TECHNISCHE UNIVERSITÄT DRESDEN

Hauptseminar Technischer Datenschutz

Paul Walther
Chair of Privacy and Data Security

Learning Goals

- Methods and tools to familiarise with state of the art on research area
  - Finding literature
  - Efficient reading of literature
- Participation in scientific discourse
  - Writing about findings
  - Perform peer review
  - Presentation of findings

Timeline

05.04.22	13:00	Welcome meeting
11.04.22	23:59	**Deadline for sending topic preferences**
12.04.22	13:00	Defense Gregor Garten (presence APB/E006)
19.04.22	13:00	**Deadline for meeting with supervisor**
19.04.22	13:00	Introduction to literature research
10.05.22	13:00	Introduction to scientific writing
20.06.22	23:59	**Deadline for submission of written report**
21.06.22	13:00	Introduction to peer review
27.06.22	23:59	**Deadline for submission of reviews**
05.07.22	23:59	**Deadline revised version of report**
12.07.22	13:00	**Presentations**

Timeline

- Times for deadlines and meetings can be found on the website of the TU Dresden

  https://tu-dresden.de/ing/informatik/sya/ps/studium/seminars/hs-td

- Participation in defenses and colloquia announced on the website is strongly recommended.

Some numbers

- Written report
  - English or German
  - Around 8 pages (double column)
  - LaTeX template can be found on course website
  - Summarizes around 8 - 20 papers

- Presentation
  - English or German
  - 15 minutes for presentation (approx. 12 slides)
  - 5 minutes for Q & A

TABLE I

INTERNET MAPS USED IN EXPERIMENTS

Origin	Nb of nodes	Nb of links	Date
SCAN-LUCENT	284772	449228	1999
LSIIT	188347	235991	4-7/2002
route-views	13529	28060	7/2002

On a more theoretical level, it's worth noticing that several examples of theoretical studies of generated model networks, which give results quite similar to those obtained in this paper for real networks are given by Callaway *et al.* in [9], Cohen *et al.* in [10], which also contains some discussion of the diameter, and Holme *et al.* in [11], which also introduces the effect of correlations. Physicists usually plot the number of cluster of size s, as a cumulative function of s in a log-log scale. This distribution has been studied thoroughly in the physics literature and it is expected that the distribution will follow a power law with slope -2.5 at the transition point and a power law with an exponential cutoff above and below the transition point as shown by Newman *et al.* in [12] and Cohen *et al.* who also discuss in [13] the generalization of this power law for random failure in scale-free networks.

III. INTERNET MAPS

Studying Internet robustness involves knowing the Internet topology. In this section we present the data that we use in our experiments and we explain how we build an overlay in order to relate the IP nodes to their owning AS nodes.

A. Sources

As we want to obtain accurate and directly applicable results, we do not use AS level Internet maps for the basis of our study because they are too coarse-grained. Instead, we focus on the IP connectivity and therefore we prefer to work at the router level of the Internet. We use three Internet maps. The first one is a router level anonymous map which is the result of the merging of a map collected by the SCAN projet [5] and another one collected by the Lucent Internet mapping project [6]. It is the biggest router level Internet map currently available to our knowledge. It has been assembled in 1999 and has been used in [4]. The map as-is is not connected. We have removed 33 nodes in order to make this map connected. This is negligible in comparison of the size of this map. Furthermore these nodes were mostly in connected components of size 1 or 2 (i.e. single nodes or pairs of nodes). The second map is a router level map collected from our laboratory (called LSIIT and located in Illkirch, France) by using the Mercator software written by Govindan *et al.* and described in [5]. This map is connected. The collect lasted four months from April to July 2002. Unlike the '99 map, this one contains the IP addresses of the routers' interfaces. The third and last one is an AS level map collected by *route-views* [14] at the beginning of July 2002. We use it mainly to build an overlay with our '02 map but also for comparison with some router level results. Table I contains some information about these maps.

B. Building the overlay

We build a topological overlay in order to relate router level and AS level information. Our overlay creation method is quite different from the methods used in [15] and in [16] because we directly map the routers found by Mercator to the ASes found in the BGP table through the use of the IP interfaces and the BGP prefixes. Thus we do not have to generate the AS graph by a collapsing algorithm such as the one in [15] and we avoid the potential errors brought by the cases where many disjoint clusters of nodes belonging to the same AS have to be reassigned.

We use a BGP routing table dump from *route-views* created in July, 1st 2002 to build this overlay as well as an AS level map of the Internet containing 13529 nodes. For the overlay construction, we associate every prefix found in the table to its advertising AS (i.e. the AS at the right end of the AS path). This AS is not necessarily the originating AS of the prefix because the originating AS can be masked by AS path aggregation [17] (so errors can be introduced here). In the case where a prefix can be associated to more than one AS (because of protocol or database errors), we keep the first AS having the "i" (i.e. internal) flag set if one is found, otherwise we keep the first AS found (11 cases in our table, also sources of errors). The table contains 119814 prefixes (consistent with the results found by Bu *et al.* in [18]).

Then we use our IP level information collected by using Mercator to build a router level map of the Internet. The description of the Mercator software and its limitations can be found in [5]. Mercator can perform interface disambiguation and thus can properly assign multiple interfaces to their corresponding router. The resulting router level map contains 203854 interfaces and 188347 nodes. This yields an incidence rate of multiple interfaces of 8.2% which is nearly half the value observed in [19]. A first explanation for this difference is that our map, with an average degree of 2.5, is probably lacking an important number of redundant links that may potentially be multiple interfaces to any one node. Then for each interface, we search the longest prefix matching it and associate the originating (or advertising) AS of this longest prefix to the interface.

In this process, 1296 interfaces could not be mapped to an AS. 57 of these interfaces were class A addresses, 405 were class B and 834 were class C. Unresolved interfaces represent 0.64% of all the interfaces which is comparable to the 0.48% rate measured in [15]. Unlike their method, we have not used Internet Routing Registries (IRR) as additional sources of information because they are not accurate enough at least for our usage. Indeed Chen *et al.* have shown in [20] that about 62% of the records in the RIPE database are either void or obsolete despite the fact that RIPE is actively maintained up to date. Among the unresolved interfaces, many do belong to ASes (as a few requests to an IRR shows) but some of them such as the 188.1.x.x German research network (called DFN) are configured not to belong to any AS. We mark all the unresolved interfaces as belonging to the AS number 0. We define the meaning of the AS number 0 as: "an IP address with AS number 0 does not belong to any AS". Despite the

Figure 1: Example for scientific article with two column layout

Grading

- Weighting:
  - Report 60%
  - Presentation 25%
  - Review 15%
- Pass required for report and presentation
- Contacting your supervisor is required

- Common grading criteria:
  - Quality of literature research (coverage and relevance of papers)
  - Quality of discussion (identify and discuss commonalities, differences, and limitations)
  - Working style (autonomy and individual initiative)

Grading

- Core grading criteria for report:
  - Logical structure
  - Citation style and bibliography
  - Grammar and spelling
- Core grading criteria for presentation:
  - Slide quality (logical structure, usage of figures, conciseness of bullet points)
  - Talk quality (Duration, Q & A)
- Core grading criteria for review:
  - Thoroughness, constructiveness, specificity, politeness

Topics

Supervisor	Topic
Sebastian Rehms	Security Automation
Sebastian Rehms	Context-Awareness and Adaptiveness for Access Control
Stefan Köpsell	Machine Learning based Linkability Attacks
Stefan Köpsell	Confidentiality and (Location) Privacy in V2X Communication
Stefan Köpsell	Secure Computation based on Homomorphic Encryption
Stefan Köpsell	Broadcast/Multicast Encryption
Paul Walther	Wireless Identification using RF fingerprints
Paul Walther	Practical Challenges in Quantum Key Agreements
Paul Walther	Machine Learning in Physical Layer Security
	Your own topic

A short topic introduction and initial literature references can be found on the TUD website

Submission of topic preferences

- Send an email to paul.walther@tu-dresden.de
- The email should contain:
  1. Your first and last name
  2. Your first and second preference of the listed topics (title)
     - or -
     A short description of an arbitrary security or privacy-related topic that you find interesting (e.g. from recent news, from lectures)

Deadline for email: Monday, 11.04.2022, 23:59

Further notes

- OPAL will be used only for announcements (participants mailing list)
- The authoritative source for up-to-date notices as well as slides, links etc. is the course website
- Don’t forget to get in contact with your supervisor as soon as a topic has been assigned to you