MECHANIZM SESJI

Waldemar Korłub

Wytwarzanie Aplikacji Internetowych
KASK ETI Politechnika Gdańska

Protokół HTTP a potrzeby witryn internetowych

- Protokół HTTP jest bezstanowy
  - Niezależne żądania – serwer otrzymuje zapytanie, generuje odpowiedź i zamyka połączenie
    - $ telnet 192.168.166.20 80
  - Brak powiązania pomiędzy kolejnymi zapytaniami
  - Każde zapytanie musi zawierać komplet informacji, koniecznych do jego przetworzenia
- Potrzeby witryn internetowych
  - Zapamiętanie wyborów użytkownika na jednej podstronie, w celu zrealizowania funkcjonalności na innej
    - Koszyk produktów
    - Składanie zamówienia

HTTP Cookies?

- Zapisywanie informacji *po stronie przeglądarki*
  - Dane przesyłane do klienta w celu zapisania
  - Potem dołączane do *każdego* kolejnego żądania
- Nie nadaje się do zapisywania informacji wrażliwych
  - Łatwość odczytania przez napastnika
    - np. numer karty kredytowej użytkownika
- Nie można ufać wartościom *cookies*, które serwer otrzymuje od klienta
  - …ani żadnym innym wartościom w żądaniu HTTP
  - Łatwość modyfikacji w przeglądarce
  - Łatwość spreparowania złośliwego żądania

Po stronie serwera

- Bezpieczeństwo, kontrola nad danymi
- Wiele możliwości
  - Pliki na dysku, XML
  - Baza danych
  - Pamięć operacyjna (szybki dostęp, np. Memcached)

- Ale zanim cokolwiek zapiszemy/odczytamy:
  jak rozpoznać różnych klientów aplikacji?

Wielodostęp w aplikacjach internetowych

- Sklep internetowy obsługuje wielu klientów równocześnie
- Każdy klient sklepu ma własny koszyk
- Aby klienci mogli złożyć swoje zamówienia, każdy koszyk należy zapamiętać
  - Po stronie serwera
- Gdy klient dodaje kolejny produkt lub klika „Zamów”, **który koszyk wybrać?**

Jak rozpoznać, że kolejne zapytania HTTP pochodzą od tego samego klienta?

Istota mechanizmu sesji

Sesja

- Sesja
  - Ciąg kolejnych zapytań HTTP, wysyłanych przez tego samego klienta
- Stan sesji
  - Dane przechowywane pomiędzy kolejnymi zapytaniami HTTP, składającymi się na sesję

Identyfikator sesji

- Każda sesja posiada unikalny identyfikator
- Identyfikator jest przekazywany do klienta (przeglądarki) w momencie otwarcia sesji
- Każde żądanie HTTP w ramach sesji musi zawierać jej identyfikator
  - Przeglądarka dołącza identyfikator do kolejnych żądań HTTP, gdy użytkownik nawiguje po stronie
- Serwer, otrzymując żądanie z dołączonym identyfikatorem, wyszukuje właściwą sesję i udostępnia jej stan (dane) na czas obsługi żądania

Identyfikator sesji

- Cechy dobrego identyfikatora sesji:
  - Długi
  - Losowy
  - Generowany nieliniowo
    - Brak możliwości odgadnięcia kolejnego identyfikatora na podstawie aktualnego lub kilku poprzednich
- Trudny do odgadnięcia!
  - Ile warty jest identyfikator sesji?

Przekazywanie identyfikatora sesji

- Podstawowe metody przekazywania danych w protokole HTTP:
  - Parametry żądania
    - GET – dołączone do URI
    - POST – w ciele zapytania
    - Zawsze obsługiwane
  - *Cookies*
    - Dołączane przez przeglądarkę do każdego zapytania
    - Mogą zostać wyłączone

Identyfikator w parametrze GET

- Niebezpieczeństwo ujawnienia identyfikatora
  - Przesłanie linka (łącznie z identyfikatorem sesji) do innej osoby
  - Zapisanie adresu w historii przeglądarki lub w zakładkach (*bookmarks*)
    - Wchodzenie na stronę zawsze z tym samym id sesji
  - Ujawnienie identyfikatora w nagłówku *Referer*
- Mechanizm używany w ostateczności, jeśli pliki *cookies* nie są dostępne
  - …lub całkowicie wyłączony w serwisach operujących na wrażliwych informacjach

Długość sesji

- Przy braku aktywności sesja powinna zostać zakończona
  - Z punktu widzenia bezpieczeństwa
    - Przechwycony identyfikator staje się bezużyteczny, gdy sesja wygaśnie
    - Zabezpieczenie dla użytkowników, który zapomnieli się wylogować na publicznie dostępnym komputerze
  - Z punktu widzenia zasobów
    - Zwolnienie zasobów zajmowanych przez sesję użytkownika (np. pamięć operacyjna, pamięć dyskowa)
- Czas trwania sesji określony w plikach konfiguracyjnych serwera bądź aplikacji

Ataki na mechanizm sesji

→ *wykład na temat bezpieczeństwa*