

Masterarbeit “ Peer-to-Peer Spam-Filter ”

Dieses Dokument gibt den Rahmen der Masterarbeit von Simon Schlachter im SS 2004 vor. Abweichungen oder Änderungen sind in gegenseitiger Absprache möglich. Die vorgegebenen Zeiten sind als ungefähre Richtlinien gedacht.

Arbeitsplan:

- 1) Ziel der Arbeit ist die Erstellung eines P2P Spam-Filter Systems auf Grundlage des vorhergehenden „Spamato“-Projektes. Bestandteil der Aufgabe ist es, einen öffentlichen Beta-Test durchzuführen, in dem sich die Qualität der Arbeit ermitteln lässt.
- 2) Einarbeitung in das Projekt „Spamato“. Identifikation von erweiterbaren bzw. fehlenden Komponenten. Recherche: Welche anderen Projekte gibt es in diesem Gebiet? [1W]
- 3) Das vorhandene System soll um weitere Filter ergänzt werden, dabei sind folgende Punkte zu berücksichtigen [6W]:
 - Implementierung einer Spamnet/Razor-Anbindung als weiteren Filter neben dem bestehenden URL-Filter.
 - Implementierung weiterer Filter, z.B. Bayesian-, „Google“- und/oder „RTT“-Filter.
 - Wie können neue Filter dynamisch in das System hinzugefügt werden?
 - Lokal adaptierbare Verrechnung der zurückgelieferten Spam-Werte der verschiedenen Filter.
- 4) Entwicklung eines (verteilten) Trust-Systems. Wie kann trotz der Anonymität der Benutzer dafür gesorgt werden, dass die Spam-Erkennung ein hohes Mass an Integrität aufweist? Wie können „ehrliche“ Benutzer von „bösen“ unterschieden werden? Dieser Aufgabenteil beinhaltet sowohl eine theoretische Betrachtung und Modellierung des Problems als auch eine praktische Implementierung. [8W]
- 5) Das vorhandene System soll so erweitert werden, dass der derzeit eingesetzte zentrale Server des URL-Filters durch ein verteiltes (P2P-)System ersetzt wird. Dabei ist zu bestimmen, welche Komponenten in welcher Form verteilt werden können. Können weitere Filter in die Verteilung einbezogen werden/in ähnlichem Masse davon profitieren? Welche (neuen) Probleme ergeben sich durch die Verteilung? [6W]
- 6) Zum Abschluss werden zwei Dokumente erstellt: [5W]
 - Ein Bericht (30 bis 50 Seiten, Sprache wählbar), welcher über die Arbeit und die Resultate Auskunft gibt. Dieser Bericht soll unter anderem auch eine kritische Beurteilung der eigenen Arbeit enthalten.
 - Einen Forschungsbericht (10 Seiten, in Englisch), welcher die Arbeit und die Resultate kompakter und unter einem wissenschaftlichen Gesichtspunkt darstellt.

Allgemeines:

- Selbstständiges Arbeiten ist Voraussetzung. Im HRS steht ein Arbeitsplatz zur Verfügung. Es besteht jedoch auch die Möglichkeit, zu Hause zu arbeiten.
- Während und nach der Arbeit ist eine Zwischen- bzw. Abschluss-Präsentation vorgesehen.
- Pro Woche findet mindestens ein Treffen mit Keno statt.

Kontaktpersonen:

1. Keno Albrecht kenoa@inf.ethz.ch
2. Roger Wattenhofer wattenhofer@inf.ethz.ch