Get Off My Prefix! The Need for Dynamic, Gerontocratic Policies in Inter-domain Routing

Edmund L. Wong and Vitaly Shmatikov
Department of Computer Science
The University of Texas at Austin
{elwong,shmat}@cs.utexas.edu

***Abstract*—Inter-domain routing in today's Internet is plagued by security and reliability issues (*e.g.*, prefix hijacking), which are often caused by malicious or Byzantine misbehavior. We argue that route selection policies must move beyond static preferences that select routes on the basis of static attributes such as route length or which neighboring AS is advertising the route.**

We prove that route convergence in the presence of Byzantine misbehavior requires that the route selection metric include the dynamics of route updates as a primary component. We then describe a class of simple dynamic policies which consider the observed "ages" of routes. These *gerontocratic* policies can be combined with static preferences and implemented without major infrastructural changes. They guarantee convergence when adopted universally, without sacrificing most of the flexibility that autonomous systems enjoy in route selection. We empirically demonstrate that even if adopted unilaterally by a single autonomous system, gerontocratic policies yield significantly more stable routes, are more effective at avoiding prefix hijacks, and are as responsive to legitimate route changes as other policies.

I. Introduction

The Internet consists of domains or autonomous systems (ASes) which rely on the Border Gateway Protocol (BGP) [27] to establish inter-domain routes connecting them. BGP was designed under the assumption that ASes are trusted. ASes are free to advertise arbitrary routes to any destination (identified by an IP address prefix), and these advertisements, which take the form of route updates, are not authenticated.

Unfortunately, there has been an increasing number of incidents in which *Byzantine* (malicious, misconfigured, or otherwise faulty) entities have taken advantage of this trust. Because it is difficult for the recipient of an update to verify that the AS advertising a route to a prefix has the right to do so, prefix "hijacks" [2], caused by either misconfiguration or malicious actions (*e.g.*, those associated with spam activity [22, 26]), have resulted in a number of very visible incidents, including the AS 7007 incident in 1997 [3] and the hijacks of Yahoo's prefixes by a Malaysian ISP in 2004 [22], of over 106,000 prefixes by a Turkish ISP in December 2004 [18, 22], of Google's prefix by Cogent in May 2005 that lasted for almost 2 days [18, 35], of 41 prefixes by RCN in January 2006 [18], of a sub-prefix of YouTube by Pakistan Telecom in February 2008 [31], and of 15% of all prefixes by a small Chinese ISP in April 2010 [32].

Many techniques have been proposed to improve the security and stability of inter-domain routing. S-BGP [16], soBGP [37], psBGP [34], and SPV [14] enable cryptographic validation of route updates. They have not found wide adoption due to their infrastructural requirements (*e.g.*, global public-key infrastructure in the case of S-BGP), the need for ASes to cooperate with each other, and other logistical challenges. Cooperation between ASes is required by other BGP security mechanisms [23, 25, 30, 40, 41]. Local methods for detecting anomalies in route updates [17, 29, 42] require ASes to maintain an accurate model of the AS connectivity graph and/or prefix ownership and can suffer from false positives.

Even if route updates were secured, Byzantine ASes could still cause disruptions. A key feature of BGP is the complete autonomy it gives to an individual AS in selecting routes from among those advertised by its neighbors. Common route selection policies are often based on *static preferences* over a route's (static) attributes, *e.g.*, path length or the AS's relationship to the neighbor advertising the route. Policies that prefer particular neighbors (*local preferences*) are widely believed to be based on the business agreements between ASes regarding the transit of each other's traffic [8]. Because route convergence is guaranteed only if policies satisfy strong constraints [6, 7, 8, 12], Byzantine ASes can prevent convergence by violating these conditions [20], by advertising desirable routes in order to attract traffic [10, 11], or by switching between different routes they are authorized to advertise, thereby causing other ASes, too, to oscillate between routes.

As a consequence, any inter-domain routing protocol that gives ASes full freedom to make routing decisions is susceptible to Byzantine misbehavior. We argue that stronger security mechanisms are not enough to solve this problem. ASes must also adopt smarter, *dynamic* route selection policies in order to deal with Byzantine ASes. In this paper, we show that dynamic policies are *necessary* for route stability by proving that a single Byzantine AS can prevent route convergence with any combination of static policies in which at least one AS prefers an indirect route to a particular destination, even if they satisfy the conditions that are sufficient for convergence in an all-rational environment. Therefore, robustness against Byzantine faults requires the route selection policy of each AS to incorporate the observed behavior of other ASes.

Fortunately, even simple dynamic policies yield large benefits. We describe a family of dynamic route selection policies that are *gerontocratic* in nature: unlike existing policies, which choose among advertised routes using static preferences, these policies incorporate information about each route's "age" (how long it has been continuously available) as a fundamental

component of route selection.

ASes can combine gerontocratic policies with static preferences, including those based on business relationships, and choose how much importance to give to the latter. Note that real-world relationships between ASes often include both peering/transit agreements and service-level agreements regarding the stability and performance of routes. Gerontocratic policies are purely local: adopting them does not require any global infrastructure or inter-AS cooperation. Thus, gerontocratic policies provide a simple, low-overhead, flexible policy routing solution that can be adopted by almost any AS with minimal modifications to BGP routers. Moreover, a universally adopted gerontocratic policy guarantees route convergence in the presence of Byzantine ASes, even if the policy otherwise prefers a specific neighbor or an indirect route and thus almost always selects routes with a desired static characteristic.

Even if a gerontocratic policy is adopted unilaterally, we show, using simulation based on Route Views data [1], that it selects very stable routes, lasting more than five times as long as the routes chosen by shortest-path policies and more than thirty times as long as those selected by local preferences. Even when combined with heavily weighted local preferences, gerontocratic policies greatly outperform policies based on local preferences alone. Furthermore, we show that gerontocratic policies are much less vulnerable to prefix hijacks while being as responsive to network failures as other policies.

The rest of the paper is organized as follows. We survey related work in Section II and address common objections to stability-based routing in Section III. We describe our model of inter-domain routing in Section IV and show how static policies fail in the presence of Byzantine ASes in Section V. We describe gerontocratic policies in Section VI. Section VII demonstrates that they greatly increase route stability while avoiding hijacked routes. Section VIII concludes.

II. Related Work

Stability-based approaches. Techniques intended to improve route stability include route flap damping [33], minimum route advertisement interval (MRAI) timers, and withdrawal rate-limiting (WRATE). Route flap damping assigns to every neighboring AS and prefix a penalty which increases when a route flaps (*e.g.*, is advertised or withdrawn) and decays exponentially over time. If a route flaps too quickly, route updates from this AS/prefix are suppressed until the penalty decays past some low watermark. MRAI/WRATE suppress advertisements/withdrawals of new routes from a particular AS or prefix until a certain interval elapses.

Route flap damping, MRAI, and WRATE are all-or-nothing: they either completely suppress the selection, advertisement, and withdrawal of a route, or they do nothing to depreference it. They do not guarantee route convergence and provide little benefit beyond reducing the number of repeated advertisements and withdrawals of unstable routes. We show in Section VII that, in contrast to gerontocratic policies, adding route flap damping to a shortest-path policy does not substantially improve the stability of the resulting routes. Furthermore, in pathological cases, a few flaps may cause routes to be suppressed for a long time, possibly resulting in a loss of connectivity [21].

Stability of popular inter-domain routes was observed in [28]. Stable Route Selection (SRS) [9] attempts to avoid short-lived disruptions by considering route stability—how long a route has been advertised and whether it is currently selected—*if and only if two routes have the same local preference*. While SRS and other previous techniques for route stability focused on empirical performance arguments, we both demonstrate the practical benefits of our approach and prove that route convergence in the presence of Byzantine misbehavior cannot be achieved unless route dynamics are considered as part of the primary route selection metric (*i.e.*, not as a tie-breaker). Therefore, in contrast to previous work, gerontocratic policies consider route age as a fundamental component of route selection.

Defenses against prefix hijacks. Pretty Good BGP (PGBGP) is based on the observation that prefix hijacks are often short-lived [15]. PGBGP temporarily depreferences a route advertisement for a given prefix if the advertised path does not contain an AS that recently originated advertisements to this prefix or any of its sub-prefixes. PGBGP guarantees neither convergence nor stability in the presence of Byzantine ASes. As we show in Section VII, gerontocratic policies outperform PGBGP in terms of route longevity and are equally effective in avoiding transient hijacks. Unlike gerontocratic policies, PGBGP cannot avoid hijacks that last longer than its preset interval, such as Cogent's hijack of Google's prefix [18, 35].

PHAS [18] collects control-plane data from BGP feeds and logs in order to detect suspicious changes in a prefix's origin AS. Other approaches use multiple vantage points in the data plane to acquire and cross-check fingerprints of selected destinations, such as route hop counts [13, 41]. These techniques require cooperation between ASes. If a hijack affects a significant fraction of the Internet, the victim network may be able to detect the attack by probing various destinations and checking how many replies are routed back correctly [39]. These data-plane techniques are vulnerable to intelligent adversaries who can recognize and re-route probes. They also do not prevent other ASes from selecting hijacked routes to the victim. Most importantly, they focus strictly on detecting prefix hijacks rather than improving route stability in general. They are thus complementary to our policy-based approach and can be used alongside it.

Cryptographic validation of route updates. Several proposals aim to secure BGP against invalid route advertisements [14, 16, 34, 37]. Some consider route stability, but only to help reduce authentication costs [4] rather than a fundamental principle of route selection. The use of cryptography has also been proposed as a way to tolerate Byzantine misbehavior in link-state routing [24]. In general, cryptographic validation of route updates is not sufficient for convergence because oscillation can be caused by policy conflicts even if all advertised routes are legitimate. Furthermore, these techniques require cooperation between ASes (*e.g.*, shared keys) and/or a global public-key infrastructure or a "web of trust."

By contrast, gerontocratic policies are local and provide

Technique	Guaranteed convergence	Local	Prefix hijack defense	Disadvantages (vs. gerontocratic)
Gerontocratic policies	**X**	**X**	**X**	—
Route flap damping [33]		X		May severely exacerbate convergence [21], no hijack defense.
PGBGP [15]		X	X	Significantly less stable routes, cannot handle longer hijacks.
SRS [9]		X		No hijack defense.
PHAS [18]			X	Requires global BGP feeds, manual intervention when hijacked.
Crypto-based approaches [14, 16, 34, 37]			X	Requires crypto infrastructure, not local.

TABLE I
COMPARISON OF OUR APPROACH TO EXISTING TECHNIQUES.

significant benefits even if adopted by a single AS. Moreover, gerontocratic policies are compatible with and complementary to virtually any proposed method for securing route updates.

Convergence and incentive-compatibility. There has been much research on conditions under which static route selection policies ensure convergence, such as the Gao-Rexford conditions [8] and the "no dispute wheel" [12], which is implied by the Gao-Rexford conditions [7]. Stability can improve if ASes are able to advertise and use different routes for different neighbors [36]. Gerontocratic policies give ASes a different kind of flexibility by allowing them to consider route stability along with (or instead of) static preferences without sacrificing convergence.

If every AS is rational, their route selection policies do not conflict, and ASes do not falsely advertise routes which were not advertised to them, BGP is guaranteed to converge to a stable set of routes [10, 19]. In this setting, BGP is incentive-compatible, *i.e.*, ASes derive no rational benefit from advertising routes other than those they actually prefer.

The assumption that every AS is rational may not hold in today's Internet. Unintentional misconfiguration may cause an individual AS to misbehave in a Byzantine fashion, preventing convergence [20]. Prefix hijacking and, in general, advertisement of non-existent or unavailable routes invalidate one of the conditions needed for incentive-compatibility to hold. We leave formalization of incentive-compatibility in the setting where route stability is explicitly considered as part of the route selection policy to future work.

III. FREQUENTLY ASKED "QUESTIONS"

In this section, we survey several objections to stability-based route selection policies and explain how they are addressed by the gerontocratic approach proposed in this paper.

"Longevity as a route selection metric is an old idea." Unlike previous longevity-based approaches, gerontocratic policies use route age as a fundamental component of route selection (as opposed to, say, a tie-breaker after local preferences [9]), while allowing ASes the flexibility to consider other factors in their decision process. For example, an AS may prefer a certain neighbor because of a business relationship or shorter routes. Incorporating age into the route selection metric yields better routes even in this case (see Section VII). Furthermore, we demonstrate the necessity of dynamic route selection policies via theoretical analysis.

"Gerontocratic policies will select routes containing failed links." By design, gerontocratic policies prefer longer-lived routes. Although a gerontocratic policy will never select an unavailable route, a potential concern is that if a network link fails in a route that is not promptly withdrawn, a gerontocratic policy may continue using this route instead of switching to an alternative route that does not contain the failed link.

We contend that this concern is misplaced. With the exception of direct links—whose failure is promptly handled by any policy—BGP per se does not provide information about whether or which links have failed. Without external help, no BGP route selection policy can tell the difference between an advertised route that contains a failed link and one that does not. In Section VII-D, we show that, on average, gerontocratic policies are just as likely to select a valid route as any plausible alternative policy. In particular, eagerly switching to newly advertised routes does not offer significant benefits over gerontocratic policies even in the case of network failures.

"Gerontocratic policies ignore business relationships between ASes." Route selection policies of real-world ASes may be governed by economic considerations and relationships with neighboring ASes rather than technical route characteristics such as stability. This is compatible with our approach. Our proposed route metric allows an AS to strongly bias its policy towards some static preferences (*e.g.*, those taking business agreements into account), should it desire to do so.

Incorporating static preferences does not affect the theoretical guarantees of our approach. Moreover, we argue that many of these business relationships have service-level agreements regarding the stability and performance of advertised routes; as we show in Section VII, considering route lifetime in addition to static preferences improves route stability and helps avoid prefix hijacks. Finally, if an AS follows the Gao-Rexford conditions [8], it will favor routes advertised by customers over those advertised by providers or peers, but gerontocratic policies may still be used to select one of the preferred routes. We discuss how this affects theoretical guarantees in Section VI and route stability in Section VII.

Even if an AS's route selection policy is strongly biased towards static preferences, a route's age must still be considered to guarantee convergence, placing some minor restrictions on the policy (Theorem 1). Furthermore, all ASes depend on the policies of their neighbors, neighbors of neighbors, *etc.*, and may suffer from instability and hijacks if neighbors select routes using static preferences. Therefore, even if business considerations override the need for convergence, an AS still benefits from gerontocratic policies adopted by its neighbors. Moreover, an AS may be inclined to establish business relationships with stable ASes.

"Gerontocratic policies do not prevent long-lived prefix hijacks." In Section VII-C, we show that gerontocratic policies would have been very effective at rejecting hijacked routes during previous hijack incidents. Depending on the configuration parameters, they might have eventually accepted a very long-lived hijacked route. Gerontocratic policies make long-lived advertisements *necessary* for a successful hijack, forcing hijackers to advertise incorrect routes for a long time before they have any effect. This makes malicious activity far more visible in the network, "smoking out" the hijackers.

"Gerontocratic policies do not prevent sub-prefix hijacks." Currently deployed route selection policies operate on a per-prefix basis: they consider routes advertised for a given prefix independently from routes advertised for its sub- or super-prefixes. Furthermore, in the data plane, packets are almost always forwarded along the route to the most specific prefix. Therefore, sub-prefix hijacking—when an unauthorized AS advertises a more specific prefix than those advertised by the legitimate AS (this was the case, for example, during Pakistan Telecom's hijack of YouTube in 2008)—is a serious problem.

According to the BGP RFC [27], the policy decision process runs separately for each prefix and chooses among the routes advertised for that prefix without considering routes to other, related prefixes. Therefore, without some external mechanism, no RFC-compliant policy can prevent sub-prefix hijacks such as the YouTube hijack, in which only malicious routes were advertised for the more specific prefix. Because gerontocratic policies are designed to be compliant with the BGP RFC, they cannot defend against sub-prefix hijacks alone.

If BGP RFC compliance is not important to an AS, several previously proposed techniques for preventing sub-prefix hijacks are compatible with gerontocratic policies. We discuss this further in Section VII-C. Also, any defense against sub-prefix hijacks that operates outside policy-based route selection is complementary to gerontocratic policies.

IV. Model

We use a simplified model of inter-domain routing. Consider a graph of N ASes. We assume that for each destination d, every AS has a route selection policy that prefers some routes to d over others. Unlike previous work, which assumed that route preferences are static, we allow preferences to evolve in response to observed route update dynamics.

Formally, we say that at time t, every AS v has some preference ρ_v over the routes to d. ρ_v is a function that assigns a value to each route R; $\rho_v(R,t) > \rho_v(R',t)$ iff v prefers R to route R' at time t.[1] Without loss of generality, we assume that an AS is connected to each of its neighbors by a single link (we discuss the case of multiple edge routers in Section VI).

A Byzantine AS may arbitrarily deviate from BGP. In particular, it may advertise or withdraw any route, including routes that do not exist or have not been advertised to it by its neighbors. It may also advertise different routes to different neighbors. We assume that any AS except the destination (this does not restrict the generality of our approach) may be faulty in this way, and the identity of the faulty AS is not known to other ASes. In reality, even fairly large ASes can suffer from Byzantine faults due to misconfiguration [35].

[1]We drop the subscript v when the AS is clear from the context.

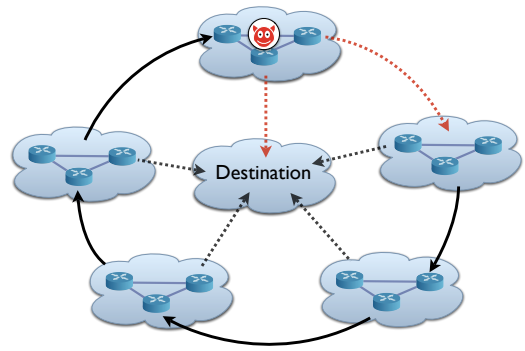

Fig. 1. Standard policy dispute wheel. Solid/dotted lines denote the preferred/alternative path. The Byzantine AS (denoted by the devil icon) has no preferences, but the path it advertises can prevent convergence.

Since any AS may be faulty, we do not consider the case when some AS v completely relies on another AS w to reach d (*e.g.*, when v is connected to the rest of the Internet via w). In this situation, no route selection policy can protect v from w's Byzantine misbehavior. Therefore, we focus on ASes that have multiple disjoint paths to d. We assume, for simplicity, that a non-Byzantine AS advertises its selected routes to all neighbors.

The main theoretical property we are interested in is *eventual convergence*: the routing protocol eventually produces a set of stable routes [12] in which no non-Byzantine AS wants to switch from its route, even if there exist Byzantine ASes that are trying to cause divergence.

We define a class of route selection policies which have this property if adopted by all non-Byzantine ASes. While convergence cannot be guaranteed by any unilaterally adopted policy, we show that gerontocratic policies yield a substantial improvement in the stability and security of chosen routes even when adopted by a single AS (see Section VII).

Our goal is to avoid oscillation caused by policy conflicts and/or faulty route advertisements inserted by a Byzantine AS. Thus, our theoretical results in Sections V and VI assume that no link failures occur during the convergence period, *i.e.*, the only sources of instability are policy conflicts and/or the behavior of Byzantine AS(es). Unavoidably, "genuine" link failures may necessitate route changes even in the absence of a Byzantine AS. In the presence of link failures not caused by a Byzantine AS, the best one can hope for is that the non-faulty ASes converge to a set of routes that remains stable until a "genuine" link failure occurs.

V. Static Policies

A route selection policy is *static* if the preference order it imposes on routes does not depend on the observed dynamics of the routing protocol. We give a simple argument that no combination of non-trivial static policies can converge in the presence of a single Byzantine AS. The intuition is simple: if the next hop in a preferred indirect route continuously

advertises and withdraws this route, the AS that prefers this route will never converge.

Below, we use parentheses to indicate concatenation, *e.g.*, (v, R) is AS v prepended to route R, and (R, R') is the concatenation of routes R and R'.

Theorem 1: In the presence of a single Byzantine AS, the only network topology and static route selection policy that converge to a stable set of routes are those in which every AS has a direct link to the destination and prefers this route over any indirect route.

Proof: Let d be the destination, and pick some AS $w \neq d$ such that there exists some AS v which prefers an indirect route R_w to d that passes through w over all other routes. Let R_v be the next highest-ranked route for v that does not contain w. Fix w to be the Byzantine AS. w can prevent convergence as follows. Suppose that v initially selected R_v and is advertising (v, R_v). w then advertises R_w to v, which v adopts. When v advertises (v, R_w), w either withdraws its route or advertises (w, v, R_v) to v, thus creating a routing loop for v. v then abandons R_w, switches back to R_v, and begins advertising (v, R_v) again. This simulates a dispute wheel, which is known to cause continuous oscillation [12].

Therefore, convergence cannot be guaranteed whenever some AS has and prefers an indirect route. A route selection policy which prefers the direct route, along with a network topology in which every AS has a direct link to d, is trivially stable, completing the proof. ■

Theorem 1 shows that a single Byzantine AS can easily cause policy dispute wheels when ASes use static policies. This includes policies which prefer a particular neighbor, existence or non-existence of a particular AS or link in the path, route's length, *etc.* Figure 1 shows an example.

As consistent with previous work, we consider only a single destination. In reality, every AS is a possible destination. If any AS could potentially be Byzantine, it follows from Theorem 1 that every AS must have a direct link to every other AS, *i.e.*, the only real-world topology in which convergence with static policies can be guaranteed is a clique.

Divergence may result from Byzantine ASes advertising non-existent, short routes (*e.g.*, [3]). It is important to note, however, that cryptographic approaches do not completely solve the problem: even if route updates are cryptographically authenticated, a Byzantine AS can cause oscillation solely by switching between valid, authorized, authenticated routes or withdrawing a previously advertised route.

The actual impact of a Byzantine AS w on the network depends on the route selection policies of those close to w. For example, w has more impact if many ASes prefer routes that use w or if w is close to many ASes that prefer shortest routes. Large ASes failed in this way before [35], and even the failure of smaller ASes often causes significant impact [3, 32].

VI. Gerontocratic Policies

We showed in Section V that no static route selection policy guarantees convergence when Byzantine ASes are present. Therefore, any Byzantine-tolerant policy must be *dynamic*: it must take into account the observed history of route advertisements when selecting routes.

We now prove that dynamic policies can be sufficient to ensure convergence in the presence of Byzantine ASes by describing a class of simple dynamic policies that do, in fact, converge. These policies follow two design principles, which we posit are useful for tolerating Byzantine ASes:

1) Withdrawing a route should incur some penalty in the preference score assigned to it by other ASes. Otherwise, a Byzantine AS could repeatedly withdraw and immediately re-advertise a desirable route, causing other ASes to withdraw and re-advertise their own routes.
2) A route that has been available for a sufficiently long time should have a higher preference than a shorter-lived route. Without this requirement, a Byzantine AS could introduce never-seen-before, possibly even fake routes and cause previously stable, long-lived routes to be withdrawn, preventing convergence. If ASes are prevented from advertising unauthorized routes (*e.g.*, via S-BGP [16]), this requirement may not be necessary. Update validation, however, only stops Byzantine ASes from advertising invalid routes, not from introducing policy conflicts or instabilities, and is thus insufficient for convergence.

Not every dynamic policy guarantees convergence in the presence of Byzantine ASes. For example, route flap damping [33] (which does not satisfy the second condition) and SRS [9] (which satisfies neither condition and does not use route dynamics as part of its primary selection criteria) are examples of dynamic policies that do not ensure convergence.

Formally, let $\ell(R, t)$ be route R's *age* at time t, *i.e.*, how long it has been continuously propagated to the given AS (advertised by its neighbor and neither explicitly nor implicitly withdrawn since it was advertised). Let $\rho_0(R)$ be some bounded function representing this AS's static preferences, and let α be some constant weight such that $0 < \alpha \leq 1$. Gerontocratic policies select routes on the basis of the following route preference metric:

$$\rho(R, t) = \alpha\ell(R, t) + (1 - \alpha)\rho_0(R) \tag{1}$$

This metric satisfies our design principles: long-lived routes eventually achieve an ℓ value large enough to overcome any difference in ρ_0, while withdrawing a route resets ℓ.

This metric is implementable in current routers, which already maintain the AS's static preferences as well as alternative routes to any given destination [27]. A router needs to simply store an extra timestamp for each of these alternative routes, indicating when it was advertised. In ASes with multiple edge routers, the timestamp must be exchanged along with the advertisements in order for each router to come to the same policy decision.

We prove that an AS using this metric eventually converges to a stable route even in the presence of Byzantine ASes, as long as there exists some downstream neighbor which eventually converges to a stable route or is the destination itself. Consequently, if every non-Byzantine AS uses a gerontocratic policy, network-wide convergence is guaranteed.

We start by proving that a network in which all ASes follow a policy based on condition (1) eventually converges.

Theorem 2: Given route preferences based on condition (1) and a network consisting entirely of non-Byzantine ASes using these preferences, every AS in the network eventually converges to some stable set of routes.

Proof: By induction on the minimum number of hops from the destination d, *i.e.*, the number of hops in the shortest possible route to d. Let $\mathcal{N}(x)$ be the set of all neighbors of some AS x. We first consider some AS $v \in \mathcal{N}(d)$. Since v is d's neighbor and has a direct physical link to d, there exists a route R_1 of the form (d). Consider some indirect route R_2 that v prefers over R_1; if none exist, the base case is trivially complete as R_1 is considered stable. In order for R_2 to remain preferred over R_1 at any time t, $\rho(R_2, t) > \rho(R_1, t)$ and thus

$$\alpha\ell(R_2, t) + (1-\alpha)\rho_0(R_2) > \alpha\ell(R_1, t) + (1-\alpha)\rho_0(R_1)$$

Moving the terms around, we have

$$\ell(R_2, t) > \ell(R_1, t) + K \tag{2}$$

for $K = (1-\alpha)/\alpha\,(\rho_0(R_1) - \rho_0(R_2))$; note that K is constant with respect to time. Since we assume no link failures in the non-Byzantine part of the network (see Section IV) and R_1 is simply the direct link to d, $\ell(R_1, t)$ is increasing during every unit of time, *i.e.*, $d\ell(R_1, t)/dt = 1 \geq d\ell(R_2, t)/dt$. There must exist some time t_+ where the right-hand side of condition (2) is greater than 0; at this point, condition (2) holds for all $t \geq t_+$ only if R_2 is not withdrawn again, since $\ell(R_1, t) + K > 0$ and withdrawing R_2 would reset $\ell(R_2, t)$ to 0. Thus, if R_2 is withdrawn at some point $t \geq t_+$, v converges to R_1; otherwise, v converges to R_2.

Now suppose that every AS within (minimum) j hops eventually converges to a particular route. We now prove that every AS within (minimum) $j+1$ hops must converge as well. Consider some ASes v and w such that $v \in \mathcal{N}(w)$, v is (minimum) $j+1$ hops away from d, and w is (minimum) j hops away from d. By the induction hypothesis, w must eventually converge to some route R_3 to d. From this point on, v has a stable route $R_4 = (w, R_3)$. By the same argument as before, there is some point at which, for any route R_5 that v prefers over R_4, a withdrawal would cause R_4 to be preferred. If R_5 is withdrawn after then, v converges to R_4; otherwise, v converges to R_5. ■

We now show that these policies tolerate a Byzantine AS.

Theorem 3: Given route preferences based on condition (1), every non-Byzantine AS eventually converges to a stable set of routes in the presence of one Byzantine AS.

Proof: (Sketch) Fix the AS that is Byzantine. This proof is similar to that of Theorem 2, except we perform the induction on the number of hops in the shortest possible route to the destination d that does not go through the Byzantine AS (by assumption, one such route must always exist). As before, those ASes that are directly connected to d converge eventually. Consider then any AS v whose shortest, non-Byzantine route to d is $j+1$ hops and goes through some AS w. By the inductive hypothesis, w, which is (minimum) j hops away from d, must eventually converge to some stable route. Thus, it follows, as in the proof of Theorem 2, that v must converge either to a route that goes through w, or to some other stable route. ■

Theorem 3 can be generalized to show that policies that use condition (1) as the metric converge to a stable set of routes even in the presence of multiple Byzantine ASes.

Corollary 4: Given route preferences based on condition (1), every non-Byzantine AS eventually converges to a stable set of routes in the presence of multiple Byzantine ASes as long as every non-Byzantine AS has a route to the destination that does not contain a Byzantine AS.

Proof: (Sketch) Fix the Byzantine ASes. Since every non-Byzantine AS has, by assumption, a route that does not contain a Byzantine AS, the same reasoning as in the proof of Theorem 3 proves that all ASes must eventually converge. ■

If an AS follows the Gao-Rexford conditions, *i.e.*, prefers routes advertised by customers over those advertised by peers or providers, condition (1) must incorporate an additional term which ensures that $\rho(R, t) > \rho(R', t)$ if the AS has a preferred transit agreement with the neighbor advertising R rather than the one advertising R'. In these cases, gerontocratic policies still guarantee convergence in the presence of Byzantine ASes if every AS has a route through a non-Byzantine neighbor with the most preferred peering/transit agreement.

VII. Evaluation

Although gerontocratic route selection policies ensure eventual convergence, they provide no theoretical guarantees about (a) the quality of the resulting routes; (b) whether the routes correspond to actual physical paths or are consistent with what other ASes have advertised; or (c) how long convergence takes. In this section, we empirically show that gerontocratic policies, even if adopted unilaterally by a single AS without cooperation from any other ASes, (1) select routes that are significantly more stable than those selected by other policies (Section VII-B); (2) are far less likely to be hijacked, even without route update authentication or cooperation from other ASes (Section VII-C); and (3) are as responsive to network failures as other policies (Section VII-D).

A. Implementation and setup

To model various route selection policies, we implemented a multithreaded discrete event simulator, which is available for download [38]. Our simulator models an AS receiving route updates from its neighbors and selecting routes to various prefixes. We use route updates and table dumps from the Route Views project [1], specifically from route-views2.oregon-ix.net, a node that has approximately 52 neighboring ASes including AT&T, Level3, and Sprint.[2]

We implemented and tested the following policies:

[2]The results are also valid for ASes with fewer neighbors. In particular, we found that, for the destinations and time period we tested, an average of 33 neighboring ASes advertised routes in each iteration.

- *Gerontocratic*: Prefer oldest routes first; break ties with route length, then (random) local preferences.
- *Local*: Randomly generate local preferences over neighboring ASes at the beginning of a simulation. These policies are commonly known as next-hop policies.
- *Mixed*: Evaluate a route by its age and an AS's local preferences, *i.e.*, as in condition (1), where $\ell(R,t)$ is the route's age in seconds, $\rho_0(R)$ is a random static score between 0 and 1, and $\alpha \in \{10^{-9}, 10^{-8}, \ldots, 10^{-2}, 0.1, 0.5, 0.9\}$.
  To get an intuition for what these values of α mean, recall that for two routes R and R' whose static scores differ by Δ (*i.e.*, $\rho_0(R) - \rho_0(R') = \Delta \geq 0$), R' is preferred over R if and only if it is $\Delta(1-\alpha)/\alpha$ seconds older (*i.e.*, $\ell(R',t) - \ell(R,t) \geq \Delta(1-\alpha)/\alpha$). For $\alpha = 10^{-7}$ and $\rho_0(R) - \rho_0(R') = 0.5$, $\ell(R',t) - \ell(R,t) \geq 57.87$ days before R' is preferred.
- *PGBGP Lite*: Same as *Local*, except depreference, for s hours, all routes to a given prefix that do not contain an AS which originated an advertisement for this prefix in the last h days; we use $h = 10$ and $s = 24$ as in [15]. Since we only simulate route selection to a destination prefix, not the path chosen for a particular IP address, we did not implement PGBGP's sub-/super-prefix hijack checks (this simplification is discussed in Section VII-C).
- *Shortest*: Prefer shortest routes first; break ties with (random) local preferences.
- *Shortest (Age)*: Same as *Shortest*, except use route's age as the tie-breaking metric before local preferences. Hybrid of *Gerontocratic* and *Shortest*.
- *Damped Shortest*: Same as *Shortest*, except use route flap damping with Cisco's default parameters [21].
- *Local Optimal*: Prefer neighbors that, on average, advertise the longest-living routes to a particular prefix. This policy is unimplementable because it assumes that an AS can determine how long the routes advertised by a given neighbor will survive after having been chosen. It defines the upper bound on how well any policy can maximize route lifetimes if local preferences over neighboring ASes are fixed in advance.
- *Optimal*: Prefer the route that will last the longest. This unimplementable policy defines the upper bound on how well any policy, static or dynamic, can do in maximizing route lifetimes.

We also simulate the Gao-Rexford versions of the policies, which prefer routes from customers over those from peers, which in turn are preferred over routes from providers [11]. Whenever possible, we assign the relationships between our simulated AS and its neighbors so that they are consistent with CAIDA AS relationship data [5] and the Gao-Rexford conditions, *i.e.*, if our AS v receives a route from a neighboring AS w and w is not a provider to its downstream neighbor AS x, then v must be w's customer since w only advertises routes through x if v or x is a customer of w [11]. Otherwise, we randomly assign the relationship. For simplicity, we ignore sibling relationships.

We simulate the policies over one year for various years.[3] For a given year y, we start the simulation by building the initial routing table. We load the last table dump from year $y-1$ and incorporate route updates whose timestamps lie between the table's timestamp and the beginning of y. Because timestamps in the table dump do not always correspond to the time when a route was first advertised, we set all timestamps to the first timestamp of year y. We then begin the simulation by selecting a route from the initial table in accordance with various policies. We process route updates and select new routes until we reach updates from year $y+1$. We update the routing table entry if and only if a route advertised by a particular AS has changed, with the exception of *Damped Shortest* and *PGBGP Lite*, where we update route penalties as needed. We repeat our simulations 250 times with different local preferences or AS relationship assignments.

Our simulation relies on the Route Views dataset, which reflects actual updates as observed in the Internet. One limitation of this approach is that we cannot model the effects of our simulated AS advertising routes to its neighbors.

B. Route lifetimes and lengths

To measure the quality of the routes selected by various policies, we simulate route selection for several destinations in the US (Google [64.233.160.0/23], Microsoft [207.46.192.0/18], UTCS [128.83.0.0/16]); Germany (GMX [213.165.64.0/19], MPI [195.71.0.0/16]); Brazil (UOL [200.98.192.0/18]); and China/HK (HKU [147.8.0.0/16], QQ [60.28.0.0/15]) for an entire year for years 2005-2009. For every year, destination, and policy, we calculate:

- Average route lifetime. We first take the median of each individual iteration, calculate the median of the resulting 250 medians, and then report a mean of the resulting 5 medians (one per year).[4]
- Average route length. We take a weighted mean, *i.e.*, we take the number of hops in each of the routes selected by the policy, multiply it by the time during which this was the chosen route, and divide this time by the total time our simulated AS had a route to this destination.

Figure 2 shows the average route lifetimes and lengths for various destinations and policies.

Comparison with shortest-path policies. For the destinations and time periods tested, *Gerontocratic* significantly outperforms *Shortest* and *Damped Shortest*, selecting routes that last, on average, over 5.88 times as long. These policies greedily switch to the "latest and greatest" shortest route, regardless of how long it has been advertised. Although they always select the shortest routes, they frequently have to change their selection when their choices are withdrawn.

[3]The one exception is 2007, which we simulate up to Dec. 16 due to significant corruption in the Route Views data.

[4]We use the median to aggregate individual route lifetimes over multiple simulated iterations because we believe that a route selection policy that selects a series of very stable routes is preferable to one that selects many short-lived, unstable routes before stumbling upon a very long-lived route. We otherwise use the mean to aggregate results over time, multiple destinations, or multiple policies.

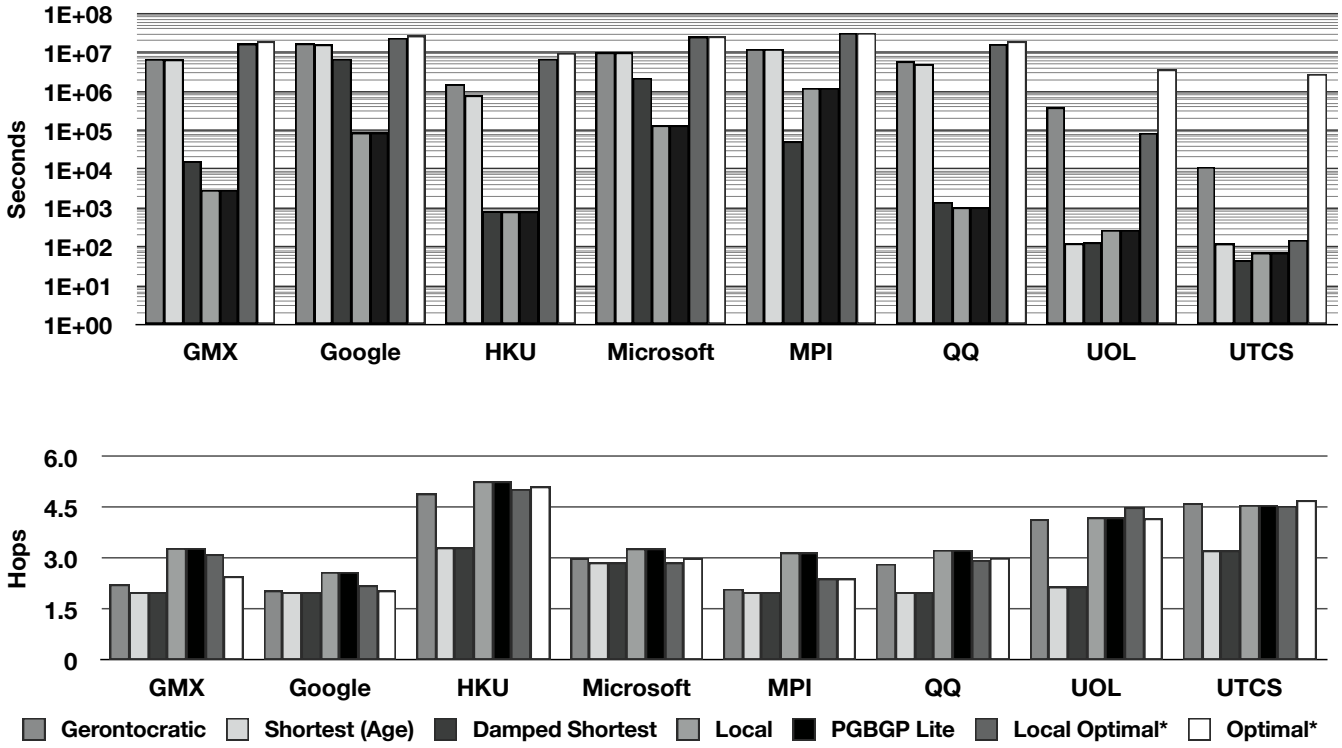

Fig. 2. Average route lifetimes and lengths. We omit *Shortest* because *Damped Shortest* selected strictly longer-living routes of similar lengths in our simulations. Asterisks (*) denote unimplementable policies.

For Google and Microsoft, *Gerontocratic* benefits from the existence of short, stable routes. Its chosen routes, respectively, last 2.48 and 4.50 times as long as *Shortest* and *Damped Shortest*, while being only 0.05 hops and 0.128 hops longer on average. *Gerontocratic* improves route stability even more when selecting routes to international destinations and UTCS. *Gerontocratic* selected routes to GMX, MPI, and UTCS that lasted more than two orders of magnitude as long; for HKU, QQ, and UOL, *Gerontocratic* selected routes that lasted more than three orders of magnitude as long. For these destinations, policies that base route selection on route length often select very unstable routes. Excluding Google and Microsoft, the destinations for which *Shortest* and *Damped Shortest* performed the best were GMX (3.68 hours on average) and QQ (20.1 minutes on average). Although *Gerontocratic* selected routes that were 1.02 hops longer on average, the shortest-length routes in our simulations were often the shortest-living ones, given that *Local Optimal* and *Optimal* both selected routes of similar length to *Gerontocratic*.

Although *Damped Shortest* considers dynamic information to some extent to avoid short-term route instability, it is vulnerable to longer-term instability, causing it to select shorter-lived routes on average. Our simulations indicate that route flap damping provides only a marginal benefit over *Shortest*: an extra 9.4 minutes of route lifetime on average.

Applying gerontocratic metrics on top of shortest-path selection greatly improves route lifetimes. By considering the age of a route as a tie-breaker between two equally long routes, *Shortest (Age)* was able to avoid long-term route flapping that plagued *Shortest* and *Damped Shortest*, while always selecting the shortest routes. Overall, *Shortest (Age)* performs similarly to *Gerontocratic* with respect to route lifetimes when selecting routes to GMX, Google, Microsoft, MPI, and, to a lesser degree, QQ. By selecting routes with fewer hops, which have fewer links and thus fewer points of failure, and using gerontocratic policies to discern more stable routes, *Shortest (Age)* performs better than other shortest-path policies.

When selecting routes to HKU, UOL, and UTCS, however, *Shortest (Age)* performed significantly worse than *Gerontocratic*. These destinations have both longer and more diverse routes available; as described above, the shorter routes turn out to be less stable than the longer ones. As a result, the gerontocratic component of *Shortest (Age)* is used less frequently, and the pitfalls of optimizing for the shortest path are once again prominent. For HKU and UTCS, *Gerontocratic* selected routes that last 1.94 and 94.8 times as long compared to *Shortest (Age)*; for UOL, *Gerontocratic* selected routes that last over three orders of magnitude as long.

Comparison with network engineering and oracle policies. Many ASes perform network engineering in order to prioritize certain neighbors over others. We used *Local* to simulate next-hop policies with different local preferences over neighboring ASes. Averaging over 250 different random local preferences, we observe that through careful choices of which neighbor to prefer, local preferences can sometimes achieve long route lifetimes. Unfortunately, many ASes advertise short-lived

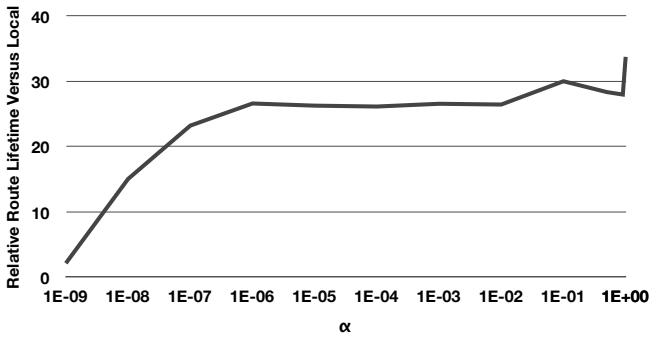

Fig. 3. *Mixed*'s route lifetime relative to *Local*, with varying α values.

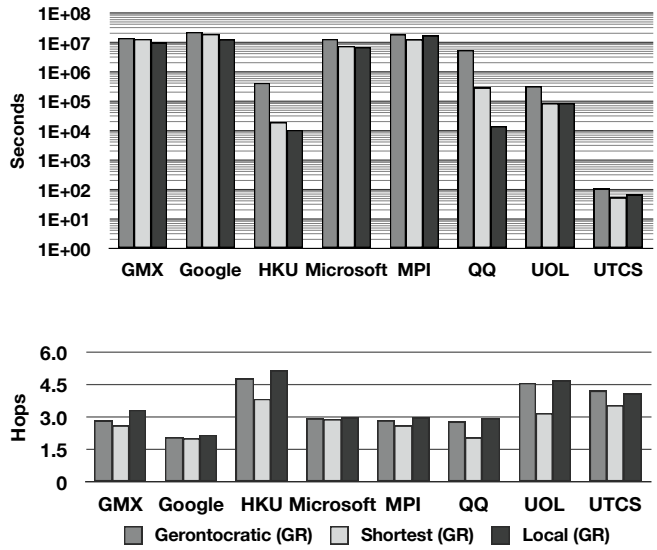

Fig. 4. Average route lifetimes and lengths for the Gao-Rexford-compatible policies that we tested.

routes as well; on average, we found that *Local* performed the worst of all policies we tested. Compared to *Local*, *Gerontocratic* selected routes that lasted 37.3 times as long while being 0.47 hops shorter. *Gerontocratic* selected routes to MPI and Microsoft that lasted, respectively, 10.0 and 75.1 times as long; for all other destinations, *Gerontocratic* selected routes that lasted 2-3 orders of magnitude as long. The only destination to which *Gerontocratic* selected longer routes was UTCS (0.05 hops longer); for all other destinations, *Gerontocratic* selected routes that were 0.54 hops shorter, on average.

Compared to *Local*, *PGBGP Lite* provides almost no benefit in terms of route lifetimes and length. The reason is that most prefix hijacks are short-lived and relatively rare, and in the absence of hijacks, *PGBGP Lite* and *Local* use the same selection criteria. In some cases, using *PGBGP Lite* can even be detrimental. For example, the Google prefix we tested did not have routes for approximately the first third of 2005. When routes began to appear on April 5, *PGBGP Lite* depreferenced all of them and selected routes solely on the basis of the tie-breaking metric. On April 6, *PGBGP Lite* began accepting routes it suspected the previous day. This effectively caused *PGBGP Lite* to repeat the same route selection process as the day before, except this time with accepted routes.

Local Optimal does beat *Gerontocratic* during most simulations, which is not surprising since *Local Optimal* assumes perfect knowledge of future lifetimes of all advertised routes. It may seem odd that *Gerontocratic* occasionally outperforms *Local Optimal* (*e.g.*, for UOL and UTCS). The reason is that *Local Optimal* is based on local preferences over neighbors, whereas long-lived routes may be advertised by different ASes at different times. In these scenarios, *Local Optimal*, with its fixed preferences, does not have the flexibility to adapt to the AS which is currently more stable, whereas *Gerontocratic*, which only considers how long a route has been around and not which neighbor advertised it, does.

As expected, *Optimal* outperforms all other policies. *Gerontocratic* is within a factor of 1/3-1/2 of *Local Optimal* and *Optimal*, with routes that are around 0.1-0.2 hops shorter.

Mixing local preferences with route age. Figure 3 shows the results of *Mixed* with varying α values, averaged over all destinations and years. Unsurprisingly, route lifetimes achieved by *Mixed* fell between those achieved by *Local* and *Gerontocratic*. More surprisingly, adding even a small gerontocratic component to local preferences provides a significant improvement over *Local* while largely retaining the ability to prefer certain neighbors over others. With $\alpha = 10^{-8}$, *Mixed* achieved average route lifetimes that were more than 15.1 times that of *Local*; with $\alpha = 10^{-6}$, this jumped to over 25.6 times.

Gao-Rexford-compatible route selection. Figure 4 shows the results of using *Shortest*, *Gerontocratic*, and *Local* to select a route after the relationships with neighboring ASes are considered. In our simulations, we found that, on average, approximately 93% of neighbors advertising routes were constrained to be providers based on the Gao-Rexford conditions. As a result, most of these simulations depend on whether the remaining unconstrained ASes were randomly set to be providers, too. In simulations where they are not, all policies are expected to perform similarly to each other because any policy will immediately favor routes that are advertised by these ASes and will only fall back to using route age, length, *etc.* as a tie-breaker.

As a result, the lifetimes and lengths of the routes selected by these policies are more similar than in the previous simulation. Moreover, for the destinations and time period we studied, the unconstrained ASes tended to advertise more stable routes, thus helping *Shortest* and *Local* close the gap even further. Despite these factors, *Gerontocratic* still provides a significant benefit over *Shortest* and *Local*, selecting routes that last on average 1.42 and 1.60 times as long while being 0.54 hops longer and 0.16 hops shorter, respectively.

C. Avoiding prefix hijacks

In this section, we evaluate how well *Gerontocratic* avoids prefix hijacking attacks when unilaterally adopted by a single AS and compare it with other implementable policies. We

Hijacker - Date of hijack	Destination	# routes	Median duration	Implementable policies affected (excl. *Local*)
AS 18747 - Nov. 17, 2004	Google (/24)	1	24 sec.	*Local* (non-GR and GR)
AS 9121 - Dec. 24, 2004	Google (/24)	1	49.3 min.	*Shortest* (non-GR and GR), *Damped Shortest*, *Local* (non-GR and GR)
	Microsoft	3	34.1 min.	***Shortest (Age)***, *Shortest* (non-GR), *Damped Shortest*, *Local* (non-GR), ***Mixed* (non-GR, $\alpha \leq 10^{-9}$)**
	MPI	2	6.53 min.	*Shortest* (non-GR and GR), *Damped Shortest*, *Local* (non-GR and GR), ***Mixed* (non-GR and GR, $\alpha \leq 10^{-9}$)**
	UTCS	6	24.1 min.	***Shortest (Age)***, *Shortest* (non-GR and GR), *Damped Shortest*, *Local* (non-GR and GR), ***Mixed* (non-GR and GR, $\alpha \leq 10^{-8}$)**
AS 174 - May 7-9, 2005	Google (/24)	15	1.84 days	*Shortest* (GR and non-GR), *Damped Shortest*, *Local* (GR and non-GR), *PGBGP Lite*
AS 9304 - Aug. 11, 2005	Microsoft	1	8.67 min.	*Shortest* (non-GR), *Damped Shortest*, *Local* (non-GR), ***Mixed* (non-GR, $\alpha \leq 10^{-9}$)**
AS 23520 - Apr. 9, 2006	Microsoft	3	29 sec.	*Shortest* (non-GR), *Damped Shortest*, *Local* (non-GR), ***Mixed* (non-GR, $\alpha \leq 10^{-8}$)**
AS 4761 - Nov. 30, 2006	Microsoft	1	1.67 min.	*Shortest* (non-GR), *Damped Shortest*, *Local* (non-GR), ***Mixed* (non-GR, $\alpha \leq 10^{-9}$)**
AS 17557 - Feb. 24, 2008	YouTube	36	1.67 hrs.	*All*
AS 8997 - Sep. 22, 2008	HKU	1	6.55 min.	*Shortest* (non-GR and GR), *Damped Shortest*, *Local* (non-GR and GR), ***Mixed* (non-GR, $\alpha \leq 10^{-9}$; GR, $\alpha \leq 10^{-8}$)**
	QQ	1	6.68 min.	*Local* (non-GR), ***Mixed* (non-GR, $\alpha \leq 10^{-9}$)**
	UTCS	1	6.48 min.	*Shortest* (non-GR and GR), *Damped Shortest*, *Local* (non-GR and GR), ***Mixed* (non-GR and GR, $\alpha \leq 10^{-8}$)**

TABLE II
HIJACKS (KNOWN AND SUSPECTED) OBSERVED IN SIMULATIONS. POLICIES THAT USE A GERONTOCRATIC COMPONENT ARE HIGHLIGHTED IN BOLD.

tested all destinations and policies from Section VII-B with the exception of *Local Optimal* and *Optimal*. In addition, we extended the simulation by testing year 2004 as well as two additional destinations—Google (64.233.161.0/24) and YouTube (208.65.153.0/24)—in order to test our policies against two well-publicized hijacks. Although we initially extended our simulation to include the prefix hijack attack by a Chinese ISP in April 2010 [32], we did not observe any hijacked routes during that period of time for the destinations we tested. Thus, our results only cover years 2004-2009.

Table II shows the prefix hijacks we observed in the data and summarizes how various policies perform. Earlier, it was empirically observed that invalid or inconsistent advertisements, such as those resulting from prefix hijacks, are often transient and short-lived [2, 15, 26]. Therefore, policies which consider route age when selecting routes avoid most prefix hijacks. *Gerontocratic* avoided almost every hijack; even in the case of Cogent (AS 174)'s hijack of Google's prefix, which lasted for over a day, *Gerontocratic* avoided the hijacked routes since other longer-lived routes were available. By breaking ties using route age, *Shortest (Age)* was able to avoid most of the hijacks; however, it was still vulnerable to TTNet (AS 9121)'s hijack of Microsoft's and UTCS's prefixes.

Because false route updates tend to advertise routes with fewer hops, policies that prefer shorter routes such as *Shortest* and *Damped Shortest* are vulnerable to almost every hijack we observed, given the appropriate tie-breaking local preferences. For similar reasons, *Local* is potentially vulnerable to every hijack. On the other hand, *Mixed*, which uses a convex combination of local preferences (*e.g.*, those representing established business relationships with other ASes) and route age, was able to avoid almost every hijack. Even with an α as small as 10^{-7}, *Mixed*, when combined with any one of the 250 different random preferences we tested, avoided prefix hijacks just as well as *Gerontocratic*. This demonstrates that **even ASes whose route preferences are, to a large extent, statically determined will benefit from incorporating route age into their policies**.

Although *PGBGP Lite* is able to avoid most hijacks by depreferencing short-lived updates from unfamiliar origin ASes, *PGBGP Lite* falls victim to Cogent's hijack of Google, which lasted approximately 1.84 days, much longer than the 1-day window during which *PGBGP Lite* considers a route suspicious. Just like route-flap damping, *PGBGP Lite* is all-or nothing: either a route is suspected and heavily depreferenced, or it is considered without qualms. Although gerontocratic metrics do not explicitly suspect unknown ASes, they have a similar effect: *Gerontocratic* depreferences and effectively ignores Cogent's false advertisements. *PGBGP Lite* could have avoided this attack with a larger window size, but such an increase would increase the time it takes *PGBGP Lite* to accept legitimate route updates with unknown origin ASes.

None of the policies we tested were able to avoid the YouTube hijack. No unilaterally adopted policy can. On Feb. 24, 2008, Pakistan Telecom (AS 17557) took over a subset of YouTube's prefix (208.65.152.0/22) by advertising 208.65.153.0/24, which was more specific than any other advertised prefix. Even if a policy selected a non-hijacked route to 208.65.152.0/22, data packets would have been ultimately forwarded along the route to the most specific prefix that matched the destination IP address. In this case, all neighboring ASes that advertised a legitimate route to 208.65.152.0/22 also advertised the hijacked route to 208.65.153.0/24. Any traffic destined for 208.65.153.0/24 would have to pass through a neighboring AS, which would forward the data over the hijacked route. Thus, even the full version of PGBGP—which depreferences suspicious routes to sub-prefixes of known prefixes and avoids neighbors that would use these routes—would have been insufficient to evade the YouTube hijack if adopted unilaterally.

Gerontocratic policies as described are "drop-in" compatible with existing BGP routers. As required by the BGP RFC [27], gerontocratic policies assign a score to each route without considering the other routes for a particular prefix. Our policies also do not require modifications to the BGP decision process, which, by specification, runs separately for each prefix and chooses among the routes advertised for that prefix without considering routes to other, related prefixes. As mentioned in Section III, no RFC-compliant policy can prevent sub-prefix hijacks without help from some external mechanism.

If BGP RFC compliance is not important to an AS, then gerontocratic route selection can be generalized to defend against sub-prefix hijacks by considering routes to multiple (sub-)prefixes during the decision process or by prioritizing updates for a particular prefix (and associated sub-/super-prefixes) depending on whether the origin AS is suspicious and/or disseminates long-lived advertisements for this prefix.

D. Reacting to network failures

Because gerontocratic policies bias route selection towards older routes and react cautiously to new route advertisements, a superficially plausible objection is that gerontocratic route selection policies react slowly to network failures. In this section, we demonstrate that this concern is misplaced.

Gerontocratic policies only select routes that are still being advertised (*i.e.*, not implicitly or explicitly withdrawn). If the selected route fails and the route is simply a direct link to the destination, then any policy will promptly deal with this failure. With indirect routes, BGP gives no information regarding the reason a route has been withdrawn. Therefore, promptly switching from a route that has *not* been withdrawn (either implicitly or explicitly) on the basis of another route's status provides questionable benefit. Moreover, even if a link has failed, it is possible that the new route still contains the failed link; in these cases, a policy that promptly switches will have to switch again and re-advertise its route selections more frequently, contributing to route instability.

To measure precisely whether gerontocratic policies react more slowly to changes in link state, we would need a control-plane oracle to tell us the state of each link at any given moment and thus determine exactly which routes contain failed links. Such an oracle does not exist, but we approximate it in our simulations by examining how often individual links are advertised and (implicitly or explicitly) withdrawn in route updates for the destinations we tested. Links that initially appear in at least $n = 8$ routes and, within time $\tau = 1$ year, are withdrawn from all but at most $\beta = 2$ of them without appearing in any subsequent route advertisement are assumed to have failed.[5] A link is assumed to have been repaired when it is subsequently advertised in any route update.

For every year and destination, we use the above technique to determine the set of failed links and use this set to calculate, for each policy, the proportion of time that an AS using this policy has a route without failed links. Table III shows the results, averaged over all years, destinations, and 250 iterations. We only tested the non-Gao-Rexford versions of the policies, since the Gao-Rexford conditions limit the diversity of the routes selected.

[5]We chose these values to avoid attributing every route withdrawal to link failure while providing enough failed link events (roughly 32.3 events on average per destination/year dataset).

Policy	Proportion
Gerontocratic	0.99883
Shortest (Age)	0.99853
Damped Shortest	0.99896
Shortest	0.99858
Local	0.99536
PGBGP Lite	0.99536

TABLE III
PROPORTION OF TIME THAT EACH POLICY HAS A VALID ROUTE TO A DESTINATION.

Overall, *Gerontocratic* performs on par with other policies. This is as expected: new route advertisements per se do not imply link failure, and newly advertised routes may contain failed links, too. Although *Gerontocratic* may end up selecting older routes when newer routes exist, this does not result in worse overall route availability in comparison to other implementable policies.

VIII. CONCLUSION

Although the Internet has largely been designed under the assumption that all participating autonomous systems (ASes) are trusted, Byzantine misbehavior among ASes is increasingly common. We prove that dynamic properties of inter-domain route advertisements must be explicitly incorporated into the route selection policies of individual ASes in order to tolerate this misbehavior. We describe a class of simple, dynamic, gerontocratic policies that take into account the age of advertised routes as well as the AS's static preferences. Such policies are sufficient to guarantee convergence to a stable set of routes in the presence of Byzantine ASes. Gerontocratic policies do not sacrifice the flexibility of individual ASes in selecting routes, which is a hallmark of BGP. We model the behavior of gerontocratic policies on actual inter-domain route updates and demonstrate that they (1) select very stable routes; (2) avoid almost all prefix hijacks without cryptographic infrastructure, cooperation from other ASes, or knowledge of the Internet's topology; and (3) promptly adapt to legitimate route changes.

Acknowledgments. We are grateful to Lorenzo Alvisi for his generous advice and guidance on many of the ideas described in this paper, and to Jeremy Powell for early discussions and help with the initial implementation. The research described in this paper was partially supported by the NSF grants CNS-0720649, CNS-0746888, and CNS-0905602, Google research award, and the MURI program under AFOSR Grant No. FA9550-08-1-0352.

REFERENCES

[1] University of Oregon Route Views project. http://www.routeviews.org.
[2] H. Ballani, P. Francis, and X. Zhang. A study of prefix hijacking and interception in the Internet. In *SIGCOMM*, 2007.

[3] V. Bono. 7007 explanation and apology. http://www.merit.edu/mail.archives/nanog/1997-04/msg00444.html, 1997.

[4] K. Butler, P. McDaniel, and W. Aiello. Optimizing BGP security by exploiting path stability. In *CCS*, 2006.

[5] CAIDA. AS relationships dataset. http://www.caida.org/data/active/as-relationships. January 20, 2010.

[6] N. Feamster, R. Johari, and H. Balakrishnan. Implications of autonomy for the expressiveness of policy routing. In *SIGCOMM*, 2005.

[7] L. Gao, T. Griffin, and J. Rexford. Inherently safe backup routing with BGP. In *INFOCOM*, 2001.

[8] L. Gao and J. Rexford. Stable internet routing without global coordination. *IEEE/ACM Trans. Netw.*, 9(6):681–692, 2001.

[9] P. B. Godfrey, M. Caesar, I. Haken, Y. Singer, S. Shenker, and I. Stoica. Stable Internet route selection. NANOG 40, 2007.

[10] S. Goldberg, S. Halevi, A. Jaggard, V. Ramachandran, and R. Wright. Rationality and traffic attraction: Incentives for honest path announcements in BGP. In *SIGCOMM*, 2008.

[11] S. Goldberg, M. Schapira, P. Hummon, and J. Rexford. How secure are secure interdomain routing protocols? *SIGCOMM Comput. Commun. Rev.*, 40(4):87–98, 2010.

[12] T. G. Griffin, F. B. Shepherd, and G. Wilfong. The stable paths problem and interdomain routing. *IEEE/ACM Trans. Netw.*, 10(2):232–243, 2002.

[13] X. Hu and Z. Mao. Accurate real-time identification of IP prefix hijacking. In *S&P*, 2007.

[14] Y.-C. Hu, A. Perrig, and M. Sirbu. SPV: Secure path vector routing for securing BGP. In *SIGCOMM*, 2004.

[15] J. Karlin, S. Forrest, and J. Rexford. Pretty Good BGP: Improving BGP by cautiously adopting routes. In *ICNP*, 2006.

[16] S. Kent, C. Lynn, and K. Seo. Secure Border Gateway protocol (Secure-BGP). *IEEE Journal on Selected Areas in Communications*, 18(4), 2000.

[17] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur. Topology-based detection of anomalous BGP messages. In *RAID*, 2003.

[18] M. Lad, D. Massey, D. Pei, Y. Wu, B. Zhang, and L. Zhang. PHAS: A prefix hijack alert system. In *USENIX Security*, 2006.

[19] H. Levin, M. Schapira, and A. Zohar. Interdomain routing and games. In *STOC*, 2008.

[20] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP misconfiguration. In *SIGCOMM*, 2002.

[21] Z. M. Mao, R. Govindan, G. Varghese, and R. H. Katz. Route flap damping exacerbates internet routing convergence. In *SIGCOMM*, 2002.

[22] C. D. Marsan. Six worst Internet routing attacks. http://www.networkworld.com/news/2009/011509-bgp-attacks.html, Jan. 15 2009.

[23] A. Mizrak, Y. Cheng, K. Marzullo, and S. Savage. Fatih: Detecting and isolating malicious routers. In *DSN*, 2005.

[24] R. Perlman. Routing with Byzantine robustness. Technical Report SMLI TR-2005-146, Sun Labs, 2005.

[25] S. Qiu, F. Monrose, A. Terzis, and P. McDaniel. Efficient techniques for detecting false origin advertisements in inter-domain routing. In *NPsec*, 2006.

[26] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. *SIGCOMM Comput. Commun. Rev.*, 36(4):291–302, 2006.

[27] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). http://www.ietf.org/rfc/rfc4271.txt, 2006.

[28] J. Rexford, J. Wang, Z. Xiao, and Y. Zhang. BGP routing stability of popular destinations. In *IMW*, 2002.

[29] G. Siganos and M. Faloutsos. Neighborhood watch for internet routing: Can we improve the robustness of internet routing today? In *INFOCOM*, 2007.

[30] L. Subramanian, V. Roth, I. Stoica, S. Shenker, and R. Katz. Listen and Whisper: Security mechanisms for BGP. In *NSDI*, 2004.

[31] P. Svensson. Pakistan causes worldwide YouTube outage. http://www.msnbc.msn.com/id/23339712, Feb. 25 2008.

[32] U.S.-China Economic and Security Review Commission. 2010 Report to Congress of the U.S.-China Economic and Security Review Commission. 111th Congress, 2nd session.

[33] C. Villamizar, R. Chandra, and R. Govindan. BGP route flap damping. http://www.ietf.org/rfc/rfc2439.txt, 1998.

[34] T. Wan, E. Kranakis, and P. van Oorschot. Pretty secure BGP (psBGP). In *NDSS*, 2005.

[35] T. Wan and P. van Oorschot. Analysis of BGP prefix origins during Google's May 2005 outage. In *SSN*, 2006.

[36] Y. Wang, M. Schapira, and J. Rexford. Neighbor-specific BGP: More flexible routing policies while improving global stability. In *SIGMETRICS*, 2009.

[37] R. White. Secure Origin BGP (soBGP). ftp://ftp-eng.cisco.com/sobgp/index.html, 2002.

[38] E. L. Wong and V. Shmatikov. Edge AS policy simulator. http://www.cs.utexas.edu/~elwong/edge-policy-sim.

[39] Z. Zhang, Y. Zhang, Y. C. Hu, Z. M. Mao, and R. Bush. iSPY: Detecting IP prefix hijacking on my own. *SIGCOMM Comput. Commun. Rev.*, 38(4):327–338, 2008.

[40] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, F. Wu, and L. Zhang. Detection of invalid routing announcement in the Internet. In *DSN*, 2002.

[41] C. Zheng, L. Ji, D. Pei, J. Wang, and P. Francis. A light-weight distributed scheme for detecting IP prefix hijacks in real-time. In *SIGCOMM*, 2007.

[42] D. Zhu, M. Gritter, and D. Cheriton. Feedback based routing. In *HotNets*, 2002.