

SQL-Injektio ja siltä suojautuminen

Lalli Nuorteva

Kandidaatintutkielma
HELSINGIN YLIOPISTO
Tietojenkäsittelytieteen laitos

Helsinki, 13. huhtikuuta 2015

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen		Tietojenkäsittelytieteen laitos	
Tekijä — Författare — Author			
Lalli Nuorteva			
Työn nimi — Arbetets titel — Title			
SQL-Injektio ja siltä suojautuminen			
Oppiaine — Läroämne — Subject			
Tietojenkäsittelytiede			
Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages	
Kandidaatintutkielma	13. huhtikuuta 2015	17	
Tiivistelmä — Referat — Abstract			
<p>Tämän kandidaatintutkielman päätavoite on selvittää millaisin keinoin SQL-injektiolta voidaan suojautua. Tutkielma koostuu seuraavista osa-alueista: SQL-injektion esittely, tietoturvallisen ohjelman toteutus, ajonaikainen SQL-injektioden estäminen ja penetraatiotestaus.</p> <p>Tutkielma ei rajoitu pelkästään erilaisten suojautumismetodien luetteluun, vaan osaan metodeista perehdytään myös tarkemmin. Tämä auttaa lukijaa ymmärtämään, miten valmiit ratkaisumallit toimivat. Kun ohjelmoija ymmärtää miten hänen valitsemansa suojautumismetodit toimivat, hän tuntee myös suojauksen heikkoudet ja vahvuudet.</p>			
Avainsanat — Nyckelord — Keywords			
SQL-injektio			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Sisältö

1	Johdanto	1
2	SQL-Injektio	2
2.1	SQL-Injektio käytännössä	2
2.2	Obfuskointi	3
2.3	SQL-Injektioiden luokittelu	4
2.3.1	Inband injektio	4
2.3.2	Out-of-band injektio	4
2.3.3	Sokea injektio	5
3	Tietoturvallisen ohjelman toteutus	5
3.1	Käyttäjän syötteen käsittely	5
3.1.1	Validointi	5
3.1.2	Korvaaminen	5
3.1.3	Parametrisoidut kyselyt	6
3.2	Koodikatselmointi	7
3.3	Hienojakoinen pääsynhallinta tietokantaan	7
4	Ajonaikainen SQL-injektoiden estäminen	9
4.1	AMNESIA	9
4.2	SQL-IDS	11
4.3	SQLRand	12
5	Penetraatiotestaus	13
5.1	Penetraatiotestauksen toiminta	13
5.2	Testaussyötteiden luominen	14
6	Yhteenveto	14
	Lähteet	15

1 Johdanto

Viimeaikoina entistä useampi palvelu on siirtynyt verkkoon. Tämän seurauksena verkossa käsitellään jatkuvasti entistä enemmän arkaluonteisia tietoja. Verkossa hoidetaan asioita kuten laskujen maksaminen, hotellien varaus ja henkilökohtaisten viestien vaihtaminen. Yleensä tiedot tallennetaan relaatiotietokantoihin. Juuri tällaiset sovellukset voivat olla haavoittuvaisia SQL-Injektioille, ellei niiltä olla suojautettu oikeaoppisesti. Tästä syystä jokaisen tietokantasovelluksen ohjelmoijan on ymmärrettävä mikä on SQL-injektio ja kuinka suojautua siltä voidaan suojautua.

Tutkimusten mukaan kahdeksan prosenttia web-palveluista sisältää haavoittuvuuksia. 87 Prosenttia näistä haavoittuvuuksista on SQL-injektio haavoittuvuuksia [VAM09]. Yleisyytensä lisäksi SQL-injektio on myös hyvin vaarallinen. Onnistuneen SQL-injektion avulla hyökkääjä voi suorittaa huonosti suojatussa tietokannassa mitä tahansa operaatioita. Tämä mahdollistaa esimerkiksi arkuluontoisten tietojen lukemisen ja muokkaamisen, tai esimerkiksi sovelluksen autentikaation ohittamisen. SQL-Injektio mielletään monesti amatöörien ongelmaksi, mutta sen avulla on murrettu myös paljon ammattilaisten toteuttamia järjestelmiä. Yksi suurimmista SQL-Injektion avulla tehdyistä murroista tehtiin Guess.com:ille [Pou02]. Hyökkääjä sai käsinsä 200 000 ihmisen nimet ja luottokorttitiedot.

Sen lisäksi että SQL-injektio on yleisin tietoturva-aukko, se on myös helppoa toteuttaa ilman syvällistä ymmärrystä sen toiminnasta. Esimerkiksi penetraatiotestaustyökalut, kuten "sqlmap", ilmoittavat sivuston heikkouksista melkein napin painalluksella. Vaikka sqlmapin kaltaiset työkalut onkin tehty nimenomaan penetraatiotestaukseen, mikään ei estä hyökkääjää käyttämästä niitä apuna hyökkäyksissä. Niinpä on tärkeää, että ohjelmoija tuntee yleisimmät penetraatiotestaustyökalut, sekä käyttää niitä.

Toisaalta SQL-injektota ei ole vaikeaa estää, mikäli ohjelmoija ymmärtää kuinka SQL-injektio toimii. Nykyaikaiset web-aplikaatioiden viitekehykset, sekä ohjelmointikielet tarjoavat työkaluja SQL-injektioiden torjumiseen. Työkalujen käytöstä on tehty niin yksinkertaista, että ohjelmoijan tarvitsee vain muistaa käyttää niitä oikeissa paikoissa lähdekoodia. Esimerkiksi Ruby on Rails:issa "Model.find_by_something(parametri)" hoitaa itsestään parametrin tarkastamisen ja käsittelyn SQL-injektion varalta. SQL-Injektion estämisestä on tehty jopa niin helppoa, että jotkut saattavat erehtyä luulemaan, ettei ohjelmoijan tarvitse enää huolehtia siitä. Kuitenkin esimerkiksi Railsin "Model.where(parametri)" -metodi on altis SQL-injektoille. Niinpä ohjelmoija ei voi luottaa viitekehyksen tai ohjelmointikielen hoitavan automaattisesti kaikkea. Tästä syystä ymmärrys SQL-injektioiden toiminnasta on yhä nykyaikanakin välttämätöntä.

2 SQL-Injektio

Anleyn artikkelin "Advancen SQL Injections in SQL Server Applications"[Anl02] mukaan SQL-Injektio hyökkäys esiintyy sillon, kun hyökkääjä pääsee muuttamaan SQL käskyn logiikkaa, semantiikkaa tai syntaksia. Tämä tapahtuu lisäämällä alkuperäiseen kyselyyn uusia SQL avainsanoja tai operaattoreita. Mikäli sovelluksen tietokantaoikeuksia ei ole erikseen rajattu, hyökkääjä voi onnistuneen SQL-injektion seurauksena suorittaa mitä tahansa tietokantapalvelimen tukemia SQL-kyselyitä. Jotkut tietokantapalvelimet sallivat myös käyttöjärjestelmätason kommentojen suorittamisen. Tällöin hyökkääjän on mahdollista suorittaa myös muunlaisia hyökkäyksiä.

SQL-Injektio on mahdollinen vain silloin kun käyttäjältä tulevaa tietoa käytetään osana tietokantapalvelimelle tehtävää kyselyä. Tämä on kuitenkin varsin tavallinen tarve sovelluksissa. Tietokannassa voidaan säilyttää esimerkiksi käyttäjätunnuksia ja salasanoja. Näin ollen kirjautuessa järjestelmään käyttäjän antamaa syötettä käytetään osana SQL-kyselyä.

2.1 SQL-Injektio käytännössä

Esimerkiksi tuotteiden etsimiseen liittyvä SQL-käsky voidaan rakentaa SQL-injektiolle alttiissa sovelluksessa seuraavalla tavalla:

```
sql = "SELECT * FROM tuotteet  
WHERE nimi ='" + params[:tuotenimi] + "'"
```

Mikäli käyttäjä antaa nimekseen "'; DROP TABLE tuotet'. Valmis kysely tietokannalle näyttää seuraavalta:

```
sql = "SELECT * FROM tuotteet  
WHERE nimi = ''; DROP TABLE tuotteet
```

Kyseinen kysely etsii ensin kaikki tuotteet joiden nimi on tyhjä. Seuraavaksi suoritetaan komento "DROP TABLE tuotteet", joka poistaa koko tuotteet taulun.

Edellä mainitun hyökkäyksen sijaan hyökkääjä olisi voinut käyttää esimerkiksi jotakin Srivastavan artikkelissa "Algorithm to prevent back end database against SQL injection attacks" [Sri14] esitellyistä menetelmistä:

Tautologia

Hyökkääjä voi käyttää tuotenimeä, joka sisältää jonkin tautologian esimerkiksi "; OR 1=1". Tällöin hyökkääjä olisi saanut vastauksena kaikki tuotteet.

Kommentti

Mikäli tuote olisi vaatinut myös tuotekoodin, eli kyselyn muotoilu olisi ollut esimerkiksi seuraavanlainen:

```
sql = "SELECT * FROM tuotteet  
WHERE nimi ='" + params[:tuotenimi] + "' AND  
tuotekoodi ='" + params[:tuotekoodi] + "'"
```

Hyökkääjä voi kirjoittaa tuotenimeksi "Haluttu tuote'; –". Hyökkäyksen onnistuessa loput kyselystä muuttuu kommentoiduksi. Tällöin tuotteen tiedot palautetaan, vaikka hyökkääjä ei tietäisi tuotekoodia. Kaksi viivaa on kommenttimerkki esimerkiksi Transact-SQL:llässä ja Micorosft SQL serverissä [KT08].

Union kysely

Hyökkääjä voi kirjoittaa tuotenimeksi esimerkiksi "Haluttu tuote' UNION SELECT * FROM kayttajatiedot". Hyökkäyksen onnistuessa vastaukseen sisältyy myös koko taulun "kayttajatiedot" sisältö.

2.2 Obfuskointi

Edellisen kappaleen esimerkit eivät todennäköisesti toimi sovelluksissa, jotka on suojattu SQL-injektioilta. Tästä syystä hyökkääjien on keksittävä tapoja, joilla palomureja voidaan yrittää kiertää. Salgadon kirjoittaman "SQL Injection Optimization and Obfuscation Techniques"[Sal13] artikkelissa esitellään obfuskointia. Obfuskointi tarkoittaa koodin tahallista monimutkaistamista ja epäselkeyttämistä sen varsinaisen toiminnan piilottamiseksi. Obfuskointia käytetään esimerkiksi haittaohjelmien piilottamiseen virustutkilta. SQL-Injektoiden tapauksessa obfuskointi voi olla yksinkertaisimmillaan esimerkiksi "DROP" avainsanan muuttaminen "DroP":iksi. Tällöin yksinkertainen mustalistaukseen perustuva palomuri saattaisi päästää SQL-injektion läpi.

Monet hienostuneemmat tavat käyttävät SQL-injektoiden piilottamiseen esimerkiksi erilaisia enkoodauksia. Salgagon mukaan enkoodauksen käyttö perustuu siihen, että eri kerrokset käsittelevät enkoodauksia eri tavalla. Esimerkiksi Unicodessa merkkiä "a" vastaa merkkijono "%u0061". Voi olla että palomuri tulkitsee merkkijonon "%u0061" tavallisena merkkijonona, kun taas tietokanta tulkitsee sen kirjaimena "a". Näin ollen esimerkiksi avainsana SELECT voidaan piilottaa uniconen avulla merkkijonoon "%u0053%u0045%u004c%u0045%u0043%u0054".

2.3 SQL-Injektioiden luokittelu

Sadeghiani ja hänen kollegoidensa artikkelin "SQL-Injection is still alive" mukaan SQL-injektioita luokitellaan sen perusteella, mitä reittiä hyökkääjä saa palautteen sovellukselta. SQL-injektioita on kolmea eri päätyyppiä, inband injektio, out-of-band injektio ja sokea injektio [SZI13].

2.3.1 Inband injektio

Artikkelin mukaan inband tyyppinen injektio on kyseessä silloin, kun vastaus saadaan samaa reittiä, jota hyökkäys on suoritettu. Tällainen hyökkäys onnistuu silloin, kun kyselyn tulos palautetaan suoraan käyttäjälle. Tällainen aukko saattaa esiintyä esimerkiksi sovelluksessa, jossa hyökkääjä voi hakea ystäviensä lisäämiä kuvia. Hyökkäyksen tuloksena hyökkääjä voisi saada vaikkapa kaikki tietokannasta löytyvät kuvat.

2.3.2 Out-of-band injektio

Sen sijaan out-of-band injektiossa tuloste saadaan eri reittiä, kun hyökkäys on syötetty. Out-of-band injektioita voidaan hyödyntää, vaikka sovellus ei palauta käyttäjälle kyselyn tulosta. Tällainen tilanne voi esimerkiksi ilmetä, jos sovellus tallettaa tietokantaan käyttäjien selaintietoja. Selaintiedot saadaan HTTP-pyynnön "User-Agent" kentästä. Hyökkääjä voisi tällaisessa tapauksessa lähettää "User-Agent" kentässä haitallista koodia. Jotta hyökkääjä saisi vastauksen suoritetusta kyselystä, hän voi ohjata vastauksen esimerkiksi oman palvelimensa logeihin. Tämä onnistuu, mikäli sovellus saadaan suorittamaan esimerkiksi seuraavanlainen kysely:

```
utl_http( 'http://www.hyokkaaajansivu.fi/injections/'  
        ||  
        SELECT password  
        FROM User  
        WHERE username = 'admin'  
        )
```

Oletetaan että käyttäjän admin salasana on "password123". Injektion onnistuessa tietokanta hakee ensin käyttäjän admin salasanan, jonka jälkeen se suorittaa HTTP-pyynnön osoitteeseen "http://hyokkaaajansivu.fi/injections/password123". Tällöin hyökkääjän palvelimen logeissa näkyy seuraavanlainen merkintä:

```
GET "/injections/admininpassword", 200
```

Sadeghianin ja kollegoiden artikkelin mukaan kyselyn tulos voidaan ohjata palvelimen lisäksi myös esimerkiksi hyökkääjän sähköpostiin.

2.3.3 Sokea injektio

Artikkelin mukaan sokeassa injektiossa hyökkääjä ei saa kyselyn palauttamaa tulosta selville mitään reittiä. Hyökkääjä voi päätellä hyökkäyksen onnistumisen esimerkiksi siitä kuinka nopeasti sivu latautuu. Lisäämällä injektoituun sql-käskyyn komennon "waitfor delay 0:0:5", tietokanta odottaa 5 sekuntia ennen kuin se palauttaa tuloksen. Tästä voidaan päätellä injektion onnistuneen [TMH10]. Hyökkääjä voi myös tarkkailla vaikuttaako hänen tekemänsä hyökkäykset sovelluksen toimintaan tai ulkoasuun jollain tapaa.

3 Tietoturvallisen ohjelman toteutus

3.1 Käyttäjän syötteen käsittely

SQL-Injektioilta välttymiseksi käyttäjän syötteet on tarkastettava. Vaaralliseksi epäilty syöte voidaan käsitellä vaarattomaksi, tai jättää suorittamatta kokonaan. Tässä kappaleessa esitellään kolme erilaista tapaa käsitellä käyttäjän syötettä: datan validointi, korvaaminen ja parametrisoidut kyselyt.

3.1.1 Validointi

Validointi voidaan toteuttaa joko musta- tai valkolistan avulla tai vertaamalla syötettä erilaisiin säännöllisiin lauseisiin. Valkolistan tapauksessa kaikki sallitut syötteet on kirjattuna valmiiksi. Tällainen lähestymistapa on kuitenkin usein mahdoton, koska valkolista kasvaisi liian suureksi. Mustalistalla voidaan pyrkiä listaamaan kaikki kielletyt syötteet tai merkit, mutta mustalista saattaa on kierrettävissä obfuskoinnin avulla [Sal13].

Käyttäjän syötettä voidaan myös verrata säännölliseen lausekkeeseen, esimerkiksi puhelinnumeokenttä voidaan rajoittaa hyväksymään pelkästään numeroita. Syöte voidaan myös tarkastaa SQL-avainsanojen tai erikoismerkien varalta. Tällainen lähestymistapa voi kuitenkin rajoittaa myös päteviä syötteitä. Esimerkiksi käyttäjänimi "O'Brian" olisi estetty '"'merkin vuoksi [ANBA14].

3.1.2 Korvaaminen

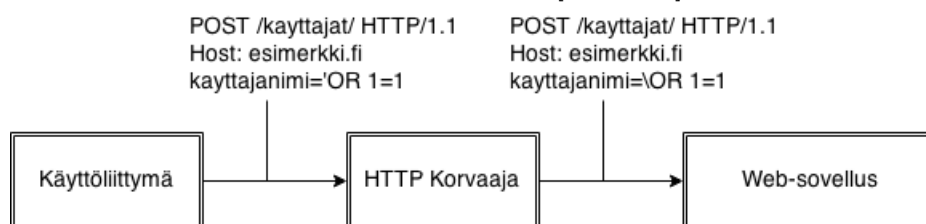
Datan validoinnin sijaan voidaan käyttää korvaamista (*engl. escaping*). Korvaamisessa vaarallisia syötteitä ei hylätä. Hylkäämisen sijaan kaikille syötteille suoritetaan korvausoperaatio. Korvausoperaatiossa haitalliset erikoismerkit, kuten ' korvataan joillakin vaarattomilla merkeillä kuten \. Vaarallisissa merkeissä on kuitenkin tietokantakohtaisia eroja. Tästä syystä ohjelmoijan tulee käyttää tietokantakohtaista korvausfunktia. Tyypillisesti korvausfunktia

ei ole tarpeellista toteuttaa itse, sillä ohjelmointikielet tarjoavat niihin usein valmiita toteutuksia. Esimerkiksi PHP:ssa on MySQL tietokantaa varten luotu "my_sql_real_escape_string()" funktio.

Sadeghanin ja Zamanin artikkelin "SQL injection vulnerability general patch using header sanitization" [SZAM14] mukaan korvaamista voidaan käyttää myös HTTP-pakettien käsittelyyn. Web-sovellukset saavat tyypillisesti viestinsä HTTP-protokollaa käyttäen. Saapuvat HTTP-paketit voidaan käsitellä korvausfunktion avulla ennen kun ne annetaan itse sovellukselle käsiteltäväksi. Tällä tavoin vanhasta sovelluksesta voidaan yrittää tehdä tietoturvallinen ilman, että sen lähdekoodia tarvitsee muokata. Artikkelin mukaan esimerkiksi PHP:n tapauksessa metodin käyttöönotto vaatii vain yhden kirjaston käyttöönottoa. Tämän ansiosta metodin käyttöönotto on todennäköisesti nopeampaa, kuin koko sovelluksen refaktorointi SQL-injektioiden varalta. Toisaalta erillisen HTTP-pakettienkäsittelijän lisääminen voi huonontaa ohjelman suorituskykyä.

Korvaaminen ei kuitenkaan ole aina toivottua. Esimerkiksi nimimerkki "O'Brian" korvautuu "O\Brian:iksi", joka ei todennäköisesti ole toivottavaa.

Kuva 1: Syöte voidaan korvata erillisessä HTTP-pakettien käsittelijässä jo ennen kuin paketti annetaan sovellukselle. [SZAM14]



3.1.3 Parametrisoidut kyselyt

Parametrisoidut kyselyt esitellään artikkelissa "SQL injection vulnerability general patch using header sanitization" [SZAM14]. Niiden avulla voidaan tallentaa käyttäjän antama syöte sellaisenaan tietokantaan, ilman että se muuttaa valmiiksi luodun kyselyn syntaksia. Tällöin "O'Brian" tallentuu "O'Brian":ina juuri kuten pitääkin. Parametrisoidussa kyselyssä luodaan SQL-kyselyistä pohjia, joihin lisätään paikanpitäjät (*engl. placeholder*). Parametrisoitu kysely annetaan tietokannalle. Tietokanta kääntää ja optimoi kyselyn pohjan vain kerran. Tällainen toimintatapa parantaa suorituskykyä. Tietokanta ei kuitenkaan suorita varsinaista kyselyä heti, koska siitä puuttuu haettavat arvot. Kun käyttäjä antaa syötteenä haluamansa arvot, paikanpitäjät korvataan arvoilla. Käyttäjän antamia arvoja ei enää tulkata SQL:lläksi, vaan niitä käsitellään sellaisinaan. Tällöin ei ole mahdollista että käyttäjän syöte rikkoisi kyselyä. Esimerkiksi jos käyttäjä antaa käyttäjänimekseen ”;

OR 1=1", niin tietokannasta haetaan käyttäjää, jonka käyttäjänimi on "OR 1=1". Parametrisoidut kyselyt ovat tuettuina lähes kaikissa yleisimmissä ohjelmointikielissä [jav15]. Oikein käytettynä parametrisoidut kyselyt suojaavat sovelluksen täysin SQL-injektioilta [SZAM14].

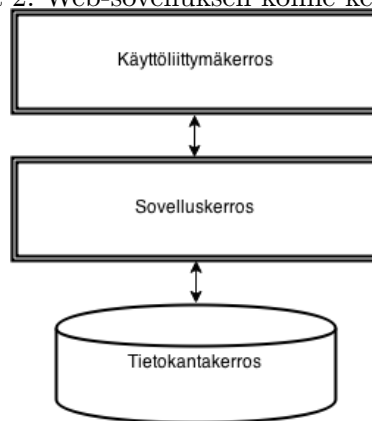
3.2 Koodikatselmointi

Antunesin ja kumppaneiden tekemässä vertailussa [AV09] vertailtiin staattisen koodianalyysin ja penetraatiotestaamisen eroja. Kumpikaan testaustapa ei löytänyt yli 51% sovelluksen tietoturva-aukoista. Tällaisten aukkojen huomaaminen on kuitenkin mahdollista, kun ohjelman lähdekoodia katselmoi useammat henkilöt.

3.3 Hienojakoinen pääsynhallinta tietokantaan

Tämä kappale keskittyy Roichmanin ja Gudesin artikkeliin "Fine-grained Access Control to Web Databases" [RG07]. Artikkelin mukaan ennen web-sovellusten yleistymistä sovelluksia ajettiin käyttäjän omalla tietokoneella. Tyypillisellä sovelluksella oli kiinteä määrä käyttäjiä. Tällaisessa sovelluksessa sovelluskerros kommunikoi suoraa tietokannan kanssa. Tämän seurauksena tietokanta tietää mikä käyttäjä sitä milloinkin käyttää. Täten on helppoa rajata käyttäjien oikeuksia.

Kuva 2: Web-sovelluksen kolme kerrosta.



Nykyään web-sovelluksissa on tyypillisesti kolme kerrosta [RG07][APG⁺12]. Käyttöliittymänä toimii käyttäjän selain, joka kommunikoi web-sovelluksen palvelimen kanssa. Palvelin välittää käyttäjän käskyt tietokannalle. Tietokannan näkökulmasta komennot antaa web-sovellus, eikä komennon käyttöliittymästä lähettänyt käyttäjä. Täten tietokanta suorittaa sokeasti kaikki saamansa komennot, ellei web-sovellukseen tietokantakäyttäjän oikeuksia ole erikseen rajattu.

Aiemmin esitellyissä menetelmissä on keskitytty ratkaisemaan ongelmaa sovelluskerroksella. Roichmanin ja Gudesin lähestymistavassa keskitytään ratkaisemaan ongelmaa tietokantatasolla parametrisointi metodin (*engl. parameter method*) avulla. Tekniikka perustuu parametrisoituihin näkymiin (*engl. parametrized views*). Parametrisoidun näkymän avulla voi suodattaa tallenteita ilman, että tarvitsee tehdä uutta näkymään jokaista eri parametria varten.

Sovelluksen tulee ylläpitää tietokantataulua johon merkitään aktiivisten käyttäjien ID:t. Roichmanin ja Gudesin esittelemä metodi toimii seuraavalla tavalla:

1. Käyttäjä kirjautuu sovellukseen ja sovellus palauttaa käyttäjälle satunnaista AS_KEY:n, mikäli kirjautuminen onnistuu.
2. Sovellus tallettaa aktiivisten käyttäjien tauluun käyttäjän ID:n ja sitä vastaavan AS_KEY:n. Tästä lähtien kaikissa käyttäjän tekemissä SQL-kyselyissä käytetään käyttäjäkohtaista AS_KEY:tä.
3. AS_KEY poistetaan kun käyttäjä kirjautuu ulos.

Kirjautumisen jälkeen käyttäjän tiedot ovat taulussa esimerkiksi seuraavalla tavalla:

KäyttäjäID	AS_KEY
20	01010101..

Nyt voidaan käyttää seuraavanlaista parametrisoitua näkymää:

```
CREATE VIEW Palkka_View WITH pAS_KEY
SELECT * FROM Palkka
WHERE Kayttaja_ID IN
(SELECT Kayttaja_ID
FROM Kayttajat_Table
WHERE Kayttajat_Table.AS_key=:pAS_KEY)
```

Näkymä ottaa parametrina AS_Key:n, jonka käyttäjä on saanut kirjautuessaan. Käyttäjän kysely tehdään näkymään "Palkka_View" eikä tauluun "Palkka". Mikäli hyökkääjä yrittäisi tehdä SQL-injektion tautologian avulla, suoritettava kysely näyttäisi seuraavalta:

```
SELECT Palkka
FROM Palkka_View(01010101..)
WHERE Palkka_pvm = '12/2015' OR 1=1
```

Hyökkääjä saisi vastauksena kaikki omat palkkatietonsa, mutta ei muiden käyttäjien, koska `Palkka_View` saa parametrikseen hyökkääjän oman `AS_KEY:n`. Myöskään UNION injektio ei ole tässä tapauksessa mahdollinen, koska hyökkääjä ei tiedä muiden käyttäjien `AS_KEY:tä`, joka tarvitaan `Palkka_Viewiin` parametrikseen.

4 Ajonaikainen SQL-injektoiden estäminen

SQL-Injektioita voidaan yrittää havaita ja estää ajonaikana. Tätä varten on luotu useita työkaluja, kuten SQLCheck, SQLProb ja Candid. [ST13]. Ajonaikaiseen SQL-injektoiden estämiseen on kehitetty monenlaisia tapoja. Tässä kappaleessa perehdytään tarkemmin AMNESIA, SQL-IDS ja SQLrand menetelmiin.

4.1 AMNESIA

Tässä kappaleessa esitellään artikkelissa "Preventing SQL Injection Attacks Using AMNESIA" [HO06] esiteltyä AMNESIA tekniikkaa. AMNESIA:n avulla pystytään estämään SQL-injektioita ajonaikaisesti. Artikkelin mukaan AMNESIA ottaa syötteenään sovelluksen lähdekoodin ja palauttaa siitä version, joka on suojattu SQL-injektioilta.

AMNESIA:ssa on ideana etsiä sovelluksen lähdekoodista ne paikat, joissa SQL-kyselyitä suoritetaan. Jokaista SQL-kyselyn suoritusta kohden luodaan äärellinen automaatti. Automaatti hyväksyy vain ne syötteet, joita se ei tunnista SQL-injektoiksi. Ajon aikana jokaista suoritettavaa SQL-kyselyä verrataan suoritusta kohden luotun äärellisen automaatin kanssa. Kysely suoritetaan vain, mikäli automaatti hyväksyy syötteen. AMNESIA tekniikka koostuu neljästä osasta:

1. Etsi suorituspajat

Ensin ohjelman koodi skannataan. Skannauksessa etsitään koodista ne paikat, joissa tietokantakyselyjä suoritetaan. Näihin paikkoihin viitataan tässä tutkielmassa sanalla "suorituspajka" (*engl. hotspot*). Esimerkiksi Javan tapauksessa etsitään koodista paikat joissa kutsutaan "`java.sql.Statement.execute(String)`" metodia.

2. SQL-kyselymallien rakentaminen

Seuraavaksi rakennetaan jokaiselle edellisessä kohdassa löydetylle suorituspajalle oma mallinsa. Tämä onnistuu siten, että AMNESIA simuloi sovelluksen toimintaa Java String Analysis (JSA) kirjaston avulla. JSA Luo analyysin tuloksena epätermistisen äärellisen automaatin (NFA), joka tunnistaa kaikki mahdolliset merkkijonot, jotka kysely voi saada arvokseen. Esimerkiksi allaoleva koodipätkä voi saada arvokseen joko: "`SELECT info FROM kayttajat WHERE kayttajanimi = β` "

tai "SELECT info FROM kayttajat WHERE kayttajanimi='vieras'".
Käyttäjän syötettä merkataan symbolilla β .

```
query = "SELECT info FROM kayttajat WHERE"
if (!kayttajanimi.empty) {
  query += "kayttajanimi =" + kayttajanimi +
    " "
} else {
  query += "kayttajanimi = vieras "
}
```

3. Instrument application

Seuraavaksi lisätään jokaiseen vaiheessa 1. löydettyyn suorituspaikkaan monitori. Monitori suoritetaan aina ennen itse tietokantakyselyä. Monitori ottaa parametriksi suorituspaikan uniikin ID:n ja merkkijonon jota ollaan suorittamassa. ID:n avulla monitori etsii kyseistä suorituspaikkaa vastaavan mallin. Alla sama koodi esimerkkinä:

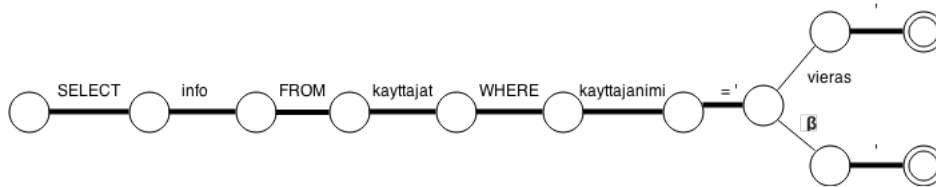
```
if (monitor.hyvaksyy(<suorituspaikan id>,
  kysely)) {
  return db.suorita(kysely);
}
```

4. Ajonaikainen monitorointi

Ajonaikana ohjelma toimii normalisti kunnes se törmää suorituspaikkaan. Suorituspaikkaan törmättyään se antaa tarvittavat parametrit monitorille. Ensin monitori käsittelee kyselyn samalla tapaa kuin tietokanta sen käsittelee. Tämän asiasta esimerkiksi erikoismerkit evaluoituvat niiden oikeaan arvoonsa. Tämä estää SQL-avainsanojen piilottamisen erikoismerkeillä. Kun kysely on käsitelty, tarkastetaan tunnistaako malli sen. Mikäli malli hyväksyy kyselyn se suoritetaan, muulloin malli tunnistaa sen SQL-injektioksi.

Oletetaan että kyselymme olisi "SELECT info FROM users WHERE kayttajanimi=" OR 1=1. Vaiheessa 1. kuvatun koodipätkän automaatti jakautuisi kahtia. Koska automaatti tunnistaa vain sellaiset kielet, jotka loppuvat merkkiin " " heti käyttäjän syötteen jälkeen, kyseinen kysely huomataan SQL-injektioksi.

Kuva 3: Vaiheen 2. koodia vastaava automaatti



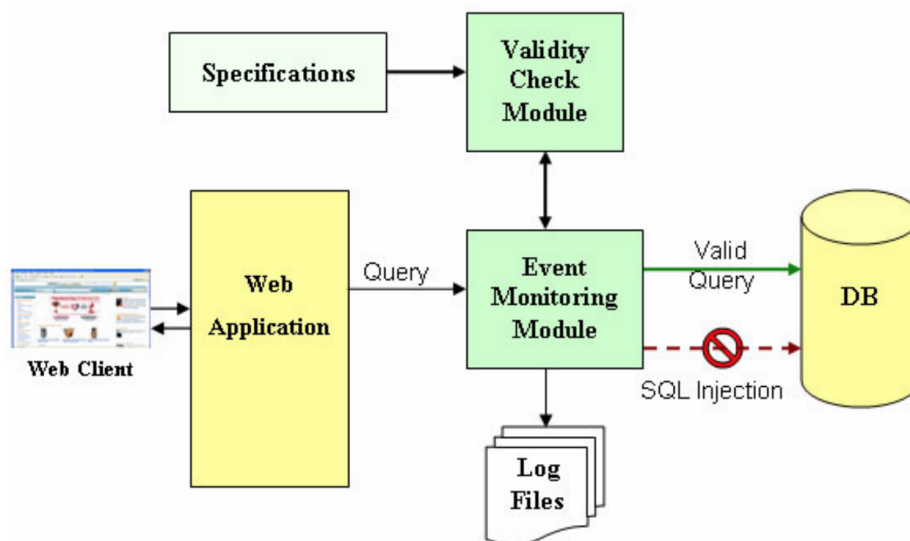
4.2 SQL-IDS

SQL-IDS (SQL-Injection Detection System) esittelee Kemalksen ja Theodroksen artikkelissa "SQL-IDS: A Specification-based Approach for SQL-Injection Detection" [KT08]. Se muistuttaa toimintatavaltaan AMNESIA:a, mutta siinä on joitakin eroavaisuuksia. AMNESIA vaatii muutoksia ohjelman lähdekoodiin, kun taas SQL-IDS ei vaadi.

Artikkelin mukaan SQL-IDS menetelmässä sovelluksen ja tietokannan väliin asetetaan erillinen tapahtumaseurantamoduuli (*engl. Event Monitoring Module*). Seurantamoduuli käsittelee sovelluksen lähettämiä SQL-kyselyitä antamalla ne oikeellisuuden tarkastusmoduulille (*engl. Validity Check Module*). Mikäli tarkastusmoduuli hyväksyy SQL-kyselyn, se lähetetään tietokannalle. Muulloin kysely havaitaan SQL-injektioiksi eikä sitä suoriteta. SQL-Injektioiksi havaituista kyselyistä kirjataan tietoja lokiin.

Oikeellisuuden tarkastusmoduuli vastaanottaa SQL-kyselyn merkkijonona. Moduuli tulkkaa merkkijonon SQL:läksi, jonka jälkeen sitä verrataan

Kuva 4: SQL-IDS lisättynä web-sovellukseen. SQL-IDS menetelmään liittyvät moduulit merkittynä vihreällä. [KT08]



ennaltamäärättyyn spesifikaatioon. Spesifikaatiot ovat määriteltynä EBNF kieliopilla. Mikäli annettu lause ei riko kyseisen lauseen spesifikaatiossa määriteltäviä syntaktisia sääntöjä, se hyväksytään, jolloin se etenee seuranta-moduulin kautta tietokannalle suoritettavaksi.

Kemalis ja Theodoros ovat testanneet SQL-IDS menetelmää lähettämällä esimerkkisovellukselle 2450 SQL-kyselyä, joista 420 oli haitallisia. SQL-IDS onnistui tunnistamaan kaikki hyökkäykset. Artikkelin mukaan menetelmä lisää keskimäärin 1ms viiveen per SQL-kysely.

4.3 SQLRand

Boyd ja Keromytis esittelevät SQLRand menetelmän SQL-injektioiden torjumiseksi artikkelissaan "SQLRand: Preventing SQL Injection Attacks" [BK04]. SQLRandissa on ideana lisätä jokaiseen SQL-avainsanaan jokin satunnainen numero. Alkuperäinen kysely voi näyttää esimerkiksi seuraavalta:

```
SELECT *  
FROM Oppilaat  
WHERE opintopisteet > 180
```

Kun SQLRand avaimena käytetään numeroa 123, valmis kysely näyttää seuraavalta:

```
SELECT123  
FROM123 Oppilaat  
WHERE123 opintopisteet > 180
```

Mikäli hyökkääjä yrittäisi suorittaa SQL-injektion tavallisilla SQL-avainsanoilla, se epäonnistuihi, koska sovellus ei tunnistaisi sitä SQL:läksi.

Tavallinen tietokanta ei ymmärrä satunnaistettuja SQL-avainsanoja. Artikkelissa tähän ongelmaan oli luotu kaksi eri ratkaisua. Ensimmäinen on muokata tietokannan SQL-tulkkiä siten, että se ymmärtää muokattuja SQL-avainsanoja. Tämä voi kuitenkin olla vaikeaa toteuttaa, sekä se saattaa sotkea muiden sovellusten toimintaa, jotka käyttävät samaa tietokantapalvelinta. Toinen ratkaisu on luoda erillinen tulkki-sovellus tietokantapalvelimen ja web-sovelluksen välille.

Boyd ja Keromytis esittelevät artikkelissaan myös SQLRandin tehokkuutta. Heidän mittaamiensa tulosten mukaan SQLRandin lisääminen tuo keskimäärin 183-316 mikrosekunnin viiveen per kysely.

SQLRandia on kritisoitu sen monimutkaisuudesta [SZAM14]. Kritisoiijat pitävät SQLRandin toteuttamista olemassa olevaan sovellukseen aikaavievänä. Sen katsotaan myös hankaloittavan sovelluksen käyttöönottoa.

5 Penetraatiotestaus

Vaikka sovellusta ohjelmoitaisiin hyvien ohjelmointikäytänteiden mukaisesti, siihen voi silti jäädä tietoturva-aukkoja. Tämän takia sovellusta on jatkuvasti tietoturvatestattava.

Penetraatiotestauksessa yritetään etsiä sovelluksesta tietoturva-aukkoja. Kun penetraatiotestaus on automatisoitua, ohjelmoijan ei tarvitse suorittaa samoja testausruutuneja käsin jokaisen muutoksen jälkeen. Penetraatiotestaus ei kuitenkaan ole virheetöntä. Artikkelissa "Using web security scanners to detect vulnerabilities in web services"[VAM09] kerrotaan, että penetraatiotestaus voi aiheuttaa turhia hälyytyksiä, tai olla hälyyttämättä kun pitäisi hälyyttää. Artikkelin mukaan jopa yli 30% havaituista virheistä osoittautuivat virheellisiksi. Tämän takia penetraatiotestauksen tulee olla vain yksi testaustyökaluista, eikä siihen voida sokeasti luottaa.

Mustalaatikko -testauksessa testataan sovellusta erilaisia syötteitä vastaan. Tällaiseen testaamiseen ei vaadita pääsyä itse koodiin [TGZ14]. Tämä on hyödyllistä esimerkiksi sellaisissa tapauksissa, kun osa ohjelman komponenteista on kolmannen osapuolen koodia, eikä siihen päästä käsiksi. Mustalaatikko testauksessa on ongelmana se, että tulos perustuu sovelluksen tulosteesta tehtyyn analyysiin, eikä esimerkiksi tietokannan todelliseen tilaan. Tästä syystä kaikkia tietoturva-aukkoja ei välttämättä löydetä, tai jotain turvallista kohtaa koodista voidaan luulla tietoturva-aukoksi. Mustalaatikkotestaukseen löytyy useita valmiita työkaluja, kuten "Acunetix Web Vulnerability Scanner" ja sqlmap.

Valkolaatikko testauksessa testaajalla on pääsy ohjelman lähdekoodiin. Valkolaatikko testaustyökaluihin kuuluu esimerkiksi staattiset lähdekoodin analysointi työkalut. Tällaisilla työkaluilla voidaan huomata mahdollisia tietoturva-aukkoja jo ennen kun ohjelmaa on ajettu ensimmäistäkään kertaa [GGS14].

5.1 Penetraatiotestauksen toiminta

Haixia edottaa artikkelissaan "A database security testing scheme of web application"[HZ09] seuraavanlaista testausmallia.

Ensiksi etsitään kaikki mahdolliset paikat sovelluksesta, joista käyttäjä voi syöttää dataa. Tämä onnistuu leveyssuuntaista hakua (*engl. Breadth-first search*) käyttämällä. Algoritmi toimii seuraavasti:

1. Alustetaan lista jossa on ainoana jäsenenä etusivun URL. Etusivu merkataan käsittelemättököks.
2. Käydään listalta läpi kaikki käsittelemättömiksi merkatut sivut. Kunkin sivun kohdalla tehdään seuraavat vaiheet:
 - Otetaan talteen kaikki paikat joista käyttäjä voi syöttää dataa.

- Etsitään sivulta kaikki linkit ja lisätään listalle ne jotka eivät vielä ole siellä.
- Merkataan sivu käsitellyksi.

3. Mikäli listalla on käsittelemättömiä linkkejä, palataan vaiheeseen 2.

Tämän jälkeen luodaan mahdollisimman kattava lista erilaisista hyökkäyksistä. Kaikkiin mahdollisiin paikkoihin joista voi syöttää dataa kokeillaan haitallisia syötteitä. Tietokannan palauttamasta arvosta voidaan päätellä onko injektio onnistunut vai ei. Esimerkiksi jos vastauksen HTTP statuskoodi on 200, kyseessä on haavoittuvuus.

5.2 Testaussyötteiden luominen

Artikkelissa "Automated Testing for SQL Injection Vulnerabilities: An input Mutation Approach"[ANBA14] ehdotetaan penetraatiotestauksessa käytettävien syötteiden luomiseen automatisoitua tekniikkaa nimeltään μ SQLi. Tekniikassa on ideana manipuloida kelpaavaa syötettä erilaisilla mutaatio-operaatioilla. Mutaatio-operaatiot jaetaan artikkelin mukaan seuraavalla tavalla kolmeen eri osioon:

1. Käyttäytymistä muuttavat operaatiot

Esimerkiksi operaatiot jotka lisäävät AND tai OR lauseen, tai operaatiot joissa lisätään puolipiste ja kokonaan uusi SQL lause.

2. Syntax-Repairing Operators

Lisää kyselyyn esimerkiksi sulut, kommenttimerkin tai heittomerkin.

3. Obfuskointi operaatiot

Esimerkiksi muuttaa kyselyssä käytettävää enkoodausta tai muuttaa totuuslauseketta ilman että sen arvo muuttuu.

Toimivaan syötteeseen voidaan lisätä yksi tai useampi mutaatio-operaatio. Artikkelin mukaan useasti yksittäinen operaatio huomataan, mutta yhdistelmät saattavat silti jäädä huomaamatta. Mutaatioiden tekeminen aloitetaan toimivasta syötteestä, koska sillä vältetään se, että syöte hylättäisiin välittömästi. Lisäksi toimivat syötteet täyttävät todennäköisemmin syötevalidoinnit.

6 Yhteenveto

SQL-Injektioit ovat olleet web-sovelluksien uhkana jo kauan. Niiltä suojautumiseen on kehitetty monenlaisia toimivia suojautumismenetelmiä. Osa menetelmistä suojaa sovelluksen täydellisesti, mutta osa vain hankaloittaa hyökkäystä. Esimerkiksi parametrisoidut kyselyt suojaavat sovelluksen SQL-injektoilta täydellisesti, kunhan niitä käytetään oikein. Sen sijaan esimerkiksi

syötteen validointi ei takaa suojaa SQL-injektioilta, mutta saattaa hankaloittaa hyökkäystä.

Osa suojautumismenetelmistä täytyy olla mukana ohjelmointityylissä alusta loppuun asti, sillä niiden lisääminen myöhemmin on työlästä. Esimerkiksi parametrisoidut kyselyt tai Roichmanin ja Gudesin esittelemä hienojakoinen pääsynhallinta tietokantaan vaikuttavat moneen paikkaan koodissa. Osa tekniikoista sen sijaan voidaan ottaa käyttöön vaivattomasti myöhämmässäkin vaiheessa. Esimerkiksi Sadeghanin ja Zamanin esittelemä HTTP-pakettienkäsittelijä, joka voidaan asettaa käsittelemään HTTP-paketteja ennen kun ne annetaan sovellukselle, on helppoa lisätä olemassa olevaan sovellukseen.

Jotkin menetelmistä heikentävät suorituskykyä, tai vaikeuttavat ohjelman käyttöönottoa. Esimerkiksi AMNESIA tekniikkaa käyttäessä sovellus täytyy jokaisen muutoksen jälkeen käyttää AMNESIA työkalun läpi. Toisaalta AMNESIA työkalua käyttäessä SQL-injektioilta suojautumiseen ei tarvitse juurikaan käyttää aikaa ohjelmoinnin aikana, sillä AMNESIA hoitaa suojautumisen automatisoidusti.

Jotta suojautumisen onnistumisesta voitaisiin varmistua, sovellusta on testattava. Sovelluksen testaamiseen on luotu erilaisia valmiita työkaluja. Osa työkaluista, kuten staattiset koodin analysointityökalut vaativat pääsyn ohjelman lähdekoodiin. Osa toimii pelkästään ohjelman antaman palautteen perusteella. Tämän hetkisiä työkaluja yhdistää korkeat "false-positive" prosentit. Tästä syystä testaustyökalujen antamista tuloksista ei voida päätellä varmaksi, onko sovellus turvallinen vai ei. Koska teustaustyökalujen antamien tulosten oikeellisuudesta ei voida olla varmoja, ohjelmoijan on myös itse ymmärrettävä mikä on SQL-injektio.

Erilaisia suojautumismenetelmiä on niin laaja kirjo, että jää epäselväksi milloin mitäkin suojautumismenetelmää kannattaa käyttää. Tutkielmaa voisi jatkaa selvittämällä milloin mitäkin menetelmää kannattaa käyttää.

Lähteet

- [ANBA14] Appelt, Dennis, Nguyen, Cu Duy, Briand, Lionel C. ja Alshahwan, Nadia: *Automated Testing for SQL Injection Vulnerabilities: An Input Mutation Approach*. Teoksessa *Proceedings of the 2014 International Symposium on Software Testing and Analysis*, ISSTA 2014, sivut 259–269, New York, NY, USA, 2014. ACM, ISBN 978-1-4503-2645-2. <http://doi.acm.org/10.1145/2610384.2610403>.
- [Anl02] Anley, C.: *Advanced SQL Injection In SQL Server Applications*. Teoksessa *Next Generation Security Software Ltd. White Paper, 2002.*, 2002.

- [APG⁺12] Avireddy, S., Perumal, V., Gowraj, N., Kannan, R.S., Thinakaran, P., Ganapathi, S., Gunasekaran, J.R. ja Prabhu, S.: *Random4: An Application Specific Randomized Encryption Algorithm to Prevent SQL Injection*. Teoksessa *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*, sivut 1327–1333, June 2012.
- [AV09] Antunes, N. ja Vieira, M.: *Comparing the Effectiveness of Penetration Testing and Static Code Analysis on the Detection of SQL Injection Vulnerabilities in Web Services*. Teoksessa *Dependable Computing, 2009. PRDC '09. 15th IEEE Pacific Rim International Symposium on*, sivut 301–306, Nov 2009.
- [BK04] Boyd, Stephen W. ja Keromytis, Angelos D.: *SQLrand: Preventing SQL Injection Attacks*. Teoksessa *In Proceedings of the 2nd Applied Cryptography and Network Security (ACNS) Conference*, sivut 292–302, 2004.
- [GGS14] Gupta, M.K., Govil, M.C. ja Singh, G.: *Static analysis approaches to detect SQL injection and cross site scripting vulnerabilities in web applications: A survey*. Teoksessa *Recent Advances and Innovations in Engineering (ICRAIE), 2014*, sivut 1–5, May 2014.
- [HO06] Halfond, William G. J. ja Orso, Alessandro: *Preventing SQL Injection Attacks Using AMNESIA*. Teoksessa *Proceedings of the 28th International Conference on Software Engineering, ICSE '06*, sivut 795–798, New York, NY, USA, 2006. ACM, ISBN 1-59593-375-1. <http://doi.acm.org/10.1145/1134285.1134416>.
- [HZ09] Haixia, Yang ja Zhihong, Nan: *A database security testing scheme of web application*. Teoksessa *Computer Science Education, 2009. ICCSE '09. 4th International Conference on*, sivut 953–955, July 2009.
- [jav15] *Using Prepared Statements*. Tarkasteltu: 02.03.2015. <http://docs.oracle.com/javase/tutorial/jdbc/basics/prepared.html>.
- [KT08] Kemalis, Konstantinos ja Tzouramanis, Theodores: *SQL-IDS: A Specification-based Approach for SQL-injection Detection*. Teoksessa *Proceedings of the 2008 ACM Symposium on Applied Computing, SAC '08*, sivut 2153–2158, New York, NY, USA, 2008. ACM, ISBN 978-1-59593-753-7. <http://doi.acm.org/10.1145/1363686.1364201>.
- [Pou02] Poulsen, Kevin: *Guesswork Plagues Web Hole Reporting*. March 2002. <http://www.securityfocus.com/news/346>.

- [RG07] Roichman, Alex ja Gudes, Ehud: *Fine-grained Access Control to Web Databases*. Teoksessa *Proceedings of the 12th ACM Symposium on Access Control Models and Technologies*, SACMAT '07, sivut 31–40, New York, NY, USA, 2007. ACM, ISBN 978-1-59593-745-2. <http://doi.acm.org/10.1145/1266840.1266846>.
- [Sal13] Saldago, Roberto: *SQL Injection Optimization and Obfuscation Techniques*. 2013.
- [Sri14] Srivastava, M.: *Algorithm to prevent back end database against SQL injection attacks*. Teoksessa *Computing for Sustainable Global Development (INDIACom), 2014 International Conference on*, sivut 754–757, March 2014.
- [ST13] Shar, L.K. ja Tan, Hee Beng Kuan: *Defeating SQL Injection*, nide 46. March 2013.
- [SZAM14] Sadeghian, A., Zamani, M. ja Abd Manaf, A.: *SQL injection vulnerability general patch using header sanitization*. Teoksessa *Computer, Communications, and Control Technology (I4CT), 2014 International Conference on*, sivut 239–242, Sept 2014.
- [SZI13] Sadeghian, A., Zamani, M. ja Ibrahim, S.: *SQL Injection Is Still Alive: A Study on SQL Injection Signature Evasion Techniques*. Sept 2013.
- [TGZ14] Thomé, Julian, Gorla, Alessandra ja Zeller, Andreas: *Search-based Security Testing of Web Applications*. Teoksessa *Proceedings of the 7th International Workshop on Search-Based Software Testing*, SBST 2014, sivut 5–14, New York, NY, USA, 2014. ACM, ISBN 978-1-4503-2852-4. <http://doi.acm.org/10.1145/2593833.2593835>.
- [TMH10] Tajpour, A., Massrum, M. ja Heydari, M.Z.: *Comparison of SQL injection detection and prevention techniques*. Teoksessa *Education Technology and Computer (ICETC), 2010 2nd International Conference on*, nide 5, sivut V5–174–V5–179, June 2010.
- [VAM09] Vieira, M., Antunes, N. ja Madeira, H.: *Using web security scanners to detect vulnerabilities in web services*. Teoksessa *Dependable Systems Networks, 2009. DSN '09. IEEE/IFIP International Conference on*, sivut 566–571, June 2009.