

Understanding Incentives for Prefix Aggregation in BGP

Costas Kalogiros
Athens University of
Economics and Business
Department of Informatics
76 Patission, Athens 10434, Greece
ckalog@aueb.gr

Marcelo Bagnulo
University Carlos III of Madrid
Department of Telematic Engineering
30 Avda de la Universidad,
Madrid, Spain
marcelo@it.uc3m.es

Alexandros Kostopoulos
Athens University of
Economics and Business
Department of Informatics
76 Patission, Athens 10434, Greece
alexkosto@aueb.gr

ABSTRACT

Over the last few years, a significant amount of the effort of the Future Internet architecture is devoted in order to improve the scalability of the next generation routing architecture. In this paper, we study providers' incentives to perform prefix aggregation or deaggregation of non-customers routes. This is essentially a trade-off between reduced router memory and reduced capacity of attracting customer traffic. We study the case where two ISPs compete for attracting traffic, by using game theory. In particular, we propose a game-theoretic model and we analyze the properties of the equilibrium. In a symmetric case, if a single Autonomous System (AS) is found to be deaggregating a given prefix, then all others will have the incentive to do the same, even if they end up with lower benefits. We find that pure equilibria do not always exist and we derive the conditions based on two model parameters. These findings suggest that BGP instability can be a common problem in a competitive scenario.

Categories and Subject Descriptors

C.2.1 [Computer Communication Networks]: Network Architecture and Design, J.4 [Social And Behavioral Sciences]: Economics

General Terms: Design, Economics

Keywords: Prefix aggregation, BGP, Multihoming, Incentives, Game theory, Future Internet

1. INTRODUCTION

Internet is one of the world's most remarkable engineering and social phenomena. Despite its success, it has become a playground where stakeholders with conflicting interests interact with each other. These ongoing "tussles" may constitute a threat to the Future Internet [6].

The BGP global routing table is growing at a super-linear rate and it is uncertain whether the available technology will be able to cope with the expected growth rate, both in terms of size and update rate [9]. Even if it is feasible to produce routers that are able to handle the future global routing tables, the expected lifetime of deployed

equipment may be significantly reduced and periodic router refreshes with new technology will be required to keep the Internet global routing system working [9]. This may significantly affect the economic viability of the Internet as we know it, since this reduction in the life-cycle of the deployed equipment may significantly increase the operator's capital expenditure required.

One may argue that Internet is a victim of its own success and this phenomenon is the consequence of the increasing number of new users. This can be partially attributed to the fact that the currently widely-used BGP-based multihoming technique [11] imposes that each multihomed network contributes with an additional entry to the global routing table. While this is certainly true, there seems to be more to it than the growth due to more networks being reachable over the Internet. In particular, it can be observed that the BGP global routing table is growing much faster than the number of address blocks that are being allocated [8]. This basically means that when a new block is allocated, the increase of the global routing table is more than one route. So, when an address prefix is allocated from an Internet Registry, multiple routes covering different parts of the newly allocated address block are announced in BGP and populate the global routing tables. For example, it is possible to observe that about 50% of the entries of the current BGP global routing table correspond to the so-called more specific prefixes that are contained in other less specific prefixes, which are also present in the global routing tables [8]. Such behavior, that is called *deaggregation*, bloats the global routing table size, preventing the aggressive address aggregation capabilities of the current *Classless Inter Domain Routing* (CIDR) architecture, and severely reduces the routing system scalability.

The reasons for the presence of these more specific prefixes in the global routing table have been studied, for instance in [8] and [1], and they include multihoming, traffic engineering and local routing policies. By announcing both the less specific prefix and a set of more specific prefixes, an AS can achieve powerful traffic engineering capabilities [11]. So, when a multihomed AS needs that a certain block of addresses is reachable over a certain ingress path, it can achieve so by announcing a route containing the more specific prefix corresponding to that address block through the target ingress link. The effect is that the traffic corresponding to that address block flows through the ingress link through which the route for the more specific prefix is announced.

While the current practice of deaggregation has a significant detrimental effect on the routing system's scalability as a whole, it does provide important benefits to the origin ASs that inject the more specific routes. Moreover, as this is a widely used technique, a given origin AS would hardly see any benefit in the form of routing table size reduction by suppressing its own routes, as long as the rest

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ReArch '09, December 1, 2009, Rome, Italy.

Copyright 2009 ACM 978-1-60558-749-3/09/12...\$10.00.

of the other ASs still inject the routes corresponding to the more specific routes. So, it is pretty straightforward to conclude that an origin AS has incentive to continue announcing the more specific routes. A similar argument can be constructed for the transit providers of the origin AS that is injecting the more specific routes. Namely, the transit provider is charging the origin AS for its services, and due to the competition it has a strong incentive to honor the announcements of the origin AS. So, while the announcement of more specific prefixes bloats the transit provider routing table, it does obtain an economic profit out of it. Thus, the transit provider has a clear incentive to propagate the routes corresponding to the more specific prefixes to the rest of the Internet.

However, as the relationship with the originating AS is more distant, the direct benefits of storing and propagating the more specific prefix routes are heavily reduced. In particular, if we consider an ISP that has not direct nor indirect commercial relationship with the originating AS (i.e. the ISP is not selling traffic to the originating AS, nor to any of its providers), the nature of the trade-off between storing and propagating the routes, containing the more specific prefixes and aggregating them, changes significantly. These ISPs do not get directly paid by a customer for propagating these more specific routes. The benefit they obtain is the indirect benefit of attracting more traffic, thanks to the longest-prefix match rule that will make the route through a more specific prefix more attractive than a less specific.

A widely used approach to estimate the usage level of resources for long timescales (i.e. a month) when demand is bursty is the *95th percentile* rule. In that case, the bandwidth used by a customer during a specific period (i.e. 10 minutes) is estimated and divided by the time period (in seconds) resulting in a single bps (bits per sec.) transfer measurement. At the end of the billing cycle period, all measurements are sorted in decreasing order and the top 5% of these measurements are thrown out. The next highest measurement is the 95th percentile, and the customer will be billed based on that rate.

The fact that providers' revenues are based on the level of traffic creates an incentive to make the routes they advertise more attractive. This situation creates a trade-off between routing table scalability and service attractiveness; an AS has to make the choice between aggressive aggregation (which implies reduced routing table, hence reduced costs), and to propagate the routes containing the more specific prefixes (which increases the transit attractiveness, so increased revenues). In this paper, we attempt to characterize this trade-off and to provide some insight about when it is better for an ISP to aggregate and when it is better to propagate the more specific prefix routes.

A significant amount of the effort of the Future Internet architecture is devoted to improve the scalability of the next generation routing architecture. It is essential to this task to have a deep understanding of the aggregation incentives that exist in the current Internet and use that knowledge as an input in the design of future architectures.

This paper makes several contributions. First, we aim to provide some insights on the incentives for (de)aggregation in the current Internet routing system. Furthermore, we study the case where two ISPs compete for attracting traffic, by using game-theoretic modeling. We examine the properties of the game equilibria when providers decide what routes to propagate selfishly, and we derive

conditions for the game equilibrium. Finally, we evaluate providers' behaviour by estimating the variables of the above conditions.

The paper is structured as follows. We give an overview of related work in Section 2 and give an estimation of possible benefits due to aggregation in Section 3. Sections 4 and 5 describe the scenario we are interested in and the game theoretic model. We try to understand the current market situation in Section 6. Finally, we conclude and outline our future work in Section 7.

2. RELATED WORK

A recent strand of work in Internet routing treats ISPs as selfish economic entities who execute BGP so that they maximize their utility. The purpose is to study the properties of equilibria resulting from providers' interaction and, if necessary, to suggest conditions and mechanisms that can achieve desirable outcomes.

[4] and [2] identify conditions for stability of BGP when ISPs select outgoing paths towards a destination, based on a preference function of path desirability. Continuing this model, [7] proved that if the previous conditions hold and a domain can check whether packets follow the advertised path, then BGP is incentive-compatible. However, [3] concludes that this may not be the case if the preference function of ISPs is sensitive to the level of attracted traffic as well. For example, ISPs can have the incentive to advertise non-existing paths if traffic attraction and best path selection must be considered simultaneously. Furthermore, the authors of [3] look for policy conditions or additional tools that would render the option of lying useless, or in other words conditions for making truth-telling the dominant strategy.

[3] studies ISPs' incentives to use BGP advertisements, in order to attract traffic from customers. The main difference from our work is that we focus on ISPs' decision to aggregate (or not) a set of more specific prefixes they learned from a neighbour into a less specific one. This issue is explicitly out of the scope of the previous work. As mentioned before, prefix aggregation can significantly lower Tier-1 ISPs' cost and at the best of our knowledge this is the first work to study providers' incentives to use it. Besides, the *longest-prefix rule* makes announcing more specific prefixes in competitive scenarios much more advantageous than simply manipulating path attributes (e.g. the *AS-PATH*). However, aggressive aggregation can be seen as a special type of lying; for example aggregating routes with different Origin AS may result in some destinations being unreachable.

3. QUANTIFYING THE BENEFITS OF AGGREGATION

Aggregation reduces the size of the routing tables. Such reduction, not only results in reduced memory requirements on the routers but also, as the more specific prefixes tend to be much more flappy than the aggregated prefixes [8], aggregated routing tables also require less update processing from the routers. As a result, an AS that performs aggressive aggregation will see its capital expenditure costs reduced. In order to understand the benefits that can be obtained in terms of reduced routing table, we need first to precisely define how aggregation is performed.

A conservative approach would follow the one used by the CIDR report aggregation summary that aggregates routes containing the

exact same AS path, in order to honor transit policies. According to the CIDR report¹, only doing that would reduce the size of the routing table in 40% (in July 2009). However, it would be possible to perform an even more aggressive aggregation strategy that would not respect the transit policies, but would preserve reachability. For instance, it would be possible to also aggregate routes which prefixes are susceptible to be aggregated in a less specific one and that contain the same origin AS (even if the routes differ in the AS path). According to [8], 50% of the routing tables correspond to more specific prefixes contained in less specific prefixes also present in the routing table. Of those, 44% of them the more specific prefix and the less specific prefix share the same origin AS. It would be possible to find even more aggressive rules that can preserve reachability, such as also aggregating routes that the AS path of the more specific prefix is fully contained in the AS path of the route of the less specific prefix. These correspond to 12% of the more specific prefixes contained in less specific prefix also present the routing table, according to [8].

Thus, using all the proposed aggregation rules would result in a reduction of the 50% of the BGP table size². Since the current BGP table has about 300.000 entries³, the net reduction would be of 150.000 entries. It is possible to actually quantify this cost reduction. According to [5], the cost of a route in the global routing table is 4 US cents per route per router per year⁴. The cost reduction that corresponds to this aggregation would be of 6.000 US per router per year. According to [5], the cost of the router capable of supporting the current global routing table is around 40.000 US, so aggregation would imply a 15% decrease in the router cost (if assumed linear).

4. AN AS-ORIGIN DEAGGREGATION SCENARIO

In Figure 1, we see a simple network topology composed of 6 ASs. AS1 and AS6 are retail multi-homed ISPs. AS2 and AS3 are Tier-2 providers of AS1. AS2 and AS3 have peering relationships with Tier-2 providers AS4 and AS5. Finally, AS6 gets connectivity from both AS4 and AS5.

Dashed arrows indicate the direction of BGP messages (blue rectangles) so that routing tables can be formed. We will describe the procedure for the prefix 163.144.0.0/23, which will allow data packets from customers of AS6 to be received by customers of AS1. Following a similar approach, AS1 would learn how to reach customers of AS6 (not described here). We should note that this is one realistic scenario, out of the many possible ones where the model of Section 5 applies.

Figure 1 shows that in the beginning, only AS1 knows how to reach its customers. AS1 wants to perform traffic engineering and deaggregates his prefix 163.144.0.0/23 into 163.144.0.0/24 and 163.144.1.0/24. The BGP advertisement message to AS2 contains

the route for the more specific prefix 163.144.0.0/24. Similarly, the BGP advertisement message to AS3 contains both the route for the more specific prefix 163.144.1.0/24, as well as the aggregated one. Namely, half of the more specific prefixes are advertised to AS2 and the rest half of them are advertised to AS3. Finally, the original prefix 163.144.0.0/23 is being advertised through both paths, as a backup solution.

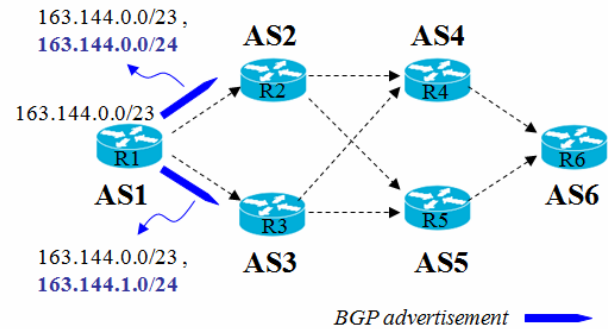


Figure 1. Topology of our scenario.

AS2 (and AS3 respectively) will run the route selection process and advertise the two routes learned from the previous step to each upstream provider AS4 and AS5. In this way, both AS4 and AS5 will get the two routes containing the more specific prefixes and the route for the less specific prefix and they will have the option of aggregating them into a single route.

In the following, we will concentrate on Tier-1 providers' incentives (AS4 and AS5) to aggregate or not. This decision will depend on the trade-off between the marginal increase in cost and the marginal increase in revenues, due to deaggregation. Note that these costs and benefits may be stochastic.

5. THE GAME AMONG TWO SYMMETRIC ISPS

The setting where two transit ISPs, AS_i and AS_j , compete for attracting traffic can be studied by using game-theoretic modeling. In the simplest case, we can assume that all ISPs of the same tier are *symmetric*. This means that all competitors i) face similar routing management costs (for example they use similar router technology) and ii) have similar revenues (for example they have equal bargaining power and thus sign similar interconnection agreements with their customers and providers).

For simplicity, we will assume that the Origin-AS AS_o has just been allocated a single block of contiguous IP addresses, represented as $B/mask$. Furthermore, we will consider the case that traffic destined to this block has its source at a single Tier-1 ISP, called AS_s . We plan to extend our model and capture bidirectional traffic, or traffic between several pairs of providers. However, we believe that the results will not differ significantly, under the assumption of symmetry.

Let us suppose that AS_o performs deaggregation of $B/mask$ into more specific prefixes, in order to achieve better traffic engineering. AS_i and AS_j must decide what they will propagate to ASs, as soon as they receive the related BGP messages from their peers. Each one's decision will be based on the incoming BGP messages (that are

¹ <http://www.cidr-report.org/as2.0/>

² This number includes the customer routes, so the actual reduction will be lower than this.

³ <http://bgp.potaroo.net/>

⁴ This is calculated as "the portion of the router's cost attributable to the routing table divided by the expected deployment lifetime of the router in its DFZ (Default Free Zone) application divided by the total route count in the DFZ".

considered to be “common knowledge”) and on the possible action of the opponent. For simplicity, in this paper we study the “single-shot” simultaneous-move game, where both providers act once and at the same time.

5.1 Game Setup

We aim to examine how much aggregation or deaggregation the players of the game will select to perform, for a given prefix B . This is called *aggregation level* and is denoted by l_i and l_j respectively.

In case of IPv4, the minimum aggregation level l_{min} for a provider is to propagate /32s. On the other hand, the maximum aggregation level l_{max} is variable, but for simplicity we will assume that it has a lower bound at 8, since the aggregation of multiple /8 does not seem likely. For the prefix $B/mask$, l_{max} is restricted to mask and thus $l_i, l_j \in [l_{min}, l_{max}]$. More aggressive behaviour (i.e. aggregate routes by violating the first condition) even though is possible, may result in some destinations being unreachable. In game-theoretic terms, l_i is the action space of AS_i .

We assume that transit providers AS_i, AS_j aim to maximize their profits and thus the “reaction function” of AS_i (and respectively for AS_j) is $\Pi_i(l_i, l_j) = R_i(l_i, l_j) - C_i(l_i)$. In particular, the profit $\Pi_i(l_i, l_j)$ of AS_i depends on the actions of both providers, and is the difference of the revenues $R_i(l_i, l_j)$ and the routing table cost $C_i(l_i)$. More specifically, the revenues of AS_i can be written as:

$$R_i(l_i, l_j) = rt2^{(l_{min} - l_{max})}P_i(l_i, l_j), \text{ where}$$

- r is the average revenue from delivering one unit of traffic to a specific IP address. If an upstream provider charges his customers based on volume (i.e. by using the 95th percentile rule), r represents the price per traffic unit (i.e. cents per Mbps). We assume that the network is “neutral” which means that providers do not charge unilaterally, based on the recipient of traffic.
- t is the average traffic delivered to a specific IP address. We assume for the time being that all IP addresses are symmetric with respect to the traffic these accept.
- $2^{(l_{min} - l_{max})}$ refers to the number of IP addresses that are contained in the original prefix $B/mask$.
- $P_i(l_i, l_j)$ denotes the probability of AS_i being selected by AS_j to carry traffic for IP addresses belonging in a specific prefix.

We suppose that if a provider decides to deaggregate more (less aggregation level), he will be the one selected to transit traffic for the contained IP addresses. Contrarily, if a provider decides to aggregate more than the other one, he will lose all the traffic. This assumption can be attributed to the “longest-prefix” rule of BGP. Moreover, we assume that each provider will have equal chances of being selected when they select the same aggregation level. This means that:

$$P_i(l_i, l_j) = \begin{cases} 0 & \text{if } l_i < l_j \\ \frac{1}{2} & \text{if } l_i = l_j \\ 1 & \text{if } l_i > l_j \end{cases} \quad \text{and} \quad P_j(l_i, l_j) = 1 - P_i(l_i, l_j).$$

Furthermore, the cost related to the routing tables for AS_i depends on the number of entries $2^{(l_i - l_{max})}$ being stored for the original prefix $B/mask$ and on the cost for each entry k . We assume that the memory cost attributed to each stored route k is constant, for any given router. For simplicity, we also assume that there is enough free memory to store routes for the original prefix $B/mask$, even if the $l_i = l_{min}$, and that all entries have equal memory requirements. Note that $C_i(l_i)$ depends only on the aggregation level of AS_i , and can be written as $C_i(l_i) = 2^{(l_i - l_{max})}k$.

5.2 Game Equilibria

Continuing our scenario, we suppose that $B/mask = 163.144.0.0/23$ and AS_o deaggregates into $163.144.0.0/24$ and $163.144.1.0/24$. For illustrative purposes, let us examine a simpler case and try to generalize later. Suppose that transit providers AS_i and AS_j have a restricted action space; that is $l_i, l_j \in [23, 25]$. Thus, we have the following cases:

- $l_i < l_j$, for example participants play $(l_i, l_j) = (23, 25)$
- $l_i = l_j$, for example $(l_i, l_j) = (23, 23)$
- $l_i > l_j$, for example $(l_i, l_j) = (24, 23)$

Assuming that $t = 1$, the payoff matrix of this simple game is given by the Table 1. In this table, each cell (l_i, l_j) has two values separated by a comma. The first value is the net benefit of participant AS_i when he plays l_i and his opponent plays l_j . Respectively, the second value is the net benefit of participant AS_j .

Table 1. Example payoff matrix

		AS_j		
		23	24	25
AS_i	23	$2^8r - k, 2^8r - k$	$-k, 2^9r - 2k$	$-k, 2^9r - 4k$
	24	$2^9r - 2k, -k$	$2^8r - 2k, 2^8r - 2k$	$-2k, 2^9r - 4k$
	25	$2^9r - 4k, -k$	$2^9r - 4k, -2k$	$2^8r - 4k, 2^8r - 4k$

In that case, the best strategy⁵ for both transit ISPs is the same (due to symmetry) and depends on the actual values of r and k . More specifically, the best pure strategy l_i^* of AS_i can be written as:

if $r < k2^{-8}$ **then**

$$l_i^* = 23$$

else if $k2^{-8} < r < 2k2^{-8}$ **then**

$$l_i^* = 24$$

else if $r > 3k2^{-8}$ **then**

$$l_i^* = 25$$

end if

⁵ Strategies are detailed planned actions that take into account previous or anticipated responses from other players. At each stage of the game, players will pick actions from their strategy, depending on the information they gathered.

The above conditions emerged by solving for the values of r that make a specific aggregation level the best reply for both players. When both providers decide to choose the same aggregation level (the diagonal of the payoff matrix), then they would expect to have the same revenues. However, the more they deaggregate, the less profits both providers have. That happens because if the aggregation level equals 23 the cost equals k . On the other hand, if the aggregation level equals 24 or 25 the cost will be $2k$ and $4k$ respectively.

Thus, if $r > 3k2^{-8}$ or $k2^{-8} < r < 2k2^{-8}$, both transit providers have the incentive to deaggregate. This behaviour is optimal for the selfish providers, because if a single one decided to aggregate, he would not attract any traffic for that prefix. But it is not socially optimum since, e.g. in the first case $2(2^{-8}r - 3k) < 2(2^{-8}r - k)$. This situation is known as the *prisoners' dilemma*⁶. On the other hand, if $r < k2^{-8}$, the routing memory unit cost is too high and both providers are better off by aggregating. In that case, the pure strategy equilibrium outcome is socially optimum also.

Note that there are some combinations of r and k where no pure strategy Nash Equilibrium exists. This happens either when $r = k2^{-8}$ or when $2k2^{-8} \leq r \leq 3k2^{-8}$. In the former case, each ISP is indifferent between the two aggregation levels, regardless of the opponent's choice. In the latter case, there is no dominant strategy because any possible outcome is outweighed by at least another one. This situation is depicted in the Figure 2 where AS_j has to decide his best aggregation level l_j by considering any possible action from AS_i . Even though we study the single-shot game, participants are assumed to be clever enough to take into account anticipated opponent actions in the future. Each circle represents a possible outcome (best responses from a player), where the first value refers to the l_i and the second to l_j chosen respectively. The aggregation level which is shown as bold represents the best response of the respective player. Slim arrows represent the first move of AS_i , while thick arrows denote how the game would be played by transitioning from one outcome to another.

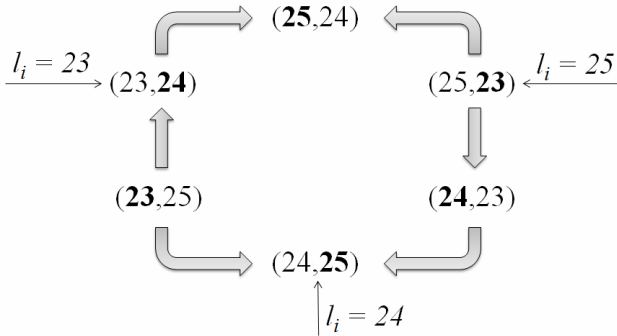


Figure 2. An example of no Pure Nash Equilibrium.

Assuming in our example that $r = k2^{-8}$, if $l_i = 23$, the best response from AS_j would be $l_j = 24$. Expecting this rational response from AS_j , ISP AS_i would respond by setting $l_i = 25$. Applying this logic

⁶ In this game, the “dilemma” faced by the prisoners (the ISPs respectively, in our example) is that, whatever the other does, each is better off confessing than remaining silent. However, the outcome obtained when both prisoners confess (both ISPs deaggregate) is worse for each than the outcome they would have obtained if both had remained silent (both ISPs aggregate).

repeatedly, we see that no player will ever stop responding. Thus, there is no equilibrium because each provider is willing to deaggregate at the maximum allowed level, in order to attract traffic. However, if both providers act aggressively, their cost will exceed revenues (since traffic will be split and cost will be increased). Consequently, whenever one ISP selects the maximum deaggregation level, the opponent's best response is the minimum deaggregation level. Intermediate levels only increase the cost, due to longest-prefix rule. But, in that case $23 < l_i < 25$ since saving router memory increases profit. This in turn allows AS_j to attract traffic by selecting the maximum deaggregation level, resulting in an unstable environment.

These results can be generalized for the case where $l_i, l_j \in [l_{max}, l_{min}]$. In particular, following a similar procedure, we find that for any $\alpha \in [0, \varphi + 1]$, where $\varphi = l_{min} - l_{max} - 1$, the pure strategy equilibrium is:

$$\begin{aligned} &\text{if } 0 \leq \alpha < \varphi + 1 \text{ then} \\ &\quad (l_i^*, l_j^*) = \begin{cases} (l_{max} + a, l_{max} + a) & \text{if } \frac{(2^\alpha - 1)k}{2^\alpha t} < r < \frac{2^\alpha k}{2^\alpha t} \\ (\emptyset, \emptyset) & \text{if } \frac{2^\alpha k}{2^\alpha t} \leq r \leq \frac{(2^{\alpha+1} - 1)k}{2^\alpha t} \end{cases} \\ &\text{else if } \alpha = \varphi + 1 \text{ then} \\ &\quad (l_i^*, l_j^*) = (l_{max} + a, l_{max} + a) \quad \text{if } \frac{(2^\alpha - 1)k}{2^\alpha t} < r \\ &\text{end if} \end{aligned}$$

where (\emptyset, \emptyset) denotes that no pure equilibrium aggregation level exists. An example scenario when $\varphi = 8$ and $l_{max} = 23$ is shown in Figure 3.

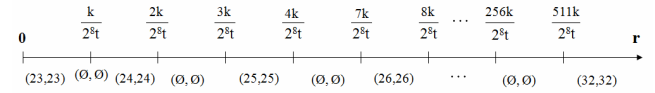


Figure 3. An example of possible equilibria.

6. ESTIMATING THE PARAMETER r

So far, we have proposed a model that captures the trade-off between the added revenue resulting from announcing a route containing a more specific prefix and the associated cost in terms of capital expenditure in router equipment. In this section, we will propose a model for estimating the parameter r of the model described in Section 5. r is the average revenues per month from delivering one unit of traffic to a specific IP address in a given prefix contained in a route. The goal is to apply the results of Section 5 and determine the aggregation level at the equilibrium in a scenario that approaches reality. Note that our aim during estimations is to find the order of magnitude, instead of the exact calculation.

As a preliminary observation, it should be noted that the revenue associated to a particular route announcement depends, among other parameters, of the charging model of the ISP and the traffic that each client sends to the particular set of destinations contained in the prefix of the announced route. All this data is available to the ISP itself, so it should be possible for each ISP to apply the proposed model and determine for each prefix that is susceptible to be aggregated, i.e. whether the revenue resulting from announcing the route is worth the cost of storing the route in the routing table. In

order to do so, the ISP needs to quantify the amount of traffic sent to the target prefix, calculate the traffic reduction that would imply not to announce the route for the more specific and evaluate the cost / revenue trade-off.

We will estimate r by relying on the fact that most ISPs charge their customers based on volume. We assume the traffic of an ISP is uniformly distributed across all the IP addresses. While this assumption is clearly unrealistic, it may be useful as a first approximation. Moreover, since we are dealing with prefixes rather than with single IP addresses, this approximation is likely to improve significantly with respect to the single IP address case.

So, if an ISP is selling b units of transit bandwidth each month, and if we call n the total number of IP addresses that are reachable using routes contained in the global routing table, then each address is sinking b/n units of transit bandwidth. If we call t to the price charged by the ISP per month per unit of transit traffic, then the revenue r per month resulting from announcing such address in the more specific route can be calculated as: $r = tb/n$.

Following [10] let us suppose that the Internet market is composed of 10 Tier-1 ISPs, 1,000 Tier-2 ISPs and 5,000 Tier-3 ISPs, that $t = 10$ US dollars per month per Mbps and $b = 1680$ Mbps. Furthermore, we assume that each Tier-2 ISP has 3 upstream providers and that all Tier-1 ISPs have the same market share. This means that each Tier-1 ISP transits $b/3$ Mbps for each of his 300 customers. The number of allocated number of IPv4 addresses has been estimated to be $n = 3,706,650,624^7$, (July 2009). Thus $r = 0.0062$, which means that no pure equilibrium would exist since $2^{9rt} = 15.79$ that lies between $[2^ak, (2^{a+1} - 1)k]$ for $a = 8$ and when $k = 0.04$. This suggests that stability of BGP due to prefix deaggregation can be a realistic problem.

7. CONCLUSIONS AND FUTURE WORK

We have analyzed providers' incentives to perform aggregation and deaggregation of non-customers routes. This is essentially a trade-off between reduced router memory (which results in reduced costs) and reduced capacity of attracting customer traffic (which results in reduced revenues). We proposed a game theoretic model and we analyzed the properties of the equilibrium, when there is one. We have found that, in a symmetric case, if a single AS deaggregates a prefix, then all others ISPs will have an incentive to do so, even if they end up with lower benefit. We found that pure equilibria do not always exist and we have derived the conditions that are based on two model parameters; the cost associated to additional memory caused by deaggregation and the revenue obtained from announcing a route containing a more specific prefix. We conclude that BGP instability can be a frequent problem in a competitive scenario.

We plan to extend our model by introducing asymmetry in terms of traffic volume that is being originated or accepted by IP addresses. Since ISPs have the necessary information, they could solve a profit maximization problem, in order to select the optimum aggregation level for each IP address.

Exchange of BGP messages usually happens asynchronously and iteratively. Thus, a repeated-game formulation would achieve a more realistic representation of the situation. We would like to examine if there is a trade-off between profit maximizing strategies

and BGP stability. For example, if strategies like “*Deaggregate one level if opponent is not maximum aggressive, otherwise aggregate at the maximum level*” proved to be optimal, they could lead to instability.

In addition, we are working in improved means to estimate the revenue r that removes the assumption that traffic is uniformly distributed across the IP address space announced in the BGP routing table. In particular, we are considering a model that proposes an inverse relationship between the amount of traffic that a given route sinks and the AS path length attribute. The intuition behind this assumption is that popular destinations are closer, as a result of peering incentives and content distribution networks.

8. ACKNOWLEDGMENTS

This research was supported by Trilogy (<http://www.trilogy-project.org>), a research project (ICT-216372) partially funded by the European Community under its Seventh Framework Programme.

9. REFERENCES

- [1] T. Bu, L. Gao, and D. Towsley. On Characterizing BGP Routing Table Growth. *Computer Networks*, 45(1):45–54, 2004.
- [2] L. Gao and J. Rexford. A Stable Internet Routing without Global Coordination. *IEEE Transactions on Networking*, 9:681–692, December 2001.
- [3] S. Goldberg, S. Halevi, A.D. Jaggard, R.N. Wright, and V. Ramachandran. Rationality and Traffic Attraction: Incentives for Honest Path Announcements in BGP. *SIGCOMM 2008*, 267–278, New York, USA.
- [4] T.G. Griffin, F.B. Shepherd, and G. Wilfong. The Stable Paths Problem and Interdomain Routing. *IEEE/ACM Transactions on Networking*, 10(2):232–243, 2002.
- [5] W. Herrin. What does a BGP Route Cost? <http://bill.herrin.us/network/bgpcost>, February 2008.
- [6] C. Kalogiros, A. Kostopoulos, and A. Ford. On Designing for Tussle: Future Internet in Retrospect. *EUNICE 2009*, LNCS 5733, 98–107, Barcelona, Spain, September 2009.
- [7] H. Levin, M. Schapira, and A. Zohar. Interdomain routing and games. *STOC '08 - 40th annual ACM Symposium on Theory of Computing*, 57–66, New York, USA, 2008.
- [8] X. Meng, Z. Xu, B. Zhang, G. Huston, S. Lu, and L. Zhang. IPv4 Address Allocation and the BGP Routing Table Evolution. *SIGCOMM Computer Communication Review*, 35(1):71–80, 2005.
- [9] D. Meyer, L. Zhang, and K. Fall. Report from the IAB Workshop on Routing and Addressing. *RFC 4984 (Informational)*, September 2007.
- [10] S. Nemo. Inter-ISP Bandwidth Markets. <http://web.invisiblehand.net> (White Paper), 2001.
- [11] I. van Beijnum. *BGP - Building Reliable Networks with the Border Gateway Protocol*. O'Reilly, Sebastopol, CA, first edition, September 2002.

⁷ <http://www.bgpexpert.com/addressespercountry.php>