

---

## - Bachelor- / Masterarbeit - Flow-basierte Spam-Erkennung

---

<b>CASED</b>	<p>Das neue Forschungs- und Entwicklungszentrum für IT-Sicherheit CASED (Center for Advanced Security Research Darmstadt) stellt sich mit seinen starken Partnern – der Technischen Universität Darmstadt, dem Fraunhofer Institut für Sichere Informationstechnologie und der Hochschule Darmstadt – einer immensen Herausforderung: In einer einzigartigen Kooperation von Informatikern, Ingenieuren, Physikern, Juristen und Betriebswirten, mit internationalen Forschungszentren und Industriepartnern, werden zukunftsweisende IT-Sicherheitslösungen erforscht, entwickelt und in die Anwendung und wirtschaftliche Verwertung gebracht.</p>		
<b>Motivation</b>	<p>Email Spam stellt ein zentrales Problem im Internet dar. Unter Spam versteht man das unerwünschte Versenden von Emails mit kommerziellem Inhalt oder anhängender Schadsoftware an eine willkürliche Menge von Empfängern. Schätzungen nach stellen Spam-Emails 89-97% des jährlich anfallenden Emailaufkommens dar. Derzeit eingesetzte Verfahren zur Erkennung von Email Spam basieren im Wesentlichen auf der Analyse der Email Nutzdaten und erfordern somit die Verarbeitung großer und insbesondere sensibler Datenmengen, wodurch datenschutzrechtliche Bedenken und Skalierungsprobleme auftreten können.</p>		
<b>Aufgabenstellung</b>	<p>Im Rahmen dieser Abschlussarbeit sollen Möglichkeiten der Verwendung von Netzwerk-Flowdaten – 5-Tupel bestehend aus Quell- und Ziel-IP Adressen, Quell- und Ziel Ports, Layer 4 Protokoll ID sowie Informationen über die Intensität des Datenverkehrs (Bit/s, Pakete/s) – zur zentralisierten Spam-Erkennung untersucht werden. Die Verwendung von Flowdaten zur Spam-Erkennung führt unmittelbar zu einer Datenreduktion. Darüber hinaus enthalten Flowdaten keinerlei Nutzdaten der Email und sind somit datenschutzfreundlicher. Die/der Studierende soll sich im Rahmen dieser Arbeit mit der Verarbeitung und Analyse von Flowdaten zur Klassifikation von Email Spam beschäftigen.</p>		
<b>Voraussetzungen</b>	<ul style="list-style-type: none"><li>• Kreatives, eigenständiges und engagiertes Arbeiten</li><li>• Kommunikationsbereitschaft</li><li>• Gute Programmierkenntnisse (z.B. C, Perl, Ruby)</li><li>• Gute Kenntnisse der Internetprotokolle IP, TCP und SMTP</li><li>• Kenntnisse im Bereich Modellbildung und Simulation sind vorteilhaft</li><li>• Linux-Kenntnisse sind vorteilhaft</li><li>• Sehr gute deutsche und/oder gute englische Sprachkenntnisse</li></ul>		
<b>Starttermin</b>	<p>Ab sofort</p>		
<b>Kontakt</b>	<table><tr><td><b>Sebastian Abt</b> <a href="mailto:sebastian.abt@h-da.de">sebastian.abt@h-da.de</a> 06151.16-8416</td><td>CASED Mornwegstraße 32 64293 Darmstadt</td></tr></table>	<b>Sebastian Abt</b> <a href="mailto:sebastian.abt@h-da.de">sebastian.abt@h-da.de</a> 06151.16-8416	CASED Mornwegstraße 32 64293 Darmstadt
<b>Sebastian Abt</b> <a href="mailto:sebastian.abt@h-da.de">sebastian.abt@h-da.de</a> 06151.16-8416	CASED Mornwegstraße 32 64293 Darmstadt		