

УДК 004.492; УДК 004.896; УДК 004.912

СИСТЕМА ФИЛЬТРАЦИИ НЕЖЕЛАТЕЛЬНОЙ КОРРЕСПОНДЕНЦИИ

А.М. Малафеев, М.А. Ананьин

Научный руководитель – к.т.н., доцент А.О. Корепанов
Самарский государственный аэрокосмический университет
имени академика С.П. Королёва

Рекламные, мошеннические и вредоносные сообщения, массово рассылаемые по электронной почте, называются нежелательной корреспонденцией или «спамом». Несмотря на значительные усилия, прилагаемые Интернет-сообществом по пресечению нежелательной корреспонденции, она продолжает преследовать пользователей электронной почты. По существующим оценкам в настоящее время 60-80% всей электронной почты является «спамом». Ежегодно количество «спам-сообщений» увеличивается почти вдвое. В связи с интенсивным развитием и массовостью использования Интернета в последние несколько лет проблема распространения «спама» стала особенно актуальна. Среднестатистическому пользователю электронной почты приходит по 50-100 писем ежедневно, более половины из них являются нежелательной корреспонденцией, что значительно снижает эффективность работы, отнимая время на обработку входящей почты.

На сегодняшний день существуют три основных подхода для борьбы со спамом. Использование спам-фильтров – это наиболее распространенный подход. Спам-фильтр на основе анализа содержимого сообщения электронной почты, рассчитывает вероятность принадлежности письма к нежелательной корреспонденции. Другой подход – использование так называемого «белого списка» – набора адресов, которым доверяет пользователь. Письма, полученные с адресов, не вошедших в «белый список» считаются «спамом». Третий общеизвестный подход, основан на использовании так называемого «черного списка» – набора адресов, письма, полученные с которых, считаются спамом. Также существуют различные методы, основанные на комбинации перечисленных подходов. Однако все они обладают существенными недостатками, которые во многих случаях делают неэффективным их применение.

В настоящей работе предлагается многоуровневый механизм контроля входящей корреспонденции. На первом этапе применяется оригинальный подход, позволяющий определить, что отправителем письма является человек, а не робот. На следующем этапе используется подход, основанный на модификации методов «белого» и «черного» списков, позволяющий устранить значительную часть недостатков своих предшественников, сохранив при этом их преимущества. На последнем этапе применяется оригинальный спам-фильтр, основанный на использовании семантического и корреляционного анализов.

Применение разработанного многоуровневого механизма позволяет достичь значительного повышения эффективности фильтрации нежелательной корреспонденции, а также делает нерентабельным и неэффективным использование автоматических систем массовой рассылки «спама». Высокая эффективность разработанных методов подтверждена результатами экспериментальных исследований.

Проект представляется на рассмотрение экспертному совету по отбору инновационных научных разработок в рамках программы У.М.Н.И.К. (участник молодёжного научно-инновационного конкурса) в связи с возможностью дальнейшей коммерциализации.