# Bezpečnosť webovských aplikácií (1. časť)

Richard Ostertág

Katedra informatiky
FMFI UK, Bratislava
ostertag@dcs.fmph.uniba.sk

2011/12

# Éra web sídiel (web sites – Web 1.0)

- statické dokumenty
- prehliadač slúžil iba na stiahnutie a zobrazenie týchto dokumentov
- jednosmerný tok „zaujímavej“ informácie (server $\rightarrow$ prehliadač)
- väčšinou bez autentifikácie používateľov
- každý používateľ: rovnaké zaobchádzanie, rovnaký obsah
- používatelia nevytvárajú prezentovaný obsah
- bezpečnostné hrozby využívali hlavne zraniteľnosti v softvéri na serveri
- kompromitovaný server
  - neznamenal únik citlivých informácií (všetky boli aj tak na webe)
  - modifikácia obsahu
  - využitie úložného priestoru a konektivity servera na šírenie warez-u

# Éra web aplikácií (web applications – Web 2.0)

- interaktívne dynamické aplikácie
- prehliadač sa stáva operačným systémom, v ktorom beží web aplikácia
- obojsmerný tok „zaujímavej“ informácie (server $\leftrightarrow$ prehliadač)
- obvykle existuje autentifikácia používateľov (registrácia, prihlásenie)
- každý používateľ: rôzne zaobchádzanie, iný prispôsobený obsah
- používatelia vytvárajú prezentovaný obsah
- bezpečnostné hrozby využívajú aj zraniteľnosti vo web aplikácii
- kompromitovaný server / web aplikácia
  - únik citlivých informácií (osobné údaje, čísla kreditných kariet)
  - modifikácia obsahu (defraudácia peňazí, útoky na iných používateľov)
  - využitie prenosovej šírky pásma, výpočtového výkonu alebo úložného priestoru servera (vytváranie botnetov pre rozosielanie spamu alebo DDoS útoky)

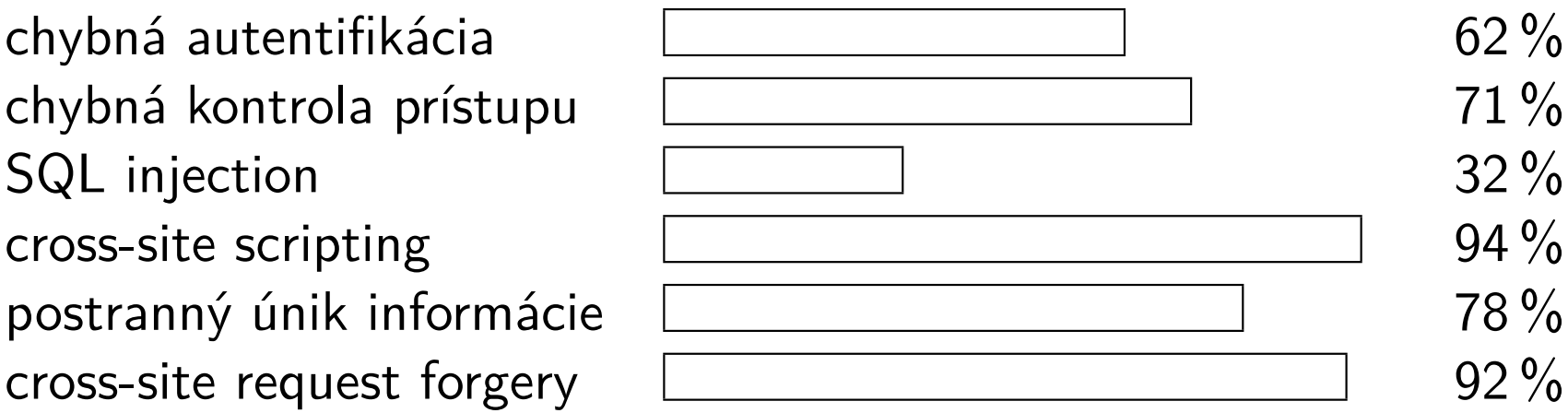
# Funkcie internetových web aplikácií

- Internetové obchodné domy
  - Amazon
  - Hej
  - Alza
- Sociálne siete
  - Facebook
  - Twitter
  - Second life
- Internet banking
  - Tatrabanka
  - VÚB
  - Slovenská sporiteľna
- Internetové vyhľadávače
  - Google
  - Bing
  - Baidu
- Internetové aukcie
  - eBay
  - Aukro
- Internetové kasína
  - bwin
- Web logs
  - Blogger
- Web mail
  - Gmail
  - Hotmail
- Internetové médiá
  - CNN
  - SME
- Internetové encyklopédie
  - Wikipédia

# Funkcie intranetových web aplikácií

- Aplikácia pre oddelenie ľudských zdrojov
  - platové podmienky
  - nábor nových zamestnancov
  - disciplinárne procedúry
- Administratívne rozhrania kľúčovej infraštruktúry
  - web server
  - mail server
  - server pre virtuálne stroje
- Softvér pre spoluprácu v tíme
  - zdieľanie dokumentov
  - správa riadenia projektov
- Obchodné aplikácie
  - Enterprise Resource Planning
  - Customer Rel. Mangement
- Softvérové služby
  - e-mailový klient
- Tradičné desktopové aplikácie
  - textový procesor
  - tabuľkový procesor

# Bežné zraniteľnosti web aplikácií




Zdroj: Dafydd Stuttard, Marcus Pinto: The Web Application Hacker's Handbook

# SSL, PCI – falošný pocit bezpečia

- používatelia sú si vedomí bezpečnostných nedostatkov web aplikácií
- veľa web aplikácií tvrdí, že sú bezpečné, lebo používajú SSL[1]:
  *This site is absolutely secure. It has been designed to use 128-bit Secure Socket Layer (SSL) technology to prevent unauthorized users from viewing any of your information. You may use this site with peace of mind that your data is safe with us.*
- organizácie tiež zvyknú uvádzať, že vyhovujú Payment Card Industry (PCI) štandardom:
  *We take security very seriously. Our web site is scanned daily to ensure that we remain PCI compliant and safe from hackers.*
  - používajte a udržiavajte firewall
  - nepoužívajte výrobcom predvolené heslá
  - šifrujte dáta prenášané cez verejné siete
  - udržiavať bezpečnostnú politiku
  - ...

[1] Zdroj: Dafydd Stuttard, Marcus Pinto: The Web Application Hacker's Handbook

# Základný bezpečnostný problém

- útočník môže odoslať akúkoľvek odpoveď
  - bezpečnostné kontroly vykonávané na strane klienta je možné obísť
  - pre prístup k aplikácii sa nemusí používať webový prehliadač
  - útočník (na strane klienta) môže:
    - čítať, pozmeňovať, odstraňovať alebo opakovane používať všetky dáta, ktoré dostane od serveru alebo web-prehliadač posiela na server
    - generovať nové dáta a vkladať ich do komunikácie
    - zmanipulovať (potlačiť, modifikovať, opakovať) akékoľvek operácie vykonávané na strane klienta
  - útočník môže manipulovať so všetkými dátami, s ktorými klient pracuje:
    - URL (časti cesty, GET parametre)
    - polia formulárov (aj skryté)
    - zakomentované časti HTML dokumentov
    - skripty (prezerať, meniť)
    - cookies a iné údaje v hlavičkách HTTP (napr. session ID, User-Agent)
- je nutné predpokladať, že všetky vstupy sú potenciálne nebezpečné

# Kľúčové faktory základného bezpečnostného problému

- nedostatočné bezpečnostné povedomie
  - problematika zabezpečenia webových aplikácií je mladá (vs. siete, OS)
- vlastný vývoj
  - veľa webových aplikácií je vyvinutých zamestnancami rôznej kvalifikácie
  - každá aplikácia je iná a môže obsahovať svoje vlastné jedinečné vady
- zavádzajúca jednoduchosť
  - začínajúci programátor môže vytvoriť funkčnú webovú aplikáciu od nuly
  - produkovať funkčný kód a produkovať bezpečný kód je rozdiel
- nové hrozby pre web aplikácie pribúdajú rýchlejším tempom ako pre staršie technológie
- obmedzené zdroje (čas, peniaze, ľudia, . . . )
  - funkčnosť aplikácie má prednosť pred jej bezpečnosťou
- „zneužitie“ technológií
  - mnohé zo základných technológií vo webových aplikáciách sú využívané ďaleko za pôvodne účely, čo vedie k ich bezpečnostnej zraniteľnosti

# Základné obranné mechanizmy

- kontrola prístupu používateľov
  - zabrániť používateľom v získaní neoprávneného prístupu
- kontrola používateľských vstupov
  - zabrániť nežiaducemu správaniu aplikácie pri „zákernom“ vstupe
- zvládnutie útoku
  - korektná funkčnosť aj v prípade priameho útoku
  - obranné a útočné opatrenia pre odrazenie útoku
- monitorovanie aplikácie
  - správca môže okamžite zasiahnuť

# Kontrola prístupu používateľov

- základné komponenty:
  - identifikácia a autentifikácia
  - správa relácií (session management)
    web aplikácia vydá užívateľovi token identifikujúci jeho reláciu
  - kontrola prístupu
- chyba v ktorejkoľvek komponente môže viesť k získaniu neoprávneného prístupu

# Identifikácia a autentifikácia

- identifikácia
  - najčastejšie pomocou prihlasovacieho mena
- autentifikácia
  - najčastejšie pomocou hesla
- utočník môže
  - získať prihlasovacie mená
  - získať heslá
  - obísť prihlasovaciu funkciu
    - vďaka chybe v návrhu

# Identifikácia a autentifikácia (klasické problémy)

- slabé heslá (krátke, malá abeceda, uhádnuteľné)
  - krátke
  - malá abeceda
  - slovníkové
  - uhádnuteľné (dátum narodenia)
    - kontrola kvality (napr. aj slovníkový útok)
- rovnaké heslá do rôznych systémov
- neobmieňanie hesiel
  - kontrola veku hesla
- keylogger
  - virtuálna klávesnica
  - jednorázové heslá
  - mechanizmy výzva-odpoveď (zero knowledge proofs)
- neskorý prechod z HTTP na HTTPS
- pri neúspešnom prihlásení nerozlišovať zlé meno od zlého hesla

# Správa relácií – session hijacking[2]

- prezradenie identifikátoru relácie
  - odpočúvaním komunikácie
  - hlavička Referer pri prechode na iné stránky
  - história prehliadača
  - prístup k používateľovmu počítaču a prečítanie uložených cookies, ...
- uhádnutie identifikátoru relácie
  - identifikátor je generovaný predikovateľným spôsobom (napríklad jednoduchá aritmetická postupnosť)
  - identifikátor má malý rozsah možných hodnôt
- možná ochrana:
  - ochrana nie je jednoduchá a 100 % účinná
  - identifikátor nie je len náhodne vygenerované číslo
  - má k sebe pripojený aj hash IP adresy serveru a klienta, User-agent hlavičky klienta a nejakej tajnej hodnoty
  - útočník potom nemôže jednoducho použiť ukradnutý identifikátor

[2] útočník zistí ID inej relácie a cezeň do nej vstúpi a pracuje pod inou identitou

# Správa relácií – session riding

- session hijacking vyžaduje, aby útočník ukradol/uhádol session ID
- session riding nevyžaduje od útočníka znalosť session ID

- útočník presvedčí používateľa aby poslal ním skonštruovanú žiadosť
  - presvedčí ho aby klikol na nejakú ním zostavenú linku (napr. v diskusii)
  - využitie obrázkov na odoslanie žiadosti

- možná ochrana:
  - nepoužívanie cookies na ukladanie session ID
  - vkladať session ID priamo do URL
  - náhodne generovať a vkladať autorizačné tokeny do každej akčnej URL

# Kontrola prístupu

- prihlásení používatelia môžu mať prístup iba k určitej časti webu
- útočník môže získať neoprávnený prístup využitím chybného predpokladu programátorov o tom, ako budú používatelia komunikovať s aplikáciou
- URL tampering – pozmeňovanie častí (najmä GET parametrov) existujúcich URL
  - ak v aplikácii útočník narazí na URL v tvare: https://www.app.sk/zaznam.php?id=1234
  - môže skúsiť zadať URL v tvare: https://www.app.sk/zaznam.php?id=1235
- Forceful browsing – vytváranie nových URL a vynútený prechod na ne
  - útočník môže skúsiť zadať URL v tvare: https://www.app.sk/zaznam.php.old
  - prípadne môže skúsiť zadať napríklad: https://www.app.sk/admin.php
  - server sa môže spoliehať na to, že keď takú URL neposlal, tak ani nemôže byť taká URL zo strany klienta požadovaná

# Kontrola vstupov – rôzne prístupy

- odmietni nebezpečné vstupy
  vyhľadáva známe reťazce používané pri útokoch
- akceptuj bezpečné (povoluje len určite neškodné vstupy)
- ošetrenie/očistenie vstupu
  - odstráň potenciálne nebezpečné postupnosti znakov (napr. <script>)
  - '<scr<script>ipt>' ?
  - '+ADw-script+AD4-' ?
- bezpečná práca so vstupom
  - parametrizované dotazy do databázy
- sémantické kontroly
  - vstupné dáta sú syntakticky korektné (napríklad číslo účtu)
  - sémanticky však nie (cudzie číslo účtu)
- vstup musí byť kontrolovaný na strane servera
  - aj keď sa pre lepší UX kontroluje aj na strane klienta
  - najlepšie opakovane kontrolovať v každej časti aplikácie

# Vkladanie riadiacich dát – code injection

- každý jazyk má špecifickú syntax a špecifické riadiace znaky
- pri implemtnácii web aplikácie sa často používajú rôzne jazyky (napr.: SQL, HTML, JavaScript, XML, HTTP, . . . )
- ak sa neošetrené dáta od klienta vložia do konštrukcie v niektorom z týchto jazykov, môžu nadobudnúť neočakávaný význam
  - SQL injection
  - XPath injection
  - HTML injection (XSS: cross-site scripting)

# SQL injection

- select * from Pouzivatelia where Meno='Janko Hrasko'
- vstup: ' or ''='

  select * from Pouzivatelia where Meno='' or ''=''
- vstup: ' and 1=0 union all select * from Tabulka--

  select * from Pouzivatelia where Meno='' and 1=0 union all select * from Tabulka--'
- niekedy aplikácia vráti iba časť výsledku alebo vracia informáciu iba nepriamo
  - chybová hláška
  - dĺžka spracovania požiadavky

# XPath injection

- //user[name/text()='Janko Hrasko' and password/text()='passwd']/account/text()
- vstup: ' or 1=1 or ''='

  //user[name/text()='' or 1=1 or ''='' and password/text()='passwd']/account/text()
- vstup: NoSuchUser'] | P | //user[name/text()='NoSuchUser

  //user[name/text()='NoSuchUser'] | P | //user[name/text()='NoSuchUser' and password/text()='passwd']/account/text()

  prvá a posledná časť nič nevrátia $\Longrightarrow$ výsledkom je vyhodnotenie P
- pri prístupe k XML cez XPath nie je možné aplikovať prístupové práva (v rámci tagov)
  - preto druhým z uvedených spôsobov sa dá získať celý XML dokument
  - naproti tomu pri SQL injection bude útočník obmedzený iba na tabuľky, kam má daná aplikácia prístup

# Zvládnutie útoku

- vývojári musia predpokladať, že aplikácia bude cieľom útoku
- zvládnutie útoku vyžaduje
  - ošetrenie chýb
  - vytváranie záznamov (audit logs)
  - varovanie správcov
  - reakciu na útok

# Zvládnutie útoku – ošetrenie chýb

- v aplikácii sa nevyhnutne objavia neočakávané chyby
  - je veľmi ťažké identifikovať každý možný spôsob interakcie útočníka s aplikáciou
- aplikácia by mala zvládnuť neočakávané chyby
  - zotavením z chyby
  - prezentovaním vhodnej chybovej správy
    - správca – podrobná
    - používateľ – stručná

# Zvládnutie útoku – vytváranie záznamov

- neoceniteľný zdroj informácie pri vyšetrovaní pokusov o prienik
- pomáhajú zistiť:
  - čo sa stalo
  - aké zraniteľnosti boli využité
  - k akým dátam sa útočník dostal
  - útočníkovu identitu
- nasledovné udalosti by sa mali vždy zaznamenávať:
  - (ne)úspešný pokus o prihlásenie
  - zmena hesla
  - dôležité transakcie (napr. presun finančných prostriedkov)
  - zamietnutie prístupu
  - odpovede/požiadavky klienta pokúšajúce sa o známe útoky

# Zvládnutie útoku – varovanie správcov

- správcovia môžu online riešiť problém (namiesto spätnej offline analýzy)
  - blokovanie útočníkovej IP adresy
  - nastraženie pasce
- varovanie by malo prísť pri
  - netypickom použití
    - veľký počet požiadaviek z jednej IP v krátkom čase
    - veľký počet presunov finančných prostriedkov z rôznych účtov na jeden účet
  - odpovediach/požiadavkach klienta
    - pokúšajúce sa o známe útoky
    - so zmenenými údajmi skrytými pred bežným používateľom

# Zvládnutie útoku – reakcia na útok

- systém môže reagovať automaticky
  - ukončením útočníkovej relácie
  - spomalením odpovedí na útočníkove IP adresy
  - blokovaním útočníkových IP adries
  - varovaním napadnutého používateľa
- získa sa tým čas pre správcov