

Software-Agenten im Internet

Florian Pircher
Technologische Fachoberschule
Oberschulzentrum Fallmerayer
Brixen, Italien

11. Februar 2016

Abstract

Menschen sind schon lange nicht mehr die einzigen Nutzer des Internets. Beinahe 50 % aller Webseiten-Aufrufe werden von autonomer Software getätigt. Diese auch als Bots bezeichnete Softwares agieren in den Schatten des Netzes. Unbemerkt indexieren sie Webseiten, verbreiten Spam, legen gefälscht Profile an oder versuchen in Datenbanken einzubrechen.

Hindernisse wie CAPTCHAs oder Honeypots galten bislang als effektive Gegenmittel, allerdings verhilft der rapide Fortschritt im Feld der Künstlichen Intelligenz modernen Bots auch derartige Barrieren zu durchbrechen. Im Bereich E-Mail-Spam findet ein unablässiger Kampf zwischen Spam-Bots und Klassifizierungsalgorithmen statt.

Im ersten Teil dieser Arbeit wird das Verhalten und Vermögen von Software-Agenten untersucht. Schwerpunkte bilden dabei die Themen Spam und Sicherheit. Der zweite Teil beschreibt die Anwendung des gewonnenen Wissens in Form der Entwicklung eines eigenen Bots der autonom durch das World Wide Web navigiert und anhand von maschinellern Lernen versucht CAPTCHAs zu knacken.

Inhaltsverzeichnis

1	Einleitung	3
2	Beobachten und Verstehen	4
2.1	Spam	5
2.1.1	Definition	5

1 Einleitung

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

2 Beobachten und Verstehen

2.1 Spam

2.1.1 Definition

Unerwünschte Nachrichten im Internet, größtenteils Werbeangebote, Phishing-Attacken oder Übermittler von Schadsoftware, werden als Spam bezeichnet. Spam-Nachrichten stellen per definitionem eine leidige Kommunikationsform dar. Die alltäglichste Art, namhaft durch ihre Omnipräsenz im Leben mit dem Internet, stellt der E-Mail-Spam dar.

Der Begriff fand noch vor dem Erscheinen des *World Wide Web* anklang, beschrieb damals jedoch eine Flut an unerwünschten Nachrichten.¹ Erste Aufkommen des Phänomens wurden im ARPANET² und im USENET beobachtet.

Der Name *Spam* entspringt dem gleichnamigen Dosenfleisch, welches im Zweiten Weltkrieg in großen Mengen an Soldaten verteilt wurde.³ Der Markenname Spam® (kurz für *Spiced Ham*) setzte sich auf diese Weise rasch im Vereinigten Königreich als Deonym für Frühstücksfleisch durch. Die britische Comedy-Gruppe Monty Python griff im Sketch *Spam*⁴ der BBC Serie *Monty Python's Flying Circus* die Allgegenwärtigkeit des Fleischprodukts auf, weshalb dieses bis heute als Symbol für einen unerwünschten Überschuss fungiert.

Mit dem Begriff *Spam* werden heutzutage diverse Arten von Nachrichten bezeichnet. Werbenachrichten sind das klassische Beispiel einer Spam-Nachricht. Sie bieten zumeist dubiose Produkte oder Dienstleistungen an.

¹Templeton, *Origin of the term "spam" to mean net abuse*.

²Postel, *On the Junk Mail Problem*.

³Longmate, »How We Lived Then: A History of Everyday Life During the Second World War«.

⁴Monty Python, *Spam - Monty Python's The Flying Circus*.

Literatur

Longmate, Norman. »How We Lived Then: A History of Everyday Life During the Second World War«. In: Random House UK, 1971, S. 142, 159.

Monty Python. *Spam - Monty Python's The Flying Circus*. 2009. URL: https://www.youtube.com/watch?v=M_eYSuPKP3Y (besucht am 27.01.2016).

Postel, Jon. *On the Junk Mail Problem*. 1975. URL: <http://tools.ietf.org/html/rfc706> (besucht am 05.02.2016).

Templeton, Brad. *Origin of the term "spam" to mean net abuse*. URL: <http://www.templetons.com/brad/spamterm.html> (besucht am 05.02.2016).