Ochrona witryny internetowej przed cyberatakiem za pomocą proaktywnej detekcji Google

Artur Strzelecki
Uniwersytet Ekonomiczny w Katowicach
2018-12-19

Abstrakt

Artykuł zawiera przegląd obecnych możliwości jakie oferuje firma Google w zakresie proaktywnej ochrony witryn internetowych przed cyberatakiem. W artykule opisano dwa eksperymenty, w którym proaktywna ochrona została zastosowana. Eksperymenty zostały wykonane na rzeczywistych stronach internetowych, działających w internecie.

Słowa kluczowe: ochrona witryn internetowych

Wprowadzenie

W ciągu ostatnich 20 lat światowa sieć internetowa bardzo się rozwinęła. Obecnie dostęp do sieci posiada co druga osoba na świecie. Widoczny wzrost użytkowników obecnych w internecie spowodował również wzrost ilości treści, która jest dla nich dostępna. Treść jest publikowana na witrynach internetowych i w aplikacjach mobilnych zarówno przez samych użytkowników, czyli odbiorców treści oraz przez wydawców treści. Wydawcy treści zazwyczaj prowadzą witryny internetowe, które służą realizacji określonych celów przez wydawcę. Wśród wielu powodów, dla których zakładane i prowadzone są witryny internetowe można wymienić te, które uruchamiają przedsiębiorstwa. Zazwyczaj zawierają treści informujące odbiorców o ofercie i usługach tego przedsiębiorstwa.

Przedsiębiorstwa tworząc witryny internetowe mogą skorzystać z istniejących systemów zarządzania treścią, które ułatwiają stworzenie i prowadzenie witryny internetowej. W marcu 2018 roku, 50,1% wszystkich stron internetowych była opartych o system zarządzania treścią (https://w3techs.com/technologies/overview/content_management/all) W tej liczbie 30,2% korzysta z systemu zarządzania treścią Wordpress, a 3,1% używa systemu zarządzania treścią Joomla. Najpopularniejsze systemy zarządzania treścią tworzone są na zasadach licencji wolnego oprogramowania, gdzie kod źródłowy jest jawnie dostępny, a system rozwijany jest przez grupę programistów i developerów, którzy nie pobierają za to wynagrodzenia. Udostępniane w ten sposób systemy zarządzania są wykorzystywane w prawie połowie witryn internetowych w całym internecie.

Taka duża popularność kilku systemów zarządzania treścią sprawia, że równie łatwo można znaleźć w nich luki bezpieczeństwa, które mogą zostać wykorzystane przez osoby niepowołane. Luki bezpieczeństwa zazwyczaj powstają w modułach dodatkowych tworzonych przez twórców oprogramowania takich jak szablony graficzne lub dodatki funkcjonalne do systemów zarządzania treścią. Wykrycie nowej luki bezpieczeństwa sprawia, że systemy zarządzania treścią, które ją posiadają stają się narażone na cyberatak przez osoby znające tę lukę. Duża skala wykorzystania systemów zarządzania treścią sprawia, że każda kolejna wykryta luka bezpieczeństwa dotyczy wielu milionów stron internetowych, które mogą być narażone na cyberatak. W artykule zostanie dokonana analiza obecnych możliwości jakie oferuje firma Google w zakresie proaktywnej ochrony witryn internetowych przed cyberatakiem oraz zostanie wykonany eksperyment, w którym proaktywna ochrona zostanie zastosowana.

Analiza literatury

Analiza literatury pokazuje, że bezpieczeństwo teleinformatyczne może ma bardzo szeroki zasięg. Cyberprzestrzeń można także określić jako przestrzeń wirtualna. Ta wyodrębniona logicznie (nieistniejąca fizycznie) przestrzeń jest tworzona przez sumę zawartych w systemach danych, plików, stron internetowych, aplikacji oraz procesów, do których uzyskuje się dostęp wyłącznie poprzez systemy teleinformatyczne (Wasilewski, 2013). Jednym z obszarów, który tu został wspomniany to bezpieczeństwo stron internetowych. Jednym przejawów tego rodzaju zagrożenia były protesty wokół umowy ACTA. Włączyła się w nie grupa Anonymous, dokonując serii ataków na witryny internetowe polskich instytucji rządowych (Lakomy, 2013). Po tych wydarzeniach powołany został zespół zadaniowy do spraw ochrony portali rządowych. Opracował on wytyczne w zakresie ochrony portali informacyjnych administracji publicznej zawierające rekomendacje mające na celu podniesienie poziomu bezpieczeństwa stron internetowych oraz poczty elektronicznej, należących do instytucji publicznych. Dotyczą one m.in. zapisów w umowach z zewnętrznymi dostawcami usług prowadzenia stron internetowych w zakresie zabezpieczeń witryn i reagowania na incydenty (Grzelak i Liedel, 2013).

Autorzy (Abramowicz, Bukowska i Filipowska, 2013) stworzyli projekt Semantyczny Monitoring Cyberprzestrzeni. To inicjatywa z zakresu zapewnienia bezpieczeństwa w cyberprzestrzeni. Głównym celem projektu było opracowanie prototypu narzędzia, które pozwoli na ciągłe monitorowanie wybranych przez eksperta źródeł internetowych dla identyfikacji pojawiających się w nich określonych aktywności mogących wskazywać na działania przestępcze. Należy również podkreślić, że wyszukiwarka Google dba o jakość wyników, które prezentuje. Stąd też nie tylko dbałość o bezpieczne wyniki jest istotna, ale także o to aby nie naruszały one cudzych praw autorskich (Strzelecki, 2019).

Inni (Dziembała i Słaboń, 2008) zauważają, że czynniki odpowiedzialne za bezpieczeństwo strony internetowej są brane pod uwagę przy tworzeniu rankingów witryn internetowych. Wskazują na stosowanie bezpiecznego protokołu HTTPS. W kolejnej pracy wskazano, że bezpieczeństwo to jeden z ośmiu kluczowych czynników, które pozwalają zmierzyć wiarygodność handlowych witryn internetowych (Garnik i Basińska, 2011). Zaproponowano także model, na podstawie którego można wyliczyć jak bardzo bezpieczna jest witryna internetowa. model bierze pod uwagę takie elementy jak nazwa serwera, kraj domeny, długość rejestracji domeny, ranking domeny i typ witryny (Kim i in. 2008). Ochrona witryn internetowych silnie jest też związana Corporate Social Responsibility. Jeśli firmy mają być odpowiedzialne za treści publikowane na swoich witrynach internetowych, to muszą także edukować klientów czego nie należy robić w sieci, aby nie narazić się na atak hakerski (Tarabasz, 2017).

Systematyzacja

Luki bezpieczeństwa w witrynach internetowych mogą być wykorzystywane przez osoby kierujące się różnymi motywami, stąd nie zawsze łatwo je wykryć.

Spam. W przypadkach witryn internetowych, które padły ofiarą ataku hakerów, dystrybucja spamu polega na wyświetlaniu linków lub reklam. Linki i reklamy zazwyczaj prowadzą do pośredniczącej witryny internetowej, której odwiedzenie zainstaluje w przeglądarce internetowej użytkownika pliki cookie. Pliki cookie w ten sposób zainstalowane zazwyczaj zawierają informację o programie sprzedaży partnerskiej w jednym z dużych sklepów internetowych takich jak Amazon lub AliExpress. Osoba, która by kliknęła w link lub reklamę, a następnie zakupiła produkty w sklepie, z którego pochodzą pliki cookie programu sprzedaży partnerskiej, zapłaciłaby całą kwotę za towar w sklepie, natomiast sklep wypłaciłby prowizję, za tak wygenerowaną sprzedaż partnerowi w programie sprzedaży partnerskiej, którego identyfikator znajduje się w pliku cookie.

Malware. Witryny internetowe, których odwiedzenie może zagrozić bezpieczeństwu użytkownika i programów komputerowych, mogą zainstalować złośliwe oprogramowanie na komputerze lub urządzeniu mobilnym użytkownika.

Niechciane oprogramowanie. Witryny internetowe, których odwiedzenie spowoduje zainstalowanie oprogramowania, którego użytkownik urządzenia końcowego nie zamierzał instalować.

Phising. Shakowana witryna internetowa może sprawiać wrażenie prawidłowo działającej, natomiast pozostawiane w niej przez użytkowników dane i informacje, mogą być przesyłane do nieuprawnionych odbiorców. To wyłudzanie informacji o odbiorcach treści (Król, 2015).

Przekierowanie. Witryna internetowa, która przekierowuje użytkowników bez ostrzeżenia w niechciane przez nich miejsce, mogła paść ofiarą ukrytego przekierowania. Hakerzy wykorzystują rzeczywiste strony, aby przekierować użytkowników w inne miejsce, bez ich wiedzy i zgody. Przekierowanie może mieć miejsce tylko na urządzeniach mobilnych, a pozostawać bez wpływu na użytkowników komputerów. Użytkownik urządzenia mobilnego zazwyczaj jest nieświadomy, że został przekierowany w inne miejsce.

Proaktywna ochrona. Firma Google wypracowała dwa procesy informujące zarówno użytkowników jak i właścicieli stron internetowych o potencjalnym zagrożeniu. Pierwszy proces polega na wykrywaniu stron i oznaczeniu ich komunikatem: **Ta witryna może wyrządzić szkody na Twoim komputerze**. Ten komunikat wprowadzono do wyszukiwarki w 2009 roku (Szymanski 2009)

Działanie tego ostrzeżenia polega na tym, że po kliknięciu w aktywny link do strony, która posiada takie ostrzeżenie, użytkownik nie jest odsyłany od tej strony, ale trafia na stronę z kolejnym ostrzeżeniem. Najprawdopodobniej witryna lub serwer internetowy przesyłający stronę, posiada lukę w zabezpieczeniach spowodowaną używaniem nieaktualnego oprogramowania.

Wykorzystano lukę bezpieczeństwa i niebezpieczny kod został dodany do strony, która wcześniej była bezpieczna. Otwarcie zainfekowanej strony internetowej może spowodować wykonanie ukrytego skryptu lub otwarcie kolejnej strony pobierającej treść z innej witryny, która usiłuje zaatakować komputer użytkownika. Konsekwencją udanego ataku mogą być zainstalowane bez wiedzy lub świadomości użytkownika wirusy, programy szpiegowskie lub keyloggery. Szkodliwe oprogramowanie może posłużyć do kradzieży haseł i numerów kart kredytowych, spowalniać komputer lub zmieniać wyniki wyszukiwania.

Widząc w wynikach ostrzeżenie, że witryna może wyrządzić szkody na komputerze, należy być ostrożnym. Rozpoznanie szkodliwego kodu na stronie internetowej przez wyszukiwarkę Google jest algorytmiczne. Skanery są dosyć precyzyjne. Jeżeli witryna w wynikach jest opatrzona ostrzeżeniem o szkodliwym oprogramowaniu, jest bardzo prawdopodobne, że padła ofiarą ataku hakerskiego, nawet jeżeli jej właścicielem jest godna zaufania osoba lub instytucja. Skaner wyszukiwarki decyduje bezstronnie na podstawie potencjalnie groźnego kodu wykrytego w witrynie internetowej.

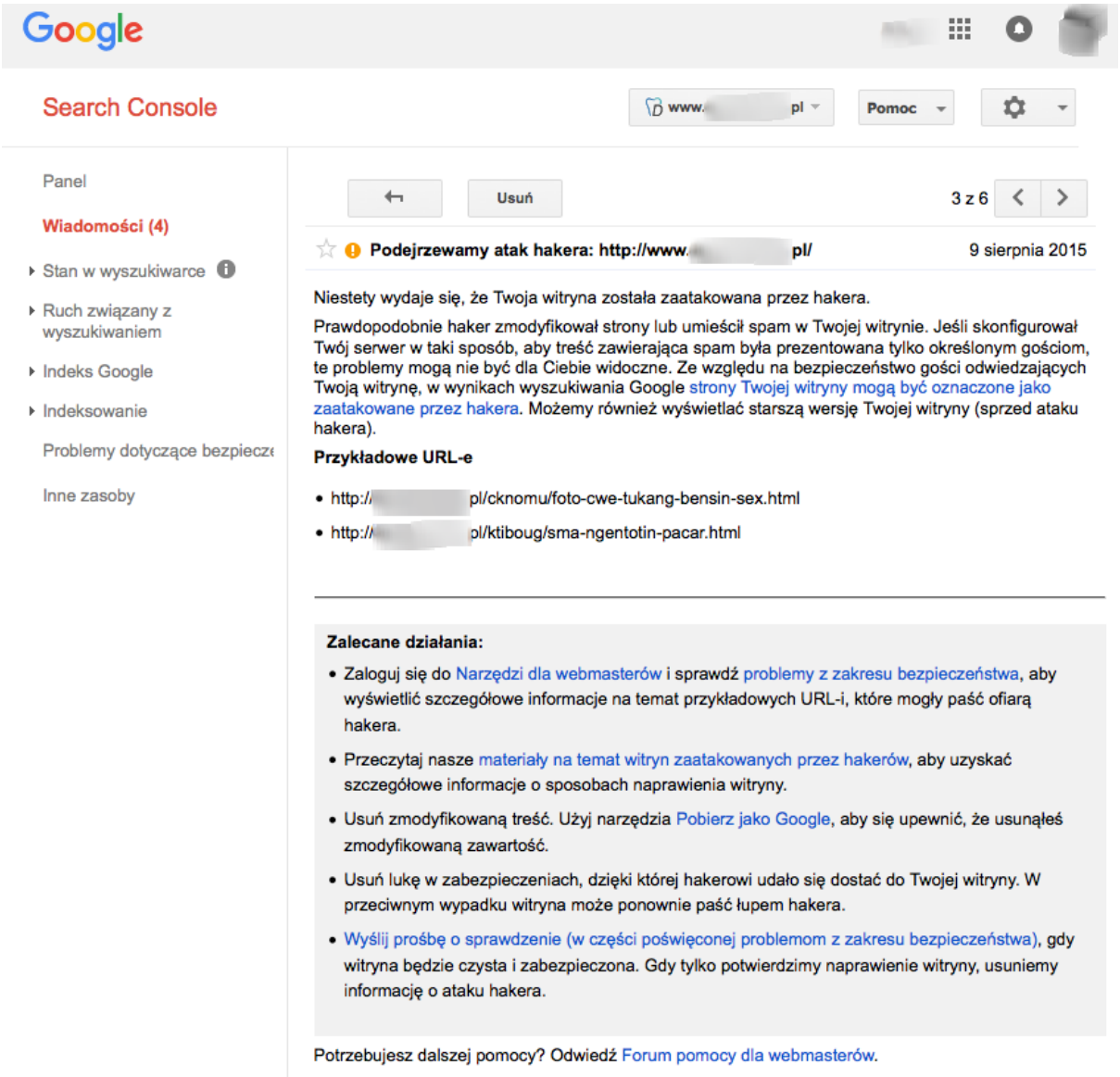
Drugi proces polega na wykrywaniu stron i oznaczeniu ich komunikatem: **Ta witryna mogła paść ofiarą ataku hakerów**. Ten komunikat został wprowadzony do wyszukiwarki w 2011 roku (Wald, 2011). Komunikat ma za zadanie pomagać użytkownikom unikać witryn, które mogły zostać zmodyfikowane w czasie ataku hakerów, którego celem zazwyczaj jest rozpowszechniania spamu. Celem wyszukiwarki jest upewnienie się, że podczas odwiedzania witryny, użytkownik ma pewność, że dostępne tam informacje pochodzą od pierwotnego wydawcy a nie od hakera. Kliknięcie w ten komunikat przeniesie użytkownika do artykułu, który zawiera więcej informacji o tym co się stało. Natomiast kliknięcie rezultatu w wynikach wyszukiwania spowoduje odwiedzenie docelowej strony.

Eksperyment

Dwa eksperymenty przeprowadzono w zamkniętym środowisku. Na czas eksperymentów autor otrzymał prawa administracyjne do zarządzania witryną internetową. Witryna internetowa został oznaczona komunikatem "**Ta witryna mogła paść ofiarą ataku hakerów**".

Jak już zostało to wcześniej przedstawione, wyszukiwarka Google dzięki własnym narzędziom wykrywa ataki hakerskie na stronie internetowej. Kiedy wykryje coś podejrzanego, publikuje powiadomienie w wynikach wyszukiwania. Wraz z publikacją ostrzegawczego komunikatu zostaje automatycznie nawiązana próba kontaktu z webmasterem witryny. Pierwszy kontakt jest za pośrednictwem usługi Google Search Console. Usługa Search Console udostępnia dane, narzędzia i dane diagnostyczne potrzebne do stworzenia i utrzymania witryny oraz aplikacji mobilnych dostosowanych do Google. Jeśli webmaster witryny posiada konto w tej usłudze i wcześniej dodał stronę którą administruje do tej usługi, to w panelu administracyjnym tej usługi będzie można zobaczyć ostrzegawczy komunikat. Ponadto, na adres email, który służy do korzystania z tej usługi zostanie wysłany analogiczny komunikat.

W pierwszym eksperymencie po uzyskaniu dostępu do usługi Search Console, wyświetlany był poniższy komunikat, przedstawiony na Rysunku 1.



Rys. 1. Podejrzewamy atak hakera
Źródło: Google Search Console

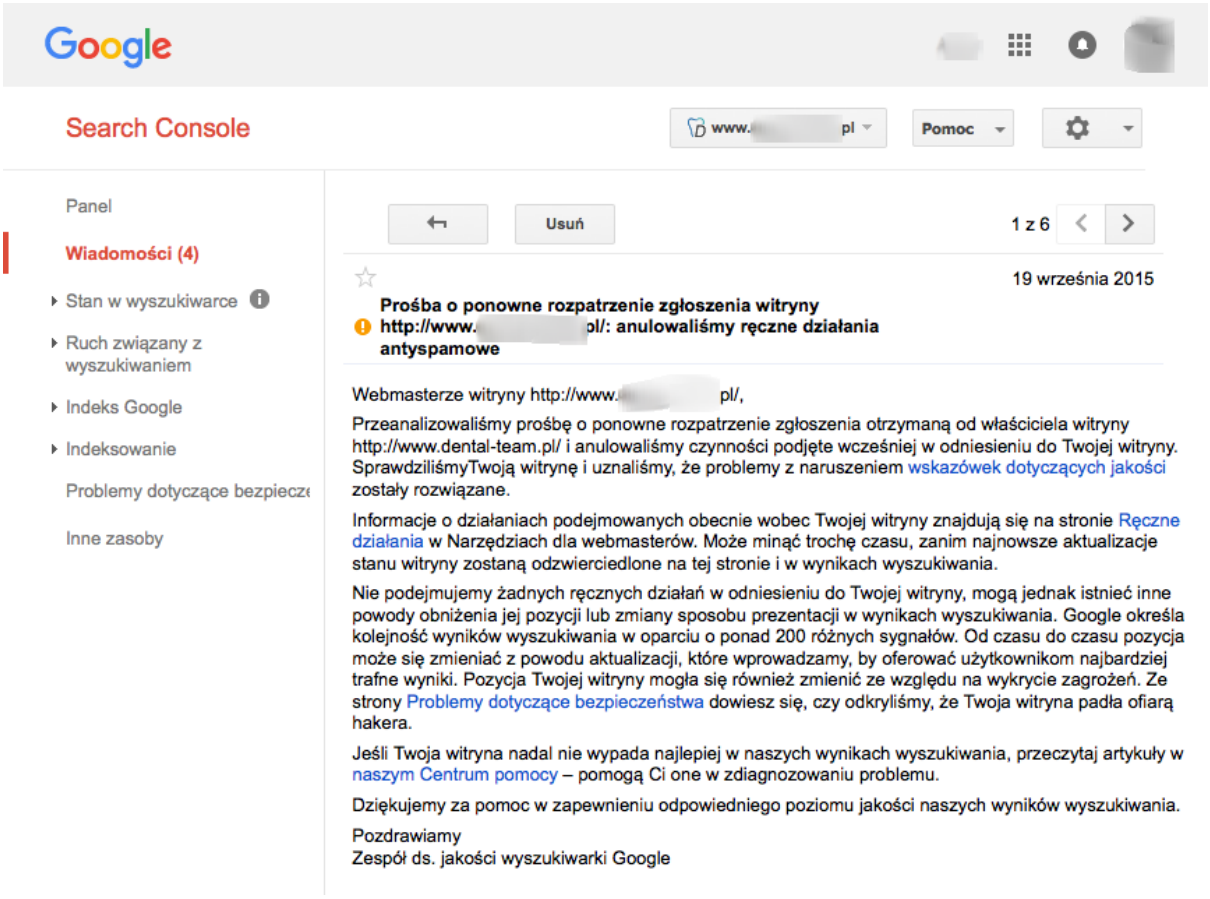
Powiadomienie o treści "Ta witryna mogła paść ofiarą ataku hakerów" ma negatywny wpływ na liczbę osób odwiedzających witrynę internetową bezpośrednio z wyników wyszukiwania. Jednak po naprawieniu luki

bezpieczeństwa, komunikat z ostrzeżeniem zostaje automatycznie usunięty z wyników wyszukiwania – zwykle w ciągu 24 godzin.

Powiadomienie nie zniknie samoistnie, dopóki właściciel witryny internetowej nie wykona odpowiednich działań. Aby naprawić stronę, należy wykonać następujące czynności:

1. Rejestracja i weryfikacja strony internetowej w usłudze Google Search Console – jeśli jeszcze konto nie jest utworzone.
2. Po uzyskaniu dostępu do usługi Search Console, w sekcji „Problemy dotyczące bezpieczeństwa" należy sprawdzić podane przykładowe adresy stron, które mogły zostać zaatakowane. Webmaster witryny internetowej na tej podstawie może zlokalizować zagrożenie i usunąć błąd w zabezpieczeniach, który umożliwił zainfekowanie witryny. Jeśli luka bezpieczeństwa nie zostanie usunięta, problem może się powtórzyć. Jeśli webmaster witryny korzysta z bezpłatnego systemu zarządzania treścią CMS jak WordPress lub Joomla, należy zaktualizować go do najnowszej wersji. Tak samo należy zaktualizować wszystkie dodatkowe wtyczki do tych skryptów.
3. Osoby posiadające pewną wiedzę informatyczną mogą same spróbować oczyścić swoją stronę internetową. Firma Google przygotowała poradnik (https://developers.google.com/web/fundamentals/security/hacked/) z którego można się dowiedzieć szczegółowo, jak naprawić stronę.
4. Po usunięciu problemu i zabezpieczeniu strony internetowej należy poprosić o ponowne jej sprawdzenie w sekcji *Problemy dotyczące bezpieczeństwa* w Search Console. Gdy Google upewni się, że witryna została naprawiona, usunie ostrzegawczy komunikat „Ta witryna mogła paść ofiarą ataku hakerów".

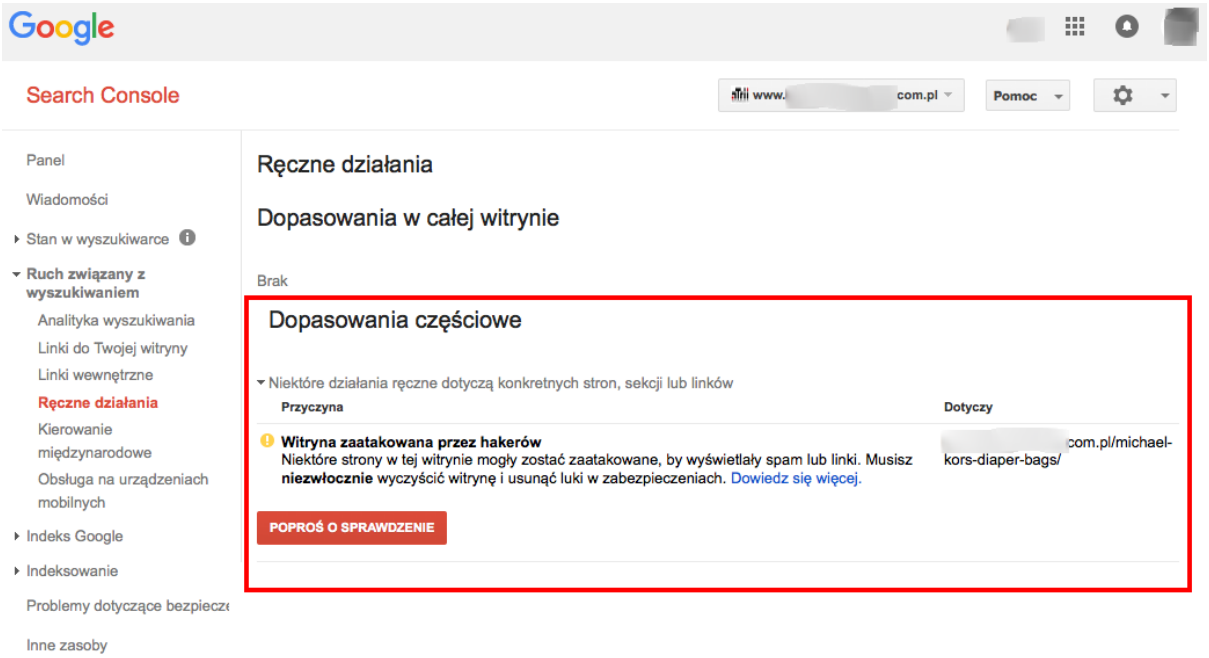
Po wykonaniu powyższych kroków pojawiła się informacja, że ręczne działania antyspamowe wcześniej nałożone zostały zdjęte, co zostało przedstawione na Rysunku 2. Na tym eksperyment został zakończony.



Rys. 2. Prośba o ponowne rozpatrzenie zgłoszenia witryny
Źródło: Google Search Console

Mogą wystąpić również takie okoliczności gdzie w wynikach wyszukiwania jest już wyświetlany ostrzegawczy komunikat, ale nie pojawiła się jeszcze o nim informacja w usłudze Search Console. Jeśli luka bezpieczeństwa zostanie naprawiona w witrynie internetowej zanim ten komunikat pojawi się w Search Console, to aby sprawić by zniknął z wyników wyszukiwania, trzeba w tej usłudze wejść do zakładki "Indeksowanie / Pobierz jako Google" i pobrać stronę główną, a następnie za pomocą przycisku Prześlij do indeksu, wysłać ją ponownie do sprawdzenia przez Google. Wtedy wyszukiwarka otrzyma sygnał aby sprawdzić ponownie zawartość witryny internetowej, już bez istniejącej luki bezpieczeństwa.

W drugim eksperymencie zaobserwowano bardziej złożony ataki hakerski na witrynę internetową, który również jest trudniejszy do usunięcia. Wtedy też w Google Search Console trzeba wykonać więcej czynności. Po pierwsze różnica wynika z tego, że na witrynę została nałożona kara ręczna pt „Witryna zaatakowana przez hakerów" i wtedy prośbę o ponowne rozpatrzenie składa się z poziomu zakładki Ręczne działania, co ilustruje Rysunek 3.



Rys. 3. Ręczne działania
Źródło: Google Search Console

Na Rysunku 3 zobrazowano działanie ręczne w dopasowaniu częściowym „Witryna zaatakowana przez hakerów„. Oznacza to, że część witryny internetowej została zaatakowana przez hakera, ale atak trwa już na tyle długo, że stał się szkodliwy dla osób odwiedzających witrynę. Najwyraźniej w takiej sytuacji nie wszystkie problemy z bezpieczeństwem zostały rozwiązane. Jeśli w zakładce działania ręczne jest przyznana kara, to istnieje możliwość odwołania się od niej. Wtedy trzeba skorzystać z tej opcji i przesłać do Google wniosek o ponowne rozpatrzenie. We wniosku należy napisać jak zostały rozwiązane wskazane przez Google problemy i podać wszystkie istotne informacje na ten temat. Jeśli wszystkie problemy z bezpieczeństwem faktycznie zostały rozwiązane, kara zostanie zdjęta. W drugim eksperymencie, po zlikwidowaniu wszystkich luk bezpieczeństwa witryna nie była już traktowana jako zaatakowana przez hakera.
W trakcie realizacji tych dwóch eksperymentów skorzystano z następujących narzędzi:

1. Google posiada swoje własne narzędzie, które informuje czy strona jest bezpieczna lub zaatakowana przez hakerów. Narzędzie dostępne pod adresem: https://transparencyreport.google.com/safe-browsing/search
2. Skorzystano także z narzędzia https://www.stopbadware.org/ z którym Google współpracuje.
3. Popularnym narzędziem jest również https://sitecheck.sucuri.net/, jednak zdarza się, że nie wykrywa ono bardziej wyrafinowanych ataków hakerskich, np. tych infekcji które sprawdzają pole User-Agent i Referer w nagłówku HTTP i działają tylko w określonych konfiguracjach.

Dyskusja

Z przeprowadzone eksperymentu wynika, że kluczowe jest posiadanie aktualnej, ostatniej wersji systemu zarządzania treścią. Wynika to z tego, że hakerzy w ogromnej części ataków używają automatycznych skryptów i programów, natomiast nie dokonują ataków ręcznie. Programy skanują setki tysięcy stron, których adresy pobierane są często automatycznie z ogólnodostępnych baz stron internetowych, firm i innych miejsc. Mogą także, na podstawie cyfrowych odcisków automatycznie wyszukiwać strony internetowe działające na Joomla! lub WordPress bezpośrednio w wynikach wyszukiwania Google. Posiadając adres odnalezionej strony w preferowanym, nieaktualnym systemie zarządzania treścią, programy korzystają ze zbudowanej przez hakera bazy luk dla danej wersji systemu. Zatem automatycznie atakują newralgiczne, znane hakerom punkty starych wersji systemów zarządzania treścią i w ten sposób infekują daną stronę internetową.

Problem ochrony strony internetowej wydaje się być także rozwojowy, ponieważ tych ataków jest coraz więcej. Spora część z nich to automatyczne próby typu „brute force", które dodatkowo mają obciążyć infrastrukturę serwerową przechowującą i udostępniającą stronę internetową. Takie próby mogą wysyłać tysiące zapytań do bazy danych w krótkim czasie. Ważne działania proaktywne w zakresie ochrony strony internetowej to posiadanie aktualnego systemu zarządzania treścią oraz wtyczek uzupełniających funkcje systemu. Dobrą praktyką jest instalowanie wtyczki tylko wtedy, gdy jest już ona niezbędna i będzie stale wykorzystywana. Kolejnym elementem jest zastosowanie silnego hasła i unikalnej nazwy użytkownika. Dobra nazwa użytkownika nie zawiera popularnych ciągów typu „admin", „test" lub „nazwa domeny" do której dostęp ma zostać uzyskany. Nie należy także przechowywać tych danych w przeglądarce internetowej lub w kliencie FTP. Natomiast jeśli, mimo tych starań stron została zaatakowana przez hakera to niezbędne jest także wykonywanie regularnych kopii zapasowych.

Podsumowanie

Obecnie rośnie udział systemów zarządzania treścią przy tworzeniu witryn internetowych. Im więcej witryn internetowych będzie opartych o powszechnie znane i dostępne rozwiązanie, tym więcej pojawi się zagrożeń ze strony hakerów, którzy będą chcieli wykorzystać te witryn w nieuprawniony sposób do swoich celów. Dlatego tak ważne jest aby korzystać z dostępnych narzędzi, które pomagają utrzymać witrynę internetową w bezpiecznej kondycji.

Literatura

1. Abramowicz W., Bukowska E., & Filipowska A., (2013). Zapewnienie bezpieczeństwa przez semantyczne monitorowanie cyberprzestrzeni. *e-mentor*, (3(50)), 11-17.
2. Dziembała M., Słaboń M., (2008). Wybrane elementy oceny witryn internetowych. *Prace naukowe, Akademia Ekonomiczna w Katowicach*.
3. Garnik I., Basińska B., (2011). Pomiar wiarygodności internetowych serwisów handlowych. *Zeszyty Naukowe Politechniki Poznańskiej, Organizacja i Zarządzanie*, 56, 23-34.
4. Grzelak M., Liedel K., (2012). Bezpieczeństwo w cyberprzestrzeni. Zagrożenia i wyzwania dla Polski– zarys problemu. *Bezpieczeństwo narodowe*, 22, 125-139.
5. Kim, Y. G., Cho, S., Lee, J. S., Lee, M. S., Kim, I. H., & Kim, S. H. (2008). Method for evaluating the security risk of a website against phishing attacks. *In International Conference on Intelligence and Security Informatics*, 21-31, Springer, Berlin, Heidelberg.
6. Król, K. (2015). Organizacyjne aspekty zarządzania bezpieczeństwem danych z perspektywy zagrożeń phishingu. *Organizacja i Zarządzanie*, 2(30), 19-32.
7. Lakomy M., (2013). Zagrożenia dla bezpieczeństwa teleinformatycznego państw–przyczynek do typologii. *Kwartalnik Naukowy OAP UW" e-Politikon",* 6, 100 – 141.
8. Strzelecki, A. (2019). Website removal from search engines due to copyright violation. *Aslib Journal of Information Management*, 71(1), 54-71.

9. Szymanski K., (2009). *Wyniki wyszukiwania z ostrzeżeniem*, Blog Google, dostępne https://polska.googleblog.com/2009/05/wyniki-wyszukiwania-z-ostrzezeniem.html
10. Tarabasz A. (2017). The role of CSR in educating consumer on cybersecurity. Comparative analysis of examples from UAE and Poland. *In International Conference "Responsible Organizations in the Global Context",* Washington.
11. Wald G., (2011), *Nowe powiadomienia w wynikach wyszukiwania dotyczące witryn zaatakowanych przez hakerów*, Blog Google, dostępne https://polska.googleblog.com/2011/02/nowe-powiadomienia-w-wynikach.html
12. Wasilewski J., (2013). Zarys definicji cyberprzestrzeni. *Przegląd Bezpieczeństwa Wewnętrznego*. 5(9), 225-234.