Archives of BGP Updates: Integration and Visualization[*]

Giuseppe Di Battista, Federico Mariani,
Maurizio Patrignani, and Maurizio Pizzonia
Dip. Informatica e Automazione
Università di Roma Tre
Via della Vasca Navale 79, 00146 Roma, Italy
{gdb,mariani,patrigna,pizzonia}@dia.uniroma3.it

Abstract

The possibility of analyzing updates exchanged between BGP talkers is crucial for several operational and research purposes. To give a few examples, they can be used to investigate the stability of specific routes, to monitor the effects of faults, and to analyze the behavior of the entire network in the presence of particular events. Several archives of BGP conversations, such as the Routing Information System of the RIPE and the Oregon Route Views database, give an answer to this need for information. We describe a work that aims at integrating the data from different BGP-update sources and at presenting such data with graph-based visualization techniques. We address both technological issues related to the data integration and user-interaction issues originated from the necessity of visualizing data that change over time.

1 Introduction

The exploration and visualization of the Internet attracts an increasing research interest, motivated by the growing size of the network and the significant impact that connectivity has on social and economic activities.

Contributions in this field can be roughly classified with respect to the granularity of the data they consider.

For example, Hermes [6, 1] visualizes the information provided by the Internet Registries about the interconnections between Autonomous Systems (ASes), showing their peerings and mutual policies.

At a more granular level, Mercator [9], Skitter [11], and Rocketfuel [15] compute maps from data collected by means of network discovery, showing interconnections among routers. Another system using traceroutes for probing the network is described in [5].

In this paper we present BGPlay, a system that shows the routing at the interdomain level acquired from communications between BPG-speaking peers [16]. BGPlay integrates the data from different archives of BGP updates and presents such data with graph-based visualization techniques.

Currently, BGPlay integrates the BGP updates collected by the Routing Information Service [3] of the RIPE and those collected by the Oregon Route Views project [4]. BGPlay aims at giving a highly intuitive visual representation of the status of the routing at a specific time, and of its evolution in a given time interval. In order to show such an evolution the system relies on an animation which shows how the BGP-paths evolve over time while preserving the user mental map [14, 7].

BGPlay has been designed to satisfy both operational and research needs. To give a few examples, it can be used to investigate the stability of specific routes, to monitor the effects of faults, and to analyze the behavior of the entire network in the presence of particular events. Instabilities and faults of interdomain routing have been the subject of recent research (see for example [13, 10, 8, 12]).

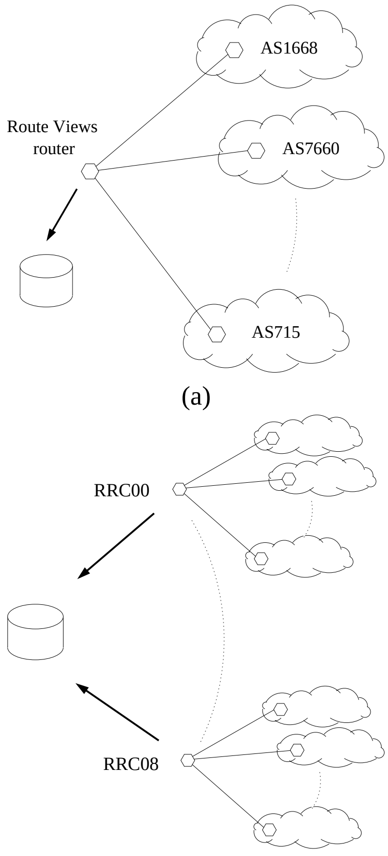
The paper is organized as follows. In Section 2 the data sources used by BGPlay are described. The overall architecture of the system is discussed in Section 3, while Section 4 and Section 5 illustrate how data are collected and visualized. Finally, Section 6 contains our conclusions and future work.

2 Archives of BGP Data

The Oregon Route Views (ORV) project (see Fig. 1.a) provides a service for Internet operators to obtain real-time

---

[*] Work partially supported by European Commission - Fet Open project COSIN - COevolution and Self-organisation In dynamical Networks - IST-2001-33555; by "Progetto ALINWEB: Algoritmica per Internet e per il Web", MIUR Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale; and by "The Multichannel Adaptive Information Systems (MAIS) Project", MIUR Fondo per gli Investimenti della Ricerca di Base.



(a)

(b)

Figure 1. The architecture of ORV (a) and of RIS (b).

information about the global routing system from the perspectives of several different locations around the Internet. Currently, the BGP Route Views router has more than 60 multi-hop eBGP peering sessions with routers of Internet service providers.

ORV collects, without providing any transit service, the BGP updates coming from its peers. Such updates allow to reconstruct the evolution of the best routes (at the AS level) adopted by each peer. ORV data archives contain both snapshots of the global routing information base of ORV at different instants of time and sequences of BGP updates. The data are made available in the MRT [2] format.

The Routing Information Service (RIS) of the RIPE (see Fig. 1.b) collects historical information about Internet routing by using Remote Route Collectors (RRC) at different locations around the world. Such information is integrated into a comprehensive view. An RRC of the RIS is a BGP speaking router that only collects BGP routing information. The collected raw data is regularly transferred to a central storage area at the RIPE NCC in Amsterdam. Each RRC is denoted by a string with format RRCxx, where xx is a two digit number. Essentially, each RRC behaves analogously to the ORV router.

The RIS offers to the users several query facilities. For example, the BGP Routing Hot Spot Utility generates lists of prefixes, originating from a specified AS, for which high BGP announcement activity has been observed by some RRC. Also, the ASInuse facility determines when an AS-number last appeared in the global routing table collected by the RIS.

In this paper, we are interested in a simpler RIS query, that can be performed using the Search by Prefix utility. It allows to specify a prefix, a time interval, and a set of RRCs in order to search the RIS database. It outputs a list of BGP updates recorded by the selected RRCs in the prescribed time interval. It also outputs the status of the Local Routing Information Base (Loc-Rib) for the selected prefix in an instant of time that is “related” to the prescribed time interval.

3 System Architecture

The architecture of BGPlay is based on the following main choices.

- The user interacts with the system by means of a browser. A query identifies a time interval and a prefix to be monitored. The data sources to be used can be selected in a set of possible alternatives.

- The results of the query, i.e. the changes in the BGP routing observed during the time interval, are visualized by an “animation.” Such animation relies on Graph Drawing techniques [7]. When the animation is started, a graph-like representation shows the routing at the beginning of the interval. During the animation the graph changes according to the observed BGP updates.

- The BGP data exploited to answer the user queries are partly fetched on-line at the moment they are needed and are partly locally stored. Namely, since the RIS provides a Web query facility we access those data online. Also, since the Route Views Project does not have a Web query facility for historical data, we periodically copy part of them locally. Currently, because of limitations on the available storage space, we maintain a copy of a limited (the most recent) period of time.

- The service is based on a client-server architecture, where the server computes the result of queries and the client is an applet running on the user’s browser. We decided of using an applet instead of producing on the server standard jpeg or gif figures because of the graphical complexity of the animation.

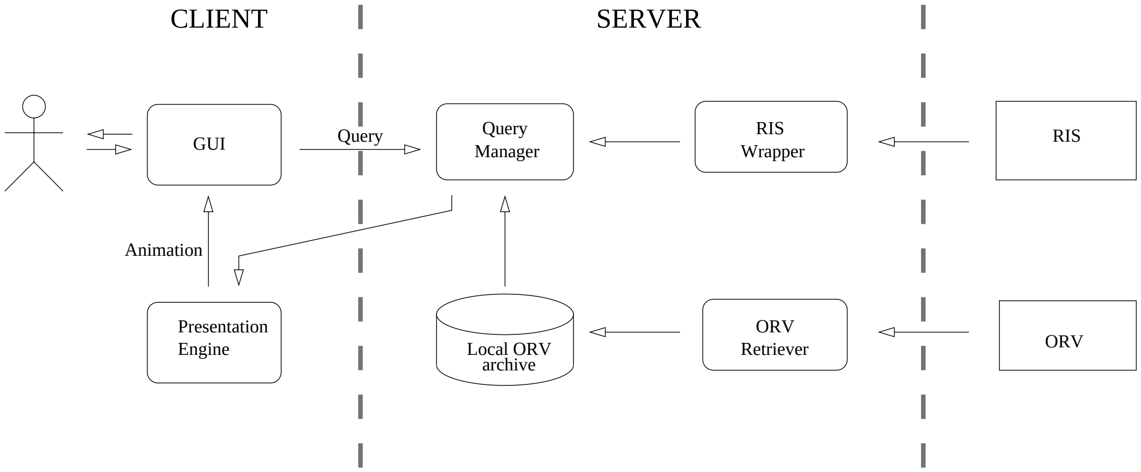


Figure 2. The system architecture.

Figure 2 illustrates the main components of the architecture:

Graphic User Interface. Provides the user with several tools for interacting with the visual representation of the routing information. For example: a *time panel* shows the time location of the most important events in the selected time interval, *animation buttons* allow to step over the events, and an *event display* gives all the available details about the currently visualized event.

Presentation Engine. Computes the layouts of the graphs representing the evolution of the routing. Further, is able to change such graph layout according to the routing events, in such a way that the user always perceives "smooth" variations. It exploits Graph Drawing methodologies and techniques.

RIS Wrapper. Queries the RIS Search by Prefix utility and retrieves the corresponding results in terms of BGP routing tables and updates.

ORV Retriever. Periodically retrieves routing tables and updates available at the ORV raw data archives in the MRT format. Updates the Local ORV Archive removing the oldest data.

Local ORV Archive. A relational database (currently, MySQL technology) storing the ORV data. Observe that, while all the RIS data refer to the UTC (Coordinated Universal Time), the ORV data refer to its local time. In BGPlay we refer to the UTC.

Query Manager. Gets a query description from the User Interface, drives the RIS Wrapper, accesses the Local ORV Archive, computes the result of the query, and delivers the result to the Presentation Engine.

An analysis of the workload and of the generated traffic lead us to a decomposition of the system in which the Presentation Engine is located on the client side, together with the Graphical User Interface.

4 Query Processing

In order to answer a query of the client, the Query Manager has to retrieve the BGP updates falling in the specified time interval. As mentioned in Section 3 two types of retrieval are involved: retrieval from the local database of ORV updates and on-line retrieval, performed by the RIS Wrapper, from the RIS Service.

The scenario is more precisely depicted in Fig. 3, where the starting and ending instants of the time interval are called t_1 and t_2, respectively. The purpose of the Query Manager is to compute:

- the status of the ORV RIB in t_1;
- the status of the RIS RIB in t_1; and
- the sequence of updates collected by ORV and RIS between t_1 and t_2.

Unfortunately, it is unlikely to have at disposal the two aforementioned RIBs at time t_1. Also, the possibilities of retrieving such information from the two data archives are quite different.

On one side, ORV offers a snapshot of its RIB every (about) two hours. Hence, we have to determine which is the last available ORV RIB before t_1. Once such RIB has been obtained (suppose it corresponds to a snapshot taken at time $t_0 \leq t_1$), we perform the following operations. First, the rows of the RIB corresponding to the given prefix p are

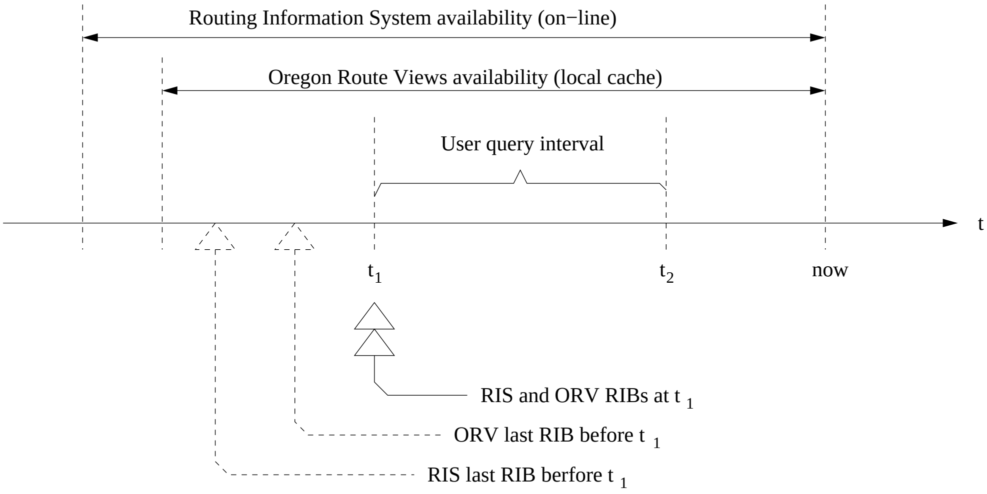


Figure 3. Available and computed RIBs.

extracted. Second, the updates occurring in the time interval t_0, t_1 are taken into account in order to compute the portion of the RIB corresponding to p at time t_1.

On the other side, the situation for RIS is more complex. Namely, as shown in Fig. 4, a query sent to the on-line RIS interface involving the time interval t'_1, t'_2 yields two types of data:

- the updates collected between t'_1 and t'_2 and
- the RIS RIB at time 23:59 of the day in which t'_1 falls.

Hence, it is not possible, for the Query Manager, to simply activate the RIS Wrapper with a query interval such that $t'_1 = t_1$ and $t'_2 = t_2$.

Because of the above discussion, the Query Manager asks the RIS Wrapper to perform a query with time interval t'_1, t'_2, where $t'_2 = t_2$ and t'_1 is the time 23:59 of the day preceding the one in which t_1 falls. Afterwords, analogously to what happens for ORV, the RIB is filtered saving only the portion concerning p. Then, the updates occurring in the time interval t'_1, t_1 are taken into account in order to compute the portion of the RIB corresponding to p at time t_1.

5 Updates Visualization

One of the purposes of BGPlay is to visualize BGP updates. However, the semantics of an update is essentially in the changes it induces in the routing. Hence, our approach relies on two types of visualization techniques. First, we exploit methods to visualize the status of the routing at a given instant of time. Second, we exploit techniques that allow the user to perceive how an update brings the routing from an initial status to a consequent one.

5.1 The Routing Graph

If we focus the attention on a prefix, the status of the routing at a given time for that prefix consists of a set of AS-paths, each representing the best route at that time to reach the prefix from a certain AS. Such a status is effectively represented with a *routing graph*. A routing graph is a decorated graph in which each vertex is an AS and edges are the pairs of ASes that appear consecutively in at least one of the paths. Each edge is labeled with the set of paths traversing it.

The AS-paths representing the status of the routing are provided to the Presentation Engine by the Query Manager. In general, part of them come from ORV and represent the routes selected by the ORV peerers to reach the prefix. Part of them come from RIS and represent the routes selected by the RRC peerers to reach the prefix. The overall information conveyed by the representation is the routing of the Internet traffic flowing toward the specified prefix at the AS level.

Visualizing a routing graph is not an easy task. In fact, even if large portions of it look like a tree, it may contain cycles and dense subgraphs. Further, some vertices may have many incident edges and a single edge may be traversed by several paths.

In designing the Presentation Engine we have identified the following requirements:

- The attention of the user should mainly be focused on the AS originating the prefix (target AS).
- An AS should appear in the drawing at a geometric distance from the target AS that is roughly proportional to the number of (AS-)hops separating them.
- Each single AS-path must be fully identified, even if traversing edges that are traversed by other paths. Ob-

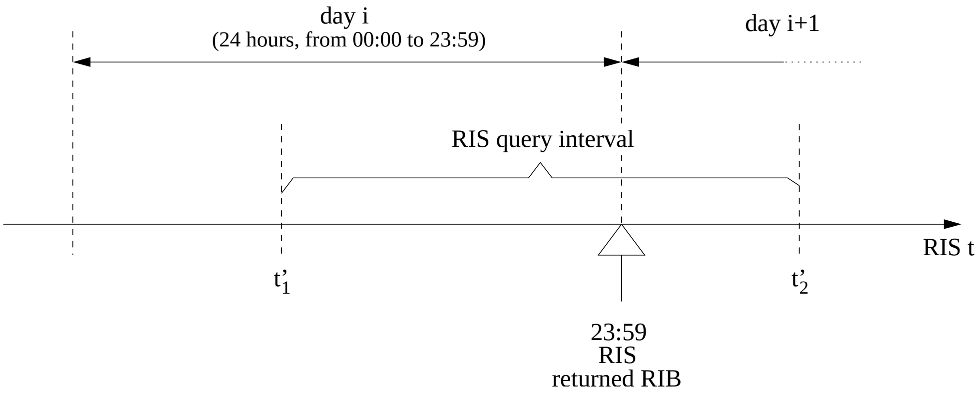


Figure 4. RIBs returned by the RIS.

serve that this requirement would be easy to meet if the graph was a tree; in fact in this case there is just one tree path from each AS to the target AS. The presence of cycles makes the problem more complex.

In order to compute drawings of the routing graphs satisfying the above requirements we used a "spring embedder." It consider the graph as a system of bodies (vertices) and forces acting between the bodies. Such forces can either attract or repel the bodies. The system is left free to oscillate until an equilibrium is reached. We made the choice to use the following types of forces:

- A repelling force is set between each pair of ASes.
- An attractive force is set between each pair of ASes connected by an edge.
- To damp the oscillations, the effect of the forces is decreased with the time. Essentially, this is obtained by progressively augmenting the "viscosity" in the system.

We also fixed the position of the target AS at the center of the area. Paths are represented with different colors and edges traversed by several paths are displayed using as many lines as the number of paths traversing that edge, where each line is colored with the color of the corresponding path.

Fig. 5 shows a drawing produced by BGPlay. It is about a prefix announced by the GARR AS (AS137). It clearly shows that the route collectors are reached by the announcements of AS137 through three main directions. Namely, part of the paths flow through AS3549 (Globalcrossing), other paths through AS1299 (TeliaNet), while others through AS20965 (GEANT). Fig. 6 shows the routing for AS7018.

5.2 Sequences of Updates

Obviously, the visualization of the routing graph before and after a BGP update occurred is sufficient to convey the information of the update. However, it is essential for the two consecutive visualizations to be "similar", while the routing change should be apparent to the user.

Consider a BGP announcement and the corresponding AS-path p'. Let AS_T be the target AS and suppose that the ending ASes of p' are AS_T and AS_S. Two cases are possible: either another AS-path p with the same pair of ending ASes was already in the routing graph or not. In the first case the user should perceive that the traffic flow from AS_S to AS_T is changing its route. In the second case it is important to make clear that a new source of traffic is becoming visible from AS_T.

A route change from p to p' involves several possible changes in the routing graph. Let us compare $p = (AS_T = AS_1, AS_2, \ldots, AS_m = AS_S)$ and $p' = (AS_T = AS'_1, AS'_2, \ldots, AS'_n = AS_S)$. We can split p and p' into sub-paths as follows. Let AS_i be the first AS of p that is equal to some AS (say AS'_j) of p'. We split p and p' into the sub-paths $(AS_1, \ldots, AS_i)$, $(AS_i, \ldots, AS_m)$ and $(AS'_1, \ldots, AS'_j)$, $(AS'_j, \ldots, AS'_n)$, respectively. We can repeat the above split process on the two sub-paths $(AS_i, \ldots, AS_m)$ and $(AS'_j, \ldots, AS'_n)$, until they are no longer decomposable. Such a process yields a decomposition of the two original paths into an equal number of subpaths pairwise starting and ending on the same vertices.

The Presentation Engine performs a graphic "morphing" where each sub-path in which p is decomposed is mapped to the corresponding sub-path of p'. Three cases are possible:

1. the two sub-paths have equal length
2. the sub-path of p is longer than the sub-path of p', and
3. the sub-path of p is shorter than the sub-path of p'.

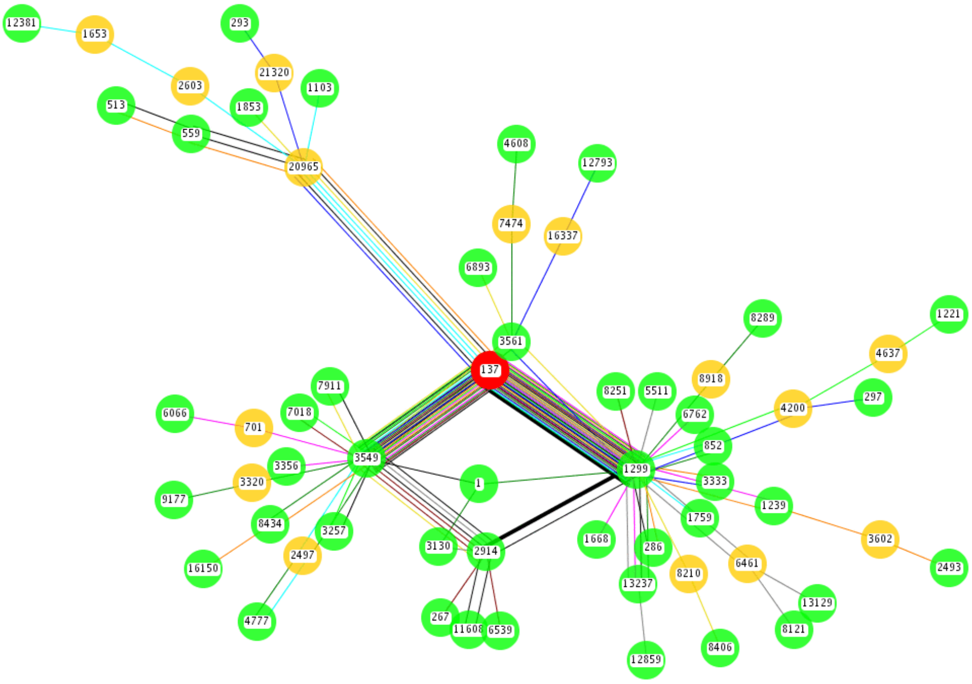


Figure 5. A drawing produced by BGPlay rapresenting the routing for prefix 193.204.0.0/15 (AS137).

In all the above cases we introduce in the routing graph a chain of additional “dummy” vertices and move them from the original sub-path to the position of the new sub-path. However, in case 2 some dummy vertices are “absorbed” into the same vertex, while in case 3 some dummy vertices are “created” from the same vertex. Of course, some ASes of the original graph may disappear since they are no longer traversed by any path and some ASes are created because they are in p' and were not in the graph.

Consider now a BGP withdrawal. It is managed by the Presentation Engine as follows. First, the involved path is highlighted, to attract the attention of the user, then the edges traversed by the path change their labels an, in case, removed.

6 Conclusions and Future Work

We have presented BGPlay, a system designed to visualize the evolution of the routing involving a given prefix at the Autonomous Systems level (BGP level). BGPlay is able to show the BGP events occurred in a given time interval through an animation illustrating the corresponding changes in the routing graph. BGPlay is available at `http://www.dia.uniroma3.it/~compunet`.

Future work will mainly focus on the following issues.

- We plan to test the effectiveness of new visualization methods, alternative to those based on the display of a routing graph.
- We would like to provide the user with more large-scale visualization facilities, that allow to track the evolution of the routing paths of several prefixes (not necessarily originated by the same AS) at the same time.
- We are interested in integrating new data sources, both publicly available and provided by private organizations.

Acknowledgements

We would like to thank Fabrizio Lombardozzi for implementing a preliminary version of the system, called Flapviewer.

References

[1] Hermes.
`http://www.dia.uniroma3.it/~hermes/`.

[2] Multithreaded Routing Toolkit (MRT) Project [Online].
`http://www.mrtd.net`.

[3] Routing information service of the ripe.
`http://www.ripe.net/ripencc/pub-services/np/ris/`.

[4] University of Oregon RouteViews project.
`http://www.routeviews.org`.

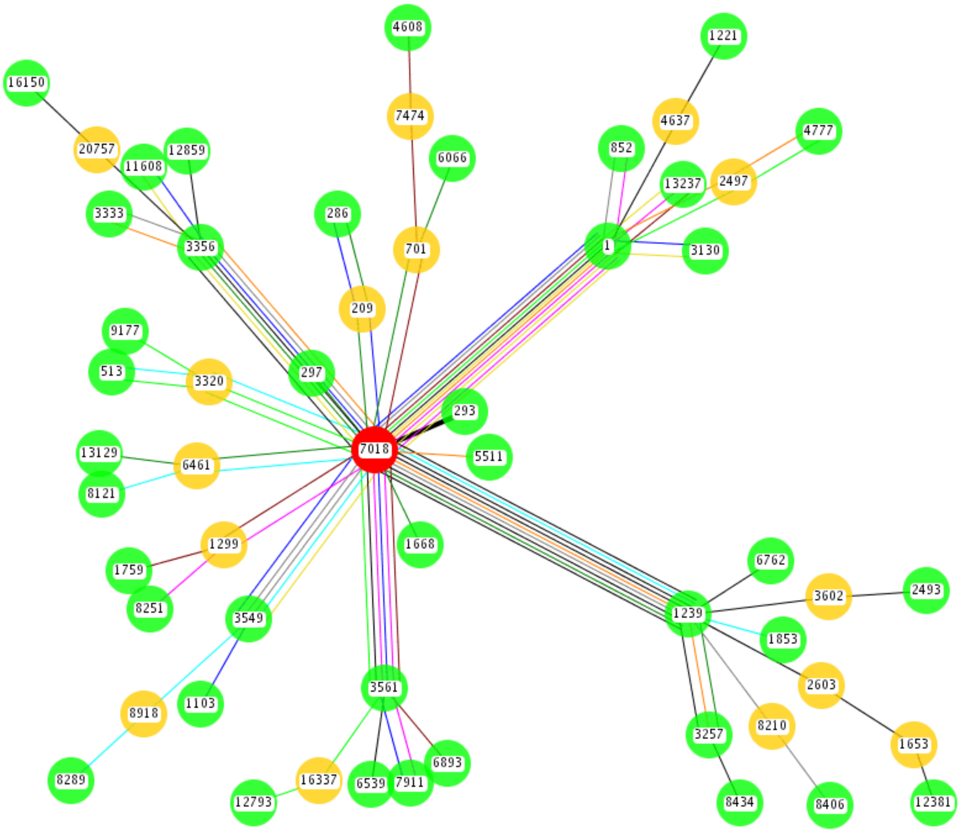


Figure 6. Flows of traffic for AS7018 represented by BGPlay.

[5] S. Branigan, H. Burch, B. Cheswick, and F. Wojcik. What can you do with traceroute? *IEEE Internet Computing*, Sept./Oct. 2001. `http://computer.org/internet/v5n5/index.htm`.

[6] A. Carmignani, G. Di Battista, W. Didimo, F. Matera, and M. Pizzonia. Visualization of the high level structure of the internet with hermes. *J. of Graph Algorithms and Applications*, 6(3):281–311, 2002.

[7] G. Di Battista, P. Eades, R. Tamassia, and I. G. Tollis. *Graph Drawing*. Prentice Hall, Upper Saddle River, NJ, 1999.

[8] L. Gao and J. Rexford. Stable internet routing without global coordination. In *Measurement and Modeling of Computer Systems*, pages 307–317, 2000.

[9] R. Govindan and H. Tangmunarunkit. Heuristics for internet map discovery. In *IEEE INFOCOM 2000*, pages 1371–1380, Tel Aviv, Israel, March 2000.

[10] T. Griffin and G. T. Wilfong. An analysis of BGP convergence properties. In *SIGCOMM*, pages 277–288, 1999.

[11] B. Huffaker, D. Plummer, D. Moore, and k claffy. Topology discovery by active probing. Technical report, Cooperative Association for Internet Data Analysis - CAIDA, San Diego Supercomputer Center, University of California, San Diego, 2002.

[12] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. Delayed internet routing convergence. In *SIGCOMM*, pages 175–187, 2000.

[13] C. Labovitz, G. R. Malan, and F. Jahanian. Internet routing instability. *IEEE/ACM Transactions on Networking*, 6(5):515–528, 1998.

[14] K. Misue, P. Eades, W. Lai, and K. Sugiyama. Layout adjustment and the mental map. *J. Visual Lang. Comput.*, 6(2):183–210, 1995.

[15] N. Spring, R. Mahajan, and D. Wetherall. Measuring isp topologies with rocketfuel. In *Proceedings of ACM/SIGCOMM '02*, Aug. 2002.

[16] J. W. Stewart. *BGP4: Inter-Domain Routing in the Internet*. Addison-Wesley, Reading, MA, 1999.