified to mean inter-AS connectivity, captured via the AS_PATH for a given prefix. This is important for multihoming and some forms of traffic engineering, where its necessary to announce more specific prefixes in addition to the covering prefix to allow traffic to be spread across multiple upstream providers instead of just taking the best route to the covering prefix. Without considering the fact that there is variation in one or more upstream ASes in the AS_PATH, these prefixes would be considered deaggregated even though they must be announced this way to achieve the desired routing policy.

Given this desire to respect routing policy, the aggregation report considers a prefix aggregable with another prefix if and only if the AS_PATHs match exactly. Aggregation is performed via the two approaches described earlier. The aggregation report also apparently aggregates across an adjacent "hole", or unannounced block, if there is no prefix below it. While this generally makes sense, it is potentially over-optimistic in that it does not consider RIR block allocations and the holes that may result from a block being allocated but unadvertised[6].

The number of advertised, aggregated, and withdrawn prefixes are totaled against the AS originating each of the prefixes, and these figures are exported to generate the CIDR Report. An example of the slightly more general CIDR Report that is emailed to network operators was shown earlier in figure 2-5, and an excerpt of the more detailed CIDR Report available on the web is shown in Figure 2-6.

In these figures, 'current' or 'netsnow' refers to the total number of prefixes currently announced by the AS. 'withdrawn' refers to the number of prefixes that were able to be withdrawn after perfect aggregation. 'aggregated' refers to synthesized aggregate prefixes created by covering two adjacent prefixes. 'announced' or 'netsaggr' refers to total number of prefixes announced by the AS after maximum aggregation, and 'reduced' or 'netgain'

[6]Unadvertised address blocks may be found in cases where a provider wishes to advertise internal infrastructure using globally unique addresses but not offer it as accessible to the public Internet

```
Aggregation Report: Aggregation using AS prepended PATH

Report prepared at Sat, 13 Nov 2010 04:11:55 UTC+1000, using data obtained within AS0.6447

The report may include routes internal to AS0.6447, and may also include routes
that are accepted from adjacent AS's and marked ``NO EXPORT''. The report also
does not take into account conditions local to each origin AS in terms of policy
or traffic engineering requirements. As an aggregation guide, this report is a
very approximate guide at best.

AS list, ordered by net reduction in advertisements

AS           AS Name                                      Current  Wthdw  Aggte  Annce Redctn       %
             Routing Table                                 347822 174358  29764 203228 144594  41.57%
AS6389       BELLSOUTH-NET-BLK - BellSouth.net Inc.          3755   3526     53    282   3473  92.49%
AS4323       TWTC - tw telecom holdings, inc.                4555   3501    461   1515   3040  66.74%
AS19262      VZGNI-TRANSIT - Verizon Online LLC              1782   1643    144    283   1499  84.12%
AS4538       ERX-CERNET-BKB China Education and Research N   1670   1428     44    286   1384  82.87%
AS6503       Axtel, S.A.B. de C.V.                           2003   1534    258    727   1276  63.70%
AS4766       KIXS-AS-KR Korea Telecom                        1866   1315    122    673   1193  63.93%
```

Figure 2-6: An excerpt from the detailed 12 November 2011 CIDR Report

refers to the total number of prefixes saved by the aggregation process; in other words, reduced is current less announced. ASes appear on the CIDR Report in order by decreasing netgain.

2.3 Coordinating Internet operations

A hallmark of the Internet, especially in contrast to comparable telephony networks, is its lack of regulation or centralized management. The very name of the fundamental entities in interdomain routing—*autonomous* systems—suggest that networks are independent and free to do as they please. Indeed, the Internet is a system where network operators, the individuals and organizations that control infrastructure, can "vote with their feet" without external intervention. Despite this autonomy, however, there is also a need for coordination in key areas in order to facilitate interoperability and globally unique identifiers—collective benefits that make the Internet better for all.

Internet coordination was performed relatively casually in the early days of the Internet as a DARPA project, and a number of these roles have evolved into more formal organizations. Formal coordination of unique and high-value resources—the Internet address space and the domain name system—has been institutionalized in IANA and the Regional Internet Registries, and the Internet Corporation for Assigned Names and Numbers (ICANN), respectively. These organizations have established policy processes and engage stakeholders in an effort to govern the resources they are responsible for in an effective and legitimate way. Stakeholders in this space include governments, domain name registry and registrar operators, trademark owners, Internet businesses and service providers, policy advocates, and consumers.

Slightly less formal are the organizations responsible for technical coordination of the Internet. The organizations involved in this space, which include the Internet Architecture Board (IAB) and the Internet Engineering Task Force (IETF), are charged with developing technical standards for the Internet in a fair and open way, in order to improve interoperability and technical coordination amongst network operators and equipment vendors. The IETF is responsible for the well known "Request for Comments" (RFC) series of Internet drafts and standards. While once the domain of researchers and network operators, the IETF participants have grown to include representatives from network hardware and software vendors who now play a large role in the IETF [Li11b].

The least formal level of Internet coordination occurs at the operations level. This refers to the coordination between interconnecting service providers that is necessary in order to implement that interconnection, as well as to solve other problems or work towards developments that are only manifest in operational Internet infrastructure. Most interconnection and bilateral dealing between ISPs is organized through contracts and business relationships (though some activities also take place through more informal channels mediated by personal identity and reputation), whereas collective coordination and operational problem solving typically occurs within Internet operator communities.

2.3.1 Internet operator communities

The North American Network Operators' Group (NANOG) is the canonical example of an Internet operator community. Emerging out of a previous group of NSFNET operators called "Regional Techs", the group holds large (400-600 person) meetings three times a year in cities across the United States and Canada for operators to meet and share knowledge ("clue") while also building business and personal relationships. In addition, and perhaps more significant, is the NANOG mailing list, which is composed of approximately

10,000 participants[7] and is commonly used for making contact with other operators about operational issues, or discussing technical problems. While not all North American networks participate in this community, it is popular and common amongst medium and large Internet networks. The norms and attitudes of NANOG and the individuals that participate in it are generally cooperative and meritocratic [MC10], in keeping with the ethos of the early Internet [Abb00].

NANOG is nominally limited to the North American continent, but it is also the pre-eminent operator mailing list and community worldwide, and so attracts participation from major operators and clueful individuals from organizations in other regions. There are also other geographically bounded communities that serve a similar role to NANOG in other regions. Many of these are coupled together with the RIR role in each locale. This suggests that shared language or culture, and the ability to meet in person are important for the activities of network operator groups, even if most participate only via mailing lists.

[7] The most recent statistics are from December 2009: `http://nanog.org/mailinglist/liststats/2009stats.php?mon=dec`

Chapter 3

Analyzing the CIDR Report

This chapter describes the methods employed to analyze the effectiveness of the CIDR Report. The chapter begins with a high-level overview of the analytical approach to impart a conceptual understanding of this thesis' analytical methods to the reader. Following this, the sources of data used to perform the analysis and the process used the gather the data are described. Steps taken to preprocess the collected data to bring it into a normalized, canonical form for analysis are explained. Following this, the implementation of the CIDR Report's aggregation report is explained. Finally, the process used to analyze the data gathered from the CIDR Report and from reimplementing the aggregation report in order to assess the efficacy of the CIDR Report is explained.

3.1 Analytical approach

The general intuition in analyzing the effectiveness of the CIDR Report is conveyed by the simple question: *do autonomous systems change their behavior, as measured by the number of aggregable routes they advertise into the routing table, after appearing on the CIDR Report?* If the hypothesis that the CIDR Report was effective in controlling routing table growth based on the social forces or reputation in the Internet operator community holds, then the behavior of ASes that appear on the CIDR Report should differ—becoming more aggregated—compared to the behavior of ASes that never appear on the CIDR Report. This can be viewed as a quasi-experiment [Bab03], with ASes that appear on the CIDR Report composing the treatment group and ASes that never appear as the control group. This cannot be viewed as a true natural experiment or controlled experiment because the ASes in the treatment group are not randomly selected. The behavior that leads to appearance on the CIDR Report is typical of large ISPs and so they are disproportionally represented on the CIDR Report. This non-random appearance of ASes in the treatment group raise potential validity concerns, and the implications of this are discussed later in section 4.3.3.

The analysis of the performance of the treatment group could be implemented very simply by comparing the behavior of a given AS over time, starting when it first appears on the top 30 aggregation report. The CIDR Report emails transmitted to mailing lists weekly contain much of the information necessary to conduct this analysis. Accordingly, we first gathered and coded the data contained in these emails from network operator mailing list archives. This information determines which ASes are in the treatment group and when they first appeared.

The CIDR Report emails have two shortcomings that demand the gathering of additional data about the routing table. The aggregation report contains no information about ASes that do not appear within the top thirty, requiring other information to be gathered in order to form a control group. Also, rank (and thus appearance) is determined by relative ranking rather than absolute number of routes, and so an AS in the treatment group may leave the top 30 without improving their route aggregation simply because other ASes announce more aggregable routes. In order to measure behavior in terms of routing table slots consumed by an AS, a non-truncated version of the CIDR Report is required. No archives of the full report are available, so we instead gathered historic routing table data and developed our own implementation of the aggregation report. The information from this generated aggregation report provides the metrics used for analysis of ASes in both the control and treatment groups for consistency.

Finally, utilizing the information from both the emailed (authoritative) CIDR Report and the full aggregation report generated from historic routing table data, we proceed with analysis, characterizing the overall behavior of ASes in the treatment and control groups, as well as behavior of ASes in both groups over time in order to observe whether the CIDR Report may have had an effect on treated AS' behavior.

3.2 Data sources & data collection

3.2.1 Authoritative CIDR Reports

The authoritative CIDR Report has been transmitted weekly on Friday afternoons to network operator communities, including the North American Network Operators Group (NANOG), starting in September 1997. The NANOG has kept a public archive, [NANa] and [NANb], of all messages sent to its mailing list since its inception in 1994, capturing all of the CIDR Report messages in its archive. This archive is an appropriate source of CIDR Reports as it contains the messages that were actually received by network operators and thus used to apply social force under the CIDR Report hypothesis.

The indexes of the archives were downloaded and parsed to obtain a set of message headers containing the sender name, subject, and date of all messages sent to the NANOG list. This message list was then filtered to create a coding candidate set of all messages with a subject line containing the string "CIDR R" in any mix of upper- and lower-case characters.

The full bodies of the messages in the candidate set were downloaded from the NANOG archives for coding. The coding process consisted of viewing each message and classifying it as one of:

- an authoritative CIDR Report that appears to be correct,
- an authoritative CIDR Report that appears incorrect (duplicates, sent on the wrong date, containing obviously invalid data, etc.),
- an email from the operator community praising, criticizing, or discussing behavior in the CIDR Report, or
- none of the above (not of interest).

Following the coding process, emails coded as authoritative and correct CIDR Reports were parsed by an automated program to extract rank and prefix count information for every AS appearing on the aggregation report. This information was then stored in a database for later use by other analysis tools.

3.2.2 Routing table data

The University of Oregon Route Views project [Rou] was used as the source of routing table data for generating a full aggregation report. Route Views gathers multiple views of the Internet routing table as seen by various major providers that peer with Route Views, and is a commonly used data source for operational and academic research. The set of networks that peers with Route Views changes over time, but typically include most of the default-free zone (DFZ) providers. The project has been storing routing table archives since November 1997.

Route Views' data sets contain the entire Route Views RIB, which contains all routes received from each provider it peers with. This differs from the Huston CIDR Report implementation [Hus11] (and possibly from the original Bates CIDR Report implementation also), which uses the best BGP route available for a given prefix, instead of all available routes.

Route Views RIB snapshots are typically generated every two hours, and so a script was developed to search the archive indexes for the most appropriate snapshot to download in accordance with the CIDR Report's weekly Friday release schedule. The script selects the RIB snapshot created most recently after midnight each Friday. For the cases where RIB files were not found for a given Friday, the most recently generated RIB from earlier in the same week was selected instead. For snapshots between November 1997 and November 2001, "Cisco CLI"-format[1] RIBs were downloaded from the `route-views.routeviews.org` router. After November 2001, MRT-format[2] RIBs were downloaded from the `route-views2.routeviews.org` router. These RIBs were stored for preprocessing and ultimately the generation of the full aggregation report.

[1]Cisco CLI is a text-based RIB format captured by copying the output of the `show ip bgp` command from a Cisco router's command-line interface.

[2]MRT is a binary RIB format that essentially encapsulates BGP UPDATE messages in a well-known format for archiving and later analysis. It is currently an Internet-Draft being developed by the IETF: `http://tools.ietf.org/html/draft-ietf-grow-mrt-14`.

3.3 Data preprocessing

Several preprocessing steps are required to extract and normalize the data from the Route Views RIB files in order to prepare the data in a canonical form for use in generating the aggregation report.

First the information required from each RIB snapshot—the advertised prefix, the IP address of the observing peer, and the AS_PATH associated with the route—are extracted from the snapshot files using RIPE's libbgpdump library [RIP] for MRT-format RIB snapshots and CAIDA's straightenRV tool [CAI] for "Cisco CLI"-format RIB snapshots. The output of these programs was a plain text representation of the information extracted from the RIB in the form of (prefix, peer IP address, AS_PATH) tuples.

Next, these simplified RIB files were processed to canonicalize the AS_PATH associated with each prefix. Since AS_PATH equality is used by the aggregation report to determine whether the routing policy of two prefixes is the same, a canonical representation of the AS_PATH is essential in order to allow AS_PATH comparison. There are four steps in the canonicalization process:

1. **Removal of non-AS_SEQUENCE elements in the AS_PATH.** The AS_PATH in a BGP UPDATE message can consist of four types of path segments [TMS07]: AS_SEQUENCE, an ordered list of ASes traversed by an update message; AS_SET, an unordered set of ASes traversed by an update message; AS_CONFED_SEQUENCE, an ordered list of private AS numbers traversed by an update message in a BGP confederation; and AS_CONFED_SET, an unordered set of private AS numbers traversed by an update message in a BGP confederation.

   AS_SETs (denoted as `{4,5}` for a set containing AS 4 and 5) are included in a set to the AS_PATH to continue to serve its loop prevention role while indicating that proxy aggregation was performed by a router along the path of the update message [KS11]. AS_SETs in the first segment of the AS_PATH make the origin of a route ambiguous because this effectively places multiple AS in the origin position of the AS_PATH. AS_SET segments are also often incorrectly used by network operators and add ambiguity to other parts of the AS_PATH [KS11]. To resolve such ambiguity, any AS_SET is collapsed to a single AS number if it contains only one AS number repeated any number of times, or is removed from the AS_PATH if it contain multiple unique AS numbers. As an example, the AS_SET `{4,4,4,4}` would be collapsed to `4`, while the AS_SET `{4,5}` would be discarded from the AS_PATH.

AS_CONFED_SEQUENCEs (denoted as `[4 5]`) and AS_CONFED_SET (denoted as `(4,5)`) should only contain private AS numbers, should not be visible on the public Internet, and contribute no Internet topological information about the route, and so are discarded without further consideration.

As an example of all of the transformations applied during this step, the path `1 1 1 2 3 (65535,65533) {4,4}` would become `1 1 1 2 3 4`.

2. **Removal of private AS numbers.** IANA has allocated private AS numbers [HB96, Hus08] for ISP internal use and documentation purposes, and these AS numbers sometimes appear in the Internet routing table. AS numbers within this range are removed from the AS_PATH.

3. **Removal of prepended AS numbers.** As discussed earlier, AS_PATH prepending is sometimes used by network operators to make a route appear less attractive to the BGP path selection algorithm. A prepended AS_PATH affects traffic engineering but not routing policy (strictly considering connectivity), and so prepended AS numbers are removed by collapsing contiguous blocks of the same AS number into a single entry in the AS_PATH.

   For example, the path `1 1 1 2 3 3 4` would be collapsed to `1 2 3 4`.

4. **Removal of simple routing loops.** Minor routing loops sometimes occur in the BGP routing table, such as when a route traverses a provider that uses multiple AS numbers for different parts of their infrastructure. These loops are removed following the algorithm used by CAIDA in straightenRV [CAI]. If a more complex loop is found that cannot be resolved using this algorithm, the entire route is discarded.

   For example, the path `1 2 3 2 5`, which may have resulted from AS 2 and 3 belonging to the same operator, is reduced to `1 2 5`. In contrast, a more complex loop such as `1 2 3 2 3 4` cannot be resolved using CAIDA's heuristic, and so the route associated with this AS_PATH is discarded.

Finally, because this canonicalization process requires a full traversal of the input RIB in order to produce the canonical RIB, other data sets can be opportunistically generated during the canonicalization process. The number of prefixes observed by each Route Views peer, for each origin AS and for the entire routing table, are recorded to allow later analysis of variation in the prefixes seen per peer.

3.4 Implementing the aggregation report

Determining which prefixes in a RIB are aggregable consists of three major steps. First, a prefix tree data structure containing all advertised prefixes is constructed in order to establish which prefixes are adjacent to and covered by other prefixes. Next, the tree is walked to recursively aggregate adjacent prefixes and to classify prefixes that may be aggregated by a covering prefix, only in cases where routing policy is not compromised. Finally, counts of aggregable and total prefixes are attributed to each origin AS. These counts are then ranked to generate the aggregation report as seen on the CIDR Report. Each of these steps is described in more detail below.

To construct the binary prefix tree (perhaps more correctly a binary prefix trie [Wu08]) containing a set of prefixes we wish to aggregate , the root of the tree must be determined. If prefixes of any length were allowed then the tree would need to be rooted at 0.0.0.0/0. However, because IANA has always allocated IP address blocks as Class A or /8 blocks, we generate the aggregation report by considering each /8 separately. Taking advantage of this fact will make the implementation of the classification algorithm more efficient as the entire routing table need not be stored in memory. Accordingly, the input RIB is sorted on the first octet of the prefix and processed by the algorithm one /8 at a time.

For each /8, a prefix tree is constructed with the /8 prefix at the root (for example, 10.0.0.0/8). Then, each prefix in the RIB as observed by each peer is inserted into the tree. Each node contains a prefix, the corresponding AS_PATH for that prefix, and pointers to

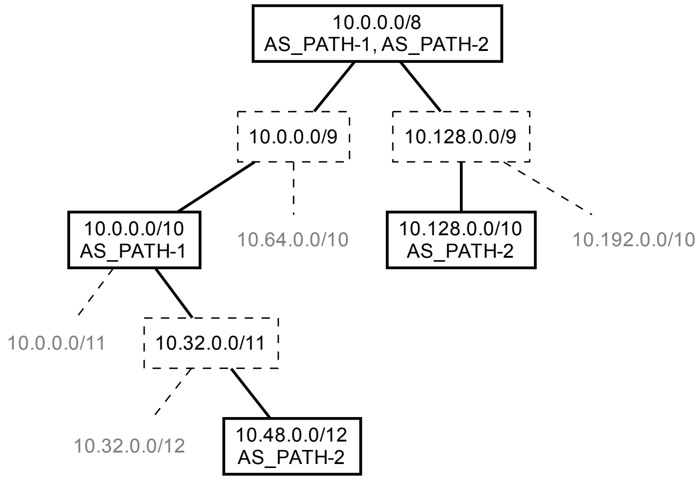
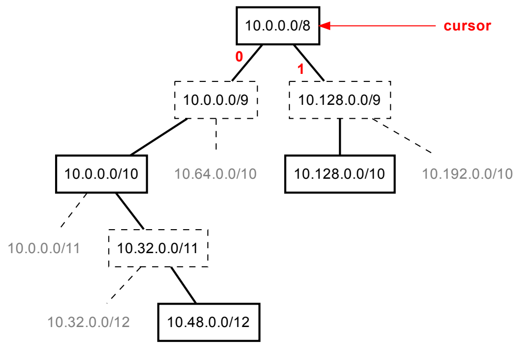


Figure 3-1: A prefix tree for 10.0.0.0/8 and some more specific prefixes. Prefixes with solid borders and AS_PATHs are announced in the routing table, and prefixes with dashed borders are placeholders.

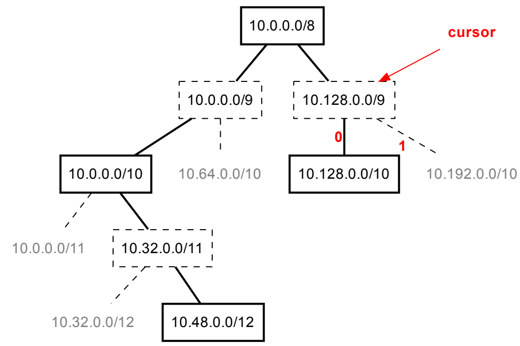
the two more specific sub-prefixes of the current prefix if they exist. An small example of a prefix tree is shown in Figure 3-1.

To insert a prefix into the tree, a cursor is placed at the root and then advanced to one of the two children of the root, depending on whether the first bit after the first octet of the IP prefix is a '1' or a '0'. This process is then repeated recursively at each node the cursor encounters for each corresponding subsequent bit in the prefix. If at any point a node does not have a child node that must be traversed to reach the insertion point, a placeholder prefix is inserted to maintain the tree structure. When the cursor's depth in the tree is equal to the length of the prefix less the original eight bits of the /8, the prefix is inserted into the tree. The insertion process is illustrated in Figure 3-2.

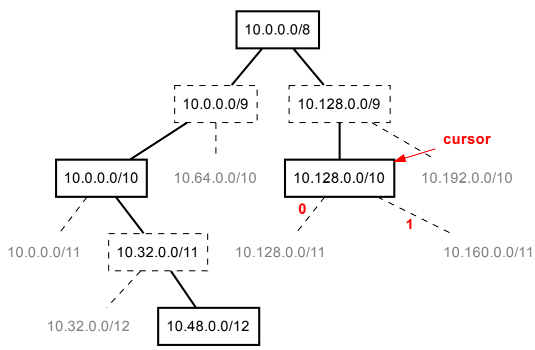
Unlike the Huston (and likely also the Bates) implementation of the aggregation report, this implementation considers all routes available from all peers for a given prefix in the routing table, instead of just the best route. Thus, for a given prefix in the tree, there are actually multiple AS_PATHS stored—one for each peer that observes the route. This can be viewed logically as several prefix trees (one for each peer) overlaid on top of each other, though the implementation uses a single tree with multiple AS_PATHs per prefix. All distinct AS_PATHs for a given prefix are gathered from the input RIB and associated with the prefix before it is installed in the prefix tree.



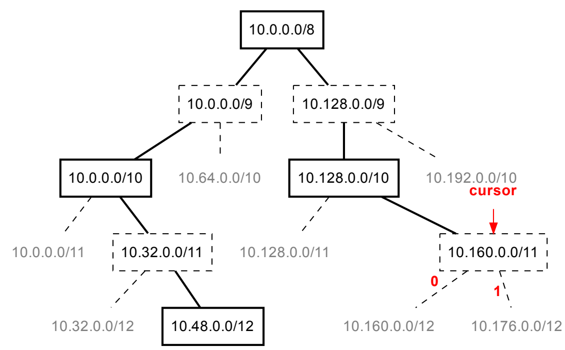
(a) Bit 9 of the prefix is a ‘1’, so the cursor proceeds to the ‘1’ child.



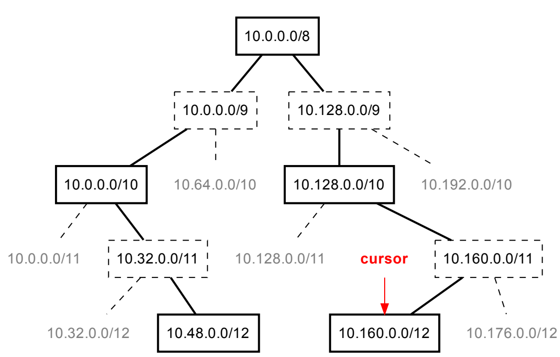
(b) Bit 10 of the prefix is a ‘0’, so the cursor proceeds to the ‘0’ child.



(c) Bit 11 of the prefix is a ‘1’, so the cursor proceeds to the ‘1’ child. The ‘1’ child does not exist, so a placeholder prefix is inserted as the ‘1’ child.



(d) Bit 12 of the prefix is a ‘0’, so the cursor proceeds to the ‘0’ child. The ‘0’ child does not exist. Since the current bit is equal to the prefix length (/12) the (non-placeholder) prefix 10.160.0.0/12 is inserted as the ‘0’ child.



(e) Insertion of 10.160.0.0/12 is complete.

Figure 3-2: Insertion of 10.160.0.0/12 into the prefix tree described in Figure 3-1. Recall that 10.160.0.0 is represented as 00001010.10100000.00000000.00000000 in binary.

Once all of the prefixes in the RIB for the current /8 are inserted into the prefix tree, the prefix tree is walked recursively to aggregate and classify aggregable prefixes within the tree. As noted in the description of the CIDR Report in Chapter 2, a prefix is considered aggregable if it has the same routing policy as a less-specific covering prefix. The aggregation and classification algorithm was implemented based on the description of the CIDR Report included in the Report's preamble, as well as detailed records of the report's operation [Husc]. Like the authoritative CIDR Report, routing policy equality in this implementation of the aggregation report is determined by comparing AS_PATH equality.

The general algorithm for the aggregation and classification process is as follows. The prefix tree is traversed using post-order recursion and the following operations are performed on each node before the function returns to its calling parent (the post-order traversal ensures that children are aggregated before their parents):

1. Attempt to aggregate children: If the current prefix is a placeholder and has two more-specific (child) prefixes announced in the routing table, check to see if their AS_PATHs match. If they do, convert the current prefix into a "real" (non-placeholder) prefix and mark the children prefixes as aggregable.

2. Attempt to aggregate by a covering prefix: Compare the AS_PATH of the current prefix with the AS_PATH of its nearest less-specific (ancestor) prefix . If they match, mark the current prefix as aggregable.

This process is again logically conducted as though there are multiple separate but overlaid trees for each Route Views peer AS. The process is actually implemented by processing all vantage points available at a given prefix and aggregating and classifying them against other ancestor and child prefixes visible from the same peer AS. Classifications of aggregability are made on a per-peer basis, and then must be generalized for the entire prefix. There is no single way to make this generalization and so a design decision must be made.

There are two obvious choices for how to generalize across each vantage point's aggregation classification for a given prefix: consider the prefix aggregable if *any* vantage point considers it aggregable, or only if *all* vantage points consider it aggregable. The latter option requires consensus that a prefix be aggregable across all views, while the former allows any claim of aggregability to stand. Our implementation of the aggregation report classifies a prefix as aggregable if it is classified as aggregable from *any* vantage point. While perhaps "pessimistic", this approach will detect and report the maximum amount of aggregation potentially possible.

Finally, when a prefix is marked as aggregable, it must be attributed to the network that announced the prefix, as this network is responsible for the redundant route announcement. This is normally performed by attributing the aggregable route to the origin AS—the first AS in the route's AS_PATH—that is presumed to be the announcing network. An ambiguous situation arises in the case of prefixes announced by multiple ASes, a condition known as a multiple origin AS (MOAS) [ZPW+01]. Again, a design decision must be made for how to attribute aggregable MOAS prefixes: the MOAS prefixes could be ignored, the route's behavior could be attributed to one of the origin ASes, or it could be attributed to all origin ASes. This implementation attributes an aggregable prefix to each AS that announces a MOAS prefix.

To attribute deaggregation behavior to each AS, the aggregation report processor maintains five quantities for each origin AS:

- Announced routes (*announced* or *netsnow*): routes that were originally announced in the input RIB. This is the "netsnow" quantity in the aggregation report
- Withdrawn routes (*withdrawn*): routes that were covered by less specific routes and thus would be withdrawn from the ideally-aggregated routing table.
- Aggregated routes (*aggregated*): routes that did not exist in the original routing table, but were synthesized by combining two adjacent prefixes into a covering prefix.
- Total reduction in routes under perfect aggregation (*netgain*): computed as $netgain = withdrawn - aggregated$
- Total advertised routes under perfect aggregation (*netsaggr*): computed as $netsaggr = announced + aggregated - withdrawn = netsnow - netgain$

If a prefix in the prefix tree is marked as aggregable and is originally from the input RIB, then this prefix is counted as a *withdrawn* route. If a prefix is synthetic aggregate (not from the input RIB) and not marked as aggregable, then it is counted as an *aggregated* route.

After all of the RIB's routes are processed, the counts of aggregable routes attributed to their origin ASes—in the form of a list of tuples containing the origin AS and the five quantities described above (origin AS, netsnow, netgain, netsaggr, aggregated, withdrawn)—must be sorted to determine the ranking of networks as found on the authoritative CIDR Report. The ranking of ASes on the aggregation report is generated by a primary sort on the netgain value, the number of aggregable prefixes announced by the AS. The AS with the greatest netgain value will be hold rank 1 on the aggregation report. To provide an ordering when networks have the same netgain value, a secondary sort is performed to rank networks with a lesser netsnow value above networks with a greater netsnow value if both have the same netgain—these networks are more deaggregated as a fraction of their total routes. Finally, a tertiary sort is performed on the AS number as a sort of last resort to produce a canonical ordering similar to that found on the full, authoritative CIDR Report [Husb]. With this final sorting, the aggregation report data is considered canonically ranked, and is ready for analysis.

3.5 Analyzing the aggregation report

Recalling the original purpose of this analysis, we must now measure the change in behavior of an AS over time once it appears on the CIDR Report. The analysis of the aggregation report component of the CIDR Report consists of separate processes to gather results for the treatment group (the group of ASes that appeared on the CIDR Report) and the control group (a sample of the ASes that never appeared on the CIDR Report). However, the overall structure of both processes is roughly similar and consists of four steps. First, the data are preprocessed to mitigate holes in the data and other inconsistencies. Next, the first appearances of each AS on the CIDR Report are located to define the start of the data sampling. From this starting point, points from the various data series are sampled at the first appearance point and various times after, in order to measure the behavior of the AS after the initial appearance on the CIDR Report. Finally, the deltas of these figures are used to generate metrics for change in AS behavior over time, which are then visualized for analysis. Each of these steps are described in more detail below.

As will be illustrated in the following chapter, there are several gaps of one or more weeks where the data from the CIDR Report or Route Views were unavailable, ostensibly due to operational problems. These gaps would potentially be problematic later when sampling data points after the first appearance, as a sampling point might fall in the gap that happens to be between valid data. Thus, it was necessary to fill the gaps in the data. While a number of reasonable approaches could have been used to fill the gap (e.g. linear interpolation between the values on either side of the gap), this implementation of the analysis simply copies the data from the week preceding the gap into the gap. This is in effect a "sample and hold" across any gaps found in the original data.

With the gaps filled to yield a continuous data set for the entire range of analysis for which we have both CIDR Report and Route Views data, we must next determine when ASes first appear on the authoritative CIDR Report. An appearance is defined by a start date—the date the AS first appears within the top 30 ranked ASes on the CIDR Report—and the length of time it spends within the top 30. Appearances were determined by linearly traversing the authoritative CIDR Report data for each AS, recording when it appears, and counting the number of weeks it appears continuously on the report. To prevent the counting of a second appearance when an AS momentarily disappears from and then reappears on the report, a gap of up to eight weeks was allowed before a reappearance would be considered a new appearance.

From the start of each appearance, a number of points are sampled to determine the ASes behavior over time following the appearance. The samples are taken regardless of whether or not the AS is still on the authoritative CIDR Report, as it is conceivable that an AS' rank may change over time and thus fall below the top 30 threshold. To maintain consistent measurements of ASes both on and off the authoritative CIDR Report, all samples are taken from the generated CIDR Report instead, with sample points determined by the appearance on the authoritative CIDR Report. Points from each data series are sampled as illustrated in Figure 3-3, starting with date of first appearance and for various time durations after the initial appearance—currently 30 days, 60 days, 90 days, 180 days, 365

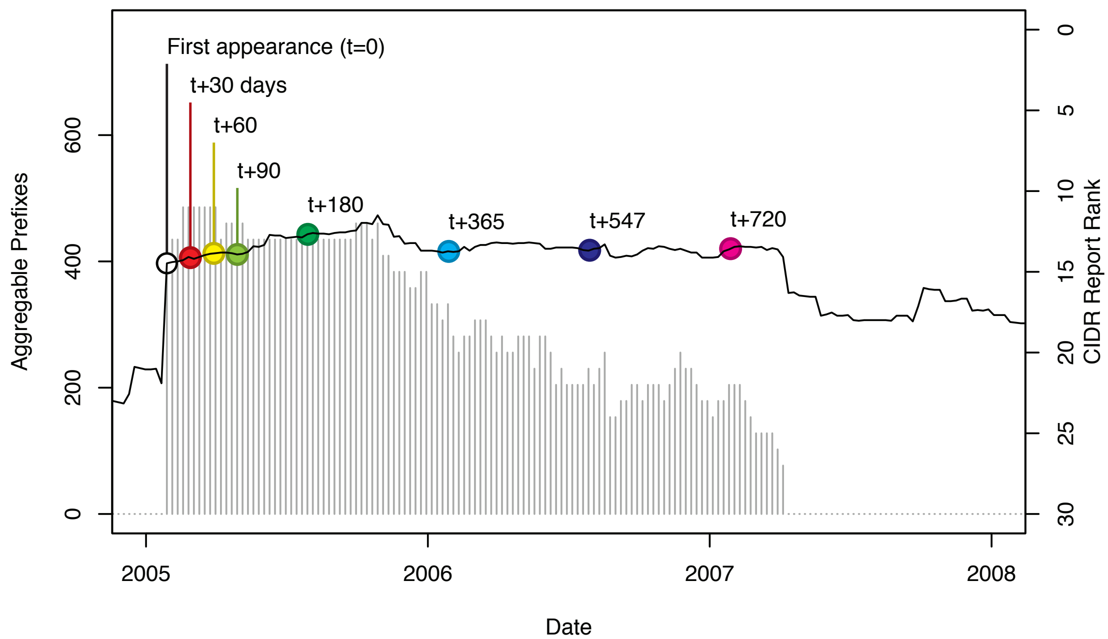


Figure 3-3: Illustration of the sampling approach used to measure AS behavior, showing the first appearance (shown by the vertical bars indicating ranking on the authoritative CIDR Report) and samples at various durations thereafter. In this case, aggregable prefixes (netgain) are sampled from AS 3602.

days (1 year), 547 days (1.5 years), and 730 days (2 years). Appearances that start within 2 years of each other are amalgamated to avoid overlapping and duplicate measurements.

With these samples collected, differences are then calculated for each of these quantities relative to the first appearance to understand the behavior change of the AS relative to its initial appearance. Visualizations and analyses of these differences and other composite measures created using them are presented in the next chapter.

3.5.1 Establishing the control group

The process above, explained in terms of the treatment group of ASes that appear on the authoritative CIDR Report, is identical for the control group with the exception of identification of "appearances" for the control group, as the ASes that compose the control group never appear on the CIDR Report. Instead, appearances in the control group must be constructed. In attempt to avoid biased construction of the control group, the control group is constructed randomly as follows.

First, a candidate set of ASes that are eligible to form the control group is established based on the following eligibility criteria:

- An AS must announce at least 10 prefixes into the routing table in order to be eligible. This is an arbitrary minimum threshold, but it is intended to exclude the large proportion of stub ASes in the routing table that announce single prefixes [Husa].
- An AS must be continuously visible in the routing table for a minimum of 2 years (720 days) to enable full sampling of its behavior.

From the set of candidate ASes matching this criteria, a number of ASes equal to the number of ASes appearing on the authoritative CIDR Report are randomly selected to form the control group. Finally, for each AS within the control group, the date of first "appearance" is randomly selected between the AS' first appearance in the routing table and 720 days before it's last appearance in the routing table. With these appearances established, the sampling and difference calculations for the control group proceeds as with the treatment group.

With data from the treatment and control groups extracted using the methods described in this chapter, we were ready to proceed with our analysis in order to answer the central question of this thesis.

Chapter 4

CIDR Report characteristics & influence on AS behavior

This chapter presents and describes observed characteristics of the CIDR Report, as well as the results of the analysis conducted to determine if the CIDR Report affects network operator route aggregation behavior. The chapter begins with a preliminary discussion about the availability of data and the quality of our aggregation report implementation. This is followed by presentation of characteristics of the CIDR report that were observed during this analysis, which will be discussed in the following chapter as clues of what may have affected the success of the CIDR Report. The chapter continues with presentation of the main results of the analysis relevant to this thesis, illustrating how individual ASes' route announcement behavior may have been affected by appearing on the CIDR Report. Finally, potential questions about the validity of the analysis are identified and discussed.

As discussed briefly in the previous chapter, there are two data sets at play in this analysis. The distinction between these two sets is important for understanding some of the figures in this section, and so they are clearly defined here:

- The authoritative CIDR Report (ACR): This is the data from the authoritative CIDR Report that was emailed out to network operators that identifies the 30 most deaggregated ASes. This is the set of ASes that would be expected to display a treatment effect if the CIDR Report was effective.

- The generated CIDR Report (GCR): This is the full aggregation report generated by our implementation of the prefix aggregation algorithm used by the CIDR Report, but containing data for every AS in the routing table instead of being truncated after the 30th AS.

Finally, to refresh the reader, some terms of art will be used in the remainder of the chapter for conciseness. These are:

- *netgain:* The number of prefixes advertised by an AS that are unnecessary (do not affect routing policy from the perspective of the CIDR Report vantage point), and could be removed from the routing table if this AS aggregated perfectly.

- *netsnow:* The total number of prefixes advertised by an AS, including both aggregable and non-aggregable routes.

4.1 Data and methodological quality

4.1.1 Data availability

The historic origins and other peculiarities of the underlying data sources for the ACR and GCR mean that the data series for two reports start at different dates and contain small gaps in their otherwise continuous record of weekly CIDR Report data. The start dates and gaps are illustrated in Figure 4-1, where a solid vertical line indicates a week of available data.

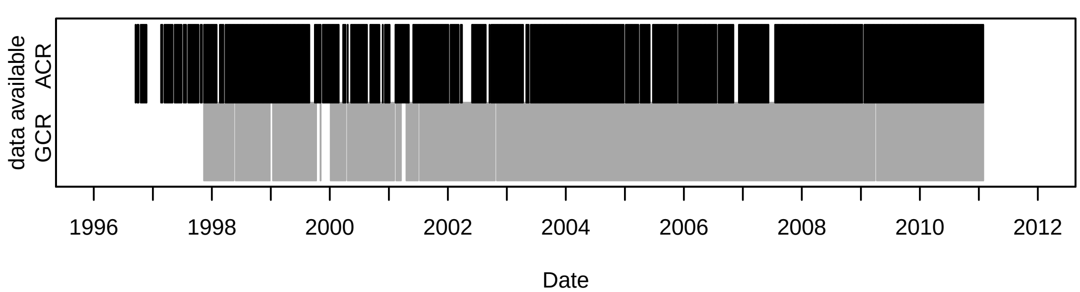


Figure 4-1: Available data

As illustrated, data from both reports is generally available from November 1997 until January 2011, with a few relatively minor gaps. The period where the ACR and GCR data overlap is the period used to analyze the effects of the CIDR Report on AS behavior.

4.1.2 GCR implementation accuracy

In addition to data availability, the accuracy of the analysis that follows also depends on the accuracy of our implementation of the CIDR Report aggregation algorithm described in Chapter 3. Recall that while the ACR is used to determine when an AS appears on the CIDR Report (and thus potentially commands attention of the community), the GCR is used to determine actual post-appearance behavior because it provides consistently-generated measures (like the number of aggregable prefixes advertised, *netgain*) for ASes both above and below the ACR top 30 threshold.

The first measure of accuracy is a comparison of the aggregable and total prefixes determined for each AS on the ACR and GCR. A plot illustrating this comparison is shown in Figure 4-2. The curves above the horizontal zero ($y = 0$) represent the sum of the differences in prefix counts between the ACR and GCR for ASes where the GCR reports more prefixes than the ACR. Similarly, the curves below the zero represent the sum of differences in prefix counts between the ACR and GCR for ASes where the ACR reports more prefixes than the GCR. The total deviation between the ACR and GCR is given by the difference between the curves above and below zero.

As can be seen in Figure 4-2, throughout the period of available data, the GCR generally claims larger quantities of aggregable and total prefixes than the ACR. This is particularly pronounced and erratic when compared against the pre-August 2002 CIDR Report (the point when the report's methodology was changed when it was re-implemented by Geoff Huston). It is difficult to determine the cause of this, but we suspect it is due to the greater number of vantage points used in generating the GCR, leaving the GCR more open to observing prefixes not visible from the ACR vantage point or observing AS_PATHs that enable classification of prefixes as aggregable.

After August 2002, agreement between the ACR and GCR improves, though there are

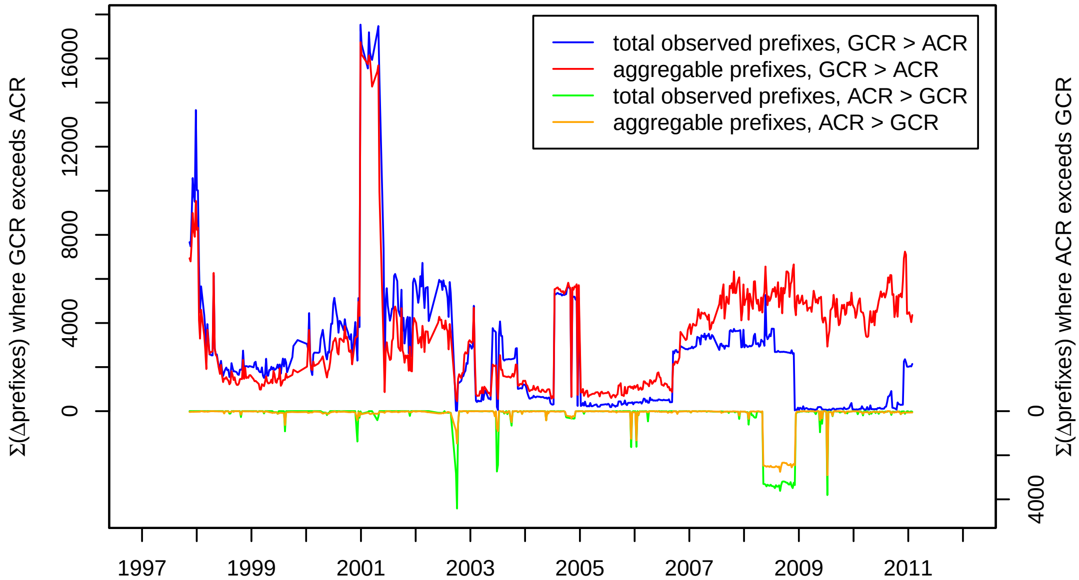


Figure 4-2: Differences in prefix counts between authoritative (ACR) and generated (GCR) CIDR Reports over time.

still inconsistencies. Most of these inconsistencies appear attributable to the GCR observing more prefixes than the ACR—this is the case when the observed prefixes (blue line) and aggregable prefixes (red line) increase simultaneously. Perhaps more concerning in terms of accuracy is the change in behavior that began in late 2008 and continues to the latest, where the ACR and GCR observe the same number of prefixes (blue line is approximately zero) but the GCR classifies 4000-5000 more prefixes as aggregable. Upon further investigation, these deviations appeared to be due to the multiple vantage points used by the GCR, which observed potential aggregation that the ACR did not observe.

In spite of these inconsistencies, which are generally minor in terms of ACR-GCR disparity for individual ASes, the GCR presents a view of potential routing table aggregability that is sufficiently consistent with the ACR to enable our use of the GCR as a data source for analyzing individual AS aggregation behavior after appearing on the CIDR Report. Further, because data related to prefix counts is always taken from the GCR, prefix data and differences calculated based on this data will always be consistent and not affected by differences between the ACR and GCR. Finally, as will be discussed next, these minor differences in prefix counts do not greatly affect the ranks of ASes on the GCR when compared to the ACR.

While it is helpful to look at differences in prefix counts, as this is the major purpose for which the GCR is used (rather than the ranking it generates), we can also compare the rankings it generates against the rankings from the ACR to compare the relative accuracy of the GCR. Ranks provided by the GCR are not currently used in this analysis, so we only briefly raise this before continuing on.

Figure 4-3 illustrates the minimum and maximum absolute differences between the ranks of ASes appearing on the ACR and the ranks of corresponding ASes on the GCR.

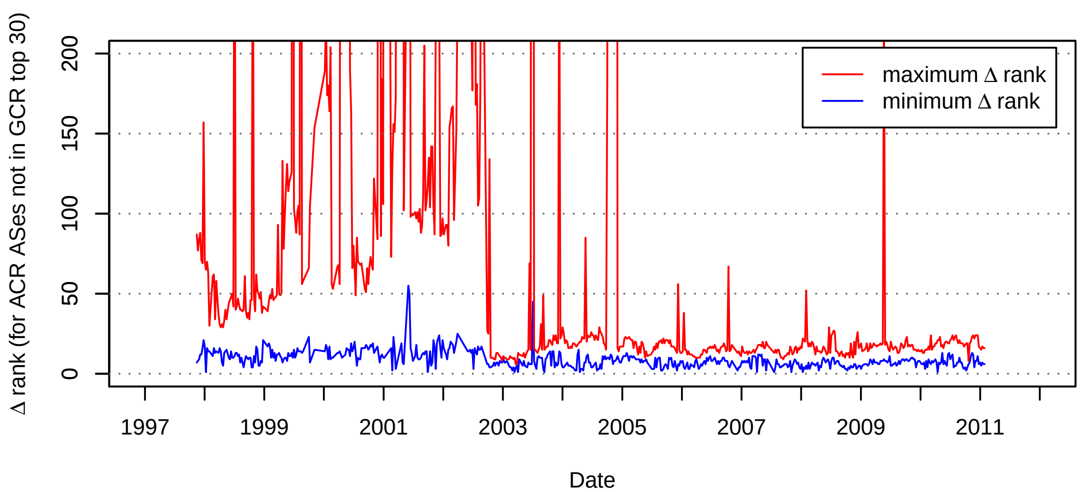


Figure 4-3: Minimum and maximum differences in aggregation report rank for ASes ranked in the top 30 on the ACR but not in the top 30 on the GCR.

As can be seen, similar to the differences in prefix counts shown in the previous figure, the rank differences are more significant and erratic from 1997 until August 2002, at which point they become more consistent. With some exceptions, likely due to input data aberrations, the rankings generally differ only a small amount, often near zero in the best case and no more than 20-30 in the worst case. Also, typically 20-25 of the ASes that appear on the ACR also appear on the GCR.

It was difficult to develop an implementation of the CIDR Report aggregation algorithm (which is used to produce the GCR) that exactly matches the output of the authoritative CIDR Report. Further, it is questionable whether an implementation of the algorithm should strive to reproduce the output of a “black box” (even if it is the black box that is used to inform and influence the operator community) instead of being based on first principles and a conceptual understanding of the purpose of the CIDR Report. Given that our implementation was developed on the basis of the principles of route aggregation, along with the fact that the GCR is reasonably in agreement with the ACR, we believe it is reasonable to use this data in our analysis.

4.2 Characteristics of the CIDR Report

In analyzing the CIDR Report, much time was spent looking at characteristics of the report and networks that appear on it. While this was not directly related to answering the question of whether appearing on the CIDR Report changes route aggregation behavior, it is useful in providing context about the CIDR Report, and also possibly offering insights about why the Report was or was not effective, or how its efficacy may have changed over time. The characteristics discussed in this section are organized into two categories: relative measures of network behavior on the CIDR Report (typically related to ranks and ranking) and absolute measures of network behavior (typically related to prefixes advertised).

4.2.1 AS appearances and rank-based observations

The first approach taken to gain an understanding of the behavior of the CIDR Report was to visualize it in a two-dimensional space, with time in one dimension and CIDR Report rank in the other dimension. An excerpt of this visualization is shown in Figure 4-4.

This visualization was not especially practical, as it was physically large and the level of detail and limited range of colors available made it difficult to visually identify and distinguish the behaviors of individual ASes. However, this visualization provided some hints regarding interesting characteristics to investigate. By visual inspection, it appeared as though the CIDR Report was volatile at lower ranks but relatively static near the top (ASes with the most aggregation potential). It also appeared that the report was more volatile in the past but become more static overall more recently. These and other aspects were then investigated using analytic techniques, which are described and presented next.

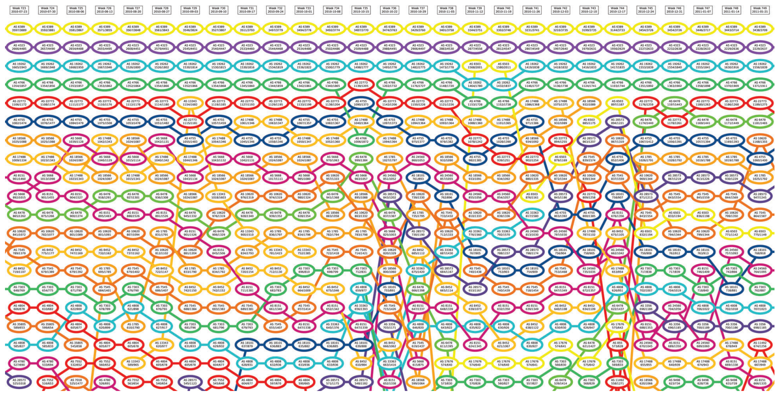


Figure 4-4: A sample of our CIDR Report visualization.

A number of interesting characteristics about the CIDR Report are of a demographic nature, such as how many ASes appear on the CIDR report, or for how long do ASes remain on the report? This first question, about how many ASes ever appear on the CIDR Report, is illuminated by Figure 4-5.

In this figure, the top plot shows the cumulative number of ASes that appear on the CIDR Report, starting with 30 ASes on the first report in 1996 and concluding with 386 unique ASes in 2011. The growth of new ASes appearing on the report appears to be greater in the late 1990s and early 2000s than in the mid-late 2000s. More remarkable is how relatively few ASes from the total set of ASes that ever appear in the Internet routing table appear on the CIDR Report. This relationship is shown in the bottom plot of Figure 4-5. In this figure, the red curve near the bottom of the graph area is the same red curve that was plotted in the upper figure, and the black curve that rises steadily is the total number of ASes visible in the routing table. This illustrates that less than 1% of all ASes have ever appeared on the CIDR Report.

While the former plot illustrated the total number of ASes that appear on the report, it does not provide any information about how long an AS appears on the report. This information is provided in the next figure, Figure 4-6, which presents a cumulative distribution function of the total number of weeks each AS appears on the CIDR Report. Noting the log scaled x-axis, we can see that 15% of all ASes that appear on the CIDR Report are on the report for only a single week, half of all ASes are on the report for approximately 10 weeks or less, and that there is a general exponential relationship between increasing fractions

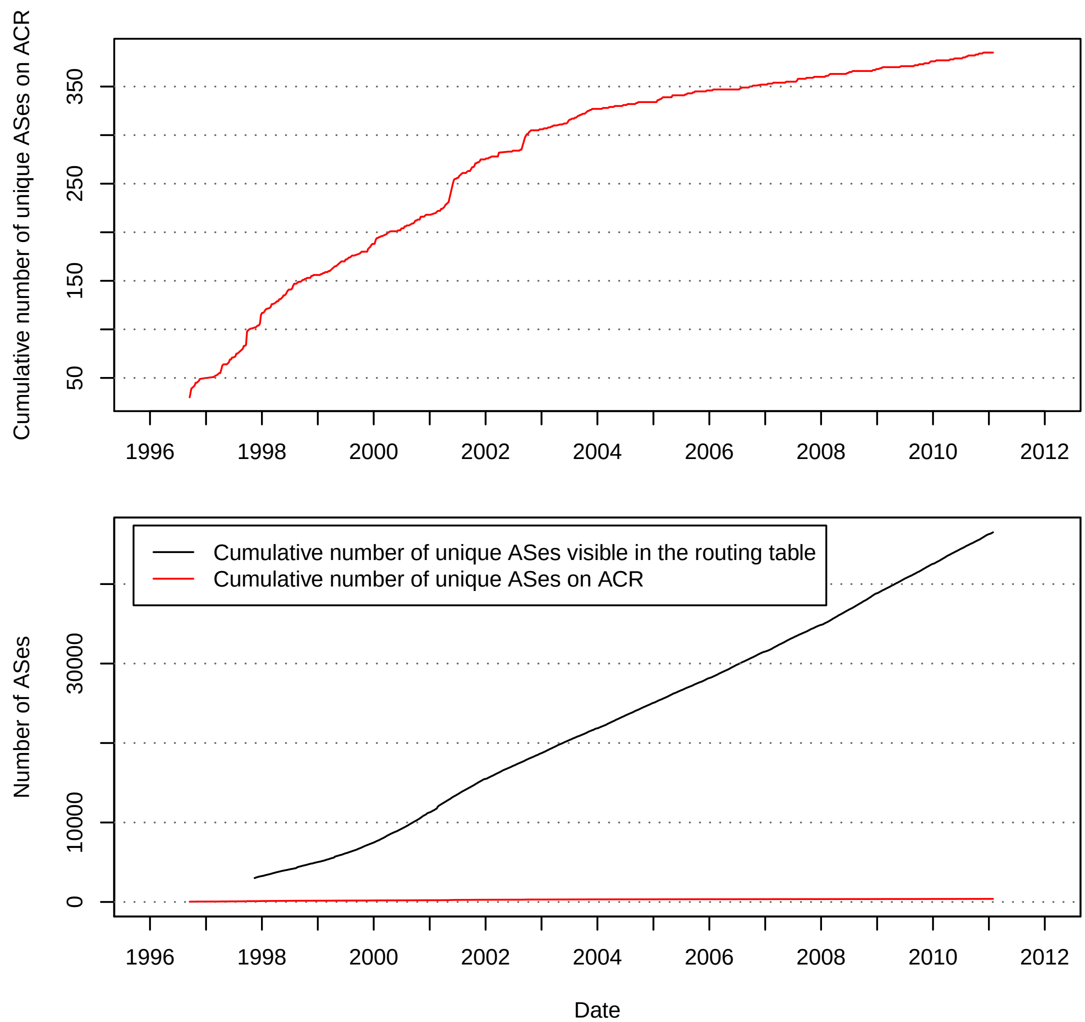


Figure 4-5: A cumulative count of the unique ASes that have appeared on the CIDR Report and compared to the cumulative count of unique ASes visible in the routing table up to the same point.

of the population and time spent on the report (i.e. for each 10% increase in population considered, the maximum duration spent on the report doubles).

This is the first indication that not all the ASes that appear on the CIDR report behave in similar ways, but instead that many ASes appear on the Report only briefly in comparison to some ASes that spend a considerable duration of time on the CIDR Report.

Returning to the observation from earlier that the top ranks of the CIDR Report are more static relative to the volatile lower ranks, we investigate this further by plotting cumulative distribution functions of the number of weeks that a particular AS occupies a given rank on the CIDR Report. These plots, for ranks 1 (most potential for aggregation) to 30, are

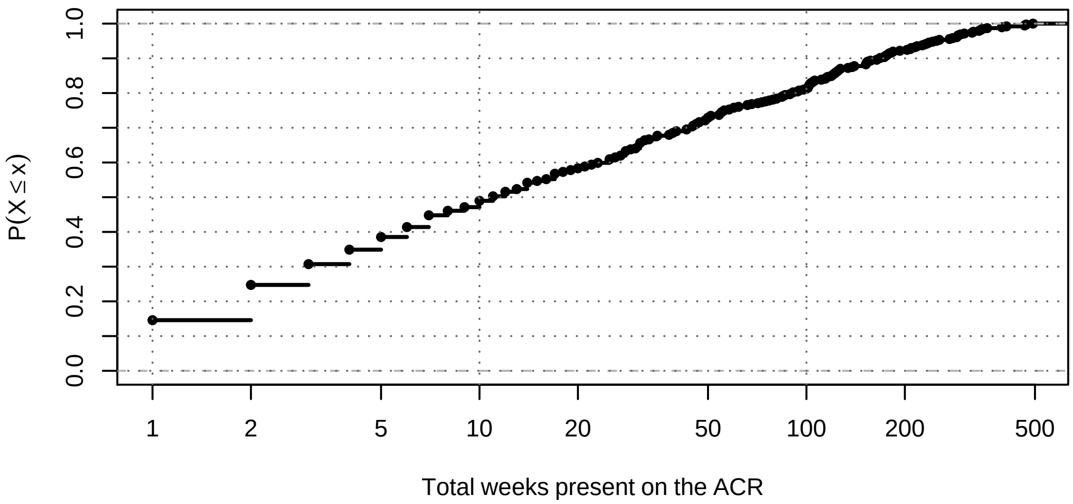


Figure 4-6: A CDF of the total number of weeks that each AS is visible on the CIDR Report.

shown in groups in Figure 4-7

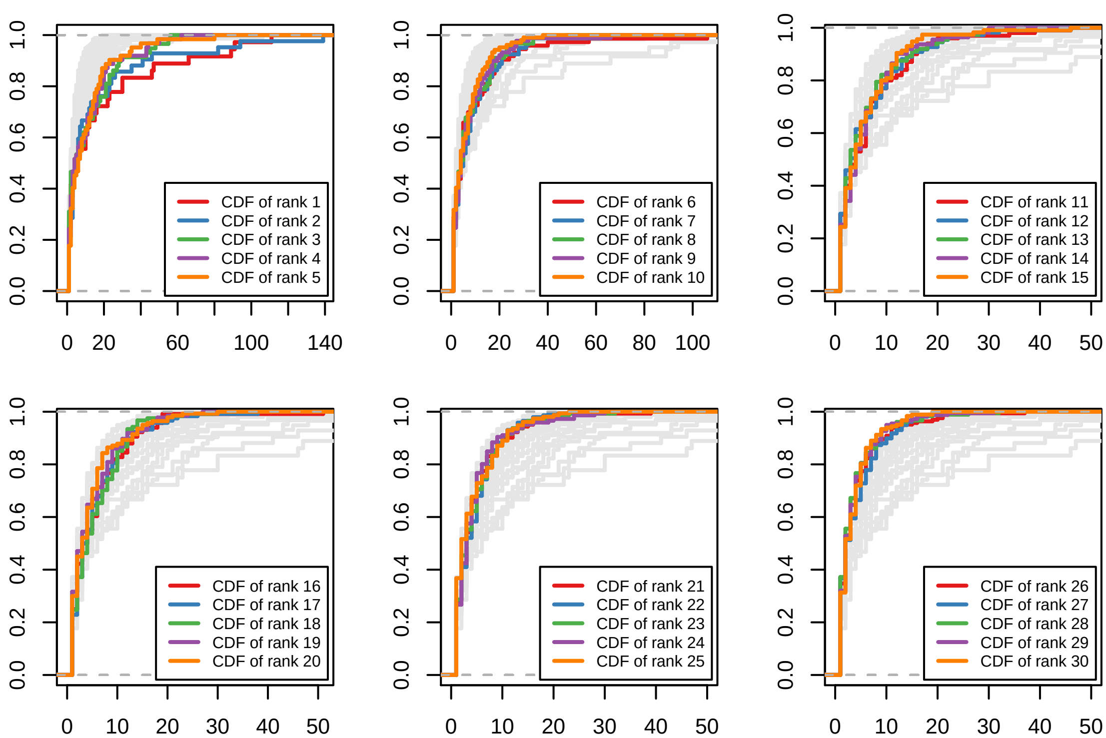


Figure 4-7: CDFs of the number of weeks that a given rank was occupied by a single AS. The x-axis is the number of weeks, and the y-axis is the fraction of the population. The light gray lines indicate CDFs of other ranks, in order to place the highlighted ranks in context.

As illustrated clearly by this figure, lower ranks on the CIDR Report are more volatile than top ranks. The top five positions, ranks 1-5, are particularly ossified, with the top 10% of ASes that ever appear occupying individual ranks for no less than 20-60 weeks. In contrast, the lower ranks, and particularly the bottom half (ranks lower than 15) are much more volatile, with around 90% of the AS population occupying a given rank for 10 weeks or less, and essentially all of the population not occupying ranks for more than 20 weeks. This is consistent with the claimed observation of operators that the CIDR Report does not appear to change much—especially the top ranks that are first visible in the email. The apparent reason for this ossification will be discussed in the next section.

4.2.2 Prefix-based observations

While the previous section presented interesting observations about AS appearances and volatility on the CIDR Report, it was focused on the relative measures of AS rank and appearance within the top 30 on the CIDR Report. It is also helpful to gain a perspective about characteristics of the CIDR report considering the number of prefixes that each AS is advertising, as prefix counts (and routing table slots) are ultimately the metric that is important in considering behavior change in the context of the routing table.

Figure 4-8[1] illustrates the fraction of the the total prefixes and aggregable prefixes visible in the Internet routing table that are advertised by ASes appearing on the CIDR Report.

This figure suggests that, in the past, the CIDR Report did focus quite effectively on most of the "worst offenders" in terms of networks advertising aggregable routes. At the beginning of the period of available data, the Report captured nearly 60% of the total aggregable routes in the top 30 list. This has decreased but, interestingly, remained approximately proportional to the growth of the routing table since 2003 or 2004, suggesting that growth of deaggregation in the ASes at the top of the CIDR Report is proportional to the growth of deaggregation across the Internet routing table. However, as the next figure will show, this proportionality has been maintained in the face of growth of the number of par-

[1] All data in this plot is from the GCR, but classification of an AS and its netgain/netsnow figures are based on the ASes present on the ACR as emailed. Data from the GCR is used to allow for consistent comparison between the entire routing table and the top 30 (ACR) ASes.

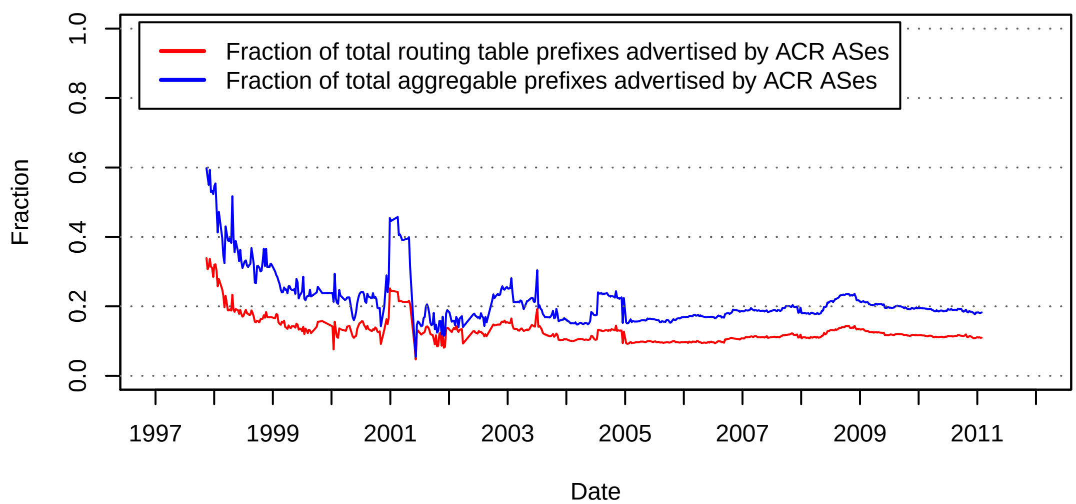


Figure 4-8: The fractions of total prefixes and total aggregable prefixes in the routing table that are advertised by ASes appearing on the CIDR Report.

ticipants in the Internet routing table, such that the ASes that appear at the top of the report have become outliers compared to most of the ASes in the routing table.

Another way of considering how the prefix advertisement behaviors of networks on the CIDR Report compares to the other networks in the routing table is to look at the distribution of deaggregation (advertisement of aggregable prefixes) in the routing table. Cumulative distribution functions of aggregable prefixes (netgain) visible in the routing table (from the GCR) over time is shown in Figure 4-9.

From these distributions we can see that while advertisement of aggregable prefixes has increased slightly across the routing table over time, as given by the downward movement of the CDF curves over time, the distribution of aggregable prefixes announced by ASes has remained roughly similar, with an increasingly long tail of outliers in the top fraction of a percentile of the population of ASes. This figure is potentially misleading, as while the distribution of aggregable prefix announcement has not changed, the total number of ASes visible in the routing table has grown over time, from 3172 ASes in the first week of 1998 to 36383 ASes in the first week of 2011. Thus, there are now approximately ten times as many ASes announcing a given number of aggregable prefixes as there were in 1998.

What is also noteworthy in this figure is the indication of the threshold points for appearing on the CIDR Report in 1998 and 2011. The fraction of the population above this

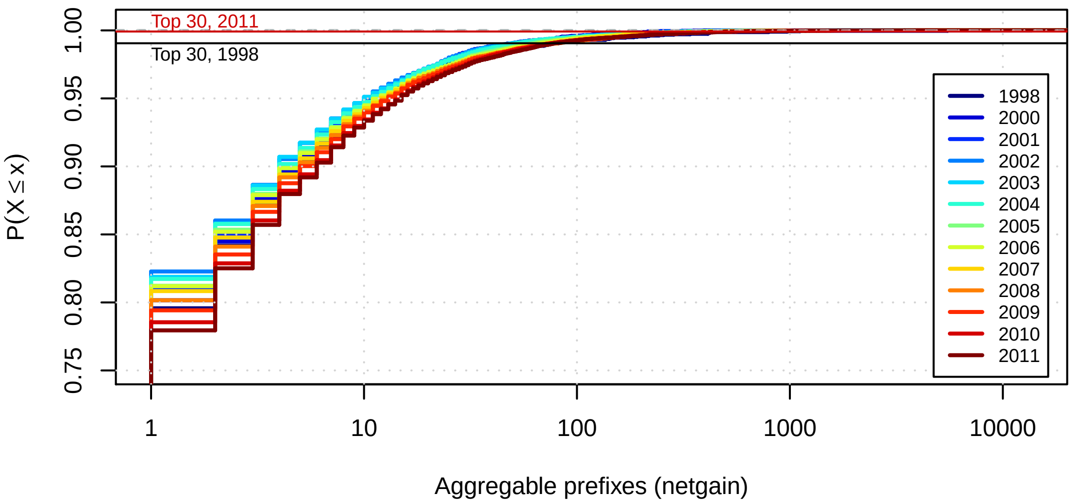


Figure 4-9: CDFs of netgain of all ASes in the routing table in the first week of the year from 1998-2011. Threshold lines indicating the cut-off point for appearing on the CIDR Report in 1998 and 2011 are indicated. Note that the graph is rescaled; approximately 70% of ASes in the routing table do not advertise any aggregable prefixes.

line is the "top 30" group that would appear on the CIDR Report. This line appears to have moved from approximately 1% of the population in 1998 to some small fraction of a percent in 2011. This conclusion follows from the fact that the number of ASes in the routing table has grown over time and yet the "top 30" threshold of the CIDR report has remained constant. However, as this figure, as well as Figure 4-8 illustrate, this also means that the CIDR Report has changed to highlight mostly outlier behavior, leaving the individually less significant but collectively more significant deaggregation below the threshold unaddressed.

The question of where the threshold to appear on the CIDR Report is in terms of netgain—how much deaggregation it takes for an AS to appear on the report—is addressed in Figure 4-10. This figure displays the minimum netgain thresholds to appear on the CIDR Report (rank 30) as well as the median (rank 15) and various other top ranks.

This figure of the CIDR Report rank thresholds shows an increasing spread in prefix thresholds for the various ranks over time, particularly in the top 5, as well as between the top 5 and the median, compared to between the median and the minimum threshold (rank 30). The increasing spread between ranks can also be viewed in a slightly different way,

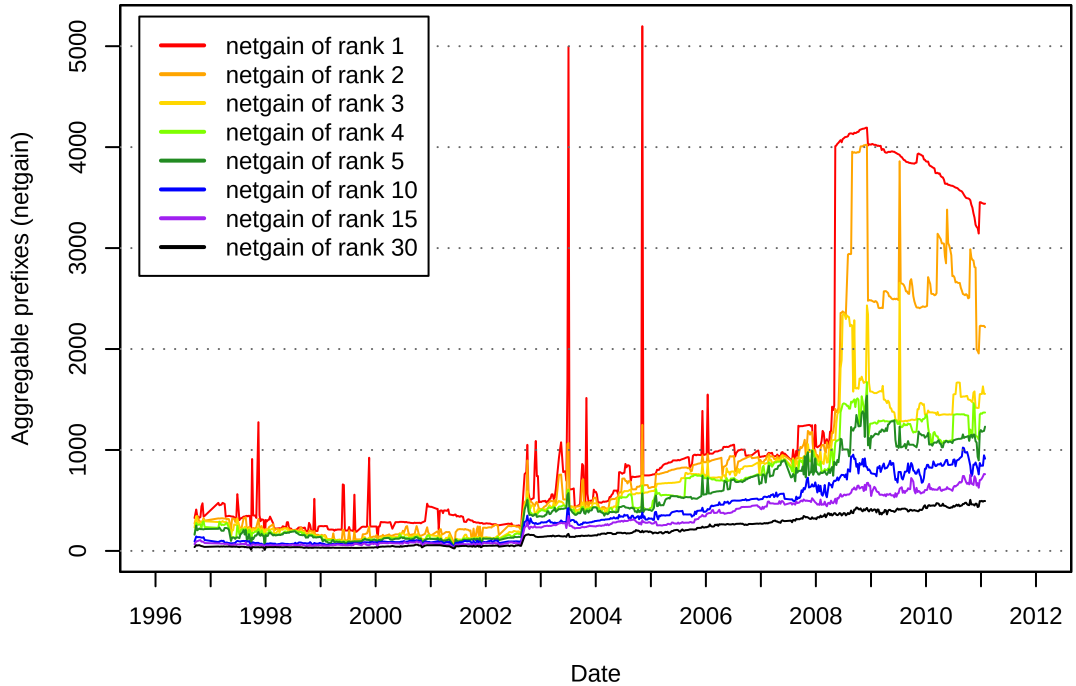


Figure 4-10: Netgain thresholds required to achieve indicated ranks on the ACR over time.

with a focus on temporal progression, in the following figure.

Figure 4-11 presents cumulative distribution functions of the number of aggregable prefixes (netgain) advertised by each AS on the ACR in the first week of the year indicated.

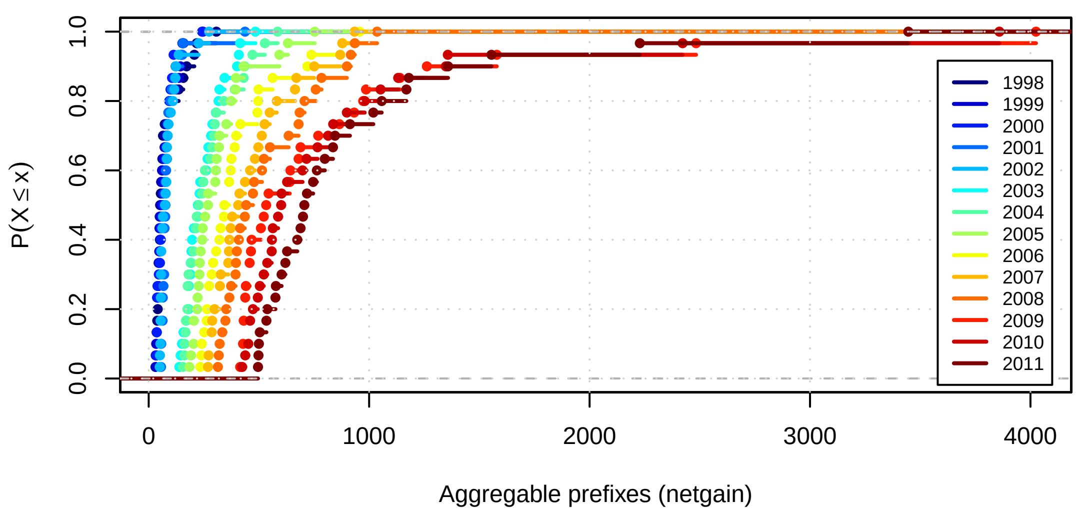


Figure 4-11: CDFs of the netgain of the ASes visible on the ACR in the first week of each year 1998-2011. Notice that the top of the population spreads more in later years.

Here again we see the growing spread between ranks and the changing shape of the distribution as ASes at the top of the report announce increasingly more aggregable prefixes in later years than in previous years. Both this figure and the previous figure suggest that the ossification in the top ranks of the CIDR report observed earlier are due to the growing spread in netgain, making it more difficult to "unseat" high-ranked ASes because they are so far from the next nearest AS (in terms of number of prefixes). This is not necessary a problem, and also our measure of volatility may be imperfect (i.e. an AS could oscillate between two ranks, appearing volatile even though its behavior does not chance perceptibly), but this does provide some explanation for the static behavior noted before.

4.2.3 Conclusions

Our observations of characteristics of the CIDR Report related to the number of prefixes advertised by each AS, as well as the previous observations about the rank of each AS and appearance of ASes on the CIDR Report, suggest that while the advertisement of redundant, aggregable prefixes in the routing table is reasonably commonplace, the ASes that appear on the CIDR Report are outliers in that they announce significantly more aggregable prefixes than most of the rest of the population of ASes participating in the interdomain routing system. In its earlier days the CIDR Report captured a larger fraction of the ASes responsible for aggregable prefixes than it does now, because of growth in the total number of ASes participating in the Internet and announcing aggregable routes. The top of the Report has become relatively static because of the extreme deaggregation of ASes at the top of the report relative to the majority of the AS population.

4.3 Analysis of AS behavior after appearing on the CIDR Report

This section presents and discusses the results of the primary question of this thesis: *do autonomous systems change their behavior, as measured by the number of aggregable routes they advertise into the routing table (netgain), after appearing on the CIDR Report?* As described in section 3.5, data from the CIDR Report is processed to determine when ASes first appear on the report, and samples are then taken from 30 to 730 days after this initial appearance (detailed in Figure 3-3) and compared to the value at the time of first appearance. Decreases in netgain would suggest that the CIDR Report does influence network operator route aggregation behavior in the expected or intended way, while no change or increased netgain would suggest that the CIDR Report has no effect. The group of ASes appearing on the CIDR Report, and thus theoretically subject to social forces to improve aggregation behavior, will hereafter be referred to as the treatment group.

In an attempt to control for normal variation in AS behavior that does not result from appearing on the CIDR Report, a "control group" was established from ASes never appearing on the CIDR Report, as described in much more detail in section 3.5. While this group will be referred to as the control group, it is more precisely the "untreated group", given that this is a quasi-experiment instead of a randomized, controlled experiment.

We measure and present three quantities in an effort to discern the behavior of ASes appearing on the CIDR Report. The first two, absolute and relative change in netgain (Δnetgain), measure directly the change in the number of aggregable prefixes announced by an AS after appearing on the CIDR Report. The third measure, deaggregation factor (DF), is the ratio of the total number of prefixes announced by an AS to the minimum number of prefixes required to be announced by the same AS in order to implement the same routing policy with perfect aggregation. Unlike netgain, DF is a relative measure of deaggregation, and considers deaggregation in proportion to the total number of prefixes a network operator must announce to operate their network. The operational definition of DF will be described in its respective section.

All of the figures that will be presented next are a series of cumulative distribution functions (CDFs) of the quantity of interest at different periods after an AS' initial appearance on the CIDR Report. In general, these CDFs can be interpreted as that the ASes in the group improved their aggregation behavior if the curve appears to move upwards or towards the left of the $x = 0$ line over time. This indicates a greater fraction of the population has reduced its advertised prefixes. While specific issues of interpretation for each plot will be discussed as appropriate in the relevant sections below, this is a very rough rule of thumb for thinking about the figures that will now follow.

With this initial explanation, we are ready to present the results of our analysis, beginning with netgain and followed with deaggregation factor, which most clearly associates a change in aggregation behavior with an appearance on the CIDR Report. Unless otherwise noted in captions of the figures or the associated text, all of the following figures utilize all appearances on the CIDR Report from 1997-2011. Additional figures of these quantities utilizing appearances from only a portion of the date range, to discern changes in behavior over time, are included in Appendix C.

4.3.1 Change in netgain following CIDR Report appearance

Netgain, the number of aggregable prefixes advertised by an AS, was the first measure we examined for behavior change in response to appearing on the CIDR Report. This seemed to be a reasonable first measure as this is the quantity most clearly of interest with regard to routing table size and the CIDR Report. Cumulative distribution functions of changes in netgain following appearance on the CIDR Report, and a corresponding untreated control group is shown below in Figure 4-12. Change in netgain is defined as $\Delta\text{netgain}_{t+k} = \text{netgain}_{t+k} - \text{netgain}_{t+0}$, where t is symbolic of the time of first appearance, and k is one of the measurement time periods from 30-730 days

From the top plot in this figure, we can observe that slightly more than half (60%) of the AS appearances on the CIDR Report were followed by a decrease in aggregation or no further deaggregation. It also shows that, in an absolute sense, there is a greater reduction in aggregable routes than advertisement of new aggregable routes (i.e. the most-reducing 10% reduce by approximately 200 prefixes or more, whereas the most-increasing 10% increase by approximately 50 prefixes or more) with the exception of outliers. There are still significant outliers on both sides, and this plot is trimmed to exclude them.

In contrast to the treatment group, the control group shown in the bottom figure exhibits different behavior, with approximately half of the population exhibiting increased announcement of aggregable prefixes, while only 20% of the population decreased announcement of aggregable prefixes. Further, the absolute amount of increased deaggrega-

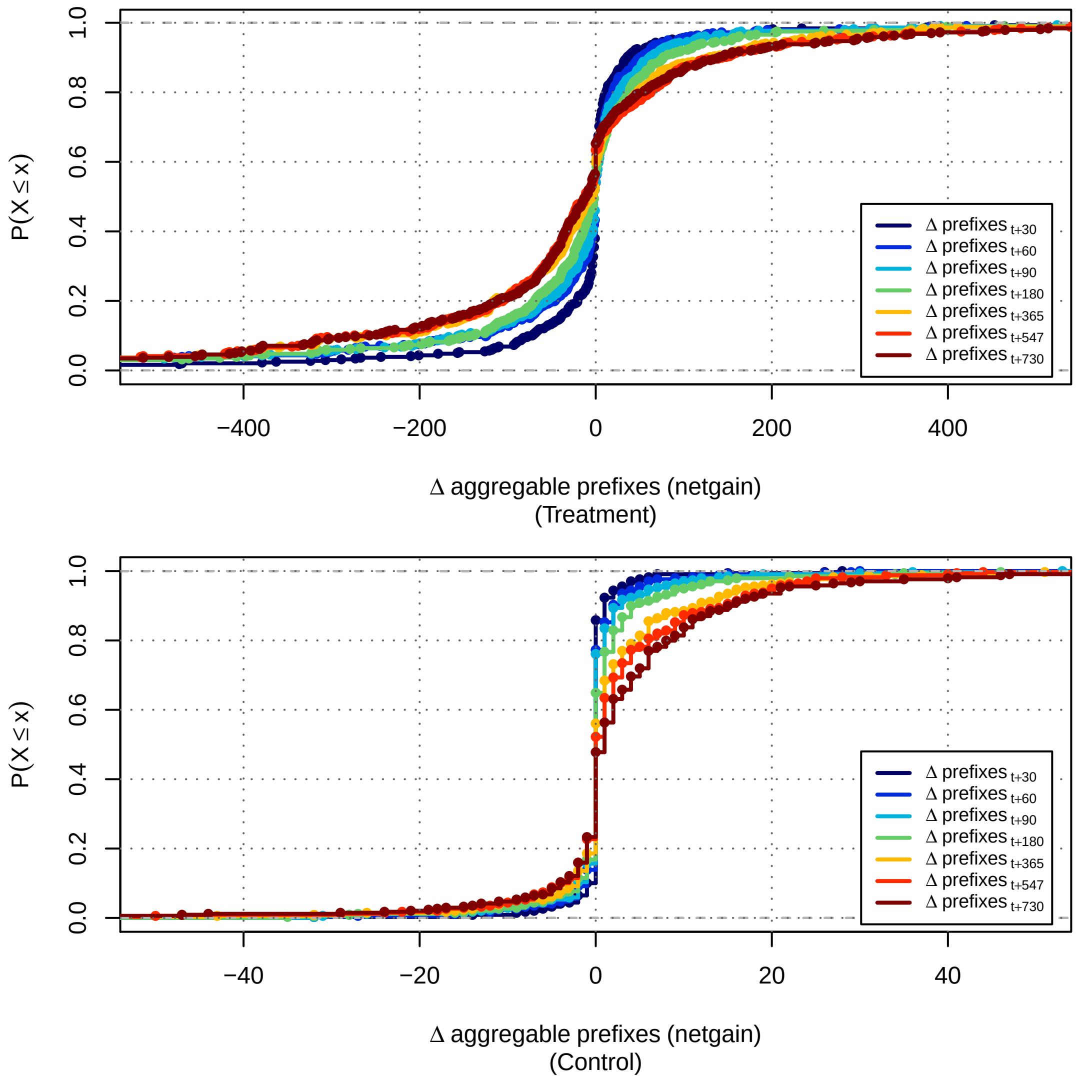


Figure 4-12: CDFs of change in number of aggregable prefixes (netgain) advertised by treated and untreated (control) ASes, for the period 1997-2011.

tion is more significant than the amount of reduced deaggregation.

While absolute measures are ultimately the metric of interest with regard to our concern about the size of the routing table and the number of slots used by each AS, they are less useful to determine the significance of increased or decreased aggregation relative to an AS' initial behavior. Thus, we present CDFs of the relative change in netgain (Δnetgain$_{t+k}$/netgain$_{t+0}$) in Figure 4-13. This figure is essentially a normalized version of Figure 4-12.

This figure makes it easier to quantify the change in netgain over time. By 730 days, ap-

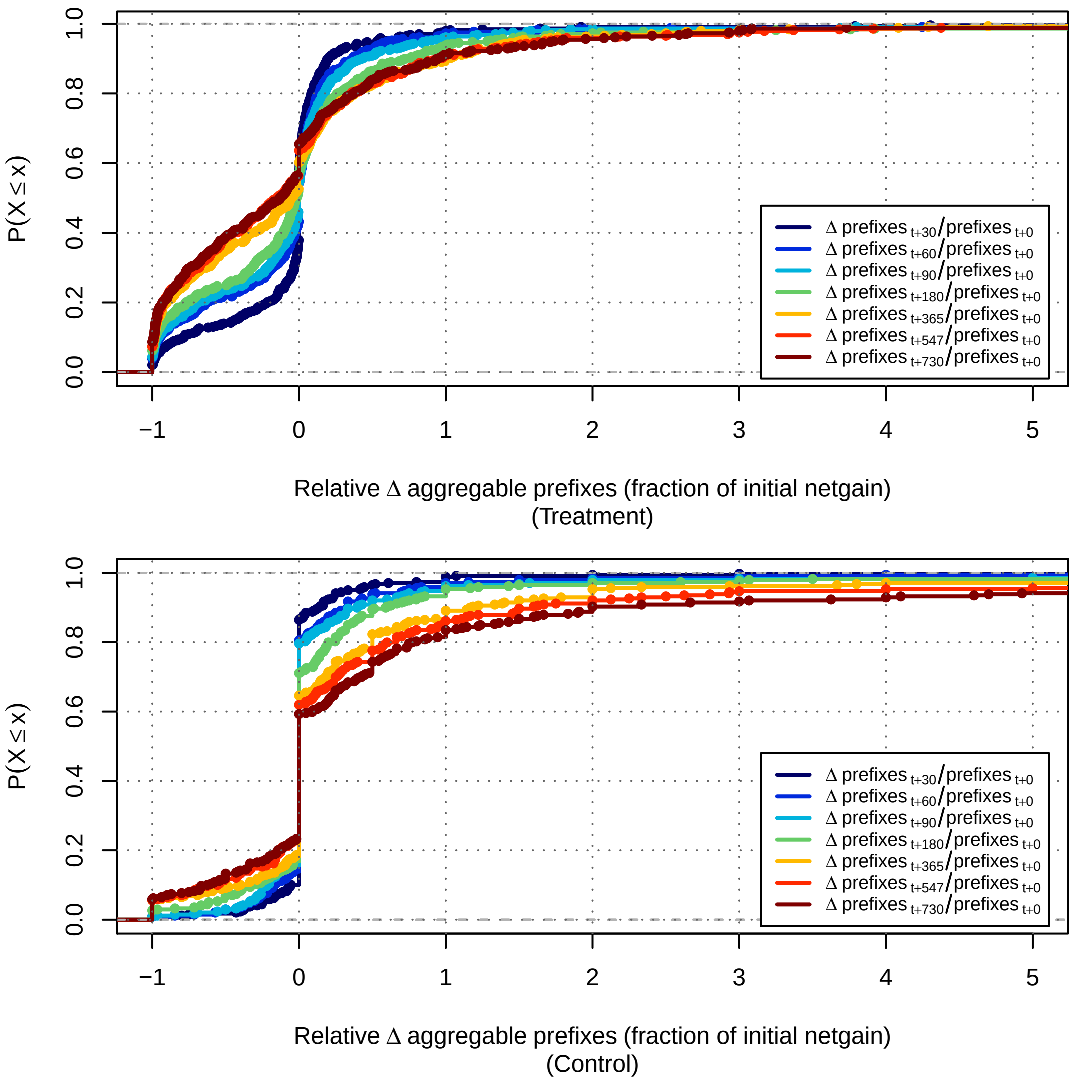


Figure 4-13: CDFs of relative change in number of aggregable prefixes (netgain) advertised by treated and untreated (control) ASes, for the period 1997-2011.

proximately 20% of the population have nearly fully aggregated, another 20% have reduced deaggregation by 50% or more, and a third 20% have held steady or reduced aggregable routes by less than 50%. In contrast, another 20% have approximately increased deaggregation by up to 50%, and the top 10% increased deaggregation by 100% or more. As noted in the previous figure, improvements in deaggregation are much less pronounced in the control group.

There appears to be a small amount of significant aggregation in the first measurement period after a CIDR Report appearance, and can be seen by looking at the $t + 30$ curve on the treatment plot above. Thirty days after appearing on the report, approximately 10% of ASes improved aggregation by at least 50%. While there was also some movement towards further deaggregation at the top of the plot, perhaps approximately 5% of the population deaggregated by at least 50%, with much more of the population remaining near the $x = 0$ line, indicating little change in behavior.

4.3.2 Change in deaggregation factor following CIDR Report appearance

While netgain provides a direct measure of unnecessary prefixes occupying the routing table, it does not take into account the total size of an AS' operations or networks, and so an AS announcing 101 prefixes which could be aggregated into one prefix would receive the same netgain score (100) as an AS announcing 2000 prefixes that could be aggregated into 1900 prefixes. While it is true that in an absolute sense, both of these configurations contribute the same number of unnecessary prefixes to the Internet routing table, the announcement of a large number of aggregable prefixes relative to one's total network is arguably less defensible or justifiable than a network that introduces a relatively small amount of deaggregation as a result of operating its large network.

To investigate this relative degree of deaggregation, we also measure the deaggregation factor (DF), the ratio of the current number of prefixes advertised by an AS to the minimum number of prefixes needed to be advertised by the same AS to implement their current routing policy with perfect aggregation[2]. More precisely, DF $=$ netsnow$_{t+k}$/netsaggr$_{t+k}$, where netsaggr$_{t+k}$ $=$ netsnow$_{t+k}$ $-$ netgain$_{t+k}$. t is symbolic of the time of first appearance, and k is one of the measurement time periods from 30-730 days. In the above example, the first network would have a DF of $100/1 = 100$, and the second would have a DF of $2000/1900 = 1.05$. Cumulative distribution functions of the deaggregation factor for the treatment and control groups is shown below in Figure 4-14. Note that unlike change in netgain, DF is not measured relative to the original point of appearance on the CIDR

[2]This configuration is preferable to a seemingly similar ratio of netgain to netsnow, our original approach, as netsnow is not independent of netgain: netgain/netsnow $=$ netgain/(netsaggr $+$ netgain), making it more difficult to reason about the causes of changes in this quantity.

Report, and so must be viewed together with the DF at the original point of appearance ($t + 0$) in order to observe changes in behavior.

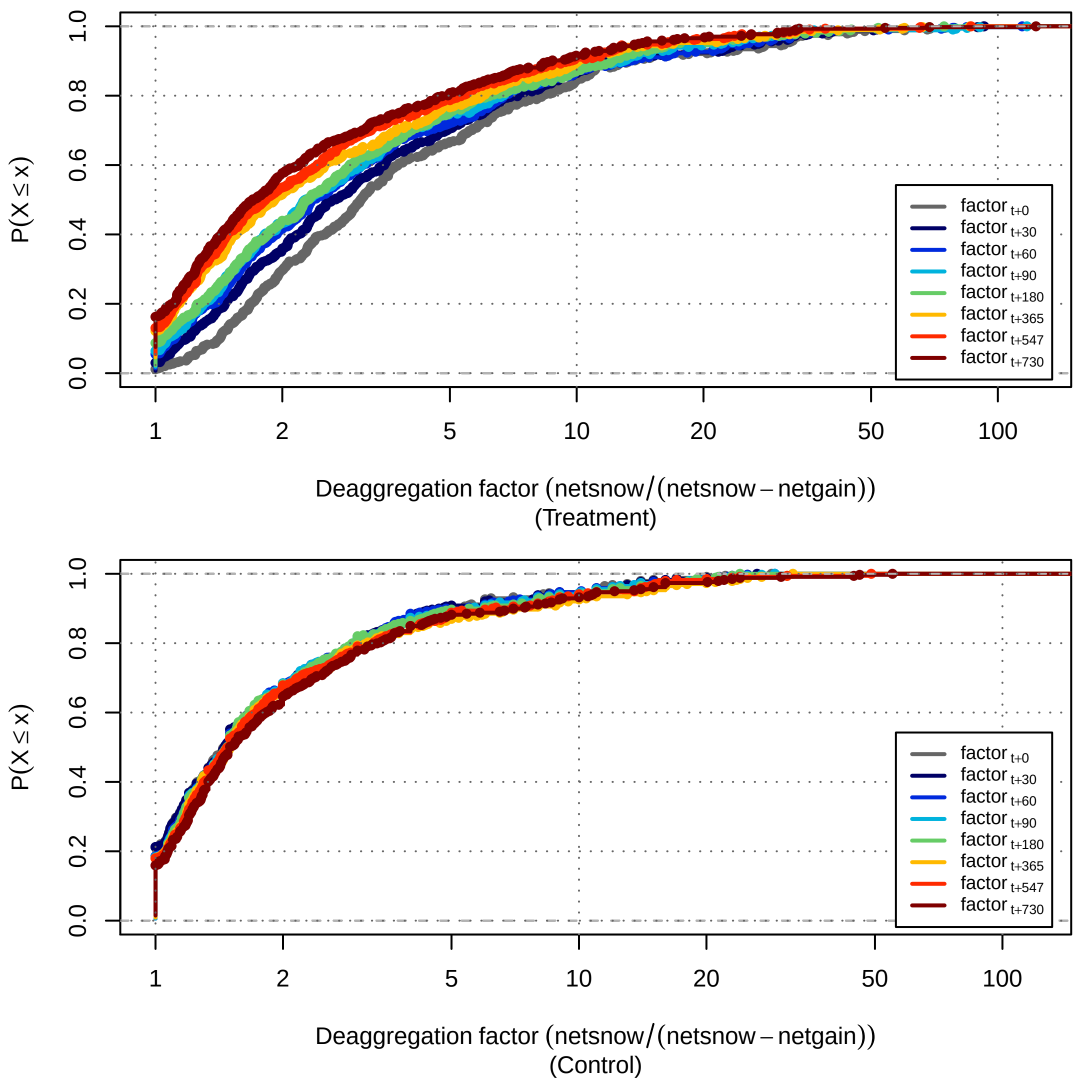


Figure 4-14: CDFs of the deaggregation factor for periods of time after an AS appears on the CIDR Report (treatment) and for untreated ASes (control), for the period 1997-2011.

This figure suggests a reduction in deaggregation factor following appearance on the CIDR Report, as indicated by upward movement of the CDF curve along vertical lines, meaning that a greater fraction of the population has a DF less or equal to the value at the vertical line than in the previous time period. In contrast to the treatment group, the control group exhibits little change, and in fact the change is a slight increase in deaggregation factor over time.

A simplified version of the previous plot, extracting only the first (t+0) and last (t+730) measurement, and presenting the curves for the treatment and control groups together is shown in Figure 4-15. In both cases, the initial measurement is the lighter color, and the latest measurement is the darker color.

This figure clearly illustrates the difference in the deaggregation factor of the ASes that appear on the CIDR Report and those that do not. It also makes it easier to observe the nature of the change over between initial appearance and the end of the post-appearance measurement period (two years following the appearance).

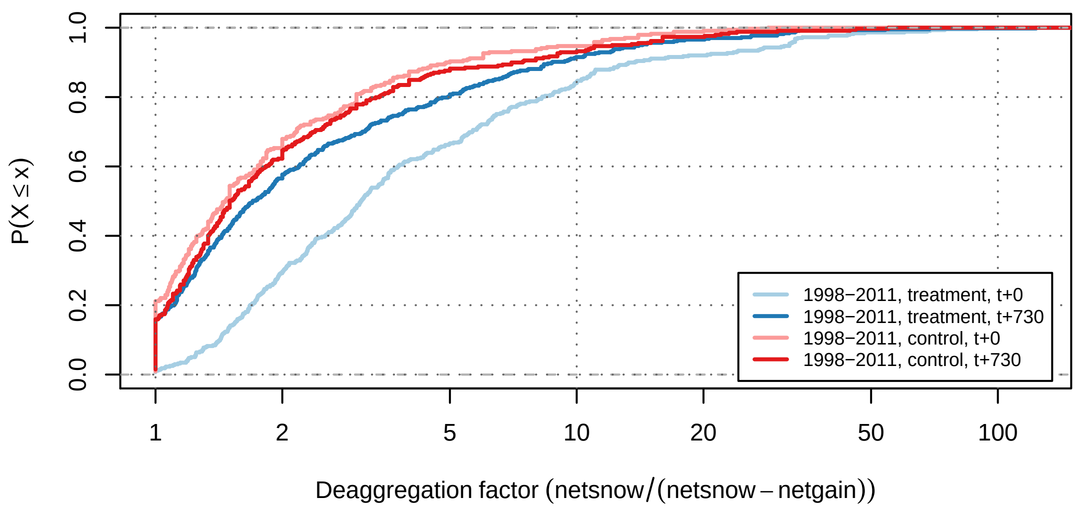


Figure 4-15: Simplification of previous figure, showing the first (t+0) and last (t+730) CDF of the deaggregation factor for both the treatment and control groups.

There is a question here about what exactly this change means, as a reduction in the deaggregation factor can be caused by one of two changes in an AS' behavior: a decrease in netgain or an increase in netsaggr (which would result from increased advertisement of non-aggregable prefixes, presumably from new address blocks). Looking back at the previous netgain figures, it is certain that some of this reduction in DF is due to reduction in netgain, though it may also result from ISPs growing their networks and bringing on new customers. The latter is an inevitable force in the commercial Internet, and while not directly encouraged by the CIDR Report (netgain will remain the same), growth of networks with a reduction in DF is still an improvement relative to growth with proportional deaggregation.

Since deaggregation factor appears to be a reasonable measure (and that clearly indicates some behavior change), we will use this measure to investigate the change in treatment effect of appearing on the CIDR Report over time. To do so, we will consider the behavior of this measure for a number of smaller time periods within the overall period for which we have data. This will allow the observation of any changes in behavior over the course of the available data. A window of three years was selected and the data set was divided into four three-year blocks, starting at the beginning of the first year and ending at the end of the last year: 1998-2000, 2001-2003, 2004-2006, and 2007-2009. Any appearances on the CIDR Report occurring in these three-year windows were analyzed and suitable untreated ASes were selected as controls, exactly as before. Plots for these time periods are shown in Figure 4-16 (similar figures for observing change in relative netgain over time are available in Appendix C).

In this figure, the light-colored curve represents the deaggregation factor at the initial measurement time (t+0) and the darker curve of the same color represents the deaggregation factor at the last measurement time (t+730 days). As we look at the way the deaggregation factor for treatment and control ASes changes over time, we can see changes in the apparent response to appearing on the CIDR Report. The treatment effect (vertical distance between dark and light lines) was most pronounced in the earliest time period, 1998-2000. The effect was roughly the same in the 2001-2003 and 2004-2006 periods, with the effect perhaps even being slightly more pronounced in the later period, especially for larger DF values. The treatment effect of appearing on the CIDR Report decreased significantly in the final time period of 2007-2009, with very little change visible over the various sampling points in the 730 day measurement window. These results would suggest that the CIDR Report may have been effective earlier in its life, but has become less effective more recently. The question of whether the CIDR Report was effective will be discussed more generally in the following chapter.

Finally, one last view of these data presents a slightly different perspective by organizing the treatment and control data from the various time periods together. Figure 4-17 shows the simplified plots for the control groups for each of the four time periods together

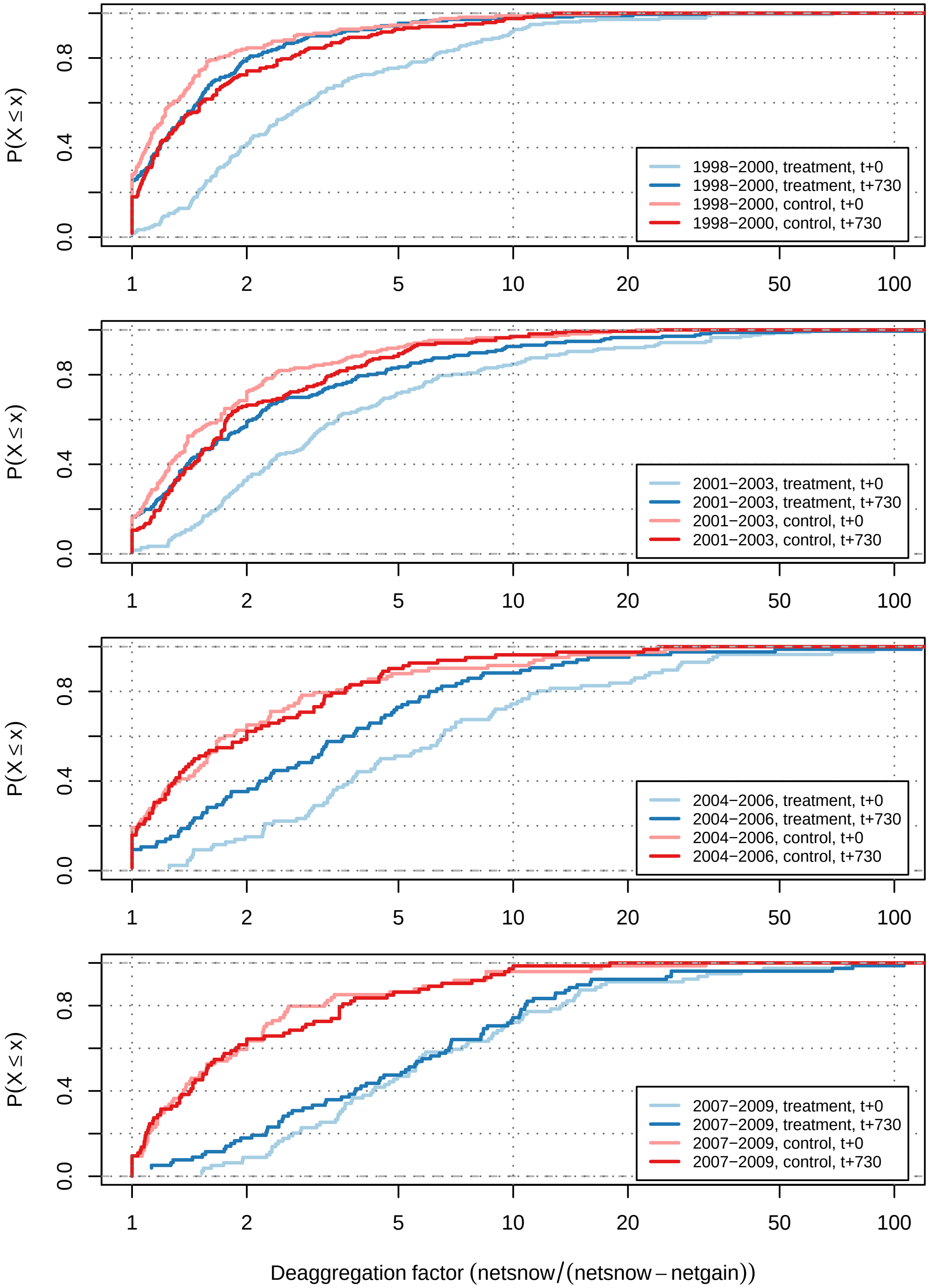


Figure 4-16: Simplified deaggregation factor CDFs for four three-year time periods over the full data availability period. In all cases, the the data ranges from the first CIDR Report of the ending year until the last CIDR Report of the ending year.

in one figure. While congested and complicated, this suggests visually that the control groups from each of the periods are in generally good agreement with each other and do not change greatly with time, except after the first period where the tendency to deaggregated decreases.

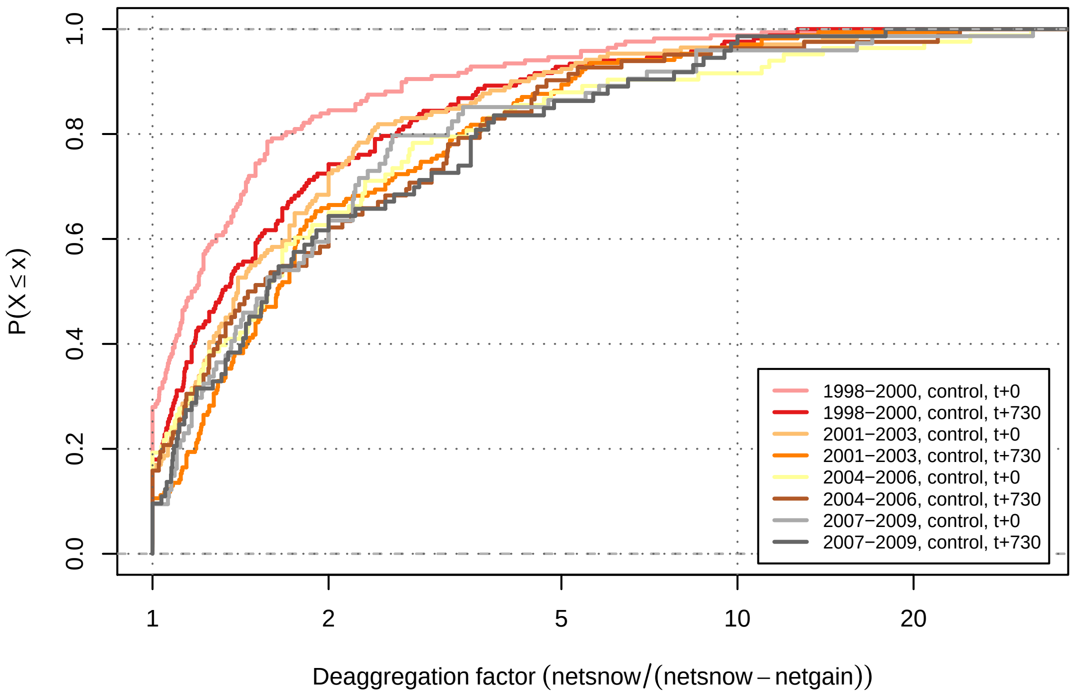


Figure 4-17: Superposition of simplified deaggregation factor CDFs for the untreated control ASes for the four time periods, showing generally similar behavior of untreated ASes throughout time.

Figure 4-18 shows the simplified plots for the treatment groups for each of the four time periods together in one figure. Unlike the control groups, this figure shows changes across time periods as well as changes in the apparent response to appearing on the CIDR Report over time. Each of the time periods is represented by a color, with the initial (t+0) measurement being in a lighter color and the last (t+730 days) measurement being in a darker color. First, in terms of the trend between the four time periods, we can see the trend of increasing deaggregation factor over time: 80% of ASes had a DF $\leq$ 2 in 1998-2000 after 730 days, and this steadily decreased across the time periods to just 20% after 730 days in 2007-2009. We can also observe the change in the treatment effect, though it is more clearly visible in Figure 4-16.

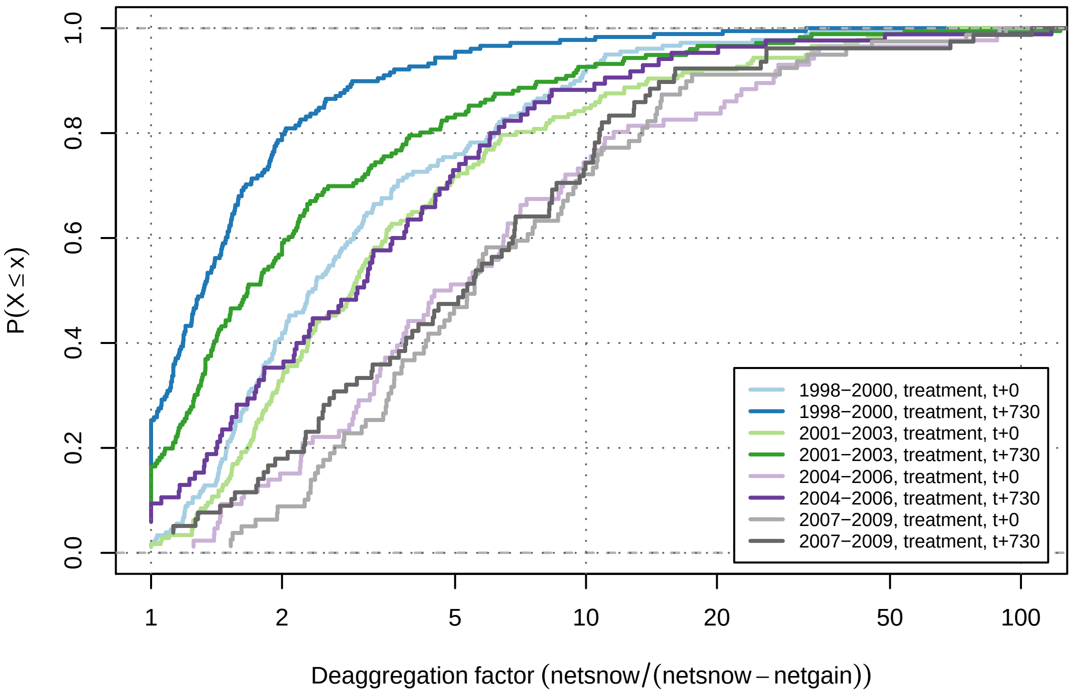


Figure 4-18: Superposition of simplified deaggregation factor CDFs for the treated ASes for the four time periods, showing a notable change in behavior of ASes appearing on the CIDR Report over time.

4.3.3 Issues of quasi-experiment validity

One of the challenges of any experiment is achieving validity—being certain that the conclusions of the experimenter that follow from the results of the experiment reflect what actually occurred in the experiment [Bab03]. This is particularly critical when attempting to make a causal inference: that a stimulus really did cause a subsequent effect. Controlling for sources of invalidity, such as characteristics of the experimental design or the situation being measured that might confound results, is relatively easy in laboratory settings, but difficult in real-world settings, and especially when a quasi-experiment is constructed after the fact using observational data. We identify and discuss a number of potential sources of invalidity, addressing the degree to which each is cause for questioning the observations made about aggregation behavior following appearance on the CIDR Report.

Selection bias Selection bias refers to the construction of a treatment and control group whose members are not from the same population, making it impossible to conclude whether the treatment caused an observed change, as opposed to some other quality inherent in the treatment or control group members. This is certainly a potential issue in this analysis of the CIDR Report, as members of the treatment group, the ASes that appear on the CIDR Report, are by definition ASes that announce more aggregable prefixes than ASes not on the report (which are used to form the control group).

It appears that there may be a difference in the population that appears on the CIDR Report compared to the population of ASes generally, given that many of the members appearing on the CIDR Report are large ISPs. They may, for instance, respond differently to the treatment effect of appearing on the CIDR Report than smaller networks. This may have also affected the deaggregation factor measure, as large ISPs are more likely to expand their network (and thus advertise new prefixes) than smaller networks. This bias was not controlled for in this study and would be generally difficult to control for because of the nature of the CIDR Report. Thus, this source of bias cannot be ruled out through experimental design. However, we may be able to make arguments about the likely behaviors of large ISPs that in turn allow us to make stronger arguments about these results.

Regression towards the mean Regression towards the mean refers to the selection of a treatment group based on their having an extreme value for the dependent variable of interest [Bab03]. Over time, members of the group tend towards the population mean for the dependent variable, appearing to demonstrate a treatment effect that in reality resulted from tending to the mean. This is again a potential issue for the CIDR Report, given that the ASes that appear on the report, and thus are included in the treatment group, are by definition the ASes that announce the most aggregable prefixes.

In the case of this study, regression to the mean would result in the appearance of improved aggregation behavior. The control group does not suffer from such a bias because of its construction by random sampling, and it suggests that the behavior of untreated ASes remains relatively static over time. However, this may not be particularly helpful because the netgain values of most ASes not on the CIDR report are low, and as in the previous discussion about selection bias, there is room to question whether the control group is representative enough because the CIDR Report selects on netgain, the post-treatment dependent variable.

Both this potential source of bias and the selection bias issue may have been better controlled with a more tightly specified control group, though it is unclear what criteria should be used to select a representative control group. One potential solution is to construct the control group from the group of ASes that are immediately below the threshold for appearing on the CIDR Report and so who are theoretically as similar as possible without actually ever appearing on the CIDR Report. Such a solution would not be as ideal as a randomized controlled experiment, but would be an improvement for this quasi-experiment.

Failure to measure pre-treatment behavior A number of the concerns above about whether the change observed in AS aggregation behavior following appearance on the CIDR Report could have been alleviated by measuring pre-treatment behavior—the behavior of ASes that would eventually appear on the CIDR Report before they appear on the Report. If a significant change in behavior was observed between pre-treatment and post-treatment, it would be reasonable to conclude that the CIDR Report did indeed influence AS behavior.

The failure to measure pre-treatment behavior was an oversight in our quasi-experimental design. However, because appearance is determined based on netgain behavior, this may not be as problematic as expected. It is not possible for an exogenous trend of decreasing netgain to have been occurring before treatment, for if an AS' netgain were larger than when it first appeared on the CIDR Report, it would have appeared earlier and thus been subject to treatment earlier. Thus, we can be reasonably confident that any decreases in netgain observed were not occurring before treatment occurred, though this does not rule out previously noted regression or selection effects.

4.3.4 Conclusions

It is difficult to causally associate appearances on the CIDR Report with reduction in AS deaggregation behavior with certainty, which would mean that the CIDR Report did indeed impact AS aggregation behavior without question. However, there appears to be a discernable correlation between appearing on the CIDR Report and decreases in the number of aggregable prefixes announced by an AS. This effect is more noticeable earlier in the study period, and decreases as time wears on, with the effect on deaggregation factor and relative netgain becomes barely distinguishable as of 2007.

Given this discernable behavior change for ASes on the CIDR Report that is both different than the control group and that decreases over time, corroborating qualitative reports of CIDR Report efficacy by network operators, we are inclined to conclude that the CIDR Report did have some effect on network operator aggregation behavior, and that this effect decreased over time. The degree to which this conclusion may be confounded by potential sources of invalidity limits our inclination to make any stronger claims, though a qualitative discussion about what this conclusion means and what may have caused the observed behavior change will proceed in the next chapter.

Chapter 5

The CIDR Report as a mechanism for inducing collective action

The figures and analysis presented thus far have provided some insight about the effects of appearing on the CIDR Report on the behavior of ASes, and general characteristics about the CIDR Report itself. However, the process of re-implementing the aggregation report and analyzing and observing AS behavior changes has also raised a number of broader and more qualitative questions about the report that we address in this chapter. We first discuss whether the CIDR Report was accurate and whether it was effective using broader definitions of these terms than were used in the analysis proper, and then discuss potential hypotheses for why the response to appearing on the CIDR Report changed over time, both as observed during the analysis and as claimed qualitatively by network operators. Finally, we consider the efficacy of the CIDR Report and Internet routing table CPR in the context of Ostrom's design principles for CPR governance institutions.

5.1 Is the CIDR Report accurate?

While part of the first section of the previous chapter was dedicated to the question of whether our implementation of the aggregation report produced similar output to that of the authoritative CIDR Report, we did not address the larger questions of whether the CIDR Report is representative of the problems observed by operators regarding deaggregation and routing table growth. These issues are important to the CIDR Report's efficacy and trustworthiness as a monitoring tool in support of the social forces influencing Internet routing and aggregation, and yet are not definitively addressed; there is simply a lack of obvious complaints or criticisms from network operators in public channels.

Is the CIDR Report's vantage point representative? The first observation we make is with regard to representativeness of the routing table and aggregation potential observed and presented by the CIDR Report as compared to what other network operators observe in their own routing tables. The nature of BGP and policy-controlled route selection and propagation make it possible for two networks to have different views and different numbers of prefixes in their "full" routing table because of the routing policies of the networks they interconnect with. Ostensibly if the CIDR Report reports behavior that does not represent what most other networks observe, the CIDR Report will be less credible and more likely to be disregarded by network operators. Unfortunately, it is difficult to characterize what a representative routing table might contain precisely because of this nature of BGP.

It would seem that the default-free zone (DFZ), the routing table maintained by the major default-free providers that form the root of our roughly-hierarchical Internet, is a reasonable place to take this measurement from. DFZ network operators must theoretically maintain the fullest routing tables in order to achieve full Internet reachability, and versions of these routing tables are shared with DFZ customers. In cases of interactions between large non-DFZ networks, such as major content provider or "eyeball" networks, routes may be observed that are not globally visible, but it is reasonable to expect problems related to excessive routes advertised between peers to be resolved via normal peering dispute resolution mechanisms.

Reasonable proxies for the DFZ have been used as vantage points for the authoritative CIDR Report since its inception, as identified in Table 2.2, as well as the analysis in this thesis. However, the DFZ routing table may still not be completely representative. Large providers receive routes from many peers and other providers, and so must maintain a RIB that is several multiples larger than the DFZ FIB, which is composed of only the best routes from the RIB. Further, the route export policies of ISPs adjacent to the vantage point for the CIDR Report may also affect the degree to which the CIDR Report represents most providers' routing tables. For example, in our analysis using Route Views as a vantage point, there were cases where the DFZ and other large providers all reported similar numbers of prefixes for a particular AS, while a smaller Route Views peer reported an order of magnitude more prefixes. These additional prefixes appeared to be in support of traffic engineering, and the small Route Views peer that was "leaking" these prefixes into Route Views was likely not respecting a NO-EXPORT policy that had been applied to these routes from their origin AS[1]. Thus, without knowing the configuration of the export policy of peers and other providers that contribute a large fraction of the routes to the BGP view of the CIDR Report's vantage point, it is possible that a non-representative view of the routing table may be generated, leading to a non-representative CIDR Report.

Is all deaggregation equally problematic? All route announcements that consume slots in the routing table incur the same cost—they consume a fraction of the router's resources that cannot be utilized by another route. However, as the Internet has changed in purpose and structure over time, the need to perform certain tasks (multihoming, TE) in BGP that result in route deaggregation and increased numbers of prefixes in the routing table have been motivated by network management and engineering. The norms of the Internet operations community appear to have recalibrated accordingly, seemingly assigning different values to the benefit and justifiability of different deaggregation-inducing behaviors.

While there are different views on the subject [Li11b], most seem to conclude that DFZ-visible route announcements due to multihoming and prefix hijacking prevention are unavoidable and justifiable, while traffic engineering is less so, especially in the case of fine-grained TE by large residential access ISPs [Ste10]. Worst of all is the failure to aggregate or announcement of "class C" blocks in the era of CIDR. Such behavior indicates a lack of "clue" and provides no benefit to anyone, and so is universally deplored.

Distinguishing between these behaviors is often difficult, as the information required to discern the intent of deaggregation is not always available from the vantage point routing tables accessible for measurements. [BGT04] present a definition for measuring multihom-

[1] This problem was confirmed to affect the authoritative CIDR Report as well, through indirect correspondence with Patrick Gilmore.

ing and [CMU+10] present some measurement techniques for observing traffic engineering. However, other forms of traffic engineering between large ASes may not be visible because the TE is based on routing parameters such as NEXT_HOP that are not visible from adjacent ASes [Ste10].

Regardless of the measurability of these various phenomena, the CIDR Report does not attempt to measure or distinguish between any of the causes or underlying intentions of deaggregation, instead simply reporting the number of prefixes announced by each AS that appear aggregable from the CIDR Report's vantage point. While this is a true representation of each AS' contributions to the routing table, it does not accurately reflect the relative values attributed to aggregation by network operators. Thus, some operators may conclude that the report is not as useful because it does not distinguish between these behaviors. Further, because some of these purposefully-deaggregated routes are difficult to withdraw, they may lead to some ASes appearing in unmoving positions on the CIDR Report, as hypothesized in [Ste10].

5.2 Is the CIDR Report effective?

While the section of the previous chapter analyzing the post-treatment behavior of ASes appearing on the CIDR Report addressed the question of whether the CIDR Report appeared to correlate with AS behavior change, it did not address the more general question of the CIDR Report's purpose and use by operators: initially to encourage aggregation following the adoption of CIDR and BGP4, and then later and more generally to encourage efficient route announcements by providing monitoring information about the greatest deviation from community norms governing route deaggregation.

While the validity concerns identified previously limit our ability to draw strong, causal conclusions, it does appear that the CIDR report had some effect on the aggregation behavior of individual ASes, especially early on in the study period. Further, in the very early days of CIDR, there is evidence of the total number of prefixes in the routing table actually decreasing over time as aggregation occurred. However, during the period studied in this thesis, the routing table generally grows continuously over time, as was shown in Figure

1-1.

Considering the counterfactual (no CIDR Report) world The ideal measure of general CIDR Report efficacy would be to observe route announcement and aggregation behavior in the counterfactual world where the CIDR Report does not exist. While we cannot actually observe this, we can construct an estimate of the counterfactual routing table under the assumption that without a CIDR Report, networks might not withdraw announcements, instead leaving them in the routing table. By observing the cumulative number of prefixes advertised for at least some period of time, we could construct an upper bound of the counterfactual routing table size. A plot of these counterfactual routing table sizes is shown in Figure 5-1.

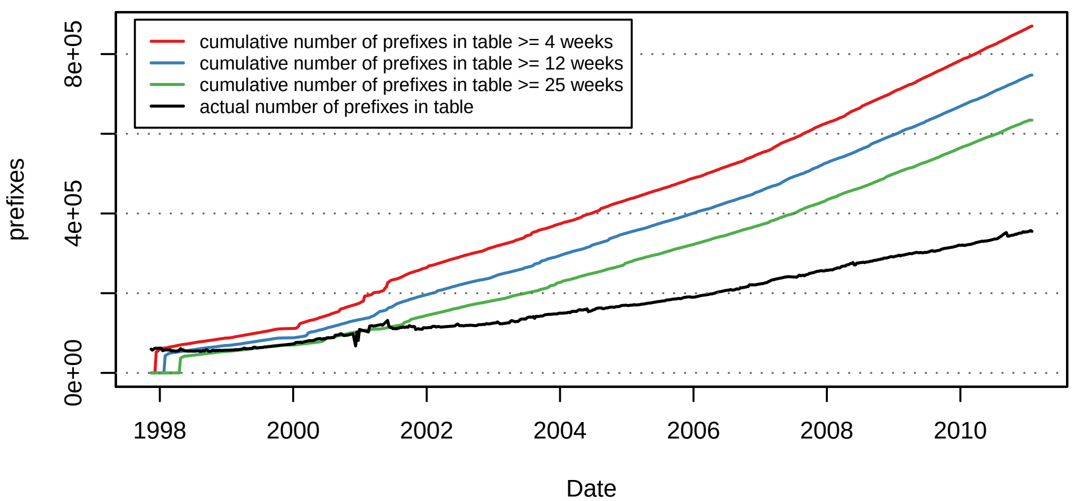


Figure 5-1: Counterfactual routing table prefix counts over time.

As illustrated in this figure, the number of prefixes in the counterfactual table size grows faster and eventually larger than the actual routing table size, even for conservative estimates of how many weeks a prefix must be in the routing table for it to constitute a "permanent" prefix. While the construction of this counterfactual is far from methodologically perfect, it provides further evidence that the CIDR Report or some other aspect of the network operations community had some effect on route aggregation behavior. The deviation between this counterfactual plot and the actual size of the routing table could alternately be interpreted as evidence that there is a tendency for aggregation or route withdrawal to occur naturally without appearing (or fear of appearing) by the CIDR Report, though this is disputed by the behavior of the control group in the previous chapter.

How quickly should the CIDR Report take effect? This is a minor observation, but there is a question that arises from observing the metrics selected for measurement in the previous chapter over time, such as in Figure 4-13. From this and other figures presented over the various measurement periods from 30-730 days after the initial appearance on the CIDR Report, we can see that some change occurs after 30 days, and often continues up to 730 days after, though most of the observed change appears to occur between 30 and 365 days rather than than 365 and 730 days (one and two years). It seems reasonable to assume that because most of the net change was observed to take place in the first year after the initial CIDR Report appearance, it was likely related to appearing on the CIDR Report and not some other phenomenon such as regression to the mean. However, this argument is somewhat speculative in that we do not understand the underlying mechanisms that cause the behavior changes observed, and thus what "time constant" they might operate under.

5.3 What caused the decrease in treatment effect over time?

If we agree that there was a behavior change by ASes in response to appearing on the Report, then as we see the decreasing difference in behavior change over time in the previous chapter, we conclude that the governance system influenced by the CIDR Report became less effective over time. This agrees with qualitative observations by network operators about the decline of the CIDR Report. However, these observations do not offer any insight about what may have occurred within the interdomain routing system and the Internet operations community surrounding it to cause these changes.

Ostrom's model of rational appropriators, introduced in Chapter 1, sets out four factors that influence the behaviors and decision-making of appropriators. These are:

- internal discount rate: the relative perceived value of future benefits versus present benefits

- internal norms: internal values that influence the selection and relative valuation of strategies (which may in part be the result of internalizing external or community norms)
- external costs: costs incurred from a particular strategy
- external benefits: benefits gained from a particular strategy

By considering external changes and events over the course of the study period that may have influenced these factors, we can develop hypotheses for the observed behavior change. While we cannot conclusively attribute the observed behavior change to any or all of these hypotheses, they are certainly potential causes of the decreased efficacy of the CIDR Report and routing table governance system that could be investigated further.

In Ostrom's work, the discount rate is typically used to describe the ease by which an appropriator may make their livelihood from another resource system, thus freeing them from considering the future consequences of opportunistic behavior in the current resource system. In the case of the routing table and the Internet more generally, there is only one resource system, and so we will not consider the internal discount rate, instead focusing on the other three factors in turn below.

Changing community norms and responses While the criterion that Ostrom identifies here is internal norms, it could be argued that most of the internal norms of participants in the Internet community are motivated by the collective norms of the community. We thus focus on these specifically.

In its very early days, the Internet operations and engineering community was relatively small and homogeneous. Consisting of groups such as the IETF or the NSFNET Regional Techs group (the forerunner of NANOG), these groups were relatively small, met in-person frequently, and otherwise kept in close contact via email mailing lists. These groups were composed of relatively consistent participants that were mostly American and affiliated with equipment vendors, service providers, contractors, and organizations that operated Internet-connected networks—mainly academic and educational institutions. Individuals and their organizations had reputations to maintain in the Internet meritocracy, and without motivation by commercial and competitive forces, this community was more focused on collective welfare than on individual benefit [Li11b].

As the Internet has grown over time, the operations and engineering community surrounding and supporting it has also grown and changed. First, as the Internet has grown in importance and deployment outside of the United States, the community of operators that participate in BGP operations has grown to include a much more diverse group of people that no longer meet together or participate on the same mailing lists, or who even speak the same language or hold the same cultural norms and values. While there are still strong and vibrant operator communities such as NANOG or RIPE (Réseaux IP Européens), these are all generally regional organizations rather than global ones. Further, and perhaps more important is that with the evolution of Internet service provision into a highly competitive commercial activity, considerations of the benefit or cost of a particular activity to the community are no longer considered, or at best considered after profit-maximizing interests [Li11b]. Under such a mindset, reducing one's impact on the routing table would at best not be in support of an ISP's core business and at worst could be viewed to harm their Internet operations, and so aggregation has become less of a concern to large, modern ISPs.

Finally, the attitude that the CIDR Report had become ineffective, espoused by many of the operators that we spoke with, may have also played a role in the change in the decrease in treatment effect of the CIDR Report. Given that it only provides information, the CIDR Report may only help improve aggregation of the routing table through social action in the Internet operations community. If leading operators deemed the report to have become ineffective, then there would be little reason for them to continue to apply pressure via social forces, as it would be a waste of time, or others would free-ride off of their efforts. Thus, if this attitude was pervasive, it may have been a self-fulfilling prophecy that changed attitudes and responses to the CIDR Report.

Changing benefits of routing deaggregation As discussed earlier, there are now greater engineering motivations to utilize deaggregation to achieve reliability or performance goals via multihoming, traffic engineering, and separate announcement of critical infrastructure addresses to avoid route hijacking. While there are questions about the degree to which this behavior has changed over time [CMU$^+$10], these behaviors are certainly considered a normal part of BGP network operation now whereas there were arguably less critical and more unusual in the era of a more hierarchical Internet [LIJM$^+$10].

Perversely, it has also been claimed (but not substantiated) that appearing on the CIDR Report has been viewed by some as a positive indication for marketing purposes [SEH06], presumably because it suggests that one operates a large network, like others appearing on the CIDR Report.

Changing costs of routing deaggregation The costs of routing deaggregation also appear to have changed over time, both in terms of the marginal cost of a routing table slot in a modern router and in terms of the community response to operating a highly deaggregated network. First, in terms of the marginal cost of a route, typical modern routers now have capacity for several million routing table entries. This demand for routing table slots has apparently been driven by the popularity of BGP/MPLS VPNs [RR99] and the business models they enable for ISPs, leading to the present where major customers of a prolific router vendor do not seem to express concerns about the Internet routing table's contributions to the total demand for routing table slots [Dav11]. This has made each routing table slot less valuable over time. While aggregable prefixes still consume a significant fraction of the total prefixes in the Internet routing table, the benefit of taking action to encourage others to aggregate their routes has likely decreased over time as slots have become less scarce.

In addition to the decreased cost of routing table slots, it appears that the community response to announcing aggregable routes has decreased over time, making it less costly in terms of reputation and criticism to announce deaggregated prefixes. It is not possible to observe a measure of all criticism in response to the CIDR Report, but we can consider a reasonable proxy—the public discussion of the CIDR Report on the NANOG mailing list. A figure of the frequency of mail messages containing "CIDR R" in the subject line over the study period is shown below in Figure 5-2.

This figure illustrates that while email discussion of the weekly CIDR Report was fairly vigorous (including both commendations for improvement and criticism for getting worse),

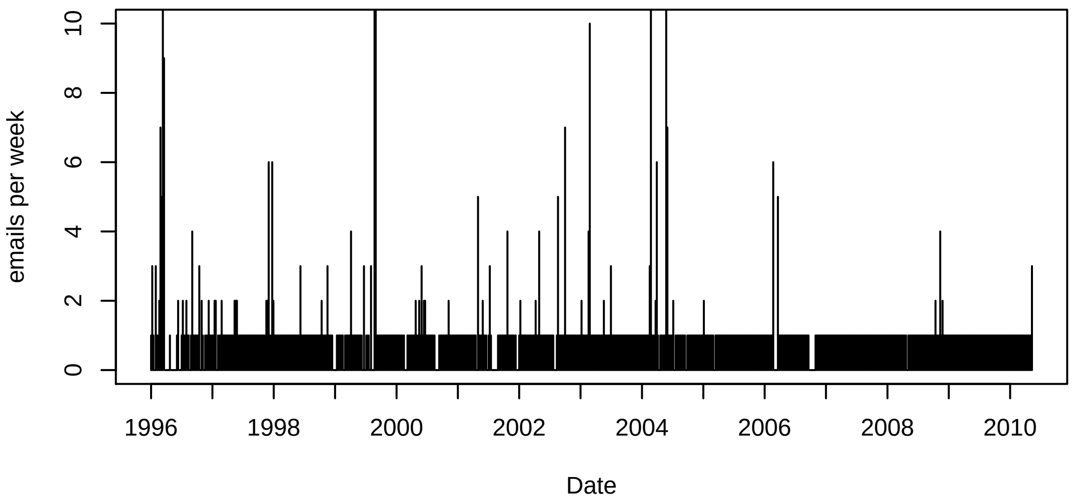


Figure 5-2: Incidence of email messages containing "CIDR R" in the subject line, transmitted to the NANOG mailing list over 1996-2011. Note that the baseline of one message per week is due to the authoritative CIDR Report.

these comments and conversations have essentially stopped completely as of late. While this is probably coupled to the changes in community norms discussed above, it also suggests that the criticism and thus reputational "cost" of appearing on the CIDR Report has decreased over time.

5.4 Learning from the CIDR Report and the routing table CPR system

In spite of our discussion of governance institutions and Ostrom's CPR framework at a number of points throughout this thesis, it is important to note that the CIDR Report itself is not a CPR governance institution as Ostrom's framework defines them. As noted in the introduction, such institutions require participants to agree to make commitments to certain behavior, and typically offer graduated sanctions to prevent opportunism in the event that monitoring activities show deviation from commitment. In this context, the CIDR Report fills a monitoring role that feeds into the loosely defined governance customs that have been developed by network operators participating in the Internet.

Ostrom [Ost90] identifies five design principles[2] necessary for participants in a CPR to make credible commitments to follow agreed-upon rules. We discuss principles particularly relevant to the CIDR Report and the routing table governance institution below, while only touching on principles that are less relevant or already satisfied.

Collective choice Perhaps the biggest challenge facing both the CIDR Report and the community norm-based approach as a method of managing routing table growth is that none of these mechanisms or the rules or principles that underlie them were ever explicitly agreed upon or developed through a collective process by participants in the interdomain routing system. Thus, when the community shared a collective set of norms and principles with regard to maintaining reputation, minimizing deaggregation, etc., the CIDR Report provided a monitoring service according to these norms and identified ASes that deviated from them most significantly. With this information, community members who shared these norms could then apply social pressure or assist the deviating ASes in improving their aggregation behavior.

However, as these norms changed and the community also changed, there was no precedent or process by which to achieve collective action or agreement to adjust the rules to meet these changes and the interests of network operators accordingly. Thus, the CIDR Report could not necessarily evolve to meet the interests of all, and there was no obvious venue for network operators to participate in order to change the CIDR Report or the rules and expectations of the community. Ostrom would suggest that these reasons lessen the incentive for individuals to commit to the rules of the routing table CPR, and ostensibly also played a role in the apparent reduced efficacy and relevance of the report over time.

Monitoring The design principle most relevant to the CIDR Report is that of monitoring, as it is the primary role of the CIDR Report. Monitoring, along with sanctions, are essential to maintaining cooperation and limiting opportunistic behavior in the long-running CPR

[2] The five identified for credible commitment are: clearly defined boundaries, congruence between CPR rules and conditions, collective choice arrangements, monitoring, and graduated sanctions. Ostrom also identifies three other principles that, when absent, have caused failures in other CPR governance institutions: dispute resolution mechanisms, recognition and non-interference of the right to organize, and the use of nested enterprises in large-scale systems

institutions that Ostrom observes, even in communities with shared norms and interest in reputation maintenance. Internal monitoring and sanctioning provide assurances to participants making commitments to abide by the rules that their trust will not be taken advantage of—that opportunists will be detected and punished.

In consideration of Ostrom's principles, the CIDR Report is a reasonable monitoring mechanism: it is provided by and can be verified by participants, and is of low cost relative to sanction mechanisms. However, it is not ideal, especially as it has remained static while the Internet has evolved. In addition to the specific issues of vantage points, accuracy, and interpretation mentioned earlier, the CIDR Report only identifies the thirty most extreme deaggregators in the network each week. While a limit on the number of networks identified each week is arguably necessary because the social response to the CIDR Report is a human process that can only cope with a limited amount of information, it has not scaled as the Internet has grown. Instead, as illustrated in Figures 4-8 and 4-9, it now focuses only on significantly deaggregated outliers—0.1% of the routing system participants with 20% of the prefixes.

This group of outliers may indeed be deserving of community attention, but the limit of 30 ASes also ignores the next 20% of the population that is collectively responsible for the remaining 80% of the aggregable prefixes. These participants may be significantly deaggregated in relative terms and may be in positions to easily improve their aggregation behavior, but because of their relative size will never appear above the threshold of the top 30. With the current and historic trend of a constantly increasing minimum threshold (in terms of aggregable prefixes announced) to appear on the CIDR Report, an AS just below this threshold could easily avoid social pressure and criticism while not committing to the broader norms that the CIDR Report embodies.

Graduated sanctions Given that the CPR governance institutions observed by Ostrom typically involved communities with frequent and repeated interaction, her studies found that graduated sanctions—small sanctions for first or infrequent offenses and larger sanctions for later or more frequent deviations—were important for achieving a stable, cooperative governance institution. Sanctions, while more expensive than norms in establishing order, help to “backstop” norms in cases where the benefits of opportunism exceed the costs of contravening norms. The CIDR Report and routing table CPR system has two significant shortcomings with regard to sanctions.

First, there are no strong, guaranteed sanctions in response to appearing on the CIDR Report. Beyond the shame or fear of reputation loss motivated by internal norms in response to appearing on the report, there is only the explicit criticism of others to motivate changes in aggregation behavior. This is not coordinated in any way, and instead relies on other participants to respond to the CIDR Report mail message by applying pressure or criticism via public email response, private communication, or in-person discussion at the next operator meeting (e.g. NANOG). It is possible that operators could use existing business relationships and contracts to apply leverage to other networks to improve their aggregation behavior in a way that is private, but this was not encountered in our study of mailing list archives or discussions with operators.

In addition, these social sanctions are not particularly graduated. They are not applied consistently across all ASes, nor are they applied in an organized fashion, but again rely on other operators to apply pressure as they choose to or are able. Further, while it is doubtful that an AS ranked in the 30th position on the CIDR Report receives the same attention as the first-ranked AS, their identification is at least publicized, compared to the AS ranked 31st who may have very similar behavior but not appear at all on the report and thus escape publication.

Other design principles The CIDR Report and associated norms of the Internet routing table do not incorporate a number of Ostrom’s other design principles:

- Clear boundaries: While the interdomain routing system is limited to BGP speakers, there is no boundary that limits routing slot appropriation to those that adhere to the community norms only.
- Rules suited to actual conditions: As alluded to earlier with regard to collective choice, the norms embodied in the CIDR Report have not evolved as routing table slots have become less scarce and deaggregation-causing activities are common

BGP operations, and so probably no longer represent the norms and concerns of the community with regard to the size of the routing table.

- Nested enterprises: This technique is used to achieve scalability in CPR governance institutions by organizing local institutions that then participate in a larger regional/etc. institutions. Such nesting may occur informally via regional operator groups and RIRs, but the CIDR Report is a global activity.

Conclusion Given these observations that the CIDR Report and the loose, norms-based governance approach used by participants in the interdomain routing system do not implement many of Ostrom's empirically-derived design principles for long-standing CPR governance institutions, it should probably not be a surprise that the report has become less effective over time. It is difficult to compare the CIDR Report to other CPR situations, but perhaps it should be considered remarkable that the CIDR Report was effective for as long as it was without a solid foundation as Ostrom would have prescribed. Regardless, it seems apparent that the lack of some of these principles, such as collective choice and sanction mechanisms, would have been important in allowing the CIDR Report to remain effective.

Chapter 6

Related work

The important central role played by interdomain routing in the Internet, along with the problems that many have seen emerge in this space over time, has caused this topic to be a focus of much work on the part of academic researchers, network equipment vendors, and network operators. This work can be classified as that which is specifically focused on the problem of growth of the routing table due to network operations, as well as work more broadly focused on the problems of our current interdomain routing architecture. Both of these topics are discussed in more detail in following sections. A section on the limited literature and other information about the Internet operations community and its norms, ethos, and characteristics is also included because of the importance of this community in producing the social forces that would be required to make the CIDR Report effective.

Note that sources from the Internet operations community (e.g. presentations from NANOG meetings, mailing list postings, etc.), while perhaps less academically credible because they are typically not published in traditional peer-reviewed forums, are included here and elsewhere in this thesis because they are extremely valuable for gaining a more thorough and comprehensive understanding of this topic.

6.1 Analyses of routing table growth & deaggregation

Some of the work in this space has been devoted specifically to the topic of the Internet routing table, arguably because it is both the immediate source of scaling problems and because it is something that researchers, vendors, and operators all think about. This work has generally focused on characterizing the growth of the routing table over time and attempting to understand the causes of growth. These works often mention the CIDR Report as a source of information but do not discuss its efficacy.

Beyond some of the ad-hoc analysis that took place during the supernetting [FLYV92] and ultimately CIDR [FLYV93] design and deployment, [Hus01] is one of the first works analyzing the Internet routing table. In this paper, Huston analyzes the size and growth of the BGP routing table from 1988-2001 and identifies four phases of growth surrounding the deployment of CIDR. He finds that while CIDR deployment did reduce the size of the routing table and limit the exponential growth preceding its deployment to linear growth shortly after, growth did ultimately return to exponential again as the Internet grew in commercial importance starting in the late 1990s. While not specifically determining the contributions of various factors to the observed growth, he does note that traffic engineering and multihoming are increasing over time. Huston concludes by casting the problem as a "tragedy of the commons", lamenting that the usual solution of a central regulator is not obviously available to a globally distributed system like the Internet. Interestingly, or perhaps tellingly, Huston does not acknowledge the CIDR Report as a regulatory mechanism for this role.

Bu et al. [BGT04] go on to analyze the routing table with a focus on determining the contributions that a number of operational behaviors have on routing table growth: load balancing (traffic engineering), multihoming, RIR address fragmentation, and failure to aggregate. Like this thesis, they also use Route Views as their data source, verifying that it provides a sufficiently comprehensive view of the Internet routing table. They perform an analysis of aggregation potential similar to both the CIDR Report and the analysis of this thesis, though they perform aggregation solely based on routing policy, rather than hierarchical prefix aggregability as well. The authors of this paper conclude that address fragmentation caused by RIR allocation policy was the greatest source of prefixes in the routing table as of 2002, with multihoming, traffic engineering and failure to aggregate contributing 30%, 25%, and 20% additional prefixes respectively. They also find that traffic engineering was the fastest-growing cause of routing table growth.

A more recent study by Cittadini et al. [CMU+10] provides an update about the state of routing table deaggregation and update dynamics, and also dispels some myths about where the blame lies for deaggregation in the DFZ. This is a useful and welcome update given the changes in Internet business relationships and operational concerns that have occurred since the previous studies. In spite of the Internet's evolution, they find that neither the fraction of the routing table that is deaggregated nor the fraction of ASes performing traffic engineering has changed noticeably since 2001, but instead has remained proportional to the size of the routing table. They also go on to analyze the distribution of deaggregation in different populations (i.e. ISPs, enterprises, etc.) and find that while a disproportionately small number of ASes advertise most of the routes in the DFZ, the proportion of these "bad guys" has not changed over time. These conclusions are interesting and present a more reasonable perspective on the data analyzed by Bu and Huston in that they consider the results relative to the routing table.

In good contrast to the more academically-oriented work in this space, [ZPW+01] is written by a trio of experienced network operators and describes operational concerns regarding routing table growth, as well as the challenges in implementing any proposed solution in a production service provider environment, such as the business case for upgrades and the typical 5-7 year depreciation cycle for routers and other equipment. In [Ste10], Steenbergen presents another operator's perspective, taking up a number of similar hypotheses regarding the cause of table growth as in [BGT04] and [CMU+10], illustrating how these factors are indeed present in the table, and discussing operational solutions to minimize these problems. Finally, [SEH06] presents the conclusions of a more qualitative study by a European network operators group, which did mention the CIDR Report and the "CIDR Police" (discussed below) as mechanisms to mitigate routing table growth.

6.2 Improving interdomain routing scalability

Much more of the work in this space, mainly by academics and researchers, has been focused on the broader problems with our current interdomain routing architecture. Along with problem statements, a number of technical solutions have been proposed, both incremental and radical (involving full protocol changes or clean-slate redesigns), but most have yet to gain any traction within vendor and operator communities. A few non-technical solutions based on economic incentives or operational practices have been proposed or attempted to resolve the more specific routing table growth problem. These appear to be far less popular than technical solutions to the problem, and while these ideas are often articulated in operator forums, they also have not gained any traction.

At a high level, [YMBB05] captures a number of challenges and proposed improvements to interdomain routing within the context of arbitrary AS topologies and a BGP-like routing algorithm. [Han06], in the context of describing a number of broader issues facing the Internet architecture, discusses the scaling and reliability challenges of interdomain routing and the even larger challenges of deploying a replacement until the incentives are sufficient to do so. In [FBR04], Feamster et al. discuss a number of specific problems in BGP introduced by the design goals of policy routing and scalability, with the intent to direct research towards designing more robust interdomain routing protocols. The IETF and IAB have also studied the problem of routing scalability more broadly in [MZF07], with the understanding that some of the scaling issues may come from our current architecture and produced broader and more forward-looking problem statements and criteria that proposed solutions to the routing scalability problem must meet.

A number of small proposals to incrementally alter BGP-speaking routers have emerged, such as routers that discard information and retrieve it later if necessary [KR06], and routers that aggregate their FIB internally, similar to the aggregation performed by the CIDR Report, to conserve memory without altering routing policy [ZLWZ10]. Operational techniques are also available to operators, including route filtering based on RIR policies to limit more specific routes used for traffic engineering [BBGR01], although this technique is likely not to be effective any more as RIRs have continued reducing their minimum allocation size with the exhaustion of IPv4 addresses. Finally, routers can also be configured to perform virtual aggregation, such as in [BFCW09], where a routing table is effectively spread across multiple routers, thus performing full lookups in two or more hops instead of one. All of these options are attractive for individual providers as immediate benefit in terms of routing table reduction can be realized without requiring coordination with others or a protocol change, and are likely to be easily implemented with a software or configuration update. However, while all of these approaches reduce the routing table size and growth rate pragmatically, they do not alter BGP's fundamental scaling characteristics, particularly with regard to update messages.

In terms of architecture-level proposals, many solutions to the problem propose to overcome the overloaded use of IP addresses as both topological locators and endpoint identifiers. It is argued that eliminating this overload would allow traffic engineering, multihoming, and other behaviors to occur efficiently while reducing the size of routing table, as aggressive hierarchical aggregation would be possible [QIdLB07]. There have been a number of proposals in this space, but one of the more well known efforts oriented at real deployment is the Locator/ID Separation Protocol (LISP) [FFML11], which maintains two separate (IP) address spaces for identifiers and locators, defines mappings between them for interdomain routing (i.e. which locator is the destination for a given identifier), and encapsulates packets with this appropriate mapping for interdomain routing. A more recent alternative that appears more favorable for incremental deployment is the Identifier-Locator Network Protocol (ILNP) [ABH10], which utilizes the current IPv6 routing architecture along with DNS to provide locator-identity separation. More radical proposals such as HLP [SCE$^+$05] and NIRA [YCB07] are not incrementally deployable but offer new architectures and models that are free from the warts of BGP and its implementation and operation.

In contrast to CIDR and many of these other proposed schemes, including locator-identifier separation, that rely on the aggregation properties of hierarchical routing, [KcFB07] claim that we need a fundamentally different approach to routing on the Internet as any hierarchical routing algorithm will not scale well when applied to scale-free "Internet-like" topologies. They argue instead for *compact routing* that by design scales sub-linearly in the worst case, at the cost of selecting paths longer than the shortest path between any two nodes. They make their argument from a mainly theoretical standpoint, rather than a more engineering-oriented perspective of pragmatically scaling the Internet to the next order of magnitude or within the bounds that are likely to be provided by Moore's Law and advancing router architectures. Also, they mainly consider shortest path routing rather than policy routing, which is of less interest in the context of interdomain routing on the Internet.

6.3 Non-technical solutions to routing table growth & deaggregation

Most efforts dedicated to the problems of routing table growth and deaggregation have focused on technical improvements to single aspects of BGP routers or full-scale, clean slate shifts to interdomain routing architectures with more favorable scaling properties. However, a few efforts have considered non-technical means to manage the growth of the routing table, including routing economies (effectively internalizing the negative externality) and encouraging more effective network operations.

In the face of routing table growth, Rekhter et al. [RRB97] argue for charging neighbors for use of slots in the routing table by establishing bilateral contracts for carrying route advertisements. They discuss the success of what they term the "spirit of cooperation" in operating the Internet thus far, but claim that it will not scale as the Internet grows and becomes more commercial, and also claim that it will become more difficult to come to agreement about policy governing appropriate behavior. In place of norms, they assert that the use of financial settlement allows for flexibility in route selection and advertisement based on the value of selecting non-default routes. As discussed previously, there are varying degrees of private benefit that are derived from route deaggregation, and financial settlement creates incentives to reduce non-valuable uses and allow for valuable uses to be compensated appropriately. Rekhter's proposal is complex, requiring bilateral contracts to be negotiated and executed between each pair of ASes that exchange routes between each other. Such complexity would likely hinder adoption, and especially incremental adoption, by ISPs. If successful, however, such a contracting system should indeed incentivize route aggregation and efficient route announcements. Ideally, in a competitive market, the price of a route carriage contract would begin to reflect the full cost of advertising a prefix into the global routing table.

A more recent consideration of use of economic disincentives to discourage routing table growth is found in [Cla10]. In this paper, Clayton describes the challenges in appropriately charging for routing table slots, as well as the realization that to fairly apportion his estimated marginal cost of $77,000 per route, some kind of mechanism to appropriately share fees with all networks burdened by the route is required—something that would likely amount to Rekhter et al.'s solution.

Others, such as Zhao et al. [ZPW+01], briefly discuss potential alternative business models for ISPs where financial settlements would be paid to advertise prefixes to DFZ providers, creating an economy for routing prefixes in addition to traffic. However, this proposal is dismissed relatively quickly for reasons of complexity and political infeasibility with the Internet community.

Like the CIDR Report emerged from the community to inform network operators about how to better operate their networks, other efforts to control routing table growth have also come from the community. First, in terms of education, presentations such as [Ste10] are relatively common at NANOG meetings, and reiterate to the community the problems associated with routing table growth while also providing operational solutions to reduce the impact of one's network on the routing table, such as advertising traffic engineering routes with a NO_EXPORT community attribute to keep TE routes out of the routing table. More interestingly, a volunteer effort called the CIDR Police [NG03] that emerged from the NANOG community acted both in an educational role and as a more direct or explicit sanction for operators to improve their behavior. The instigators of the project claim to have had a material impact on the growth of the routing table in the 2000-2002 period.

6.4 The Internet operations community and the social forces within

There is unfortunately a dearth of work available discussing the community of Internet creators and operators, and its norms and characteristics or the makeup of its membership, in spite of the arguable importance of this group in the realization and continued functioning of the Internet. Some insights can be gathered from a thesis [Mat09] and a follow-on paper [MC10] that studies the social interactions between network operators in the context of BGP operations. In these works, Mathew finds that operators organize loosely based on reputation and network stature in the Internet topology, and have traditionally expected other "clueful" operators to operate their networks competently as a point of reputation in the community.

Broader discussions of the Internet ethos and the nature of the (academically-oriented) community that was involved in the founding and growth of the ARPANET and NSFNET, the networks that later became the Internet as we know it today, can be found in some of the historic and ethnographic accounts of the Internet's origins. First among the work in this area is Abbate's *Inventing the Internet* [Abb00], and Hafner's slightly less formal *Where Wizards Stay up Late* [HL96] was also helpful.

Finally, the nature of the Internet operations community can be gleaned to some extent by reading the mailing lists that the members of the community participate in, such as [NANa].

Chapter 7

Managing future routing table growth

This thesis set out to investigate whether the CIDR Report was effective in motivating collective action to mitigate the effects of routing table growth, and to draw conclusions about how to manage the Internet routing table. By analyzing the route announcement and aggregation behavior of ASes after appearing on the CIDR Report and comparing this to the behavior of ASes that never appeared on the CIDR Report, we found that an improvement in AS aggregation behavior was likely associated with appearing on the CIDR Report, and that this treatment effect decreased in more recent time periods. The overall routing table continued to grow in spite of this treatment effect, though this was likely due to the organic growth and participation of new networks, rather than continued deaggregation of CIDR Report participants. Regardless, the routing table has not recently grown at a rate that exceeds upper bounds of our technology (Figure 1-2) or the capabilities of routers available on the market. We now turn our attention to second objective of the thesis, about how to manage the Internet routing table.

7.1 Is routing table growth a problem?

The first question that must be asked in considering the issue of managing the routing table is whether routing table growth is even still a problem affecting the scalability of our current interdomain routing architecture. The Internet routing table is no longer the driving force behind router capabilities [Dav11], and as illustrated in Chapter 1 and also discussed in

[FIRG09] and cited by [Hus11], Moore's law has outpaced routing table growth.

However, there are also evolutionary changes taking place in the Interdomain routing space that may cause the Internet routing table to grow faster than has come to be expected. The Internet will likely be using IPv4 addresses for some time, given the slow pace at which the transition to IPv6 has taken thus far. While the forthcoming exhaustion of free address space will likely hasten IPv6 adoption, it will also likely incentivize the more efficient use of IPv4 addresses in order to extend the useful lifetime of IPv4 before IPv6 adoption is deemed complete. This more efficient use will likely be accomplished through partitioning larger address blocks for trading, as well as the deaggregation of larger blocks into smaller, more numerous announced prefixes to facilitate expression of routing policy for larger blocks of machines behind NATs and other transition technologies. Both of these strategies will result in increased numbers of IPv4 prefixes in the routing table.

At the same time, as the IPv6 routing system moves from experimental and prototype implementations to production-ready networks that more closely mirror the IPv4 routing system [CH10], the number of prefixes in the IPv6 routing table will also grow to be larger than present, and indeed currently appears to be growing exponentially (though the table itself is still quite small[1]). If IPv6 is deployed via dual-stacked routers, as is often the case, this means that available router resources for storing routing tables and processing BGP updates will be contended for by two routing tables instead of just one. Eventually the IPv4 table should be able to be deprecated, but it is unknown when that will be viewed to be commercially viable—a point at which all customers have at least as good IPv6 connectivity as IPv4 connectivity.

So while routers currently have sufficient capacity to meet ISP internal needs while also adequately supporting the Internet routing table, the scaling constraints of BGP that were introduced at the beginning of this thesis have not been eliminated; Moore's law has simply allowed these limitations to be outpaced by technology. If the growth properties of the Internet routing table do change in future, our interdomain routing system may again be strained or made less robust by its own growth, as there is nothing that ultimately constrains

[1] The IPv6 routing table contains approximately 6000 prefixes, but currently appears to be growing exponentially: `http://bgp.potaroo.net/v6/as6447/`

the size of the routing table.

If we agree that there are at least potential problems of scalability or robustness affecting our interdomain routing system under certain routing table growth scenarios, the next question is: what should be done about this problem? Assuming we believe the Internet is too valuable to allow to become unreliable, then some action must be taken. A solution space is outlined below, focusing on uncoordinated solutions (solutions that can be implemented unilaterally) and coordinated solutions (solutions that require collective action). Both technical and non-technical solutions are discussed.

7.2 Uncoordinated solutions

The first set of solutions that we can consider to this problem are uncoordinated solutions—measures that require only individual action to enjoy individual benefit in terms of mitigating the effects of routing table growth. Solutions in this class are easy to deploy because they do not require the collective action that a protocol transition would entail. However, they are also somewhat limited in that they must work within the existing technical, business, and policy architecture of the Internet and its interdomain routing system. In other words, these solutions cannot change the underlying scaling properties or route aggregation incentives of our current system, and instead use other tactics.

In this class of solutions there are several technical approaches that have been proposed, as discussed in the related work. Generally these involve spreading the routing table across multiple routers (virtual aggregation) [BFCW09] at the cost of adding additional hops and packet processing latency, or performing aggregation of the RIB inside the router. The latter approach, which is similar to the aggregation performed by the CIDR Report (though almost certainly using a different criteria for the "same routing policy" invariant, such as next-hop IP address), is a solution preferred by network operators [ZPS10], and could certainly improve the longevity of the more expensive FIB routing table slots for individual networks that deploy this solution. Further, it is likely that the operators affected by the size of the routing table, such as DFZ network operators, would implement a solution like this (providing their vendors supported it), allowing other networks to continue to rely on them as well. The in-router aggregation approach, if applied with similar results to the CIDR Report could reduce the size of the FIB by approximately half[2]. Table growth induced by IPv4 transfers and IPv6 deployment would not typically be aggregable, but this approach could make more room for these new sources of prefixes by reducing the size of the routing table due to deaggregated prefixes.

Also in this class of uncoordinated solutions are non-technical solutions. Unlike the technical solutions which mitigate the effects of routing table growth, these are proposed to work by creating incentives for announcing fewer or more aggregated routes. One option in this space is to continue using the CIDR Report as it exists today. The CIDR Report is included here, instead of the following coordinated solutions section, because of the acknowledgment that it is not a CPR governance institution on its own, and so without a strong institution using the information it provides, it can at best inspire individual action to persuade others or induce oneself to reduce their impact on the routing table. Given that the CIDR Report appears to be ineffective now, both from operators' perspectives and from our analysis, it could possibly be improved based on the observations Chapter 5 in hopes that it might better activate the latent norms-based behaviors that seemed to make the CIDR Report effective in the past. However, this does not appear to be an extremely promising option given its recent history, operator attitudes towards the report, the problems identified in Chapter 5, and the fact that the technical solution above would likely be more effective.

Bilateral economic solutions may also be possible. For example, a Tier-1 provider could begin to charge its peers and customers for the number of routes they advertise. However, this would require all of the peers and customers to acquiesce to this change instead of taking other action, such as switching to a provider that does not charge for route announcements. The autonomous nature and competitive forces of the Internet might thus make such bilateral solutions untenable unless implemented by all Tier-1 providers simultaneously, and so economic solutions of this nature are probably better considered in the following section instead.

[2]The ratio of prefixes in the fully aggregated table to prefixes in the current routing table is 175224/355045 ≈ 0.49 as of 28 January 2011, using our implementation and data sources.

7.3 Coordinated solutions

Beyond individual technical solutions described previously, it appears that the most promising solutions to managing the size and growth of the routing table will require coordination and collective action to achieve. Whether technical solutions that involve adoption of a new interdomain routing protocol, economic mechanisms to settle the costs of propagating a route across the Internet, or a CPR governance institution for interdomain routing, all of these approaches will require the cooperation of most of the participants in the Internet's routing system in order to be effective and realize benefit from the solution.

Regardless of the solution or approach, there are general challenges in achieving collective action. First, as Olson [Ols82] and others point out, achieving collective action can be difficult when the benefits from individual action are not concentrated, or are not selective, creating an incentive for opportunistic free riding. Also, while not uncommon in other spheres of life, the notion that the Internet is no longer composed of actors with homogeneous interests is relatively novel, as identified by [CWSB05]. Instead of a cadre of similarly-minded technologists, there are now a much more diverse and potentially conflicted set of interests at play in shaping the architectural evolution of the Internet, as well as achieving their objectives through the architecture. Even CPR governance institutions are not guaranteed to form in situations where they would provide benefit to appropriators, as Ostrom [Ost90] identifies, because there is a constitutional collective action problem (a meta-problem) in agreeing to develop an institution in the first place, before the institution can then be used to resolve operational collective action problems. Acknowledging that collective action will be required, and may be difficult, we now sketch out some potential approaches in this space and ways in which the collective action towards these ends might be achieved.

A new interdomain routing architecture The most common approach taken thus far when approaching the limits of an Internet protocol or architecture is to replace the system with a new, more scalable version (e.g. NCP to IP(v4), IPv4 to IPv6, EGP to BGP, etc.). The adoption of new protocols is challenging; this challenge increases when a protocol is not backwards compatible with its predecessor, and increases even more at the network layer of the protocol stack where the new protocol must be coordinated along the entire end-to-end path between hosts, rather than just on the hosts themselves. This dilemma often means that there is no benefit from or incentive for incremental protocol adoption, thus making adoption difficult to justify for rational, commercially-motivated actors. This, at least at a high level, is plainly visible in the slow-moving adoption of IPv6 [CH10] that has theoretically been ongoing for the past 16 years [DH95] and has only recently started to gain traction with the coming depletion of unallocated IPv4 address blocks.

In this context, we can consider the adoption of a new interdomain routing architecture and protocol such as one of the potential candidates described in Chapter 6. The move to a new routing architecture and protocol involves many actors that are responsible for various components of the Internet and the devices and software that connect to it: router and network equipment vendors, operating system and application software developers, and Internet service providers, as well as the organizations that purchase and operate this hardware and software and purchase Internet service from ISPs. All of these parties must coordinate in order to realize the transition to a new architecture. Without clear incentives, such as has been claimed to be the case with IPv6, the transition moves slowly until there is motivation (address exhaustion in the IPv6 case or table growth scaling issues in the routing table case).

In the case of a new interdomain routing architecture, the situation may not be as discouraging as with IPv6. Not all routing architectures designed to replace BGP are equally difficult to transition, and there may be other motivations and incentives that would promote the adoption of of these replacement protocols even without immediate and real routing table scaling issues. With regard to transition difficulty, protocol designers have arguably learned lessons from the difficult IPv4-IPv6 transition and have put more effort into considering backwards compatibility and transition strategies. One interesting candidate in this light is the Identifier-Locator Network Protocol (ILNP) [ABH10]. It provides locator-identity separation utilizing existing Internet technologies—DNS, BGP, and IPv6—in a backwards-compatible way, while allowing for aggressive hierarchical aggregation following protocol adoption. Further, there may be "killer apps" and new markets that provide commercial incentives for deployment of new routing architectures that were not similarly present with IPv6 [Li11b], which has only recently been motivated by the coming end of readily available IPv4 address space. For example, the growing popularity of mobile devices and applications may motivate the adoption of a routing architecture that splits locator and identity to allow mobility and provide each device with a static identifier without the added overhead of Mobile IP [Per02].

Economic solutions Economic solutions to the problem of managing the size and growth of the routing table generally deal with imposing a cost on users of routing table slots to provide an incentive for networks to aggregate and otherwise reduce and manage their route announcements carefully to avoid incurring unnecessary costs. These costs could be based on the scarcity of routing table slots and market-based pricing to determine the value of each slot, or it could be imposed arbitrarily by individual network operators. Regardless, the intent would be to internalize the externality that ASes impose on other networks through route announcement and especially deaggregation.

Economic approaches to managing the size of the routing table are orthogonal to other approaches such as protocol changes. However, such solutions would likely be most favorable in cases where routing slots are scarce, such as might occur if a routing architecture transition failed and BGP remained. Economic solutions have the advantage that they treat all participants equally, rather than the unequal quasi-sanction for only the top 30 of the CIDR Report. However, there are potential equity issues in an economic approach compared to the current Internet. Depending on pricing of routing table slots, less wealthy participants could be disadvantaged in participating in the interdomain routing system. Further, a market-based pricing mechanism could allow the price of routing table slots to be manipulated by concentrated interests for strategic or anticompetitive purposes.

Proposals have been made for establishing bilateral contracts between providers to carry routes upstream in exchange for compensation [RRB97]. The approach of bilateral dealing is probably best with the interest of maintaining a decentralized Internet of autonomous networks. However, for such contracting mechanisms to be generally adopted, ISPs would need to collectively agree or accept a change to the current settlement model used by ISPs. Such a move would likely need to be coordinated amongst the DFZ service providers, which some fear may raise concerns about antitrust and collusion [Li11b]. Further, it is not clear that network operators currently have an appetite for the complexity introduced by route carriage contracts [ZPS10], though this could potentially change if faced with the alternative of an overloaded interdomain routing system.

Governance institutions Much of this thesis has been spent discussing CPR governance institutions, and indeed governance institutions are not often considered explicitly on the Internet, though norms and community cooperation often play roles in operating the Internet. More broadly, there is the option of developing an institution to manage the Internet routing table. A central regulator of the interdomain routing system would almost surely not be selected voluntarily by network operators, and so ignoring the option of governmental intervention, we are left with the CPR governance institution option. Again, like economic approaches, the governance institution approach is orthogonal to the routing architecture in place, but would likely be favorable only if interdomain routing remained a problem because a protocol or architecture transition was infeasible or had failed.

To develop a CPR governance institution, network operators would need to collectively devise and commit to a set of rules governing use of the routing table, as well as sanctions and monitoring mechanism (potentially like the CIDR Report) to provide assurances of commitment. Having discussed many of these issues in the context of the CIDR Report in Chapter 5, a routing table governance institution would need to be more structured than the CIDR Report, and utilize collective choice and sanctions, while also considering the need for multiple scales of operator communities to enable better communication and social enforcement.

Further, there is the constitutional collective action problem (the meta-problem) that Ostrom alludes to about how organizations come to agree upon the rules and other characteristics of the governance institution that they will then commit to. Ultimately, rational participants in the Interdomain routing system would need to believe that they would be better off under some institutional arrangement than the status quo. This would again mean that such a solution would only be likely in the case where the robustness of the interdomain routing system was at stake. However, in spite of the cooperative norms of the Internet community, there is also sometimes libertarian sentiment amongst operators in support of the belief that interconnection is voluntary between autonomous networks. Establishing some form of collective order amongst Internet operators may thus be more challenging than establishing norms of the same nature.

7.4 Conclusions

While the CIDR Report was once effective in encouraging aggregation of routing announcements to manage the growth of the Internet routing table, it now appears to mainly be a historic artifact of the previous scalability challenge facing the interdomain routing system that led to the the adoption of CIDR. As such it served a different purpose and focused on different behavior than may be problematic in the interdomain routing system today. However, it is noteworthy that it appeared to influence operator route aggregation behavior for some time even after the CIDR transition, as both claimed by operators and shown by the analysis in this thesis.

If interdomain routing does face growth-based scaling concerns again, is it likely that uncoordinated action in the form of in-router aggregation will be adopted to provide a short-term solution, while a longer-term solution is developed in the form of a new, evolved routing architecture. Indeed, this is the approach recently recommended by the IETF [Li11a], with the recommendation of the incrementally adoptable ILNP as the protocol for future study and development as BGP's successor. While ILNP is likely more easily deployed than IPv6, the coordination necessary to transition to a new interdomain routing architecture will still be difficult and incentives should be considered in hopes of avoiding the delayed transition currently being experienced by IPv6.

Given that a routing architecture transition that changes the scaling properties of interdomain routing will ideally eliminate the need for economic or institutional approaches to managing the Internet routing table, a technical solution is arguably preferable in this specific case. However, governance institutions in the style of Ostrom's CPR institutions should not be ignored or disregarded generally as a solution to other collective action problems facing the Internet, especially those that do not have obvious technical solutions like the interdomain routing system does.

For all of the critiques leveled against the CIDR Report as a governance institution, its partial efficacy is a testament to the social forces—the strength of norms and importance of reputation—in the Internet operations community. As network operators and others who rely on them work to solve the difficult problems facing the Internet that are not purely technical[3], the leverage of inter-ISP relationships and the forces within the social network of the Internet operations community should be considered. Whether through the development of CPR-like institutions, or as a part of larger and more comprehensive solutions, the social forces within the Internet operations community may be effective in cases where purely technical solutions have previously failed.

7.5 Directions for future research

This thesis on the CIDR Report and social forces at play in the Internet operations community have exposed a rich area for further study in understanding the effects of the CIDR Report and the nebulous underlying governance mechanism that enabled these effects. In terms of routing table growth more broadly and Internet collective action and governance institutions more broadly still, the use of governance institutions as an alternative to other proposed solutions for collective problems facing the Internet could use more study and exploration about potential opportunities therein.

With regard to this work, there are several methodological improvements that should appear in future iterations of this work that would greatly improve the quality of the analytic process and hopefully reduce a number of the confounding factors identified previously. First, measurement of the pre-treatment behavior of ASes that would eventually appear on the CIDR Report would allay concerns that some behavior other than appearance on the CIDR Report was causing changes in aggregation behavior. Perhaps more importantly, construction of a more representative control group (such as the ASes who are just shy of appearing on the CIDR Report) would afford greater confidence that the

[3]Problems that may be susceptible to social mediation include spam email, interactions with malicious or untrustworthy actors, or intentional "cyberattacks", etc.

observed behavior change following a CIDR Report appearance was not due to some of the biases mentioned in Chapter 4. Finally, it would better to utilize a statistical test, such as a Kolmogorov-Smirnov test, to determine whether the treatment and control behaviors actually differ significantly.

Beyond these methodological improvements, there are also a number of further analyses and new directions to explore regarding the CIDR Report. Improving our analytical tools to allow more fine-grained analysis of CIDR Report response over time and amongst classes of ASes (i.e. ISPs vs. others) could expose other behaviors not observed in this analysis. Our observation about the sensitivity of the CIDR Report to variations in prefixes received from peers at the vantage point may also merit further study to estimate the degree to which outlier prefixes have contributed to inflated rankings on the CIDR Report. Finally, given our observations about attitudes towards deaggregation and potential causes of the diminished response to the CIDR Report, it would be interesting to explore alternate constructions of a report, such as a report ranked by deaggregation factor or a report that highlighted recent increases in deaggregation, that might evoke a more effective response from network operators.

Finally, there are potentially interesting opportunities in considering the role that CPR governance institutions could play in helping to solve some of the collective action problems facing the Internet and Internet operators. This is not so much an extension of this work as suggestions for future directions based on this work. First, as Ostrom observed in her list of empirical design patterns for long-standing CPR governance institutions, there is a need for sanctions to "back up" norms and punish opportunism—something which the CIDR Report lacked. However, in a globally distributed Internet community that transcends borders and laws, many of the traditional economic or legal tools that might normally be used to implement sanctions are ineffective or unavailable. It would thus be interesting to consider technical and social mechanisms that could possibly be used to implement sanctions or otherwise disincentivize opportunistic behavior that is against the collective interests of Internet operators and users. The development of such mechanisms would likely require a degree of collective action itself, which leads more generally to the need for appropriate forums for participants to come together to propose and commit to collective action. While forums such as network operator meetings and the IETF exist today, these may not be the correct forums, both because of their existing focus and history, and because of participants fear of the appearance of collusion and the resulting consequences from antitrust law. Such work could be a valuable contribution to thinking about improvements to and altogether new Internet architectures.

Appendix A

Software produced for this thesis

A number of software tools were built in order to conduct the analysis undertaken for this thesis. This includes an implementation of the CIDR Report's aggregation report, as well as a number of small tools for processing routing table data, as well as conducting the analysis and producing the figures that go into this thesis. All of these tools can be found in the Git repository at the following URL:

`https://github.com/woodrow/cidr-report_analysis`

The source files used to generate this thesis can be found in the Git repository (commit 79eb232a718710eba75f61059d631e0da04584b0) at the following URL:

`https://github.com/woodrow/sm-thesis`

Both URLs were functional and publicly available as of the date of submission of the thesis.

Appendix B

Sample CIDR Reports

B.1 Emailed CIDR Report

```
This report has been generated at Fri Nov 12 21:11:47 2010 AEST.
The report analyses the BGP Routing Table of AS2.0 router
and generates a report on aggregation potential within the table.

Check http://www.cidr-report.org for a current version of this report.

Recent Table History
        Date      Prefixes    CIDR Agg
        05-11-10    337795      206264
        06-11-10    337843      206672
        07-11-10    338022      206620
        08-11-10    338060      207433
        09-11-10    339052      207760
        10-11-10    339893      207903
        11-11-10    340203      208173
        12-11-10    340330      208528



AS Summary
         35919  Number of ASes in routing system
         15316  Number of ASes announcing only one prefix
          4556  Largest number of prefixes announced by an AS
                AS4323 : TWTC - tw telecom holdings, inc.
      101649920  Largest address span announced by an AS (/32s)
                AS4134 : CHINANET-BACKBONE No.31,Jin-rong Street
```

Aggregation Summary
The algorithm used in this report proposes aggregation only when there is a precise match using the AS path, so as to preserve traffic transit policies. Aggregation is also proposed across non-advertised address space ('holes').

--- 12Nov10 ---

ASnum	NetsNow	NetsAggr	NetGain	% Gain	Description
Table	340755	208585	132170	38.8%	All ASes
AS6389	3751	407	3344	89.1%	BELLSOUTH-NET-BLK - BellSouth.net Inc.
AS4323	4556	1679	2877	63.1%	TWTC - tw telecom holdings, inc.
AS6503	2001	433	1568	78.4%	Axtel, S.A.B. de C.V.
AS19262	1780	316	1464	82.2%	VZGNI-TRANSIT - Verizon Online LLC
AS4766	1728	575	1153	66.7%	KIXS-AS-KR Korea Telecom
AS17488	1360	272	1088	80.0%	HATHWAY-NET-AP Hathway IP Over Cable Internet
AS22773	1242	164	1078	86.8%	ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc.
AS4755	1385	403	982	70.9%	TATACOMM-AS TATA Communications formerly VSNL is Leading ISP
AS18566	1091	158	933	85.5%	COVAD - Covad Communications Co.
AS24560	1056	201	855	81.0%	AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Telemedia Services
AS10620	1333	523	810	60.8%	Telmex Colombia S.A.
AS33363	1560	784	776	49.7%	BHN-TAMPA - BRIGHT HOUSE NETWORKS, LLC
AS18101	905	138	767	84.8%	RELIANCE-COMMUNICATIONS-IN Reliance Communications Ltd.DAKC MUMBAI
AS7545	1438	698	740	51.5%	TPG-INTERNET-AP TPG Internet Pty Ltd
AS28573	1167	514	653	56.0%	NET Servicos de Comunicao S.A.
AS8452	1073	434	639	59.6%	TE-AS TE-AS
AS4808	922	287	635	68.9%	CHINA169-BJ CNCGROUP IP network China169 Beijing Province Network
AS8151	1345	721	624	46.4%	Uninet S.A. de C.V.

```
AS17676      640      66      574    89.7%   GIGAINFRA Softbank BB Corp.
AS7303       826     256      570    69.0%   Telecom Argentina S.A.
AS22047      563      31      532    94.5%   VTR BANDA ANCHA S.A.
AS3356      1191     690      501    42.1%   LEVEL3 Level 3 Communications
AS7552       642     141      501    78.0%   VIETEL-AS-AP Vietel
                                              Corporation
AS9443       571      76      495    86.7%   INTERNETPRIMUS-AS-AP Primus
                                              Telecommunications
AS1785      1799    1320      479    26.6%   AS-PAETEC-NET - PaeTec
                                              Communications, Inc.
AS14420      571     100      471    82.5%   CORPORACION NACIONAL DE
                                              TELECOMUNICACIONES - CNT EP
AS4780       713     243      470    65.9%   SEEDNET Digital United Inc.
AS4804       540      76      464    85.9%   MPX-AS Microplex PTY LTD
AS36992      650     189      461    70.9%   ETISALAT-MISR
AS6478      1392     932      460    33.0%   ATT-INTERNET3 - AT&T Services,
                                              Inc.

Total      39791   12827    26964    67.8%   Top 30 total


Possible Bogus Routes
```

. . . list of bogons removed . . .

```
Please see http://www.cidr-report.org for the full report

------------------------------------
Copies of this report are mailed to:
  nanog at merit.edu
  eof-list at ripe.net
  apops at apops.net
  routing-wg at ripe.net
  afnog at afnog.org
```

B.2 Detailed CIDR Report

Aggregation Report: Aggregation using AS prepended PATH

Report prepared at Sat, 13 Nov 2010 04:11:55 UTC+1000, using data obtained within AS0.6447

The report may include routes internal to AS0.6447, and may also include routes
that are accepted from adjacent AS’s and marked "NO EXPORT". The report also
does not take into account conditions local to each origin AS in terms of policy
or traffic engineering requirements. As an aggregation guide, this report is a
very approximate guide at best.

AS list, ordered by net reduction in advertisements

AS	AS Name	Current	Wthdw	Aggte	Annce	Redctn	%
	Routing Table	347822	174358	29764	203228	144594	41.57%
AS6389	BELLSOUTH-NET-BLK - BellSouth.net Inc.	3755	3526	53	282	3473	92.49%
AS4323	TWTC - tw telecom holdings, inc.	4555	3501	461	1515	3040	66.74%
AS19262	VZGNI-TRANSIT - Verizon Online LLC	1782	1643	144	283	1499	84.12%
AS4538	ERX-CERNET-BKB China Education and Research N	1670	1428	44	286	1384	82.87%
AS6503	Axtel, S.A.B. de C.V.	2003	1534	258	727	1276	63.70%
AS4766	KIXS-AS-KR Korea Telecom	1866	1315	122	673	1193	63.93%
AS22773	ASN-CXA-ALL-CCI-22773-RDC - Cox Communication	1243	1184	10	69	1174	94.45%
AS4755	TATACOMM-AS TATA Communications formerly VSNL	1383	1196	29	216	1167	84.38%
AS17488	HATHWAY-NET-AP Hathway IP Over Cable Internet	1360	1210	151	301	1059	77.87%
AS6478	ATT-INTERNET3 - AT&T Services, Inc.	1392	1225	174	341	1051	75.50%
AS1785	AS-PAETEC-NET - PaeTec Communications, Inc.	1799	1146	106	759	1040	57.81%
AS28573	NET Servicos de Comunicao S.A.	1182	948	103	337	845	71.49%
AS33363	BHN-TAMPA - BRIGHT HOUSE NETWORKS, LLC	1560	1015	234	779	781	50.06%
AS20115	CHARTER-NET-HKY-NC - Charter Communications	1552	991	216	777	775	49.94%
AS24560	AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Tel	1056	860	92	288	768	72.73%
AS18101	RELIANCE-COMMUNICATIONS-IN Reliance Communica	905	813	46	138	767	84.75%
AS10620	Telmex Colombia S.A.	1333	963	207	577	756	56.71%
AS8151	Uninet S.A. de C.V.	1351	933	212	630	721	53.37%
AS3356	LEVEL3 Level 3 Communications	1204	744	31	491	713	59.22%
AS7303	Telecom Argentina S.A.	826	776	71	121	705	85.35%
AS11492	CABLEONE - CABLE ONE, INC.	1256	778	111	589	667	53.11%
AS8452	TE-AS TE-AS	1073	846	184	411	662	61.70%
AS4808	CHINA169-BJ CNCGROUP IP network China169 Beij	948	720	59	287	661	69.73%
AS18566	COVAD - Covad Communications Co.	1091	651	26	466	625	57.29%
AS7545	TPG-INTERNET-AP TPG Internet Pty Ltd	1436	823	221	834	602	41.92%
AS855	CANET-ASN-4 - Bell Aliant Regional Communicat	628	586	11	53	575	91.56%
AS17676	GIGAINFRA Softbank BB Corp.	640	576	3	67	573	89.53%
AS4780	SEEDNET Digital United Inc.	718	574	32	176	542	75.49%
AS7552	VIETEL-AS-AP Vietel Corporation	642	583	44	103	539	83.96%

. . . middle of list (35743 lines) omitted . . .

AS3.644	COLOBRIDGE-AS Colobridge gmbh	1	0	0	1	0	0.00%
AS3.641	TASJIL-PRODUCTION-AS tasjil art production an	1	0	0	1	0	0.00%
AS3.643	TEREWENKO-AS FOP TERESHCHENKO OLEKSANDR TROHU	1	0	0	1	0	0.00%
AS3.647	VSESVIT-AS ZAT Televizijni kabelni merezhi Vs	1	0	0	1	0	0.00%
AS3.650	SCB-AS LLC ICB "Sovcombank"	1	0	0	1	0	0.00%
AS3.648	METALLINVEST-AS Metallinvest LLC	1	0	0	1	0	0.00%
AS3.651	OPTIVERA jsc company optivera	1	0	0	1	0	0.00%
AS3.654	AMBIT-ASN AMBIT SYSTEMY INFORMATYCZNE BOGDAN	1	0	0	1	0	0.00%
AS3.656	LUXTELECOM Luxembourg Telecom S.A.	1	0	0	1	0	0.00%
AS3.662	VIVICOM-AS VIVICOM Polska Dariusz Latecki	1	0	0	1	0	0.00%

AS3.676	INVITELSK Invitel International SK&CZ	2	0	0	2	0	0.00%
AS3.672	NVT-AS Closed Joint Stock Company NOVOTEL	1	0	0	1	0	0.00%
AS3.668	TE-AS LLC "Telecom-express"	1	0	0	1	0	0.00%
AS3.665	OGK4-AS OJSC OGK-4	1	0	0	1	0	0.00%
AS3.666	SSM-GLIWICE-AS Slaska Siec Metropolitalna sp.	1	0	0	1	0	0.00%
AS3.667	ASLINKTELECOMNN Link Telecom NN Ltd.	1	0	0	1	0	0.00%
AS3.670	GOLTA-NETWORKS FOP Getman Valentyn Borysovych	1	0	0	1	0	0.00%
AS3.671	WIZJANET WizjaNet Pawel Strykowski Marcin Tom	1	0	0	1	0	0.00%
AS3.674	WGMAST Wood Group Management Services Ltd	1	0	0	1	0	0.00%
AS3.682	ATWORKS-AS @Works SRL	1	0	0	1	0	0.00%
AS3.677	SOFTNET-PL-AS SoftNet Sp. z o.o.	1	0	0	1	0	0.00%
AS3.683	EN-TO-TRE-HJEMMESIDE-AS 123hjemmeside ApS	1	0	0	1	0	0.00%
AS3.699	BRUMOVICE_NET Jiri Strohalm PE	1	0	0	1	0	0.00%
AS3.710	HARDCOM-AS HardCom Sp. z o.o.	2	0	0	2	0	0.00%
AS3.702	JUTRA FUH JUTRA Artur Kujawski	1	0	0	1	0	0.00%
AS3.707	ASRPK Rostovskaya proizvoditelnaya kompaniya	1	0	0	1	0	0.00%
AS3.715	B-AND-P-FUND-SERVICE-AB B & P Fund service AB	1	0	0	1	0	0.00%
AS3.731	FAKTA-AS Karlstrom & Dahlstrand AB	1	0	0	1	0	0.00%
AS3.719	LETYTA-AS STRYI ISP AS	1	0	0	1	0	0.00%
AS3.727	FREEBIT PE Sukhomlin Vitaliy Leonidovich	1	0	0	1	0	0.00%

Appendix C

Additional figures

This appendix presents figures for five measures, absolute and relative netgain, absolute and relative netsnow, and deaggregation factor, that are used and described in Chapter 4. These figures present the change in these measures following CIDR Report appearances in the treatment group and constructed "appearances" in the control group, for the four three-year time periods within the larger study period: 1998-2000, 2001-2003, 2004-2006, and 2007-2009. The purpose of these figures, which did not fit in the main section of the thesis, are to illustrate the change in these measures in the treatment and control groups over the study period.

C.1 Absolute netgain

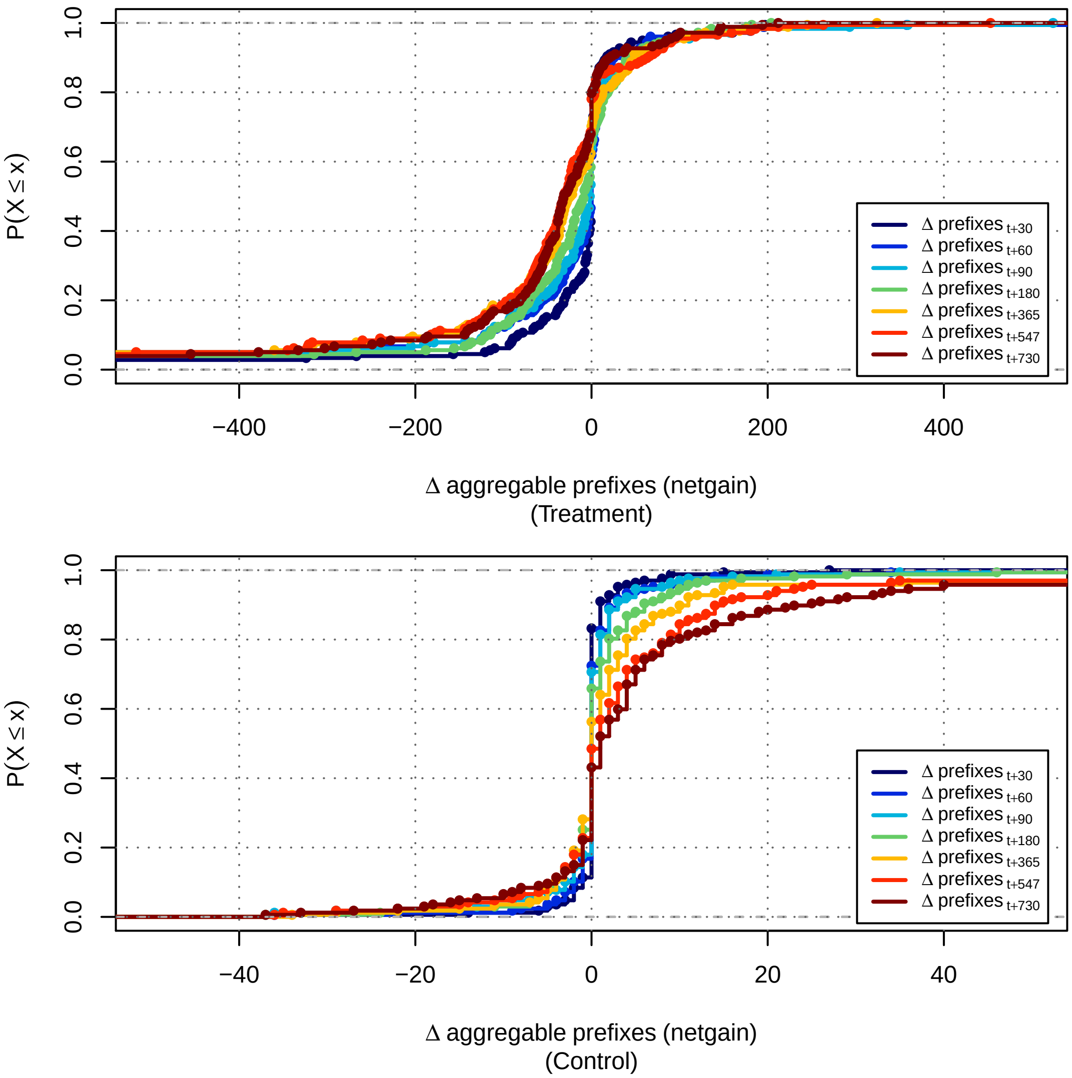


Figure C-1: Cumulative distribution function of change in number ofaggregable prefixes (netgain) advertised by treated and untreated(control) ASes, for the period 1998-2000.

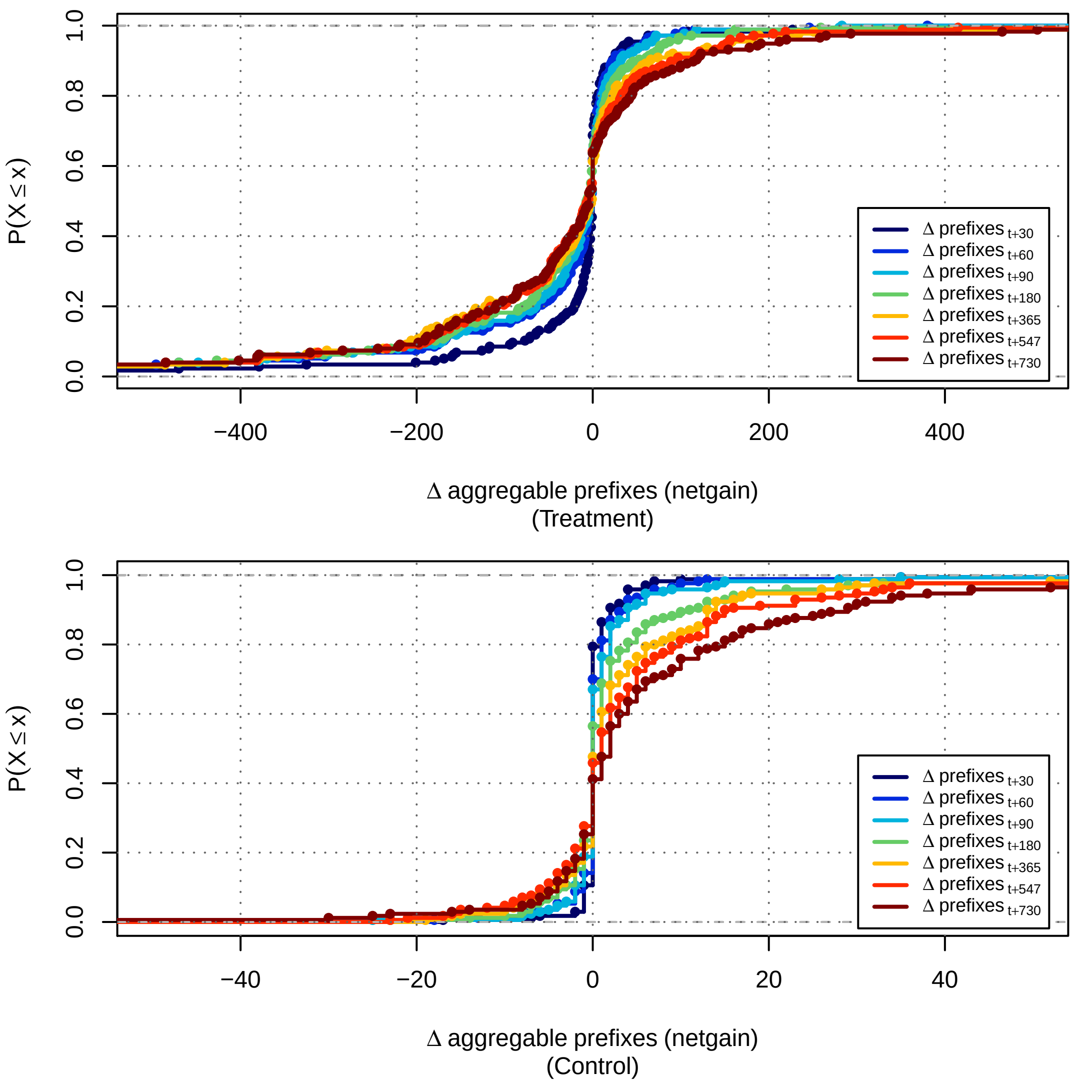


Figure C-2: Cumulative distribution function of change in number ofaggregable prefixes (netgain) advertised by treated and untreated(control) ASes, for the period 2001-2003.

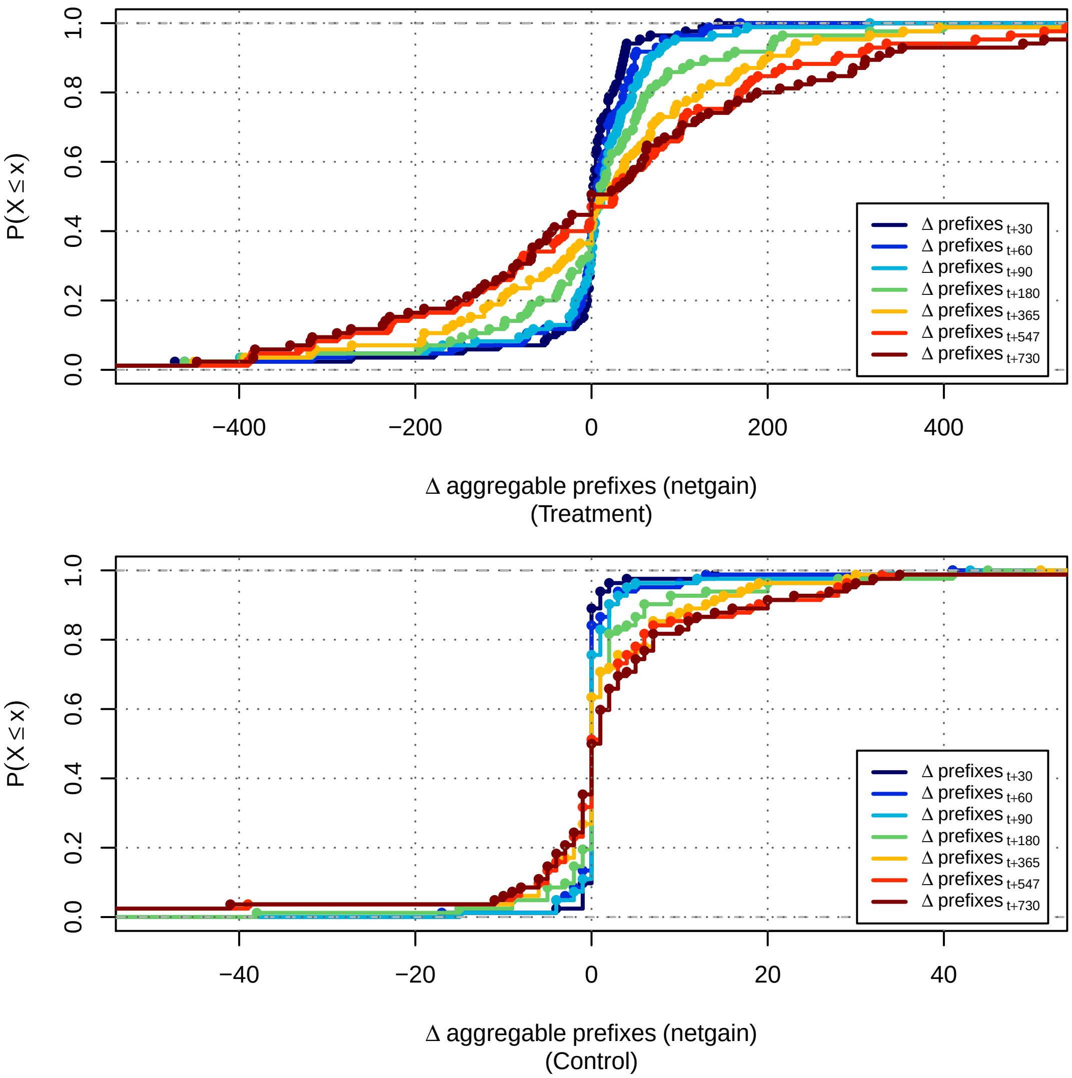


Figure C-3: Cumulative distribution function of change in number ofaggregable prefixes (netgain) advertised by treated and untreated(control) ASes, for the period 2004-2006.

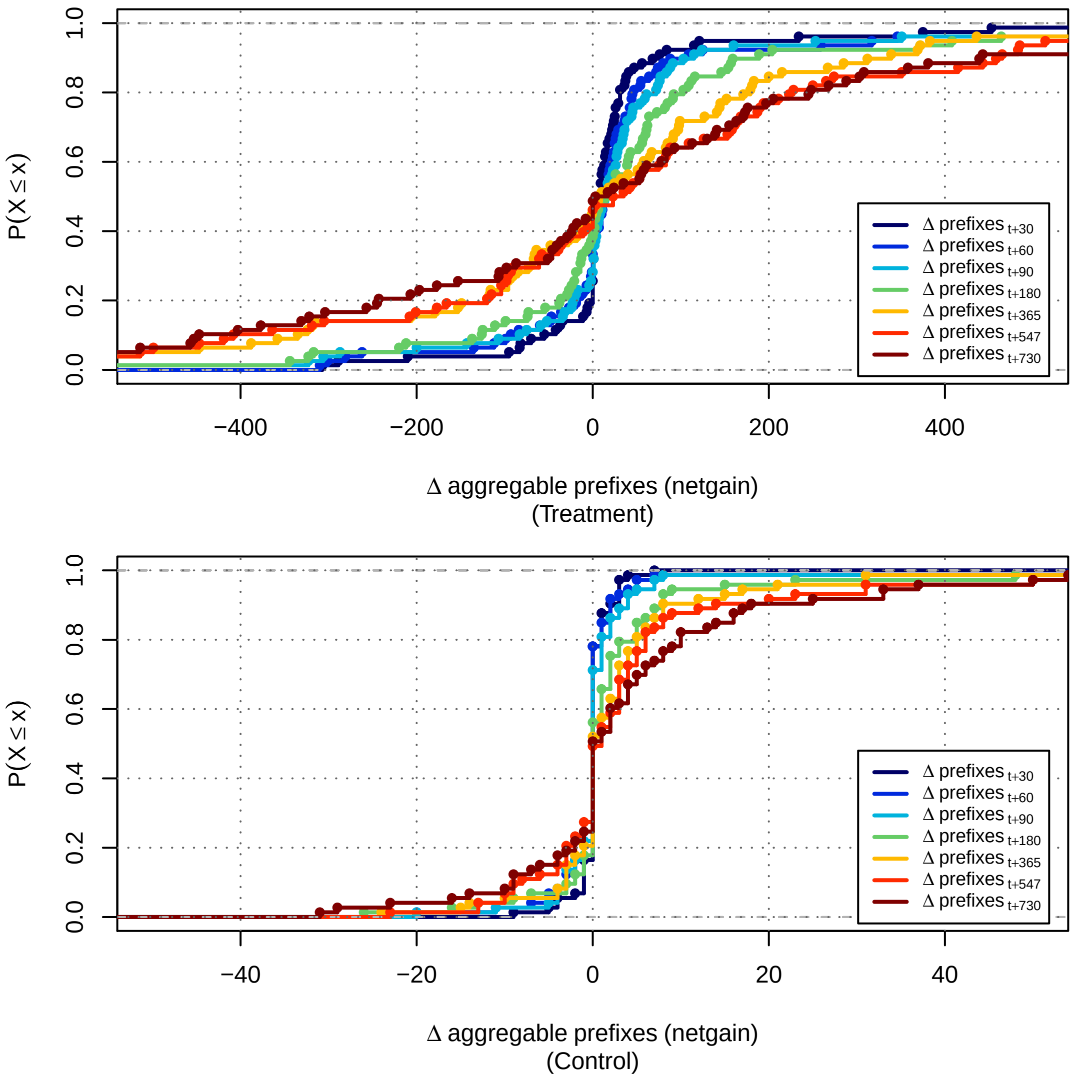


Figure C-4: Cumulative distribution function of change in number ofaggregable prefixes (netgain) advertised by treated and untreated(control) ASes, for the period 2007-2009.

C.2 Relative netgain

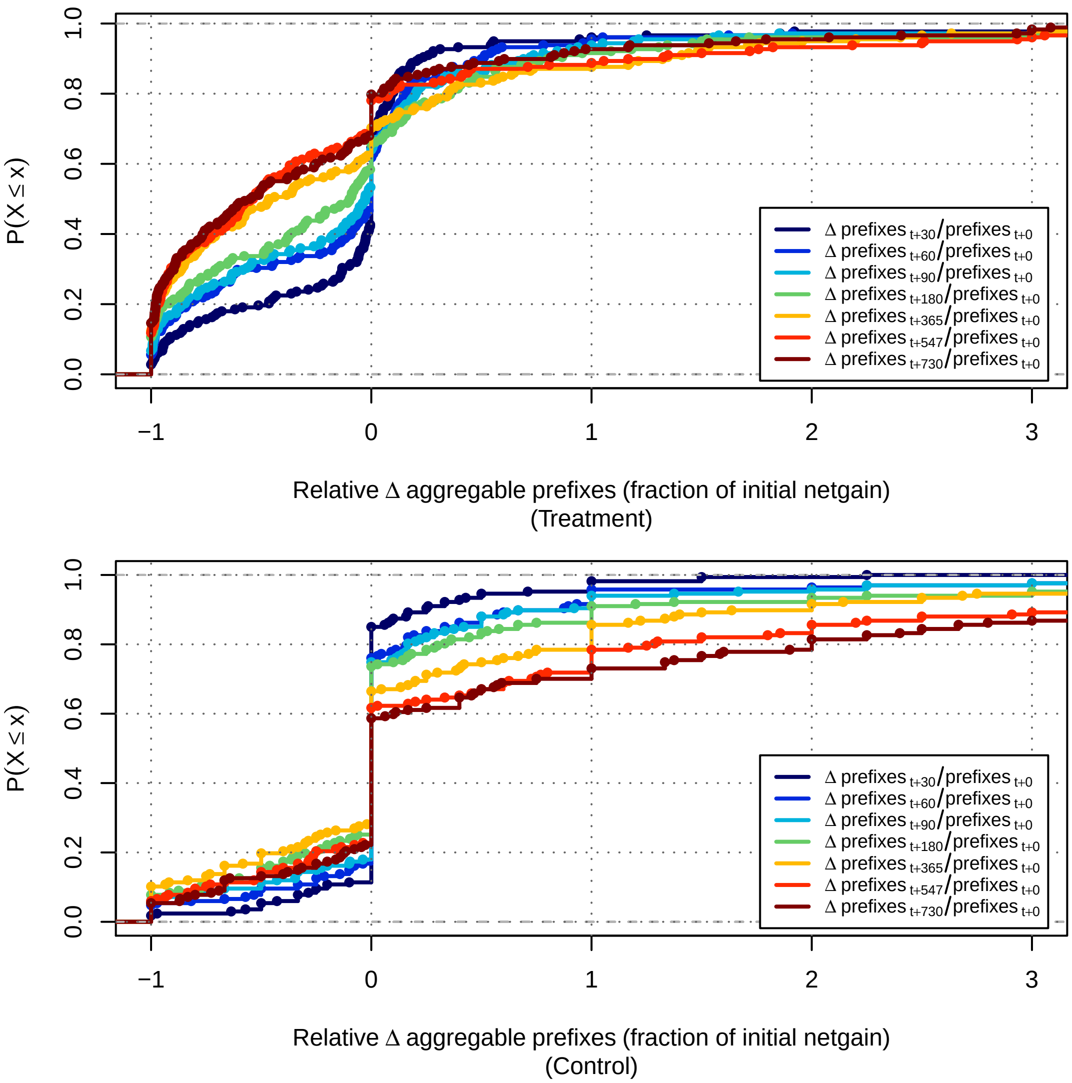


Figure C-5: Cumulative distribution function of relative change in number of aggregable prefixes (netgain) advertised by treated and untreated (control) ASes, for the period 1998-2000.

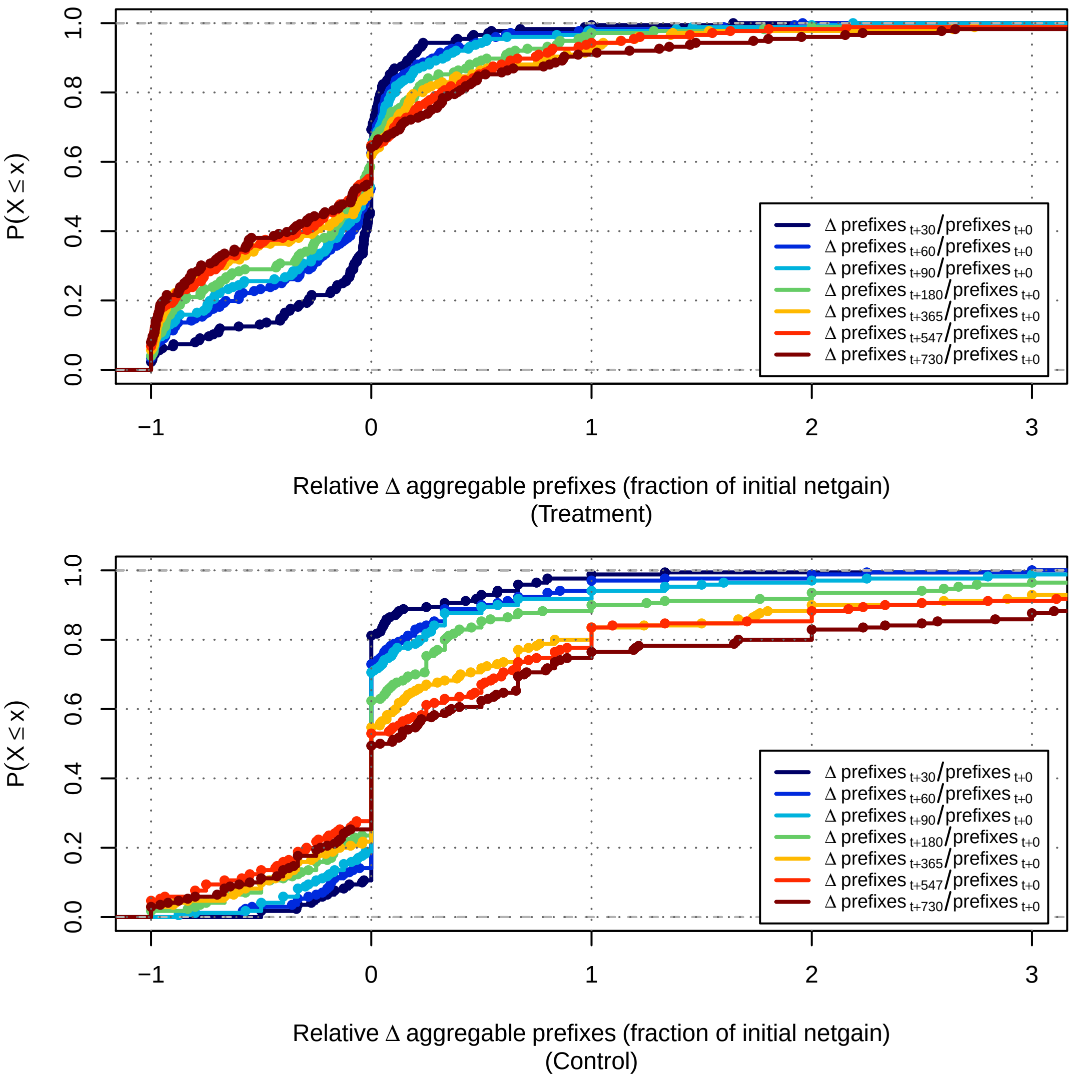
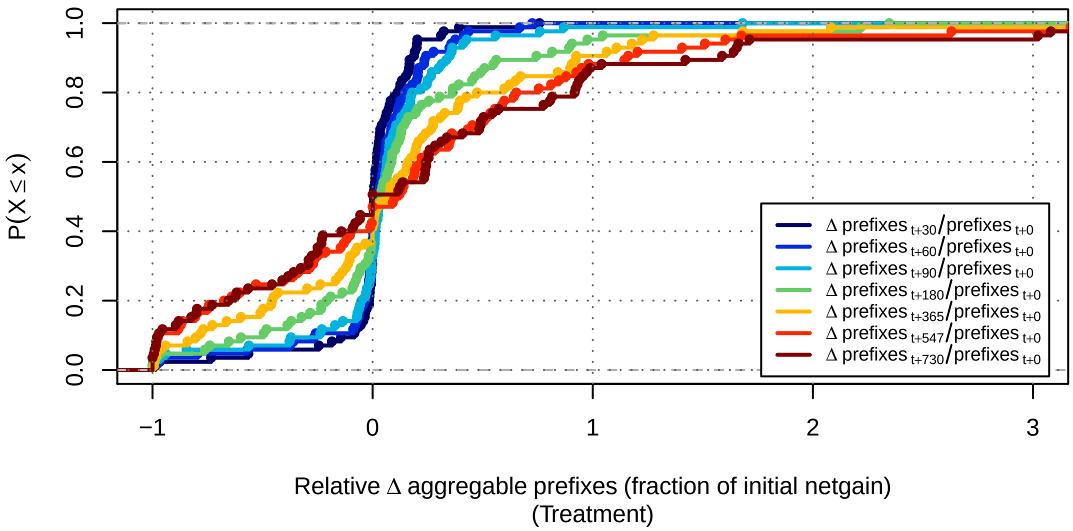


Figure C-6: Cumulative distribution function of relative change in number of aggregable prefixes (netgain) advertised by treated and untreated (control) ASes, for the period 2001-2003.



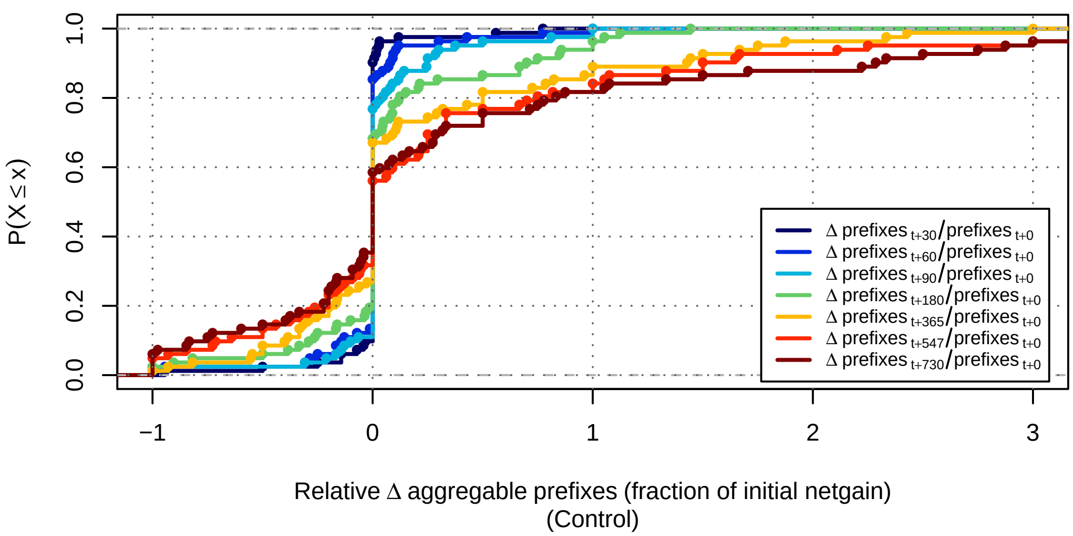


Figure C-7: Cumulative distribution function of relative change in number of aggregable prefixes (netgain) advertised by treated and untreated (control) ASes, for the period 2004-2006.

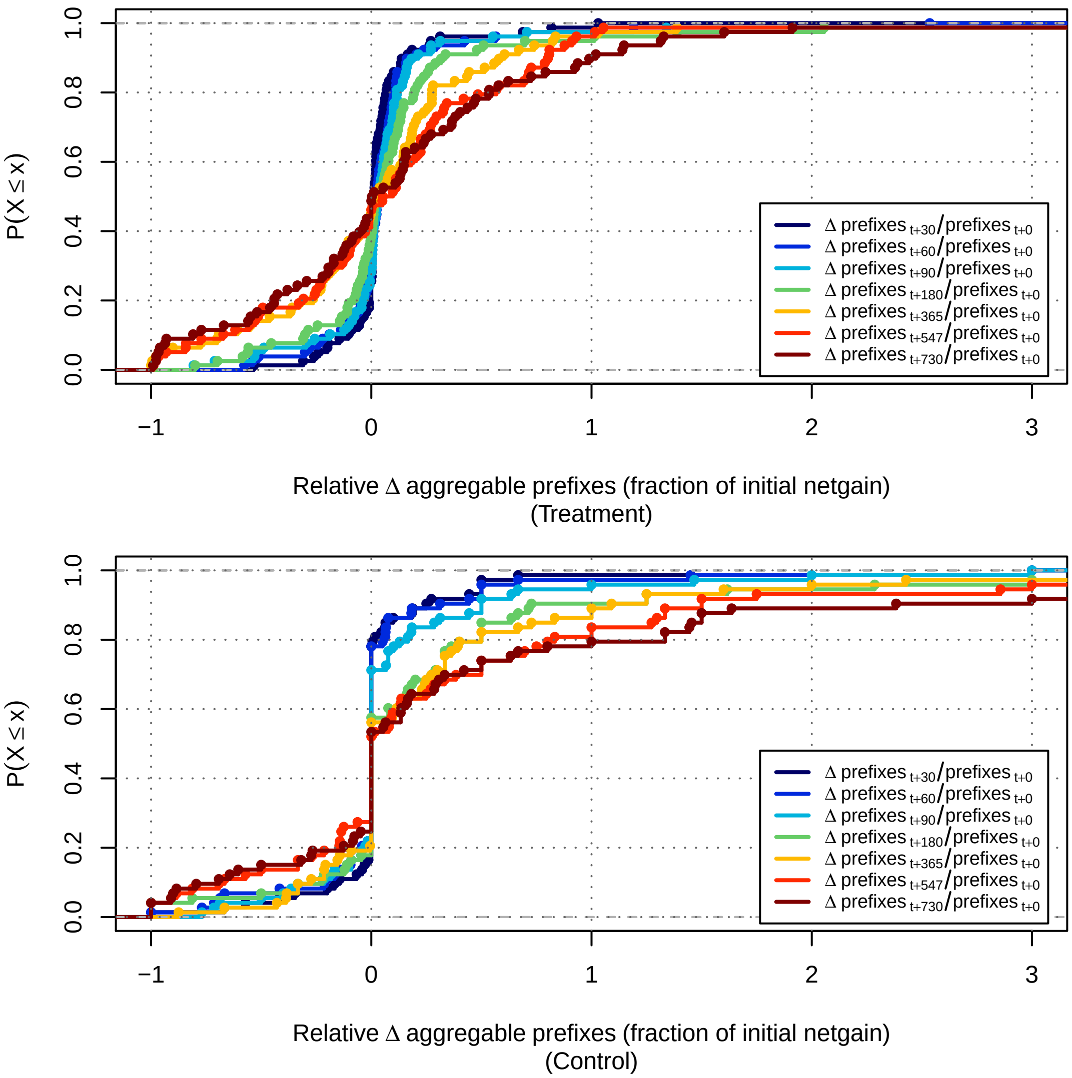


Figure C-8: Cumulative distribution function of relative change in number of aggregable prefixes (netgain) advertised by treated and untreated (control) ASes, for the period 2007-2009.

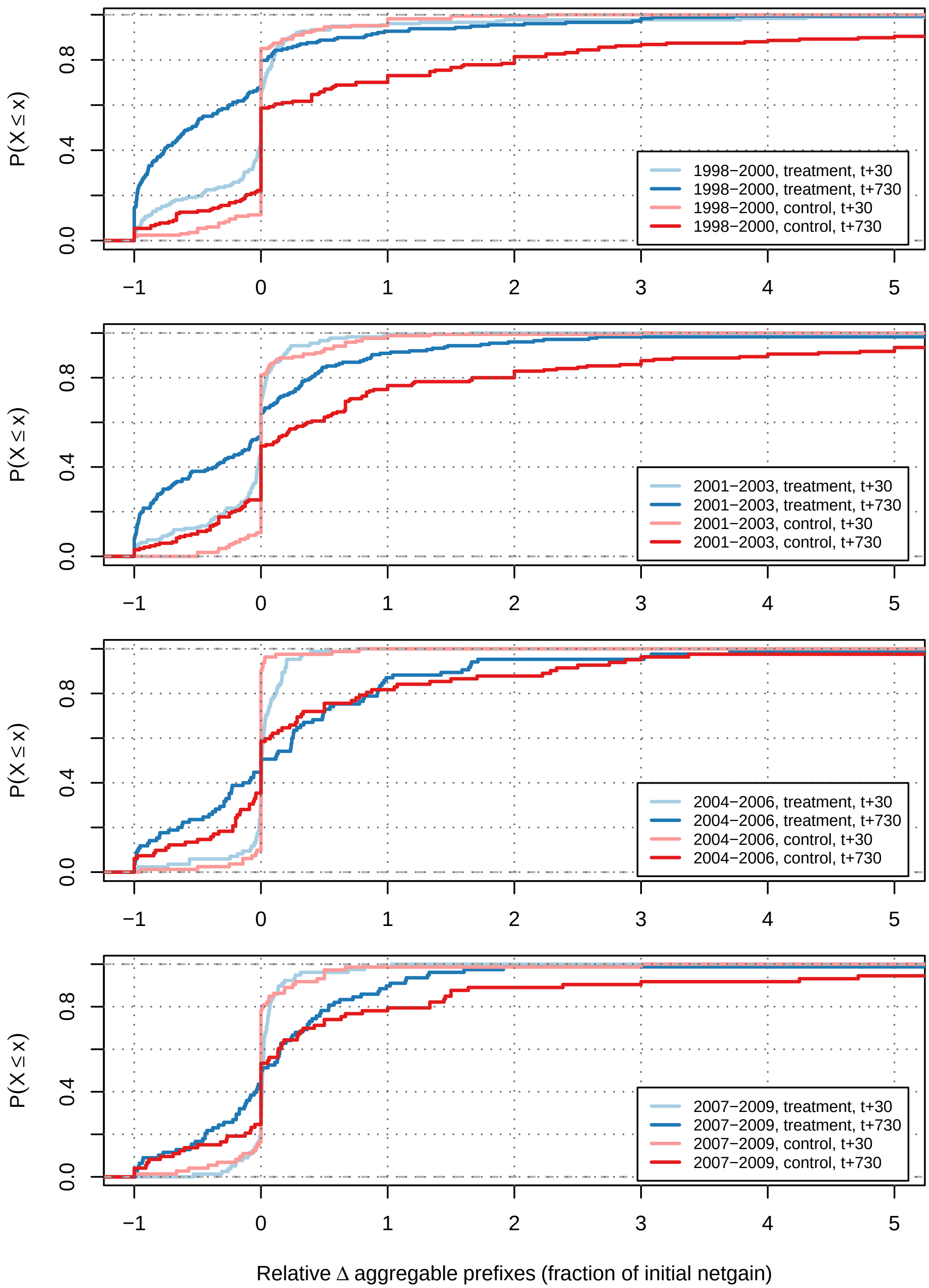


Figure C-9: Simplified relative netgain CDFs for four three-year time periods over the full data availability period. In all cases, the the data ranges from the first CIDR Report of the ending year until the last CIDR Report of the ending year.

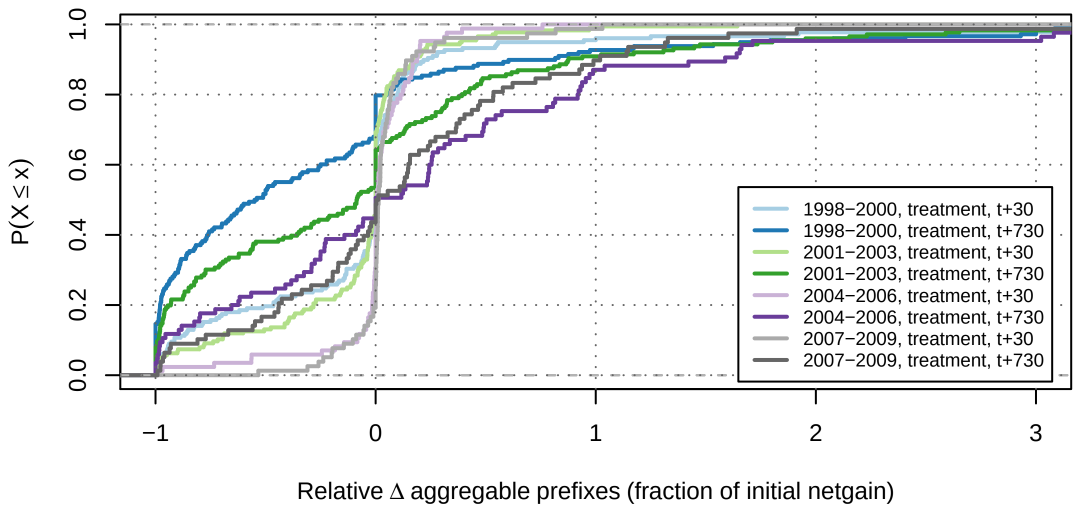


Figure C-10: Superposition of simplified relative netgain CDFs for the treated ASes for the four time periods, showing a notable change in behavior of ASes appearing on the CIDR Report over time.

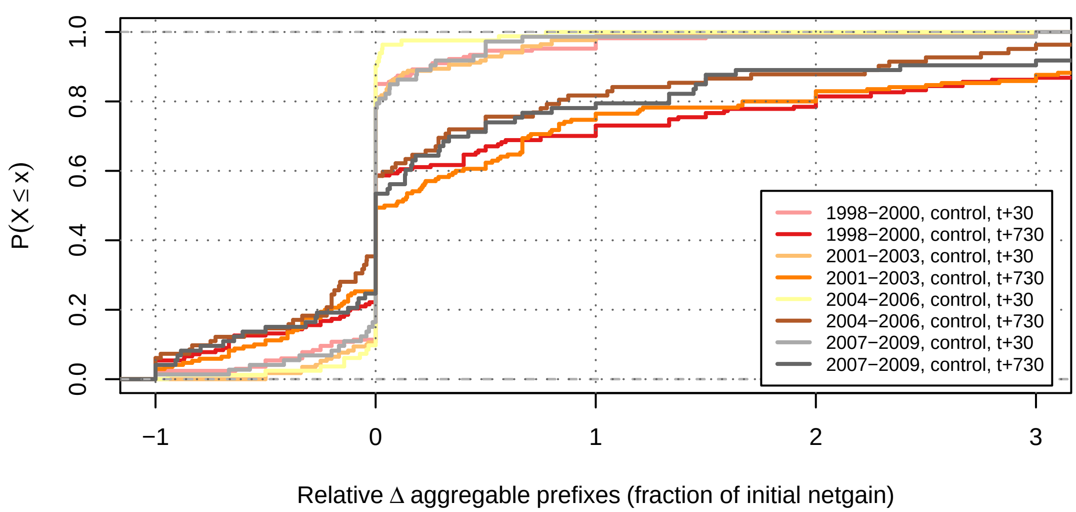


Figure C-11: Superposition of simplified relative netgain CDFs for the untreated control ASes for the four time periods, showing generally similar behavior of untreated ASes throughout time.

C.3 Absolute netsnow

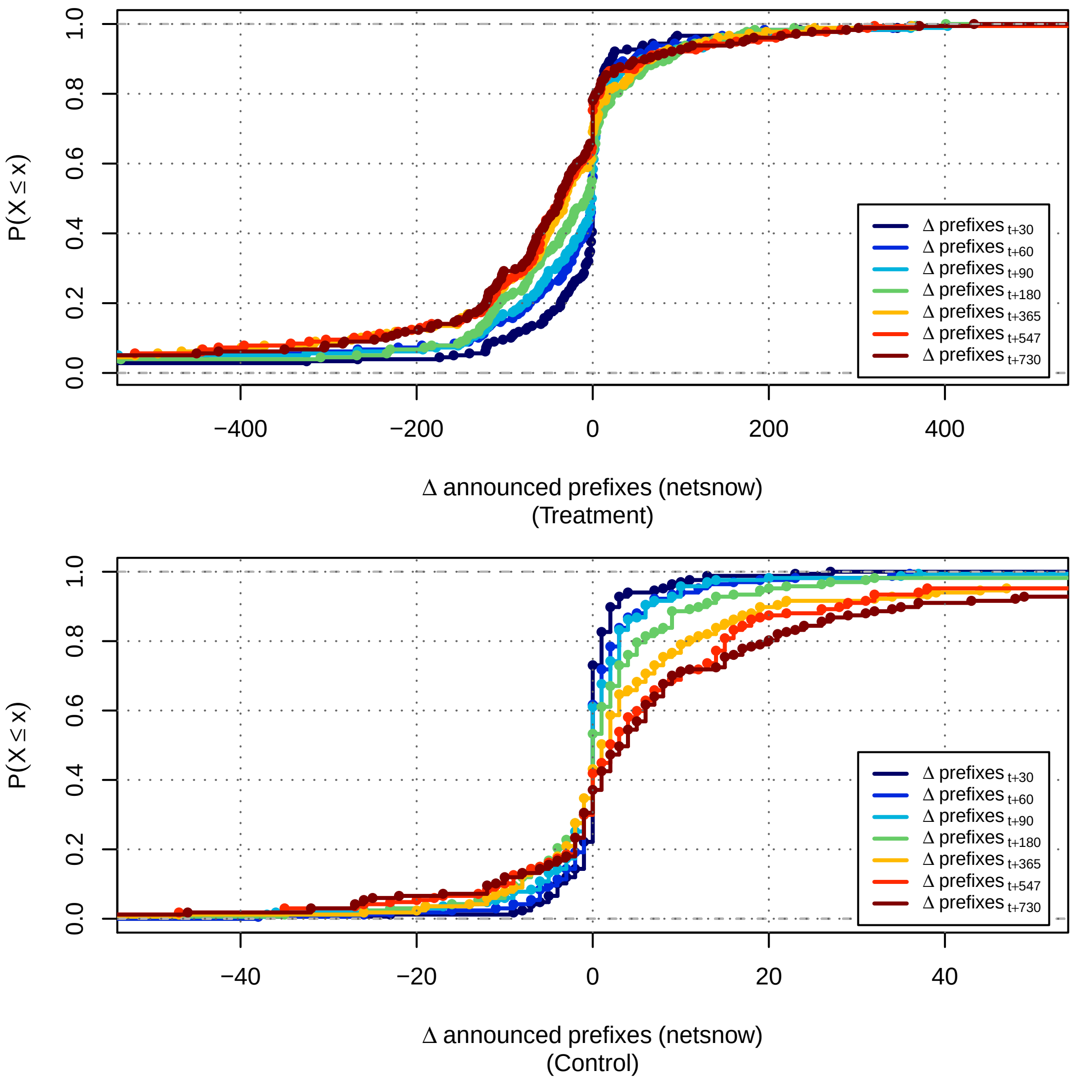


Figure C-12: Cumulative distribution function of change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 1998-2000.

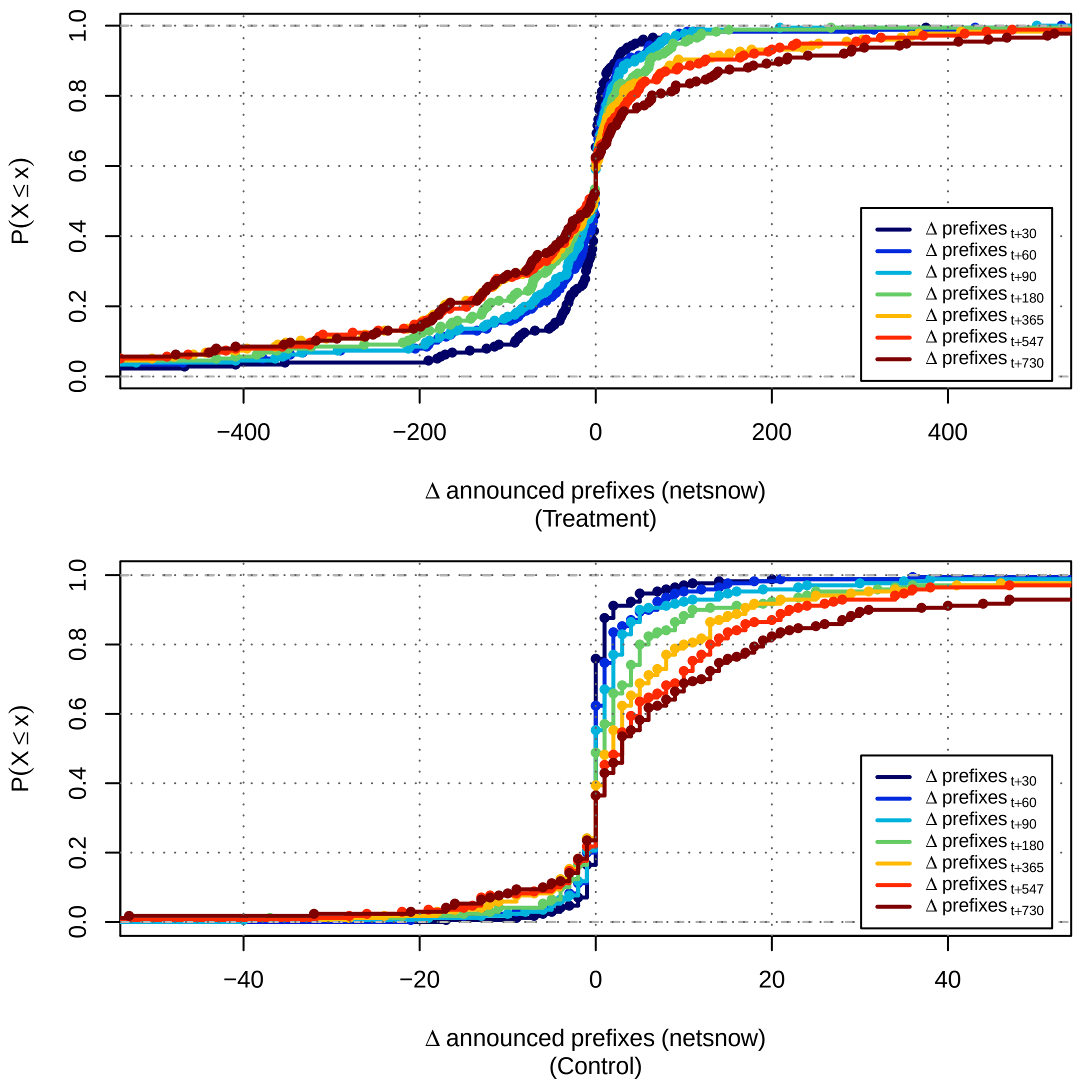
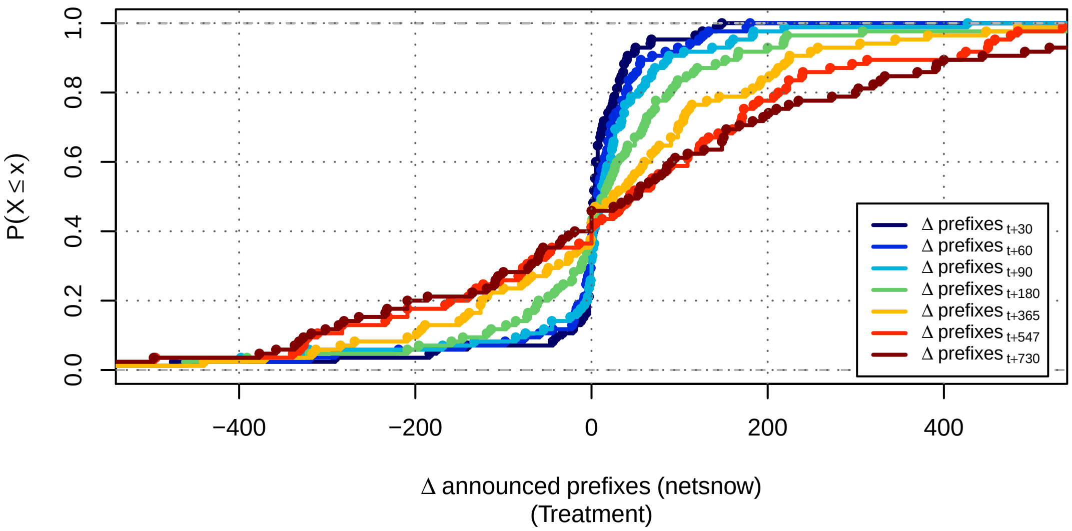


Figure C-13: Cumulative distribution function of change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 2001-2003.



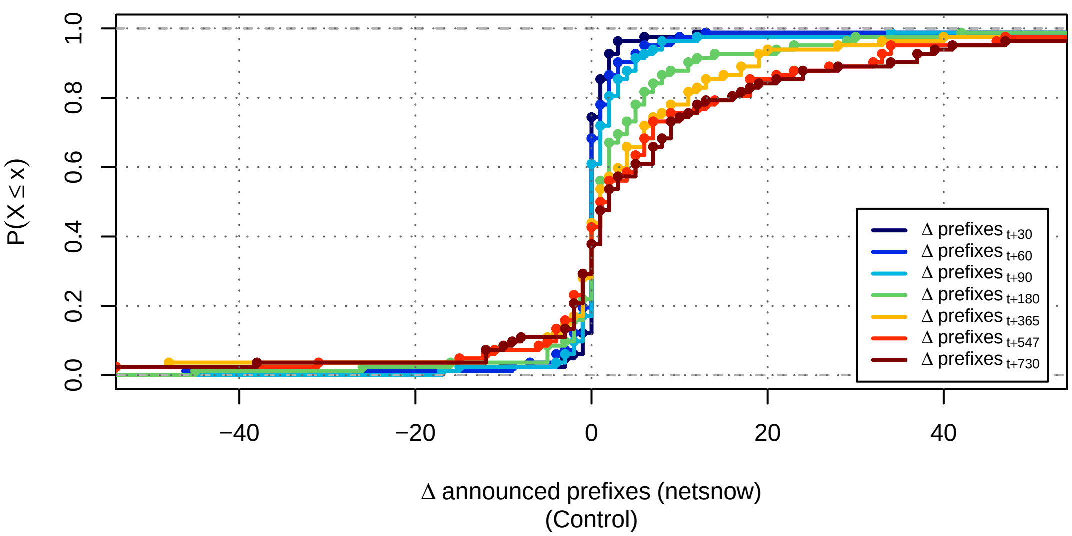


Figure C-14: Cumulative distribution function of change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 2004-2006.

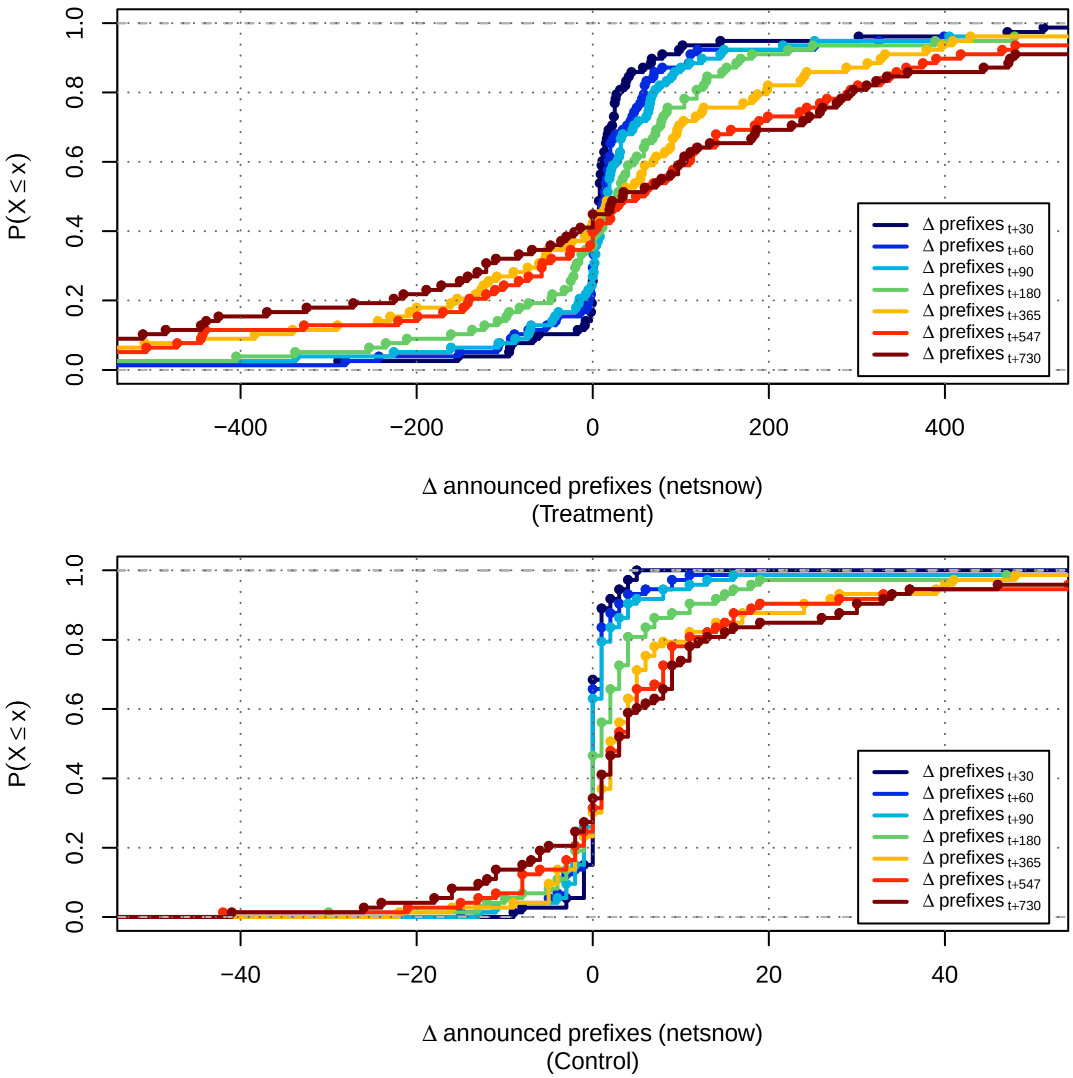


Figure C-15: Cumulative distribution function of change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 2007-2009.

C.4 Relative netsnow

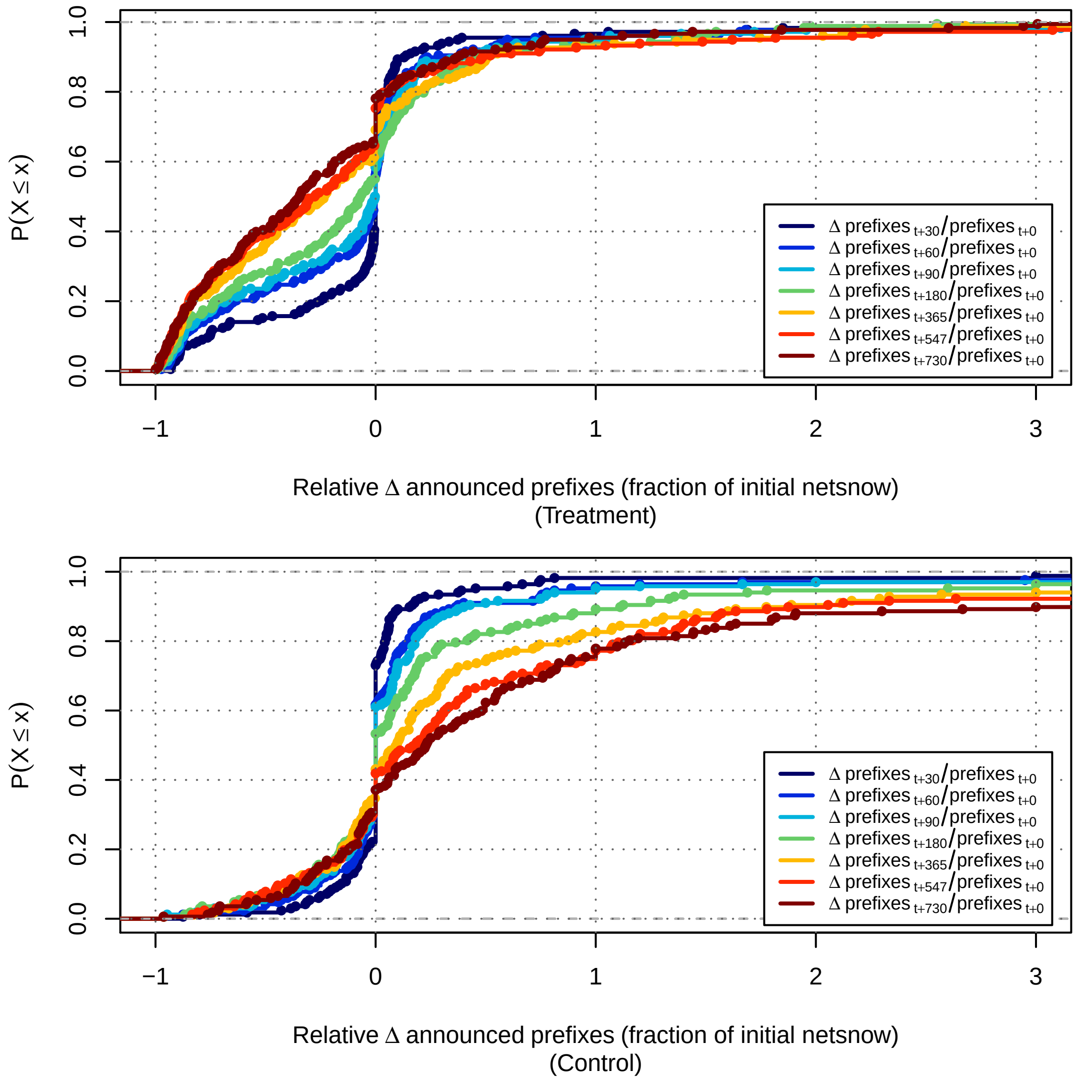


Figure C-16: Cumulative distribution function of relative change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 1998-2000.

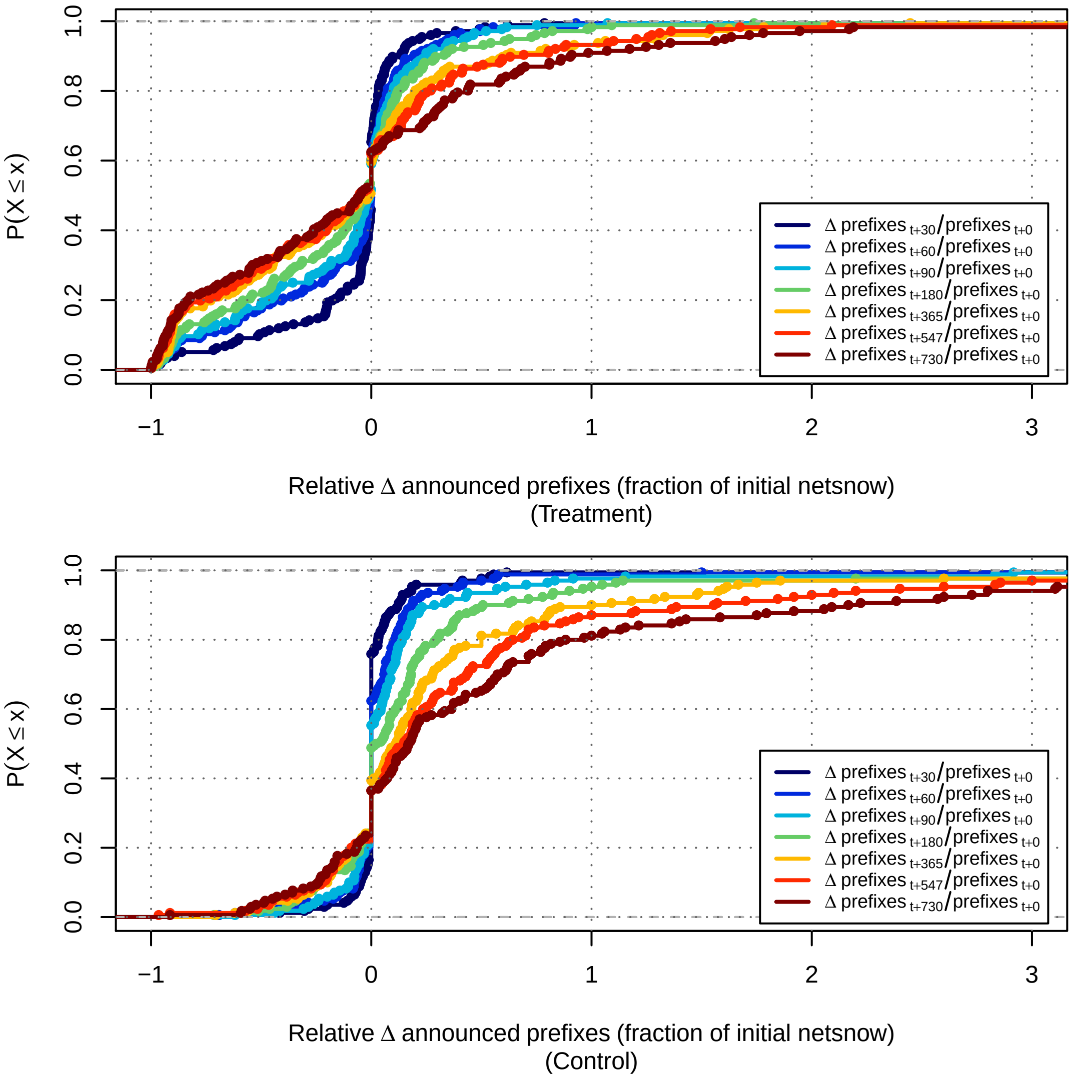


Figure C-17: Cumulative distribution function of relative change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 2001-2003.

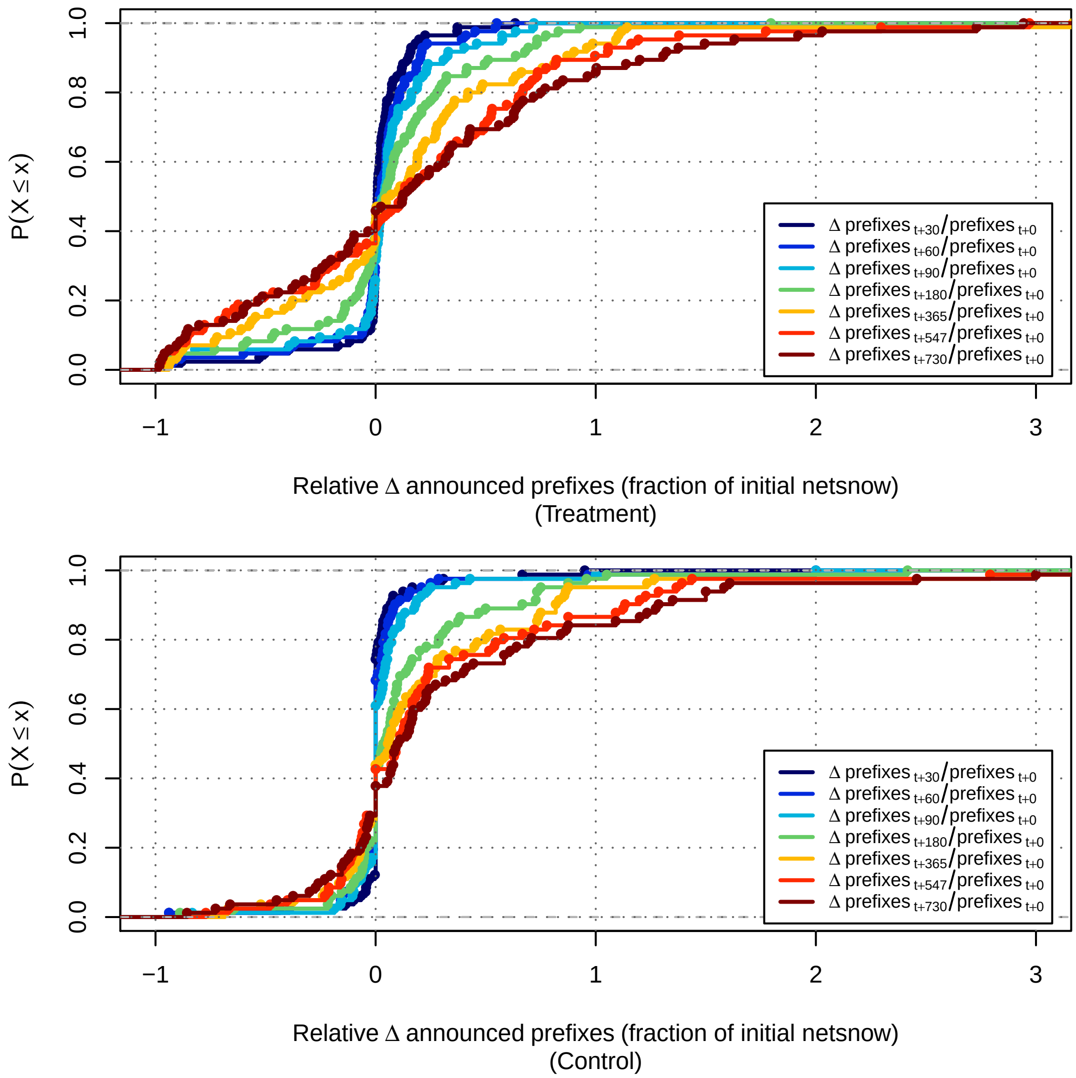
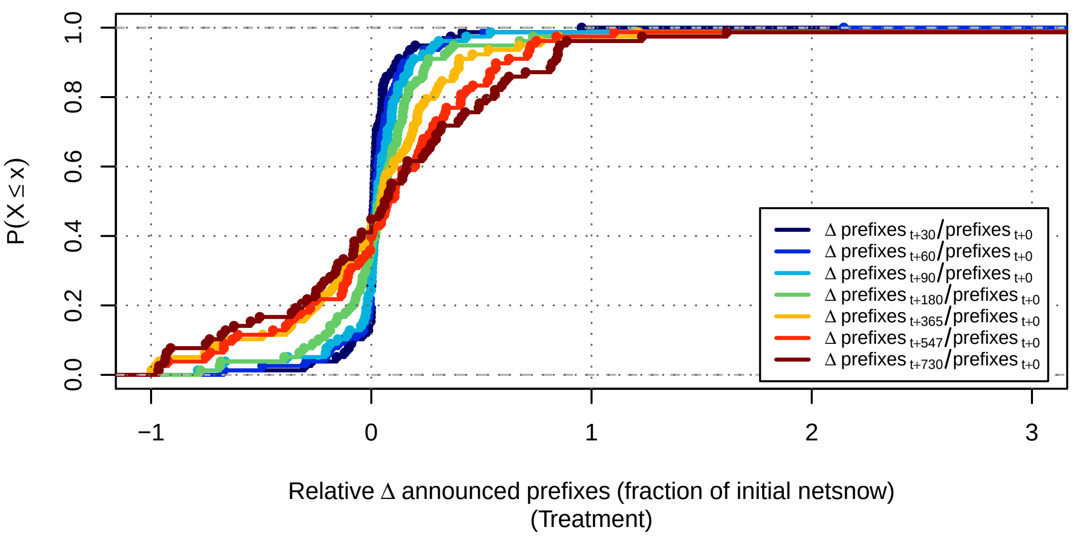


Figure C-18: Cumulative distribution function of relative change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 2004-2006.



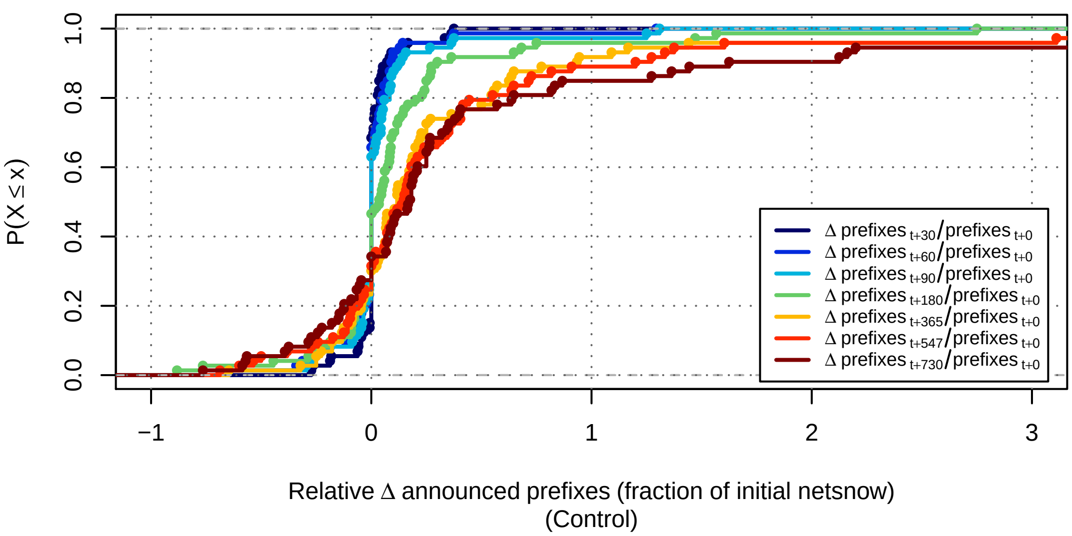


Figure C-19: Cumulative distribution function of relative change in total number of prefixes (netsnow) advertised by treated and untreated (control) ASes, for the period 2007-2009.

C.5 Deaggregation Factor

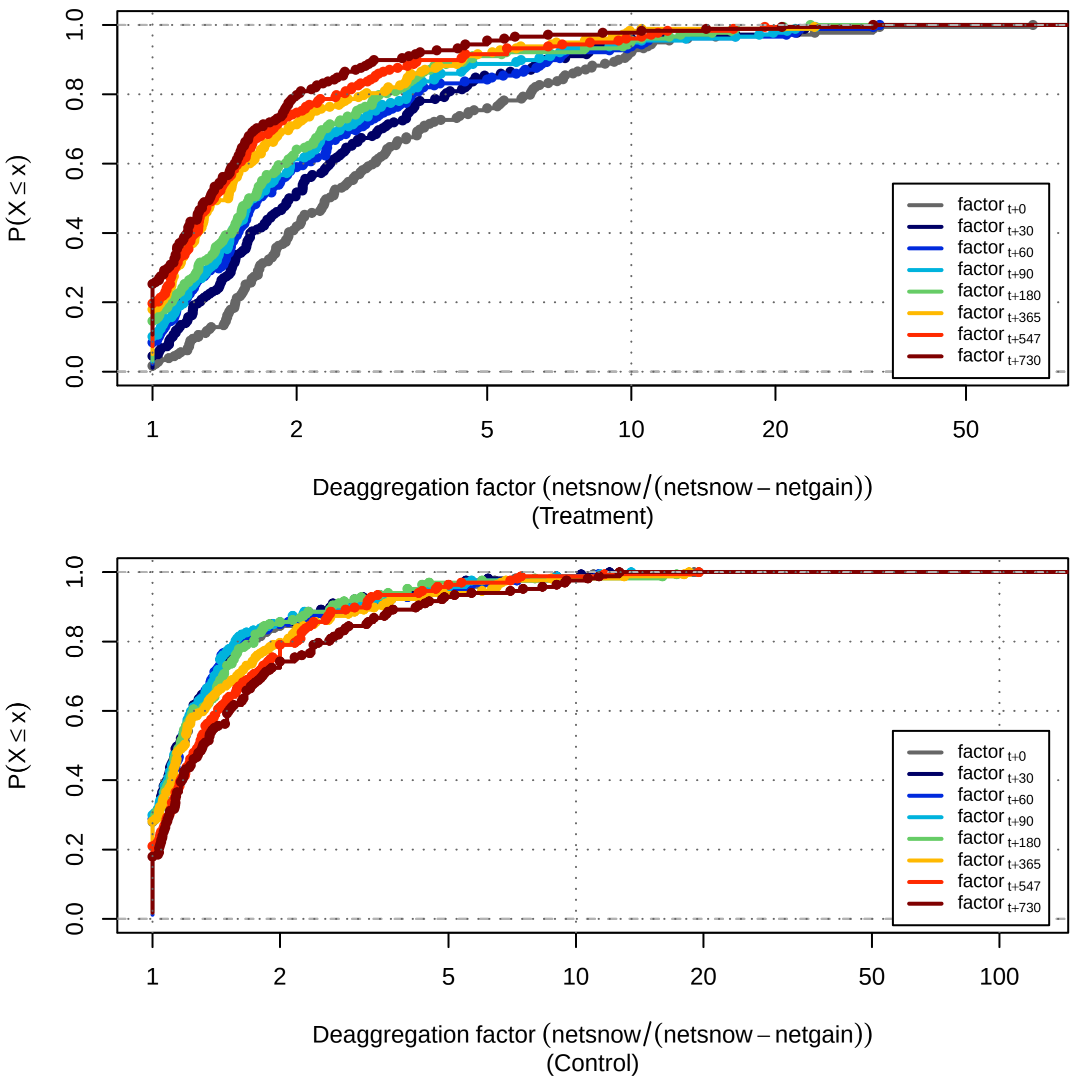


Figure C-20: Cumulative distribution function of change in the deaggregation factor, the ratio of currently advertised prefixes to perfectly aggregated prefixes, of treated and untreated (control) ASes, for the period 1998-2000.

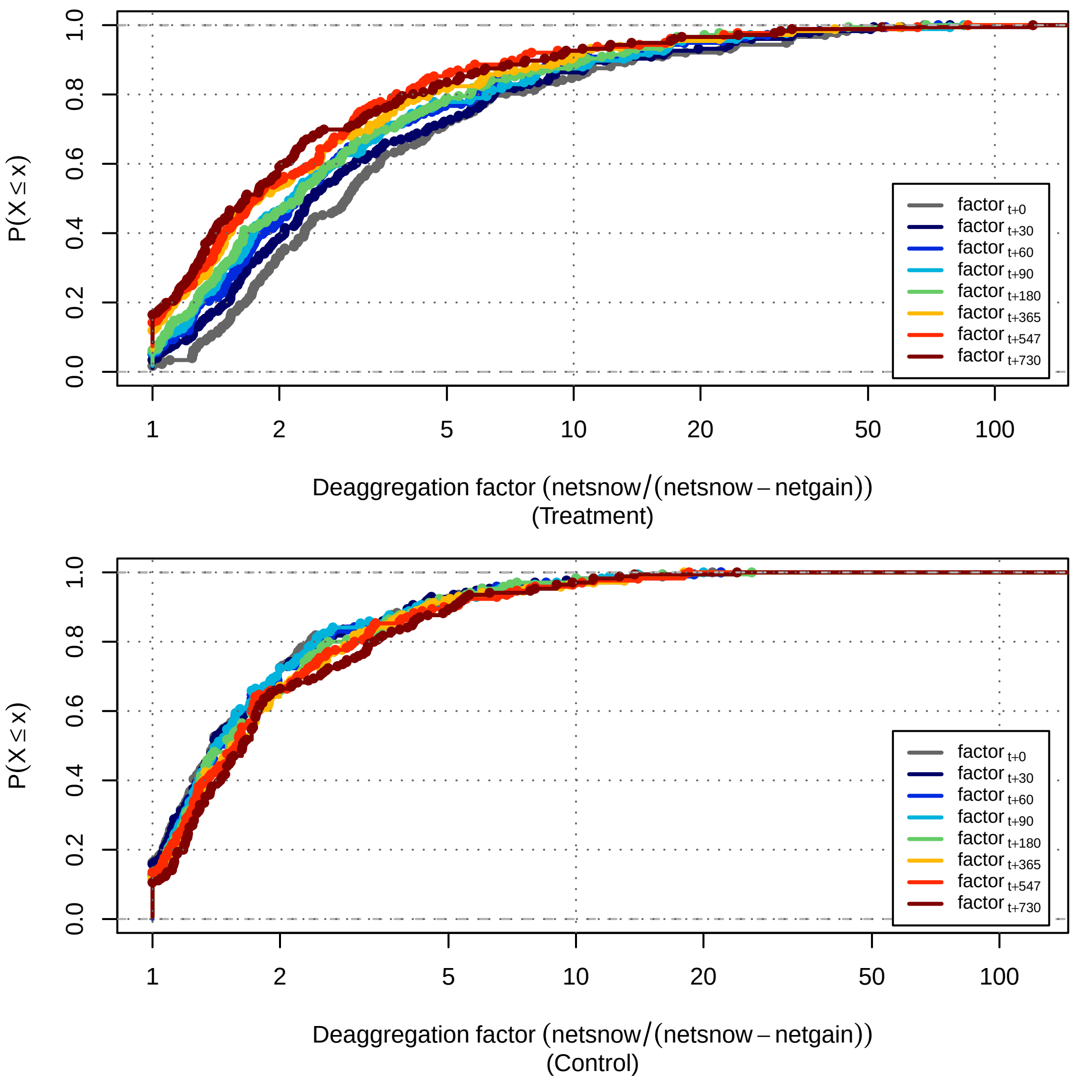


Figure C-21: Cumulative distribution function of change in the deaggregation factor, the ratio of currently advertised prefixes to perfectly aggregated prefixes, of treated and untreated (control) ASes, for the period 2001-2003.

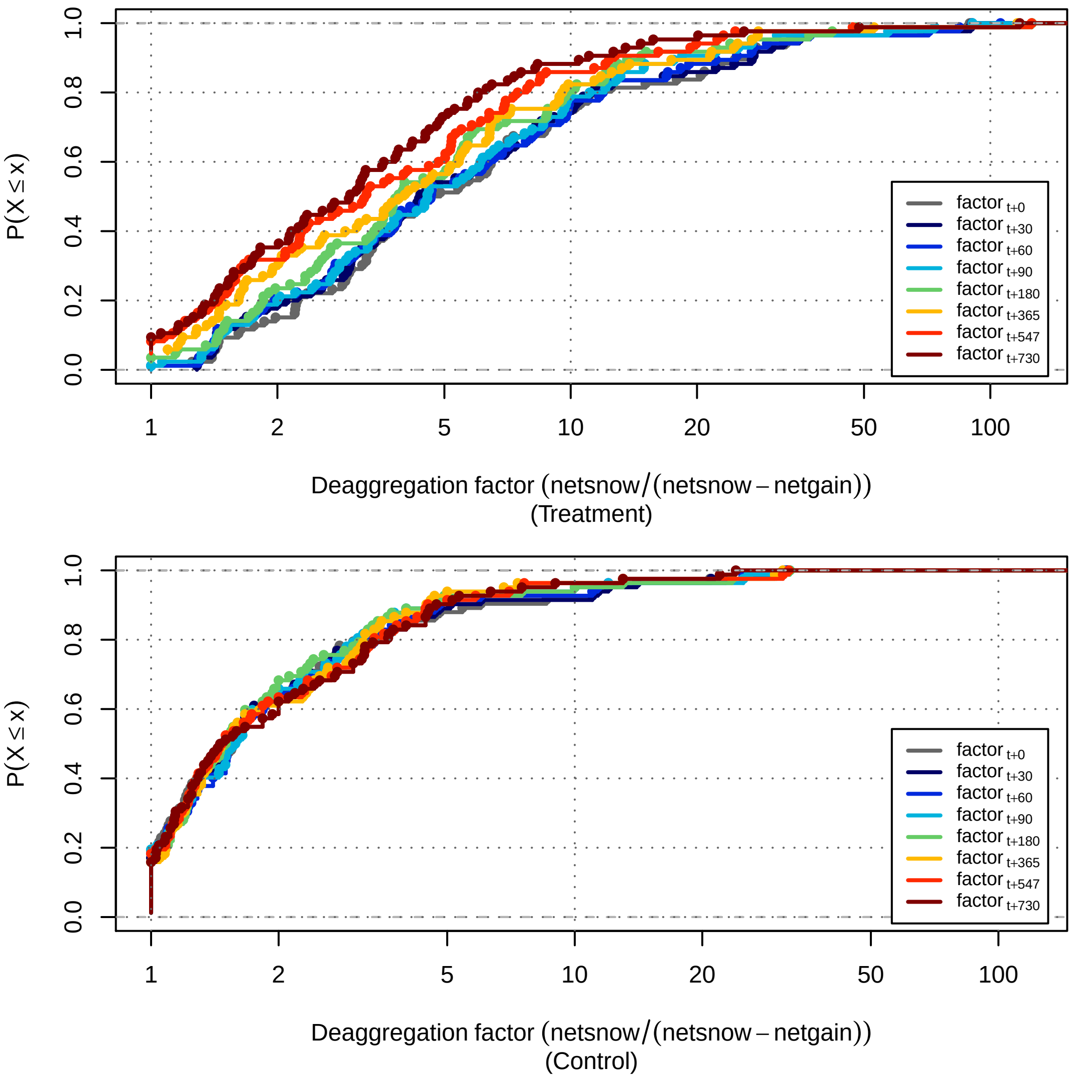


Figure C-22: Cumulative distribution function of change in the deaggregation factor, the ratio of currently advertised prefixes to perfectly aggregated prefixes, of treated and untreated (control) ASes, for the period 2004-2006.

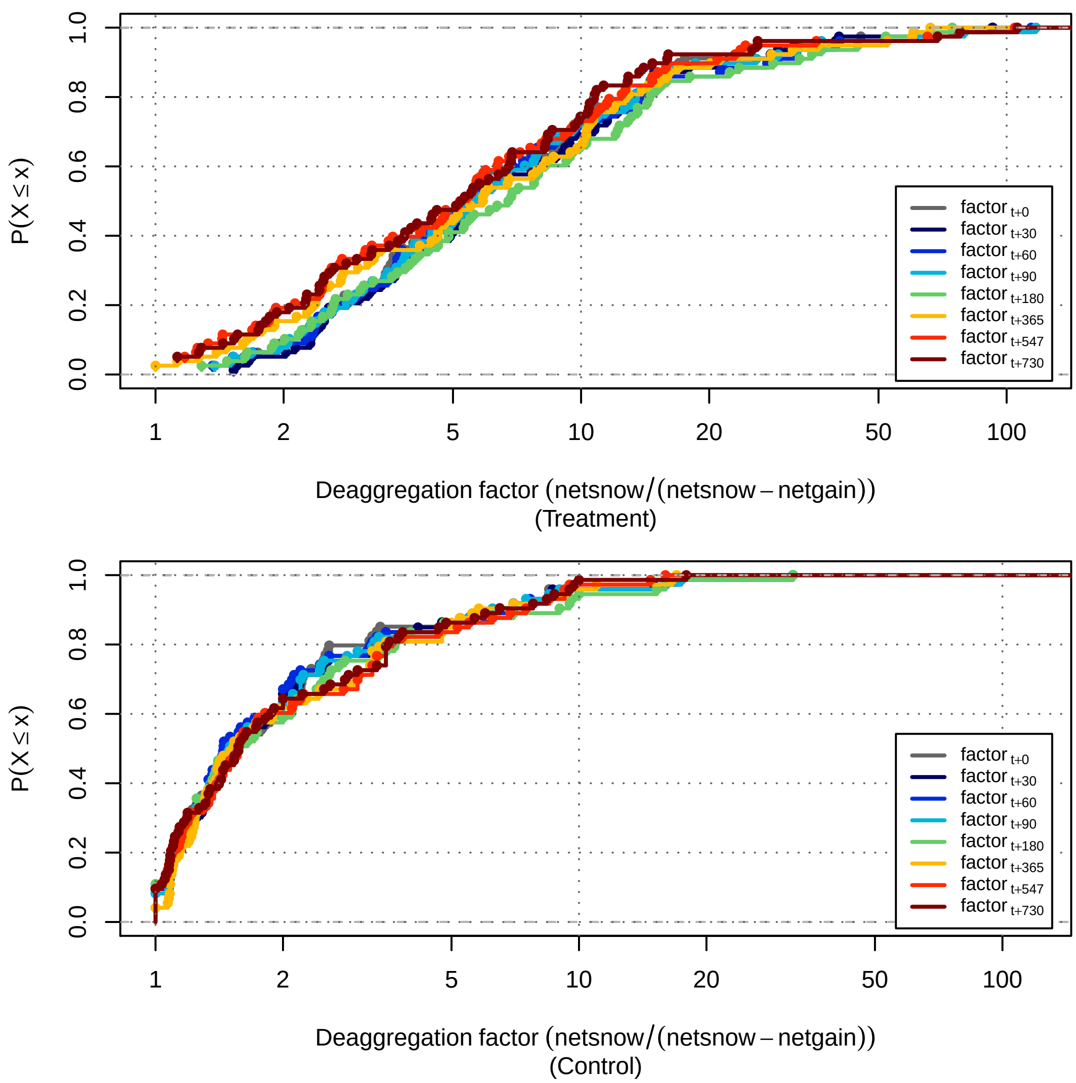


Figure C-23: Cumulative distribution function of change in the deaggregation factor, the ratio of currently advertised prefixes to perfectly aggregated prefixes, of treated and untreated (control) ASes, for the period 2007-2009.

Bibliography

[Abb00] J. Abbate, *Inventing the Internet*. Cambridge, MA, USA: MIT Press, 2000.

[ABH10] R. Atkinson, S. Bhatti, and S. Hailes, "Evolving the internet architecture through naming," *IEEE J.Sel. A. Commun.*, vol. 28, pp. 1319–1325, October 2010. [Online]. Available: http://dx.doi.org/10.1109/JSAC.2010.101009

[ARI10] ARIN. (2010) ARIN Recognizes Interop For Returning IPv4 Address Space. [Online]. Available: https://www.arin.net/announcements/2010/20101020.html

[Bab03] E. R. Babbie, *The Practice of Social Research*, 10th ed. Wadsworth Publishing, 2003.

[BBGR01] S. Bellovin, R. Bush, T. G. Griffin, and J. Rexford. (2001) Slowing Routing Table Growth by Filtering Based on Address Allocation Policies. [Online]. Available: http://www.cs.princeton.edu/~jrex/papers/filter.pdf

[BFCW09] H. Ballani, P. Francis, T. Cao, and J. Wang, "Making routers last longer with ViAggre," in *Proceedings of the 6th USENIX symposium on Networked systems design and implementation*. Berkeley, CA, USA: USENIX Association, 2009, pp. 453–466. [Online]. Available: http://portal.acm.org/citation.cfm?id=1558977.1559008

[BGT04] T. Bu, L. Gao, and D. Towsley, "On characterizing BGP routing table growth," *Computer Networks*, vol. 45, no. 1, pp. 45 – 54, 2004, the Global Internet. [Online]. Available: http://www.sciencedirect.com/science/article/B6VRG-4BT12JX-1/2/7223d197eb9974835bee39ea9fd2db59

[BUZ08] M. Brown, T. Underwood, and E. Zmijewski. (2008, June) The Day the YouTube Died. Presentation at NANOG 43 meeting. [Online]. Available: http://www.nanog.org/meetings/nanog43/presentations/Brown_youtube_light_N43.pdf

[CAI] CAIDA, "straightenRV." [Online]. Available: http://www.caida.org/funding/atomized_routing/download/

[CH10] J. Cowie and C. Hepner. (2010, June) The IPv6 Routing Table in 2010. Presentation at NANOG 49 meeting. [Online]. Available: http://www.nanog.org/meetings/nanog49/presentations/Monday/cowie-hepner-final-14Jun.pdf

[Cla88] D. Clark, "The design philosophy of the DARPA internet protocols," *SIGCOMM Comput. Commun. Rev.*, vol. 18, pp. 106–114, August 1988. [Online]. Available: http://doi.acm.org/10.1145/52325.52336

[Cla10] R. Clayton, "Internet Multi-Homing Problems: Explanations from Economics," in *Economics of Information Security and Privacy*, T. Moore, D. Pym, and C. Ioannidis, Eds. Springer US, 2010, pp. 67–78. [Online]. Available: http://dx.doi.org/10.1007/978-1-4419-6967-5_5

[CMU+10] L. Cittadini, W. Muhlbauer, S. Uhlig, R. Bush, P. Francois, and O. Maennel, "Evolution of internet address space deaggregation: myths and reality," *IEEE J.Sel. A. Commun.*, vol. 28, pp. 1238–1249, October 2010. [Online]. Available: http://dx.doi.org/10.1109/JSAC.2010.101002

[CTL96] R. Chandra, P. Traina, and T. Li, "BGP Communities Attribute," RFC 1997 (Proposed Standard), Internet Engineering Task Force, Aug. 1996. [Online]. Available: http://www.ietf.org/rfc/rfc1997.txt

[CWSB05] D. D. Clark, J. Wroclawski, K. R. Sollins, and R. Braden, "Tussle in cyberspace: defining tomorrow's internet," *IEEE/ACM Trans. Netw.*, vol. 13, no. 3, pp. 462–475, 2005.

[Dav11] B. Davie, "Personal correspondence," 2011, Personal correspondence with Stephen Woodrow, 2 March 2011.

[DH95] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," RFC 1883 (Proposed Standard), Internet Engineering Task Force, Dec. 1995, obsoleted by RFC 2460. [Online]. Available: http://www.ietf.org/rfc/rfc1883.txt

[FBR04] N. Feamster, H. Balakrishnan, and J. Rexford, "Some foundational problems in Interdomain routing," in *In HotNets, 2004.*, 2004, pp. 41–46.

[FFML11] D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, "Locator/ID Separation Protocol (LISP)," Internet Draft, Internet Engineering Task Force, 2011. [Online]. Available: http://www.ietf.org/id/draft-ietf-lisp-12.txt

[FIRG09] K. Fall, G. Iannaccone, S. Ratnasamy, and P. B. Godfrey, "Routing Tables: Is Smaller Really Much Better?" in *In HotNets, 2009.*, 2009. [Online]. Available: http://conferences.sigcomm.org/hotnets/2009/papers/hotnets2009-final156.pdf

[FLYV92] V. Fuller, T. Li, J. Yu, and K. Varadhan, "Supernetting: an Address Assignment and Aggregation Strategy," RFC 1338 (Informational), Internet Engineering Task Force, Jun. 1992, obsoleted by RFC 1519. [Online]. Available: http://www.ietf.org/rfc/rfc1338.txt

[FLYV93] ——, “Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy,” RFC 1519 (Proposed Standard), Internet Engineering Task Force, Sep. 1993, obsoleted by RFC 4632. [Online]. Available: http://www.ietf.org/rfc/rfc1519.txt

[Han06] M. Handley, “Why the Internet only just works,” *BT Technology Journal*, vol. 24, pp. 119–129, July 2006.

[Har68] G. Hardin, “The Tragedy of the Commons,” *Science*, vol. 162, pp. 1243–1248, 1968.

[HB96] J. Hawkinson and T. Bates, “Guidelines for creation, selection, and registration of an Autonomous System (AS),” RFC 1930 (Best Current Practice), Internet Engineering Task Force, Mar. 1996. [Online]. Available: http://www.ietf.org/rfc/rfc1930.txt

[Her08] W. Herrin. (2008) What does a BGP Route cost? [Online]. Available: http://bill.herrin.us/network/bgpcost.html

[HKC+96] K. Hubbard, M. Kosters, D. Conrad, D. Karrenberg, and J. Postel, “Internet Registry IP Allocation Guidelines,” RFC 2050 (Best Current Practice), Internet Engineering Task Force, Nov. 1996. [Online]. Available: http://www.ietf.org/rfc/rfc2050.txt

[HL96] K. Hafner and M. Lyon, *Where Wizards Stay up Late: The Origins of the Internet*, 1st ed. Simon & Schuster, 1996.

[Husa] G. Huston, “AS6447 BGP Routing Table Analysis Report.” [Online]. Available: http://bgp.potaroo.net/as6447/

[Husb] ——, “CIDR Report Aggregation Report: Aggregation using AS prepended PATH.” [Online]. Available: http://www.cidr-report.org/as2.0/aggr.html

[Husc] ——, “CIDR Report AS Report (Report for AS 6389).” [Online]. Available: http://www.cidr-report.org/cgi-bin/as-report?as=AS6389&view=6447

[Hus01] ——, “Analyzing the Internet’s BGP routing table,” *The Internet Protocol Journal*, vol. 4, no. 1, pp. 2–15, 2001. [Online]. Available: http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_4-1/bgp_routing_table.html

[Hus08] ——, “Autonomous System (AS) Number Reservation for Documentation Use,” RFC 5398 (Informational), Internet Engineering Task Force, Dec. 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5398.txt

[Hus09] ——. (2009, March) BGP in 2008. [Online]. Available: http://www.potaroo.net/ispcol/2009-03/bgp2008.html

[Hus11] ——, “Personal correspondence,” 2011, Personal correspondence with Stephen Woodrow, 1 March 2011.

[KcFB07] D. Krioukov, k. c. claffy, K. Fall, and A. Brady, “On compact routing for the internet,” *SIGCOMM Comput. Commun. Rev.*, vol. 37, pp. 41–52, July 2007. [Online]. Available: http://doi.acm.org/10.1145/1273445.1273450

[KR06] E. Karpilovsky and J. Rexford, “Using forgetful routing to control BGP table size,” in *Proceedings of the 2006 ACM CoNEXT conference*, ser. CoNEXT ’06. New York, NY, USA: ACM, 2006, pp. 2:1–2:12. [Online]. Available: http://doi.acm.org/10.1145/1368436.1368439

[KS11] W. Kumari and K. Sriram, “Deprecation of the use of BGP AS_SET, AS_CONFED_SET,” RFC 5398 (Informational), Internet Engineering Task Force, May 2011. [Online]. Available: http://tools.ietf.org/id/draft-ietf-idr-deprecate-as-sets-04.txt

[Li06] T. Li. (2006, October) Router Scalability and Moore’s Law. Presentation at IETF Routing and Addressing Workshop. [Online]. Available: http://www.iab.org/about/workshops/routingandaddressing/Router_Scalability.pdf

[Li11a] ——, “Recommendation for a Routing Architecture,” RFC 6115 (Informational), Internet Engineering Task Force, Feb. 2011. [Online]. Available: http://www.ietf.org/rfc/rfc6115.txt

[Li11b] ——, “Personal correspondence,” 2011, Personal correspondence with Stephen Woodrow, 23 March 2011.

[LIJM+10] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, “Internet inter-domain traffic,” in *Proceedings of the ACM SIGCOMM 2010 conference on SIGCOMM*, ser. SIGCOMM ’10. New York, NY, USA: ACM, 2010, pp. 75–86. [Online]. Available: http://doi.acm.org/10.1145/1851182.1851194

[Mar00] C. D. Marsan. (2000) Stanford move rekindles ’Net address debate. [Online]. Available: http://www.networkworld.com/news/2000/0124ipv4.html

[Mat09] A. Mathew, “Pass the Packet, Please? Explorations into Trust and the Social Organisation of Network Administration,” Master’s thesis, School of Information, UC Berkeley, 2009. [Online]. Available: http://www.ischool.berkeley.edu/programs/masters/projects/2009/passthepacket

[MC10] A. Mathew and C. Cheshire, “The New Cartographers: Trust and Social Order within the Internet Infrastructure,” in *Proceedings of the 2010 Telecommunications Policy Research Conference*, 2010.

[McK06] N. McKeown. (2006) Internet Routers: Past, Present and Future. Presentation at a British Computer Society meeting. [Online]. Available: http://tiny-tera.stanford.edu/~nickm/talks/BCSv6.0.ppt

[Mue10] M. Mueller, "Critical resource: An institutional economics of the Internet addressing-routing space," *Telecommunications Policy*, vol. 34, no. 8, pp. 405 – 416, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/B6VCC-5091YH1-1/2/f7e7761313d098520109c9500a400817

[MZF07] D. Meyer, L. Zhang, and K. Fall, "Report from the IAB Workshop on Routing and Addressing," RFC 4984 (Informational), Internet Engineering Task Force, Sep. 2007. [Online]. Available: http://www.ietf.org/rfc/rfc4984.txt

[NANa] "NANOG Historic Mailing List Archive." [Online]. Available: http://www.nanog.org/mailinglist/mailarchives/old_archive/

[NANb] "North American Network Operators Group (NANOG) Mailing List Archive." [Online]. Available: http://mailman.nanog.org/pipermail/nanog/

[NG03] H. Nussbacher and B. Greene. (2003, February) CIDR Police: Please Pull Over and Show Us Your BGP Announcements. Presentation at NANOG 27 meeting. [Online]. Available: http://www.nanog.org/meetings/nanog27/presentations/hank.pdf

[Ols82] M. Olson, *The Rise and Decline of Nations*. Yale University Press, 1982.

[Ost90] E. Ostrom, *Governing the Commons*. Cambridge University Press, 1990.

[Per02] C. Perkins, "IP Mobility Support for IPv4," RFC 3344 (Proposed Standard), Internet Engineering Task Force, Aug. 2002, obsoleted by RFC 5944, updated by RFC 4721. [Online]. Available: http://www.ietf.org/rfc/rfc3344.txt

[PK08] A. Pilosov and A. Kapela. (2008, October) Stealing The Internet: An Internet-Scale Man In The Middle Attack. Presentation at NANOG 44 meeting. [Online]. Available: http://www.nanog.org/meetings/nanog44/presentations/Tuesday/Kapela_steal_internet_N44.pdf

[Pos81] J. Postel, "Internet Protocol," RFC 791 (Standard), Internet Engineering Task Force, Sep. 1981, updated by RFC 1349. [Online]. Available: http://www.ietf.org/rfc/rfc791.txt

[QIdLB07] B. Quoitin, L. Iannone, C. de Launois, and O. Bonaventure, "Evaluating the benefits of the locator/identifier separation," in *Proceedings of 2nd ACM/IEEE international workshop on Mobility in the evolving internet architecture*, ser. MobiArch '07. New York, NY, USA: ACM, 2007, pp. 5:1–5:6. [Online]. Available: http://doi.acm.org/10.1145/1366919.1366926

[RIP] RIPE NCC, "BGPDump." [Online]. Available: https://bitbucket.org/ripencc/bgpdump/wiki/Home

[RL94] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 1654 (Proposed Standard), Internet Engineering Task Force, Jul. 1994, obsoleted by RFC 1771. [Online]. Available: http://www.ietf.org/rfc/rfc1654.txt

[RLH06] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," RFC 4271 (Draft Standard), Internet Engineering Task Force, Jan. 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4271.txt

[Rou] "University of Oregon Route Views project." [Online]. Available: http://www.routeviews.org/

[RR99] E. Rosen and Y. Rekhter, "BGP/MPLS VPNs," RFC 2547 (Informational), Internet Engineering Task Force, Mar. 1999, obsoleted by RFC 4364. [Online]. Available: http://www.ietf.org/rfc/rfc2547.txt

[RRB97] Y. Rekhter, P. Resnick, and S. M. Bellovin, "Financial incentives for route aggregation and efficient address utilization in the Internet," in *Coordinating the Internet*. Cambridge, MA, USA: MIT Press, 1997, pp. 273–287.

[SCE+05] L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica, "HLP: a next generation inter-domain routing protocol," in *Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, ser. SIGCOMM '05. New York, NY, USA: ACM, 2005, pp. 13–24. [Online]. Available: http://doi.acm.org/10.1145/1080091.1080095

[SEH06] P. Smith, R. Evans, and M. Hughes. (2006) RIPE Routing Working Group Recommendations on Route Aggregation. [Online]. Available: http://www.ripe.net/ripe/docs/ripe-399

[Ste10] R. Steenbergen. (2010, October) An Inconvenient Prefix: Is Routing Table Pollution Causing Data Center Warming? Presentation at NANOG 50 meeting. [Online]. Available: http://www.nanog.org/meetings/nanog50/presentations/Monday/NANOG50.Talk49.Steenbergen.routingtable.pdf

[TMS07] P. Traina, D. McPherson, and J. Scudder, "Autonomous System Confederations for BGP," RFC 5065 (Draft Standard), Internet Engineering Task Force, Aug. 2007. [Online]. Available: http://www.ietf.org/rfc/rfc5065.txt

[Tra95] P. Traina, "Experience with the BGP-4 protocol," RFC 1773 (Informational), Internet Engineering Task Force, Mar. 1995. [Online]. Available: http://www.ietf.org/rfc/rfc1773.txt

[vB02] I. van Beijnum, *BGP*. O'Reilly Media, Inc., 2002.

[Wu08] W. Wu, *Packet Forwarding Technologies*. Auerbach Publications, 2008.

[YCB07] X. Yang, D. Clark, and A. Berger, "NIRA: A New Inter-Domain Routing Architecture," *Networking, IEEE/ACM Transactions on*, vol. 15, no. 4, pp. 775 –788, aug. 2007.

[YMBB05] M. Yannuzzi, X. Masip-Bruin, and O. Bonaventure, "Open issues in interdomain routing: a survey," *Network, IEEE*, vol. 19, no. 6, pp. 49 – 56, nov.-dec. 2005.

[ZLWZ10] X. Zhao, Y. Liu, L. Wang, and B. Zhang, "On the Aggregatability of Router Forwarding Tables," in *INFOCOM, 2010 Proceedings IEEE*, march 2010, pp. 1 –9.

[ZPS10] X. Zhao, D. J. Pacella, and J. Schiller, "Routing scalability: an operator's view," *IEEE J.Sel. A. Commun.*, vol. 28, pp. 1262–1270, October 2010. [Online]. Available: http://dx.doi.org/10.1109/JSAC.2010.101004

[ZPW+01] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "An analysis of BGP multiple origin AS (MOAS) conflicts," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*, ser. IMW '01. New York, NY, USA: ACM, 2001, pp. 31–35. [Online]. Available: http://doi.acm.org/10.1145/505202.505207

Colophon

This thesis is set in 12 point Times New Roman, and was typeset with LaTeX using a modified version of the MIT thesis class[1]. The bibliography and citations were managed using BibTeX with the help of BibDesk. Chris Monson's LaTeX Makefile[2] was used to automate the thesis build process and to produce PDF output files.

All figures (except those included from other sources) were created with R or GraphViz and output as PDFs for inclusion by LaTeX. One figure was augmented with Adobe Illustrator.

The analysis and data presented in this thesis were generated with the help of R, PostgreSQL, tools written in Python, and a number of the GNU Coreutils utility programs, in addition to the tools mentioned in the body of the thesis (straightenRV, libbgpdump, etc.).

Source code for the analysis tools and the thesis were managed using Git and GitHub[3]. Source files were edited using SubEthaEdit, Vim and jEdit.

An Apple MacBook and a Lenovo Thinkpad X200, running Mac OS X 10.6 and Ubuntu 10.10 respectively, were the primary systems used to produce this thesis. A large multicore system with many gigabytes of memory and terabytes of disk array running Debian Squeeze was used for data processing.

[1] `http://stuff.mit.edu/afs/athena.mit.edu/reference/thesis/tex/`
[2] `http://code.google.com/p/latex-makefile/`
[3] `github.com`