10th IEEE Global Internet Symposium 2007
Paper Abstracts

8:45-10:00 Session 1: Peer-to-Peer

Distributed Social-based Overlay Adaptation for Unstructured P2P Networks

Ching-Ju Lin (National Taiwan University, TW); Yi-Ting Chang (National Taiwan University, TW); Shuo-Chan Tsai (National Taiwan University, TW); Cheng-Fu Chou (NTU, TW)

***Abstract*— The widespread use of Peer-to-Peer (P2P) systems has made multimedia content sharing more efficient. Users in a P2P network can query and download objects based on their preference for specific types of multimedia content. However, most P2P systems only construct the overlay architecture according to physical network constraints and do not take user preferences into account. In this paper, we investigate a social-based overlay that can cluster peers that have similar preferences. To construct a semantic social-based overlay, we model a quantifiable measure of similarity between peers so that those with a higher degree of similarity can be connected by shorter paths. Hence, peers can locate objects of interest from their overlay neighbors, i.e., peers who have common interests. In addition, we propose an overlay adaptation algorithm that allows the overlay to adapt to P2P churn and preference changes in a distributed manner. We use simulations and a real database called *Audioscrobbler*, which tracks users' listening habits, to evaluate the proposed social-based overlay. The results show that social-based overlay adaptation enables users to locate content of interest with a higher success ratio and with less message overhead.**

A Case for Unstructured Distributed Hash Tables

Krishna Puttaswamy (University of California, Santa Barbara, US); Ben Zhao (University of California at Santa Barbara, US)

***Abstract*—Structured peer-to-peer overlays support compelling applications such as large-scale file systems and distributed backup using the distributed hash table (DHT) interface. While unstructured file-sharing systems continue to flourish, wide adoption of structured applications has been elusive. We explore an alternative path to deployment of these applications by asking the question, can structured applications be run on top of unstructured overlays? We build an unstructured distributed hash table (UDHT) on top of state of the art search and topology management mechanisms, and evaluate whether it can sufficiently emulate properties of DHTs to support structured applications.**

Examining Graph Properties of Unstructured Peer-to-Peer Overlay Topology

Chao Xie (Georgia State University, US); Sijie Guo (Huazhong University of Science and Technology, CN); Reza Rejaie (University of Oregon, US); Yi Pan (Georgia State University, US)

***Abstract*— During the past few years, unstructured peer-to-peer (P2P) file-sharing systems have witnessed a significant increase in popularity. However, there lacks a systematic study on graph properties of the overlay topology. In this paper, we use accurate snapshots of the Gnutella overlay that span over roughly three years to explore changes in graph properties over long timescale. We investigate the effect of *network address translation* (NAT) on topology analysis. We examine a wide spectrum of graph properties characterizing the Gnutella top-level overlay topology and illustrate some interesting results. We find that the connection limit plays an important role in forming the unstructured overlay topology.**

10th IEEE Global Internet Symposium 2007
Paper Abstracts

10:30 - 12:15 Session 2: Measurement & Sampling

Importance of IP Alias Resolution in Sampling Internet Topologies

Mehmet Gunes (University of Texas at Dallas, US); Kamil Sarac (University of Texas at Dallas, US)

***Abstract*—Internet measurement studies utilize traceroute-based path traces to build representative Internet maps. These maps are then used to analyze various topological characteristics of the Internet. IP alias resolution is an important step in building a map from a set of collected path traces. In this paper, we study the impact of incomplete IP alias resolution on Internet measurement studies. Using a set of synthetic topologies and a genuine topology map, we experimentally show that the accuracy/completeness of alias resolution has an important effect on the observed topological characteristics. The results obtained in this work point out the importance of IP alias resolution and call for further research in alias resolution.**

Effect of sampling rate and monitoring granularity on anomaly detectability

Keisuke Ishibashi (NTT Corporation, JP); Ryoichi Kawahara (NTT Service Integration Laboratories, JP); Tatsuya Mori (NTT Service Integration Laboratories, JP); Tsuyoshi Kondoh (NTT, JP); Shoichiro Asano (National Institute of Informatics, JP)

***Abstract*—In this paper, we quantitatively evaluate how sampling decreases the detectability of anomalous traffic. We build equations to calculate the false positive ratio (FPR) and false negative ratio (FNR) for given values of the sampling rate, statistics of normal traffic, and volume of anomalies to be detected. We show that by changing the measurement granularity, we can detect anomalies even with a low sampling rate and give the equation to derive optimal granularity by using the relationship between the mean and variance of aggregated flows. With those equations, we can answer for the practical questions that arise in actual network operations; what sampling rate to set in order to find the given volume of anomaly, or, if the sampling is too high for actual operation, then what granularity is optimal to find the anomaly for a given lower limit of sampling rate.**

Stochastic Sampling for Internet Traffic Measurement

Tilman Wolf (University of Massachusetts, US); Yan Cai (University of Massachusetts, US); Patrick Kelly (University of Massachusetts, US); Weibo Gong (University of Massachusetts at Amherst, US)

***Abstract*— The increasing complexity of the Internet demands continued improvements to measurement techniques and data analysis methods to aid our understanding of network operation. The availability of accurate measurement data is necessary in many areas ranging from attack detection, novel pricing schemes, buffer dimensioning and switch design to general network management. In this paper, we develop a theory for accurate and unbiased Internet traffic measurement using the tools of Poisson random sampling. We show how this approach helps in storing, managing, and aggregating data from different sources with independent clocks and sampling rates. We present results that show that stochastic sampling maintains important information about network measurements that would be lost when using conventional uniform sampling.**

A Black-box Router Profiler

Roman Chertov (Purdue University, US); Sonia Fahmy (Purdue University, US); Ness Shroff (Purdue University, US)

***Abstract*— Simulation, emulation, and wide-area testbeds exhibit different strengths and weaknesses with respect to fidelity, scalability, and manageability. Fidelity is a key concern since simulation or emulation inaccuracies can lead to a dramatic and qualitative impact on the results. For example, high-bandwidth denial of service attack floods of the same rates have very different impact on the different platforms, even if the experimental scenario is supposedly identical. This is because many popular simulation and emulation environments fail to account for realistic commercial router behaviors, and incorrect results have been reported based on experiments conducted in these environments.**

In this paper, we describe the architecture of a black-box router profiling tool which integrates the popular ns-2 simulator with the Click modular router and a modified network driver. We use this profiler to collect measurements on a Cisco router. Our preliminary results demonstrate that routers and other forwarding devices cannot be modeled as simple output port queues, even if correct rate limits are observed. We discuss our future work plans for using our data to create high-fidelity network simulation/emulation models that are not computationally prohibitive.

10th IEEE Global Internet Symposium 2007
Paper Abstracts

13:45 - 15:00 Session 3: Security

The Case for Public Work

Wu-chang Feng (Portland State University, US); Ed Kaiser (Portland State University, US)

***Abstract*—Whether it is port scans, spam, or distributed denial-of-service attacks from botnets, unwanted traffic is a fundamental problem in all networked systems. Although proof-of-work has been proposed as a mechanism for thwarting such attacks, few proof-of-work systems have been successfully deployed. One of the problems in the proof-of-work approach is that the systems that issue and verify puzzles are typically located at or near the server edge. Rather than eliminate the denial-of-service problem, such approaches merely shift the problem from the service itself to the proof-of-work systems protecting the service. As a result, adversaries can disable services by flooding the issuer, by flooding the verifier, or by flooding all of the network links that lead to the issuer and verifier.**

To address this problem, this paper proposes a new approach for building proof-of-work systems based on publicly verifiable client puzzles. The system works by issuing a single "public work function" that clients must solve for each of its subsequent requests. Because the work function is publicly verifiable, any network device at the client's edge can verify that subsequent traffic will be accepted by the service. The system mitigates floods to the issuer since only a single work function needs to be given per client, thus allowing duplicate requests and replies to be supressed. The system mitigates floods to the verifier and across links leading to the server edge by allowing the verifier to be placed arbitrarily close to the client adversary.

Camouflaging Honeynets

Vinod Yegneswaran (SRI International, US); Chris Alfeld (University of Wisconsin, Maidison, US); Paul Barford (University of Wisconsin - Madison, US); Jin-Yi Cai (University of Wisconsin - Madison, US)

***Abstract*—Over the past several years, honeynets have proven invaluable for understanding the characteristics of unwanted Internet traffic from misconfigurations and malicious attacks. In this paper, we address the problem of defending honeynets against systematic mapping by malicious parties, so we can ensure that honeynets remain viable in the long term. Our approach is based on two ideas: (i) counting the number of probes received in the honeynet, and (ii) shuffling the location of live systems with those that comprise the honeynet in a larger address space after the probe count has exceeded a threshold. We describe four different strategies for randomizing the location of the honeynet. Each strategy is defined in terms of the degree of defense that it provides and its associated computational and state requirements. We implement a prototype middlebox that we call Kaleidoscope to gain practical insight on the feasibility of these strategies. Through a series of tests we show that the system is capable of effectively defending honeynets in large networks with limited impact on normal traffic, and that it continues to respond well in the face of large resource attacks.**

Inherent Behaviors for On-line Detection of Peer-to-Peer File Sharing

Genevieve Bartlett (University of Southern California, US); John Heidemann (University of Southern California, US); Christos Papadopoulos (Colorado State University, US)

***Abstract*—Blind techniques to detect network applications—approaches that do not consider packet contents—are increasingly desirable because they have fewer legal and privacy concerns, and they can be robust to application changes and intentional cloaking. In this paper we identify several behaviors that are *inherent* to peer-to-peer (P2P) traffic and demonstrate that they can detect both BitTorrent and Gnutella hosts using only packet header and timing information. We identify three basic behaviors: failed connections, the ratio of incoming and outgoing connections, and the use of unprivileged ports. We quantify the effectiveness of our approach using two day-long traces, achieve up to an 83% true positive rate with only a 2% false positive rate. Our system is suitable for on-line use, with 75% of new P2P peers detected in less than 10 minutes of trace data.**

10th IEEE Global Internet Symposium 2007
Paper Abstracts

15:30-17:15 Sess. 4: New Architecture and Models

The Virtually Invisible Internet

Suresh Singh (Portland State University, US);
Jim Binkley (Portland State University, US)

***Abstract*— We consider the question "What if we could redesign the Internet, how would we do it?" The preliminary proposal discussed in this paper is based on two observations – today, and in the future, wireless access by users will be the predominant mode of communication and communication typically follows a group-pattern where groups of people and/or entities engage in a specific communication activity. Using these two observations, we propose a novel group-based architecture with the goal of making the Internet more usable as well as invisible to users.**

Dynamic Internetworking Based on Late Locator Construction

Börje Ohlman (Ericsson, SE);
Anders Eriksson (Ericsson, SE)

***Abstract* — The legacy Internet technology is optimized for a semi-static inter-domain topology. Mobility or multihoming is handled by extending the legacy technology with new protocols. This paper describes a novel internetworking architecture with native support for a highly dynamic edge topology. The architecture distinguishes between a rather static core network on the one hand, and on the other hand edge networks forming an edge topology that may change on a short timescale due to mobility or re-homing events. The source host addresses the destination host directly with a hierarchically structured global locator. Indirection via a mobility agent is thus not needed. The name to locator resolution is based on a novel mechanism that constructs a global host locator on-demand that describes the current internetwork path from the core network to the host (late locator construction). This enables the resolution of the host name into a hierarchical and topologically significant host locator also with a highly dynamic edge topology where the path to the destination host traverses several moving and multihomed networks. The same locator construction mechanism is used to handle both node and network mobility as well as multihoming. Simulation results that verify the basic functionality of the late locator construction approach are reported.**

Packet Forwarding: Name-based Vs. Prefix-based

Craig Shue (Indiana University, US);
Minaxi Gupta (Indiana University, US)

***Abstract*— Using domain names for routing, instead of IP prefixes, has the potential to address many of the core outstanding issues in today's Internet. To initiate research in that direction, this paper compares the performance of name-based routing in the core of the Internet with that of IPv4 routing. Our analysis concludes that name-based routing is well within the scope of feasibility.**

OverSim: A Flexible Overlay Network Simulation Framework

Ingmar Baumgart (Universität Karlsruhe, DE);
Bernhard Heep (Universität Karlsruhe, DE);
Stephan Krause (Universität Karlsruhe, DE)

***Abstract*— A fundamental problem in studying peer-to-peer networks is the evaluation of new protocols. This paper presents *OverSim*, a flexible overlay network simulation framework based on OMNeT++. It was designed to fulfill a number of requirements that have been partially neglected by existing simulation frameworks. OverSim includes several structured and unstructured peer-to-peer protocols like Chord, Kademlia and Gia. These protocol implementations can be used for both simulation as well as real world networks. To facilitate the implementation of additional protocols and to make them more comparable OverSim provides several common functions like a generic lookup mechanism for structured peer-to-peer networks and an RPC interface. Several exchangeable underlay network models allow to simulate complex heterogeneous underlay networks as well as simplified networks for large-scale simulations. We show that with OverSim simulations of overlay networks with up to 100,000 nodes are feasible.**