Anonymous and Censorship-resistant Content-sharing in Unstructured Overlays

Michael Backes[1,2] Marek Hamerlik[3] Alessandro Linari[4]
Matteo Maffei[1] Christos Tryfonopoulos[3] Gerhard Weikum[3]
[1] Saarland University, [2] MPI-SWS, [3] Max-Planck Institute for Informatics, Saarbrücken, Germany
[4] Nominet UK & Oxford Brookes University, Oxford, UK
email: {backes,maffei}@cs.uni-sb.de, {mhamerli,trifon,weikum}@mpi-inf.mpg.de, alinari@brookes.ac.uk

Abstract

Semantic overlay networks are an instance of unstructured overlays, where peers that are semantically, thematically, or socially close are organized into groups to exploit similarities at query time. In this work we present Clouds, a novel P2P search infrastructure for providing anonymous and censorship resistant search functionality in such networks. Although we utilize semantic overlays to exploit their retrieval capabilities, our framework is general and can be applied to any unstructured overlay.

Categories and Subject Descriptors: H.3.4 [Systems and Software]: Distributed systems.
General Terms: Performance, Security.
Keywords: Semantic overlays, anonymity, censorship resistance.

1. System Overview

Semantic Overlay Networks (SONs) have proven a useful technology not only for distributed information retrieval, but also as a natural distributed alternative to Web 2.0 application domains such as decentralized social networking in the spirit of Flickr or del.icio.us. In this work we present Clouds, a novel system that provides anonymity and censorship resistance over SONs. Anonymity is achieved by relying on a self-organization of peers into groups that we call *clouds*. Message routing takes place among clouds instead of peers, thus hiding the identity of both the resource provider and the querying peer, while cloud size is a tunable parameter that affects anonymity and efficiency. Censorship resistance at communication level is achieved by a cryptographic protocol that guarantees the secrecy of the resource, thus avoiding censorship based on the inspection of the messages circulating in the network. The design of such a protocol needs to meet a number of challenging goals: allowing for the exchange of encrypted messages without assuming previously shared secrets, avoiding centralized infrastructures, like trusted servers or static gateways (e.g., as in [2]), and guaranteeing efficiency without establishing direct connections between peers. Clouds is the first system to guarantee anonymity and censorship resistance in SONs by addressing the aforementioned challenges.

Anonymity in Clouds is achieved by cloaking both the querying peer and the resource provider behind a group of neighboring peers, called cloud. Peers generate clouds at random, without necessarily using them, to minimize the correlation between the events of joining and using a cloud. They also non-deterministically decide to participate or not in clouds created by other peers. By design, clouds are populated by peers in the neighborhood of the cloud initiator. Communication takes place between clouds, and all peers in a cloud share the same probability of being involved in any communication which has this cloud as the start- or endpoint (*k-anonymity*). Clouds protocol is designed so that the observable behavior is the same for all peers, regardless of initiating of forwarding a message, to avoid compromising anonymity. The proposed cryptographic protocol aims at addressing the problem of censorship at the communication level, where a malicious party aims at filtering out any communication (i.e., query or resource) that contains unwanted content. The privacy of the resource is protected by cryptography, making it hard for the attacker to censor the communication by inspecting the message content.

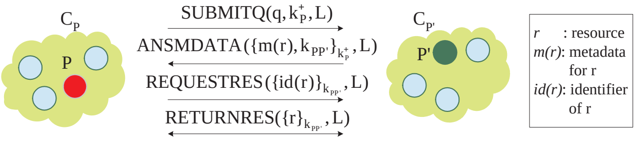
C_P
P
$\mathrm{SUBMITQ}(q, k^+_P, L)$
$\mathrm{ANSMDATA}(\{m(r), k_{PP'}\}_{k^+_P}, L)$
$\mathrm{REQUESTRES}(\{id(r)\}_{k_{PP'}}, L)$
$\mathrm{RETURNRES}(\{r\}_{k_{PP'}}, L)$
$C_{P'}$
P'
r : resource
m(r): metadata for r
id(r): identifier of r

Figure 1. Overview of Clouds protocol.

The protocol is composed of four steps summarized in Figure 1. A querying peer P uses a cloud C_P it participates in to issue a query q. This query follows a random walk in the cloud to obscure the message initiator, leaves the cloud from multiple peers to ensure higher resistance to censorship, and is routed towards a region in the network that possibly contains matching resources. A *footprint list* L is used to collect the traversed clouds, and facilitates the routing of the subsequent messages. A responder to the query P' encrypts the answer with the public key k^+_P received with the query message, and routes it towards the cloud of the querying peer C_P, as specified in L. Messages in the last three phases of the protocol specify a cloud as a destination, and when this cloud is reached, the message is broadcasted to reach the intended recipient. Notice that the cloud-based communication protocol is largely independent of the underlying network, and depends on a randomized query routing strategy and the footprint list.

Experiments with real-world data and queries show the effectiveness of Clouds under different attack scenarios (i.e, man-in-the-middle, blocking, intersection, and surrounding). For more details on Clouds protocols and a complete performance evaluation, the interested reader is referred to [1].

2. References

[1] M. Backes, M. Hamerlik, A. Linari, M. Maffei, C. Tryfonopoulos, and G. Weikum. Anonymity and Censorship Resistance in Unstructured Overlays. http://www.mpi-inf.mpg.de/∼trifon/papers/pdf/TR-MPI-Clouds08.pdf.

[2] M.K. Reiter and A.D. Rubin. Crowds: Anonymity for Web Transactions. *ACM Transactions on Information and System Security*, 1998.

PODC'08, August 18–21, 2008, Toronto, Ontario, Canada.
ACM 978-1-59593-989-0/08/08.