## Угрозы и уязвимости

1. На сервере все файловые каталоги открыты для пролистывания извне. Например, http://www.000000.ru/libreq/ или http://www.000000.ru/wstat7/daily/2013/. Это является грубейшим нарушением безопасности, позволяя злоумышленику изучить структуру сервера и обнаружить уязвимые места. В частности, с высокой долей вероятности это обстоятельство сыграло определенную роль в обнаружении и эксплуатации описанных далее уязвимостей. Для устранения ошибки достаточно положить в корень каждого каталога пустой файл index.html. Еще лучше запретить пролистывание каталогов извне на уровне сервера. Еще лучше – и то и другое.
2. Файл robots.txt выполнен по стандартному шаблону, без учета реальной ситуации на сервере. Его возможности по закрытию уязвимостей для SEO-атак не использовапны. Формально он составлен грамотно, без ошибок, однако в отрыве от реальных задач.   
     
   Disallow: /articles/libreq/  
   Disallow: /intdesign/libreq/  
   Disallow: /libreq/  
   Disallow: /interior/libreq/  
   Disallow: /ferplast/libreq/  
   Disallow: /madeaqua/libreq/  
   Disallow: /phprusearch/  
   Disallow: /price/libreq/  
   Disallow: /price/aformprice.php  
   Disallow: /selloff/libreq/  
   Disallow: /catalog.php  
     
   Например, все каталоги /libreq/ действительно надо закрыть от индесации, однако для этого не надо перечислять их все, поскольку при таком громоздком описании резко возрастает вероятность ошибок (и в нашем случае – они возникают). В данном случае достаточно было бы вместо перечисления всех строк написать:   
     
   Disallow: \*libreq  
     
   и возможность использования описанной далее уязвимости была бы перекрыта.  
   Так же в robots.txt не указано огромное количество каталогов, которые реально необходимо закрыть от индексации. Например, каталоги \*wstat7 или \* include. Отсуствие, например, в robots.txt запрета доступа для поисковых роботов к \*wstat7 привело к попаданию сайта в базы SEO-спамеров и в в итоге, к наложению на сайт санкций со стороны поисковиков (см. дальше).
3. Некорректно отрабатываются «лишние» символы в адресной строке, что с одной стороны ведет к возможности для злоумышленника сгенерить на сайте бесконечное число дублей страниц, а с другой стороны придать адресам в адресной строке «криминальный» характер.

Хотя на сайте используются SEF-адреса, не запрещено использование в SEF-адресах переменных для GET-запросов. Это позволяет «дописать» к SEF-адресу любые тексты, которые будут восприняты как php переменные и получить страницы с текстом этого запроса. Примеры:

* + 1. Подстановка в адресную строку якобы переменной «заказать проститутку домой»  
       www.000000.ru/sale.php?заказать%20проститутку%20домой – приведет к формированию страницы информирующей, что товар «заказать проститутку домой» отсуствует на складе. Злоумышленики могут нагененить большое количество таких страниц и загнать их в индекс поисковиков, размыв тематику сайта, а возможно и вызвать наложение санкций, что мы и имеем в нашем случае. Например, в индексе Яндекса находятся страницы сайта с запросами «купить героин» и «Путин козёл»
    2. Если злоумышленику известны имена переменных используемых в коде, он может присвоить им в адресной строке конкретные значения и тем самым заставить php-скрипт совершить то или иное действие. Часто имена переменных подбираются по словарю.

1. На сайте – открытая статистка, которая доступна любому посетителю извне, а так же поисковым роботам. Некоторые страницы статистики представляют собой большую ценность для SEO-спамеров.   
   В частности в статистике есть страница показывающая сайты, с которых посетитель пришел на 000000.ru. Сайты-источники представлены в виде активных ссылок на них, что является большой ценностью в плане получения тИЦ и PR. Любой злоумышленник перейдя с продвигаемого сайта на 000000.ru может дать Яндексу и гуглю ссылку на соответствующую страницу статистики и получить активную ссылку на свой сайт. Поскольку статистика 000000.ru полностью открыта и не защищена, сайт активно используется SEO-спамерами в качестве линкофармы.   
   Статистика однозначно показывает, что в настоящее время 000000.ru включен в спамерские базы для автоматического продвижения и с точки зрения поисковиков является линкофармой.  
   В качестве примера можно рассмотреть страницу статистики, отображающую реферальные сайты за один день, 30 сентября. [http://000000.ru/wstat7/daily/2013/09/30/referrers.html](http://alfaroeco.ru/wstat7/daily/2013/09/30/referrers.html)  
   Из топ 50 рефералов – 30 заходы с поисковиков и мы их исключаем из анализа.  
   Из оставшихся нативных 20 сайтов-источников, только 5 (пять!)  
   <http://mtools.pp.ua/iskusstvennye-rasteniya-dlya-akvariuma.html>  
   <http://www.echomsk.spb.ru/>  
   <http://dfgsdf.ru/foto-tumb-pod-akvarium-svoimi-rukami.html>  
   <http://www.mosaquarium.ru/index.php?t=2&n=5>  
   <http://alfarodecor.ru/price/?PHPSESSID=59efa640f24f024d0b35fc4a50e0b43f>  
   являются честными реферальными сайтами. А 15 оставшихся – «левыми» сайтами, вроде <http://dojki-porno.com/>, имитирующими переход ради получения обратной ссылки.   
   Соответственно можно констатировать что 75 (семьдесят пять!) процентов непоискового трафика состоит из имитации переходов с заведомо криминальных сайтов, продвигающихся черными методами.  
   Таким образом, очевидно, что основная роль, которую в настоящее время играет 000000.ru – он служит линкофармой для наращивания тИЦ и PR, что, безусловно, зафиксированно поисковиками. Справедливости ради надо сказать, что поисковики умеют отличать «умышленные» линкофармы от «неумышленных» и расценивают такие сайты, как 000000.ru, как пострадавшие и подвергшиеся взлому. Это избавляет сайты от бана, но не избавляет от жестких фильтров по выдаче.
2. На сайте применено очень рискованное решение для генерации страниц. Значение переменной str по адресу [http://www.000000.ru/libreq/handler/putstr.php?str](http://www.alfaroeco.ru/libreq/handler/putstr.php?str)= отображается как текст на странице, причем активное содержимое – исполняется, что позволяет зломышленику генерировать на сайте страницы любого произвольного содержания, включая активные обратные ссылки, например, [http://www.000000.ru/libreq/handler/putstr.php?str=%3Ca%20href='www.kremlin.ru'%3EPutin%20Kozel%3C/a%3E](http://www.alfaroeco.ru/libreq/handler/putstr.php?str=%3Ca%20href='www.kremlin.ru'%3EPutin%20Kozel%3C/a%3E). И даже запускать на сайте исполняемые сценарии.

## Рекомендуемые меры

1. Самую серьезную угрозу сейчас представляет открытая статистика, позволяющая использовать сайт в качестве линкофармы. Эта уязвимость прямо сейчас интенсивно используется в реальном времени и её можно отследить ежедневно.  
   Рекомендации: Закрыть паролем всю статистику. Отдавать поисковикам и браузерам 403 код ответа при заходе на статистику (саму статистику при этом отображать). Добавить строку Disallow: \*wstat7 в robots.txt
2. Прописать более общие правила в robots.txt с использованием звездочек, например:  
   Disallow: \*libreq   
   Disallow: \*search/  
   Disallow: \*price  
   Disallow: \*catalog  
   Но гораздо более предпочтительное решение запретить к индексации по умолчанию весь сайт, открыв только нужные для индексации адреса, например:  
   Allow:  
   Allow: /madeaqua  
   Allow: /interior  
   Allow: /intdesign  
   и т.д.
3. Отдавать 404 или 403 код ответа на любой лишний символ в адресной строке в тех разделах сайта, где используется SEF.
4. Использовать метод POST а не GET, везде где отсуствует очевидная необходимость применения GET. В частности для страницы [http://www.000000.ru/libreq/handler/putstr.php](http://www.alfaroeco.ru/libreq/handler/putstr.php), если уж действительно необходимо генерировать содержимое таким образом.
5. Обязательно использовать атрибут rel canonical в заголовках всех страниц.
6. Положить в корень каждого каталога пустой файл index.html. Еще лучше запретить пролистывание каталогов извне на уровне сервера. Еще лучше – и то и другое.