DOI 10.37882/2223–2966.2022.03.38

# ПРОГРАММА ОБНАРУЖЕНИЯ НЕЖЕЛАТЕЛЬНЫХ ЭЛЕКТРОННЫХ ПИСЕМ

***Чжэн Цзини***
*МГТУ им. Н.Э. Баумана*
*sofiazjy@gmail.com*

*Аннотация.* В статье рассматриваются основные риски, связанные с электронными письмами и их признаки, дается обзор основных подходов к защите электронной корреспонденции. Особое внимание уделяется методам обнаружения нежелательных писем. Предлагается новое решение фильтрации корреспонденции, расширяющее возможности байсовской классификации, на базе системного подхода и методов искусственного интеллекта. Описывается программное решение на языке Python на базе библиотеки SMTP.

*Ключевые слова:* электронная почта, почтовые сервисы, спам, политика безопасности, фильтрация, шифрование, лингвистический анализ, мешок слов, DLP-системы, SMTP, IMAP.

## UNWANTED E-MAIL DETECTION PROGRAM

***Zheng Jing***

*Summary.* The article examines the main risks associated with emails and their signs, provides an overview of the main approaches to protecting electronic correspondence. Particular attention is paid to methods for detecting unwanted emails. A new correspondence filtering solution is proposed, expanding the possibilities of Bayesian classification, based on a systematic approach and artificial intelligence methods. The software solution in Python based on the SMTP library is described.

*Keywords:* email, postal services, spam, security policies, filtration. encryption, linguistic analysis, a bag of words, DLP systems, SMTP, IMAP.

Электронная переписка представляет собой одну из системообразующих реалий сегодняшнего дня. Обойтись без нее не представляется возможным, но при этом все, что связано с почтовыми сервисами, несет в себе высочайший потенциальный риск, так как почти все компьютеры в мире в той или иной мере контролируются спамерами и чат-бот-сетями, а доля спама в обмене электронными сообщениями превышает 50% [1].

Нежелательность некоторых электронных писем определяется или их бесполезностью, сопровождающейся затратами ресурсов (память, время на обработку), или исходящими от них угрозами. К числу последних могут относиться: заражение компьютеров вирусами, червями, троянами, утрата контроля над почтой, хищение персональных данных, интеллектуальной собственности и другие негативные последствия.

В свою очередь к признакам нежелательности могут относиться:

- Сомнительный отправитель. Корреспонденция от нежелательных адресатов, например, от спамогенераторов, удаляются автоматически. Имеется возможность самостоятельного добавления в список блокировки доменов верхнего уровня целых стран и регионов. Например, если в списке установить флажок MX [Мексика], будут блокироваться сообщения с адресов, оканчивающихся на.mx.

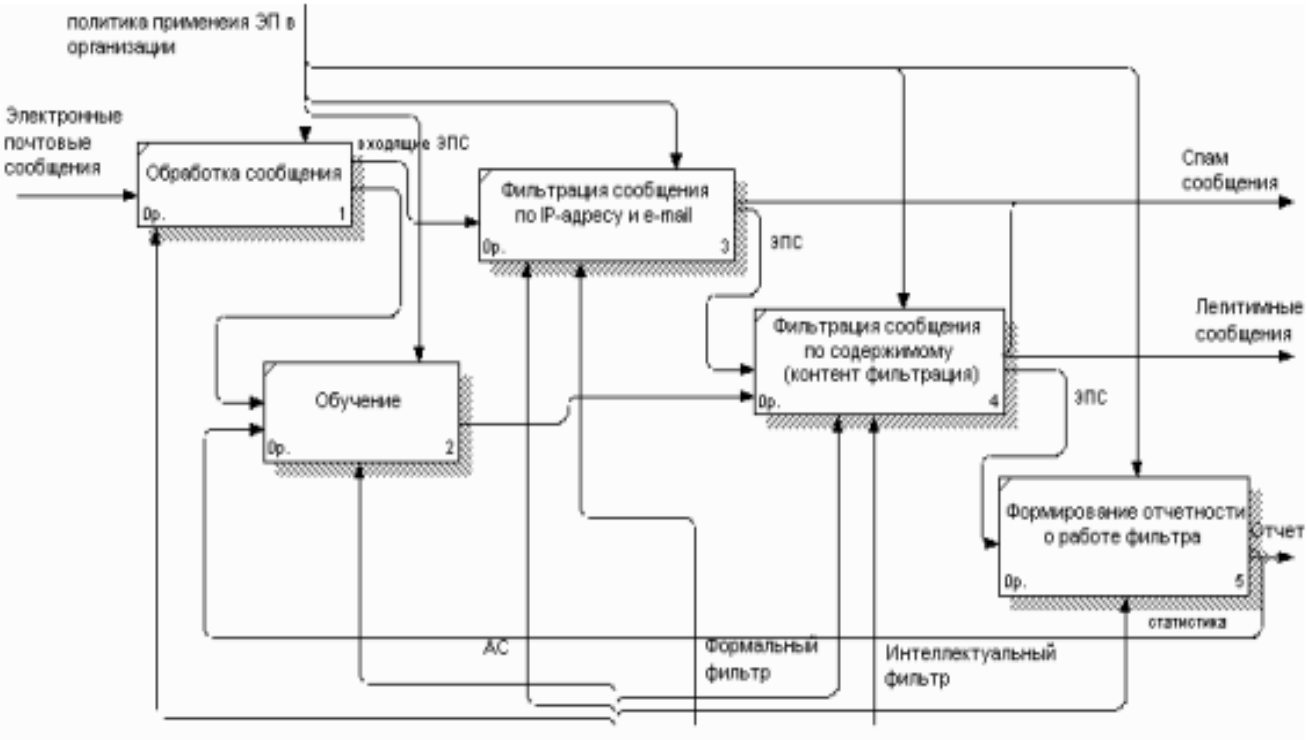


Рис. 1. Фильтрация сообщений на основе глубокого обучения

- Сомнительный заголовок. Программное обеспечение, генерирующее спам-письма часто допускает ошибки при оформлении заголовков. Благодаря этому просчету антиспам-фильтры и обнаруживают нежелательную корреспонденцию. Такая защита весьма надежна и эффективна.
- Сомнительное содержание. Классическая проверка на спам и агрессивный маркетинг.
- Сомнительные вложения. Кроме проверки текста и заголовка писем требуется проверка вложений.
- Сомнительное поведение. Требуется так же проверка на наличие вредоносного кода.
- Подозрение на доступ посторонних к переписке. Бывает так, что письмо само по себе не содержит никаких подозрительных признаков, однако нет уверенности, что доступ к корреспонденции имеет только адресат. Это одна из ситуаций, когда уместно использование шифрование переписки.

В ходе борьбы с нежелательной корреспонденцией применяются следующие основные подходы:

- Фильтрация. Основной метод, основанный на поиске корреспонденции с признаками нежелательности и ее блокировке тем или иным способом. Фильтрация нужна для того, чтобы письма, перед тем как поступить в ваш почтовый ящик, проходили автоматическую обработку в соответствии с заданными вами правилами (политиками безопасности), включая новейшие стандарты безопасности, такие как SSL/TLS, DMARC, DKIM, SPF и другие.
- Доверенные списки отправителей. Проверенные источники, добросовестные отправители. Правила фильтрации писем для них существенно мягче чем для прочих корреспондентов.
- Черные списки отправителей (DNS Black List (DNSBL)). Достаточно распространенный метод, заключающийся в блокировке всех отправлений, исходящих из почтового ящика, уличенного в распространении спама.
- Серые списки greylisting. Неизвестные отправители в режиме наблюдения и изучения вначале помещаются в серый список. В зависимости от их исхода дальнейших проверок возможен переход в доверенный или черный список.

Различные методы проверки отправителя с использованием технических и технологических средств. Проверка существования пользователя на отправляющей стороне (callback), проверка «правильности» отправляющего сервера такими методами, как наличие записи в реверсной зоне DNS, легальности имени при установке SMTP-сессии (helo), проверка SPF записи (для работы этого в DNS записи о хосте используется соответствующая запись о легальных серверах отправителей).

- Сокрытие реального адреса получателя. Применяются самые изощренные методы для сокрытия от спамеров и автоматических скриптов реальных адресов электронной почты.
- Шифрование. Метод защиты, основанный на сокрытии от нежелательных глаз всей или части переписки при помощи ее шифрования как правило при помощи криптостойких алгоритмов.
- Гибкая настройка политик безопасности. Возможность задавать для правил анализа текстов условные приоритеты, а потом использовать эти приоритеты как параметры фильтрации событий в журнале событий [8]. Наибольшими функциональными возможностями обладают системы класса DLP-системы. DLP-система — специализированное программное обеспечение, предназначенное для защиты компании от утечек информации [6]. Наиболее известными продуктами этого класса являются McAfee DLP, Sophos Endpoint Protection, InfoWatch Traffic Monitor, Solar Dozor, Oracle IRM, Microsoft RMS [9].
- Лингвистический анализ. Мешок слов. TF-IDF — это метод, который увеличивает веса слов, часто встречающихся в данном документе, и уменьшает веса слов, часто встречающихся во многих документах.
- Аддитивное сглаживание.

Вопросам противодействия спаму посвящены исследования И.С. Ашманова [10], С.С. Валеева, А.С. Катасёва [11], А.П. Никитина, М.А. Семеновой, А.А. Шварца и др. В основном, это фильтры, построенные на байесовском подходе, что, как известно, не позволяет учитывать семантику электронных сообщений.

Они просты во внедрении и удобны в использовании, при качественном обучении отсекают до 98% спама [1] и дают возможность дополнительно обучить фильтр в случае ложных срабатываний. На базе этого метода реализованы системы SpamBuster, SpamKiller, SpamEater, SpamGuard, SpamAssassin и DSPAM.

Однако, такой метод имеет три фундаментальных недостатка, которые широко используются отправителями спама, и могут свести эффективность фильтрации практически к нулю:

Во-первых, он ориентирован только на работу с текстом. Спамеры помещают информацию, которую хотят донести до получателей, в документ, который не может быть проанализирован как простой текст, например в изображение или документ формата PDF. А этот документ, в свою очередь, вставляют в тело письма. Фильтр не может классифицировать такое сообщение как спам, поэтому пропускает его.

Во-вторых, метод основан на предположении, что в спаме чаще содержатся одни слова, а в нормальных письмах другие. Если это предположение оказывается неверным, то метод утрачивает свою эффективность.

В-третьих, для обхода фильтра может быть использован метод «Байесовского отравления» — в письмо добавляется специально подобранный лишний текст, обманывающий фильтр и заставляющий его считать сообщение нормальным.

При разработке систем фильтрации входящих сообщений недостаточно полно используется системный подход и современные технологии искусственного интеллекта для решения задачи классификации. Тем самым, задача разработки эффективных методов и алгоритмов фильтрации спама в организации является актуальной.

Предлагается программное решение на языке Python на базе библиотек smtplib, imaplib.

Общая схема работы в виде IDEF0-диаграммы представлена на рисунке 1.

Рассмотрим процесс проверки писем на нежалательный контекст при помощи скриптов языка Python. Разобьем процесс на несколько шагов.

- Чтение электронной почты начинается с подключения к почтовому ящику на сервере. Определяем интерпретатор для скрипта. Явно указываем кодировку.
- Импортируем модуль imaplib для возможности подключения к почтовому ящику по IMAP.
- Создаем сессию для подключения к почтовому ящику по IMAP и заносим ее в переменную mail.
- Подключаемся к почтовому ящику по IMAP с использованием реальной учетной записи gmail.
- Выводим список папок в почтовом ящике. Выбираем для работы папку входящие (inbox). Получаем массив со списком найденных почтовых сообщений.
- Сохраняем в переменную ids строку с номерами писем.
- Получаем массив номеров писем.
- Задаем переменную latest_email_id, значением которой будет номер последнего письма.
- Получаем письмо с идентификатором latest_email_id (последнее письмо).
- В переменную raw_email заносим необработанное письмо.
- Переводим текст письма в кодировку UTF-8 и сохраняем в переменную raw_email_string.
- Запускаем скрипт — при отсутствии нежелательных признаков, он должен вернуть пустой ответ.

- Импортируем модуль email для получения заголовков и тела писем.
- Получаем заголовки и тело письма и заносим результат в переменную email_message. Обратите внимание, что мы используем переменную raw_email_string, в которую ранее занесли необработанное письмо.
- Выводим на экран заголовок To (кому отправлено письмо).
- Выводим на экран заголовок From (от кого отправлено письмо).
- Выводим на экран заголовок Date (дата отправки письма).
- Выводим на экран заголовок Subject (тема письма).
- Выводим на экран заголовок Message-Id (идентификатор письма).
- Чтение тела письма. Проверяем, является ли письмо многокомпонентным. Если да, то выводим по очереди на экран значения каждого компонента. Предварительно, перекодируем текст в UTF-8. Если письмо не многокомпонентное, выводим его содержимое.

Таким образом, происходит проверка и фильтрация всех полученных писем.

ЛИТЕРАТУРА


1. Защита почты от спама и фишинга: актуальные угрозы и передовые решения. [Электронный ресурс] Режим доступа: https://www.kp.ru/guide/zashchita-pochty-ot-spama-i-fishinga.html
2. Классификация текстов с помощью мешка слов. Руководство [Электронный ресурс] Режим доступа: http://datareview.info/article/klassifikatsiya-tekstov-s-pomoshhyu-meshka-slov-rukovodstvo/
3. Tejan Karmali. Классификатор спама в Python с нуля. Дата публикации August 2, 2017. [Электронный ресурс] Режим доступа: https://www.machinelearningmastery.ru/spam-classifier-in-python-from-scratch-27a98ddd8e73/
4. Хмельков Игорь. Мешок слов и сентимент-анализ на R. [Электронный ресурс] Режим доступа: https://habr.com/ru/post/255143/
5. Усман Малик. Python для НЛП: Создание модели мешка слов с нуля [Электронный ресурс] Режим доступа: https://pythobyte.com/python-for-nlp-creating-bag-of-words-model-from-scratch-31931/
6. Что такое DLP-системы, кому и когда они нужны [Электронный ресурс] Режим доступа: https://rt-solar.ru/products/solar_dozor/blog/2080/
7. Обзор средств защиты электронной почты [Электронный ресурс] Режим доступа: https://habr.com/ru/company/cybersafe/blog/269513/
8. Как мы DLP-систему выбирали (практический опыт) [Электронный ресурс] Режим доступа: https://habr.com/ru/post/440838/
9. Как работают DLP-системы: разбираемся в технологиях предотвращения утечки информации [Электронный ресурс] Режим доступа: https://xakep.ru/2011/05/04/55604/
10. Ашманов, И.С. и др. Технологии фильтрации содержания для Интернет. Труды международной конференции «Компьютерная лингвистика и интеллектуальные технологии-2021» [Электронный ресурс] Режим доступа: https://www.dialog-21.ru/digest/2002/articles/ashmanov/
11. Катасёв А.С., Катасёва Д.В. Разработка нейросетевой системы классификации электронных почтовых сообщений // Вестник Казанского государственного энергетического университета. — 2015. — № 1 (25). — С. 68–78.




Журнал «Современная наука: актуальные проблемы теории и практики»