



Message Analysis Technology



The design goals for Katharion are to maintain an aggressive posture against email threats, while minimizing the number of falsely identified legitimate messages. To achieve this, Katharion uses a wide spectrum of tests, including both traditional techniques for spam detection and a variety of next-generation approaches for message analysis.

Overview

Katharion's multi-layered approach to spam detection is designed such that no single element will by itself classify a message as spam - thus avoiding the false positives associated with less thorough approaches - while the breadth of analysis results in a high detection rate for continually evolving patterns of junk email.

Developed by an established engineering team who had previously developed solutions for leading technology companies including Broadcom, Digital Insight, and Lucent, Katharion's approach for detecting junk email is the result of thousands of hours of development and the company's experience with processing billions of messages.

Dynamic Feedback-Based Rules Optimization

Katharion's message analysis leverages feedback from thousands of Katharion users as well as from many monitored legitimate and 'spam trap' email addresses. Feedback can be incorporated into the message analysis algorithm in near-real time, with certain safeguards in place to ensure that spammers posing as customers cannot exert undue influence over the detection process.

Signature Analysis

Also known as Message Fingerprinting, Signature Analysis entails the comparison of digital email "signatures" for a given message to signatures of known spam messages in frequently-updated public and internal databases. While many spammers introduce random text or varying elements (From address, subject line, URL, body text, etc.) in their messages, signature analysis can often detect these minor variations.

Bayesian Analysis

A self-learning Bayesian Engine analyzes patterns of phrases in messages, and assigns mathematical probabilities for the presence of those phrases in junk mail versus legitimate mail.

Proprietary and Collaborative Blacklists

Katharion leverages both public and private 'blacklists' of both IP addresses and unique resource identifiers known to be used by spammers.

This approach extends to both network data such as individual mail servers, relays, or IP networks, and to URIs communicated in the junk mail - including URLs, phone numbers, and physical addresses known to be used by spammers. The tests also can incorporate historical data, making them a very accurate method of detecting junk mail.

Real-Time Message Source Analysis

With the benefit of processing millions of messages per day, Katharion's network of systems can analyze spikes in mail flow from a particular source in real-time or near real-time, to assess whether an increased volume is simply a legitimate high-volume mailing, or the result of a spammer hijacking or the use of a 'zombie' network. This approach is not possible with software solutions, appliances, or with service-oriented solutions whose architecture is essentially 'stacks of appliances.'

Heuristics

Used as a complement to the other detection techniques, Katharion's set of heuristic rules is updated frequently and encompasses both message headers and the message body. The heuristics are constructed to be effective for both English-language and non-English messages.

Customizable Whitelists and Blacklists

While Katharion is designed to be accurate 'out of the box', customers can choose to allow or disallow email messages from a given sender, domain, or network via so-called whitelists and/or blacklists. These can be applied for a given recipient, across a full domain, across multiple domains within an organization, or on an account-wide basis.