Сервисы проверки ссылок сами могут раскрывать личные данные: ученые обнаружили тысячи опасных URL

Дата публикации: 10.08.2026

Мы привыкли считать проверку подозрительной ссылки одним из самых простых способов защититься от интернет-мошенников. Специальные сервисы анализируют адрес сайта, ищут вредоносный код, проверяют репутацию домена и помогают определить, не скрывается ли за обычной ссылкой фишинговая страница. Однако исследователи обнаружили парадоксальную проблему: инструмент безопасности в некоторых случаях сам способен стать источником утечки данных.

Причина связана не столько с ошибкой программного обеспечения, сколько с устройством современного интернета. Сегодня URL может содержать значительно больше информации, чем просто адрес страницы. В ссылку нередко встроен уникальный токен — длинная последовательность символов, которая фактически выполняет функцию временного цифрового ключа.

Такие ссылки окружают нас повсюду. Они используются для сброса пароля,

подтверждения электронной почты, входа без пароля, просмотра документов, доступа к облачным файлам, регистрации на мероприятиях и управления бронированиями. Иногда достаточно получить такую ссылку, чтобы открыть предназначенную конкретному человеку страницу без дополнительного ввода логина и пароля.

Именно здесь возникает неожиданный конфликт между безопасностью и конфиденциальностью. Если корпоративная система автоматически отправляет ссылку из входящего письма в сервис проверки URL, тот анализирует ее на наличие угроз. Но некоторые подобные платформы одновременно публикуют результаты сканирования или сами URL в общедоступных потоках данных.

В результате ссылка, которая должна была попасть только одному человеку, теоретически может стать доступной посторонним.

Исследователи из CISPA, Института безопасности и конфиденциальности Макса Планка и Университета Ка' Фоскари в Венеции решили выяснить, насколько серьезна эта проблема. Их работа LEAKYLINKS была представлена на IEEE Symposium on Security and Privacy 2026.

В центре исследования оказались сервисы, предназначенные для анализа интернет-адресов. Подобные системы чрезвычайно полезны для современной кибербезопасности. Они позволяют обнаруживать фишинговые сайты, вредоносные скрипты и другие угрозы еще до того, как пользователь самостоятельно откроет страницу.

Особенно активно такие технологии применяются в компаниях. Например, защитная система электронной почты может автоматически извлекать ссылки из полученного сотрудником сообщения и отправлять их на проверку. Для пользователя все происходит незаметно: письмо приходит уже после анализа потенциально опасных адресов.

Проблема начинается тогда, когда результат такого анализа оказывается публичным.

Чтобы оценить масштаб явления, исследователи разработали систему LEAKYLINKS и собрали URL из открытых потоков шести крупных сервисов сканирования. Всего было автоматически обработано более двух миллионов веб-адресов.

Проверять такое количество ссылок вручную было бы практически невозможно и одновременно создавало бы серьезную проблему с точки зрения конфиденциальности. Поэтому исследователи применили большую языковую модель, которая классифицировала потенциально чувствительные URL.

При этом система искусственного интеллекта работала на защищенной внутренней инфраструктуре. Это было принципиально важно: обнаруженные персональные сведения не отправлялись внешним сервисам ИИ и другим третьим сторонам.

Результаты оказались неожиданными. Среди более чем двух миллионов проанализированных URL исследователи обнаружили свыше 4000 случаев, когда общедоступная ссылка могла предоставлять доступ к конфиденциальной информации. Точность автоматического обнаружения таких случаев достигала 97%.

Среди найденных материалов встречались визовые документы, государственные документы и сведения о бронированиях. Таким образом, речь идет не просто о технических параметрах веб-страниц, а о потенциальном доступе к реальным персональным данным.

Механизм утечки проще понять на примере. Представим, что человеку приходит письмо со ссылкой для просмотра бронирования. В конце адреса находится длинный уникальный токен. Сервер распознает его и понимает: предъявителю этой ссылки разрешено показать конкретное бронирование.

Защитная система электронной почты решает проверить ссылку и передает ее внешнему сканеру. Сканер анализирует страницу и затем публикует полный URL в открытой базе. Если другой человек или автоматическая программа обнаружит этот адрес, секретный токен уже находится внутри него.

Получается своеобразная ситуация, когда система безопасности проверяет ключ от двери на надежность, а затем оставляет копию ключа на публичной доске.

Это не означает, что любая длинная ссылка содержит секретные данные. Большинство URL совершенно безопасно публиковать. Опасность возникает прежде всего там, где владение самим адресом автоматически предоставляет определенные права.

К таким сценариям могут относиться ссылки для сброса пароля, беспарольного входа, подтверждения учетной записи, доступа к закрытому документу, просмотра заказа или бронирования. Конкретный риск зависит от устройства сайта и срока действия токена.

Исследователей заинтересовал еще один вопрос: действительно ли кто-нибудь просматривает публично опубликованные адреса или проблема остается исключительно теоретической?

Для проверки были созданы специальные страницы-приманки. Их URL передали сканирующим сервисам, после чего исследователи наблюдали за последующими обращениями к этим адресам.

Оказалось, что опубликованные ссылки действительно привлекали дополнительный трафик. Были зарегистрированы автоматические посетители, обращения к встроенному содержимому и другие действия. Авторы работы подчеркивают, что эти наблюдения сами по себе не доказывают злонамеренного поведения.

Тем не менее эксперимент продемонстрировал принципиально важный факт: публичные потоки URL действительно просматриваются третьими сторонами. Поэтому опубликованный адрес нельзя считать информацией, которую никто не заметит.

Исследование показывает и более широкую проблему современной кибербезопасности. Отдельно каждая технология может работать совершенно правильно. Сайт создает удобную ссылку для быстрого доступа, почтовая система пытается защитить пользователя от фишинга, а сервис сканирования собирает данные для обнаружения новых угроз.

Уязвимость появляется в момент взаимодействия этих систем. Секретная ссылка, безопасная внутри одного процесса, неожиданно попадает в среду, где предполагается публичность информации.

Уменьшить риск технически возможно без отказа от сервисов проверки URL. Исследователи предлагают автоматически распознавать потенциально чувствительные адреса и не публиковать их полностью. Другой вариант — сокращать срок действия токенов и требовать дополнительную проверку личности перед выполнением важных действий.

По оценкам авторов работы, даже после исключения потенциально конфиденциальных ссылок более 99% анализируемых URL по-прежнему могли бы оставаться публичными. Поэтому дополнительная защита практически не уничтожила бы полезность открытых баз для специалистов по кибербезопасности.

Разработчикам сайтов также важно помнить, что секретный URL не всегда остается секретным. Ссылка может пройти через почтовый фильтр, антивирус, корпоративную систему защиты, мессенджер, браузер или другие промежуточные сервисы. Поэтому для действительно чувствительных операций надежнее сочетать токен с дополнительной проверкой пользователя.

Обычному пользователю полностью контролировать эти процессы сложно,

поскольку большая часть проверки происходит автоматически. Тем не менее полезно относиться к персонализированным ссылкам почти так же, как к паролям. Ссылку для восстановления доступа, входа в аккаунт или просмотра закрытого документа не следует без необходимости публиковать, пересылать посторонним или отправлять в случайные онлайн-сервисы для проверки.

Особенно показательно, что обнаруженная проблема появилась внутри инфраструктуры, созданной именно для повышения безопасности. Исследование LEAKYLINKS напоминает: в сложных цифровых системах недостаточно оценивать безопасность каждого компонента отдельно. Необходимо учитывать, что произойдет с информацией после того, как несколько безопасных на первый взгляд технологий начнут обмениваться данными между собой.

Ссылка: «Измерение рисков безопасности и конфиденциальности сервисов сканирования URL-адресов» DOI: 10.1109/sp63933.2026.00130.