doi:10.16018/j.cnki.cn32-1650/n.201501002

一种通用的彻底解决 SQL 注入漏洞的编码方法

付熙徐,龚希章
(上海海洋大学 现教中心,上海　201306)

摘要:SQL 注入攻击的本质是在字符数据中插入可执行的代码。对字符串进行 16 进制 ASCII 编码即可彻底防止在数据中插入任何可执行的代码,从而达到彻底阻止 SQL 注入攻击的目的。而这种编码能保留字符数据的所有性质,不影响基于该字段的连接、比较、排序等操作。

关键词:信息安全;SQL 注入;ASCII 编码;16 进制
中图分类号:TP309.2　　**文献标识码**:A　　**文章编号**:1671-5322(2015)01-0005-04

Web 系统的安全对高校网站和应用系统至关重要。近年来对 Web 系统的攻击时有发生,而对漏洞和攻击的防范往往滞后于攻击事件的发生,系统管理人员若不能及时彻底解决导致攻击的漏洞问题,多次发生的攻击事件将会导致非常严重的后果。

SQL 注入攻击是一种常见的网络攻击方法[1],通过在输入数据(如认证界面的表单)中添加特殊字符和 SQL 命令以达到获取系统信息、篡改数据甚至控制被攻击系统的目的。近年来,基于 SQL 注入的网络攻击事件给业界带来了极大的损失[2],很多系统由于被注入攻击而导致系统被控制、数据丢失、信息泄露等严重后果。虽然大量的防范方法和措施被提出,但仍然存在一些隐患,有些方法阻止攻击的代价也较大。另外,被注入攻击过的数据库中可能会保存有可执行的编码,造成二阶注入。

通过对各种 SQL 注入攻击和防范方法的分析,可以发现输入的数据是 SQL 注入攻击的关键。在系统设计时对输入数据的合理编码,使得输入数据没有可能被执行就可以完全避免注入攻击的发生。用编码替换原数据库内容,也可解决被攻击数据库中的二阶注入问题。

文章首先对 SQL 注入攻击进行了详细的分析,在此基础上分析和评价了一些常用的 SQL 注入攻击的防范方法;之后,评价标准、编码方案和实施方法将被提出;最后,通过与禁用字符集法、黑名单法、正则表达式法、参数化法等常用方法进行比较,得出该方法容易实现且可以彻底解决 SQL 注入攻击的结论。

1　SQL 注入攻击分析

SQL 注入攻击,顾名思义,就是在 SQL 语句中注入攻击语句以获得系统的访问权、篡改系统数据以及系统控制权的攻击活动。相关文献都对 SQL 注入攻击的方式进行了介绍[1,3-6]。总的来说,SQL 攻击过程和攻击方式可用图 1 概括。

首先,攻击者可以在 SQL 语句中加入一些特殊字符和语句如"'"等,造成服务器报错探测一些系统信息,如图 2 所示。

更有甚者通过插入错误的查询条件可以从报错信息中获得表和一些字段的名称[4],这些信息对进一步实施攻击非常有用。

一些简单的 SQL 注入手段就能让攻击者获得某些用户的权限。假设用户表名为 users,以下 SQL 语句的执行结果就是选择所有的用户(--使后面的语句成为注释)。

select * from users where username =″ or 1 = 1 --′ and password =″

对于认证系统,只需在用户名输入框中输入′ or 1 = 1 --就相当于获取了第一个用户的权限。如果知道一些重要的用户名,甚至可以获得系统管理员的权限。

如果已通过攻击获取表和字段的名称则可以

收稿日期:2014-09-29
作者简介:付熙徐(1981-),男,上海人,工程师,博士生,主要研究方向为计算机系统与人工智能。

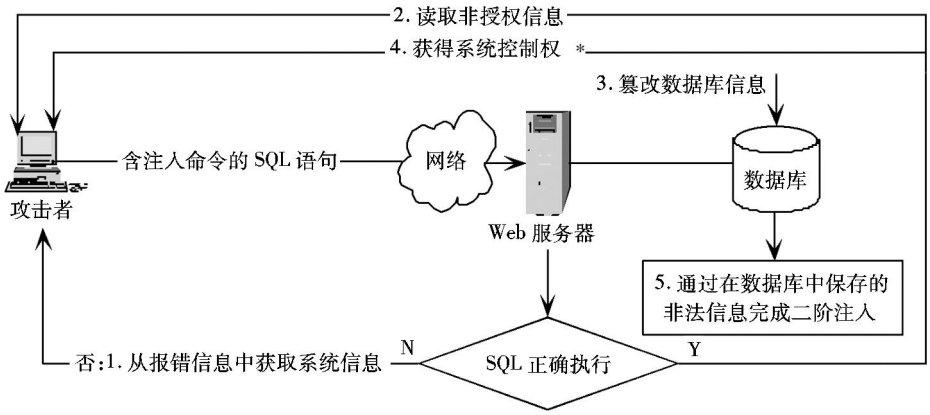

图 1　SQL 注入攻击的方式和过程

Fig. 1　SQL injection process

```
Microsoft VB Script 编译器错误 错误 '800a03f6'
缺少 'End'
/iisHelp/common/500-100.asp, 行242
Microsoft OLE DB Provider for ODBC Drivers 错误 '80040e14'
[Microsoft][ODBC SQL Server Driver][SQL Server]Unclosed quotation mark before the character string '''.
/salary/chklogin.asp, 行25
```

图 2　SQL 注入攻击造成的错误泄露了数据库的类型

Fig. 2　Error caused by injection exposed the type of database

进一步用 update、delete、drop 等语句对数据库进行篡改。

如果使用的数据库支持对操作系统的修改,而不幸系统管理员又使用了高权限用户的话,攻击者可获取系统的控制权。以 SQL Server 为例,数据库系统管理员用户可以用 xp_cmdshell 命令创建操作系统管理员用户 ,从而控制整个计算机。

还有一种情况,就是攻击者利用数据库中存储的数据进行二阶攻击[1]。这种攻击有一定隐藏性,尤其在系统被攻击 1 次后,这种情况比较常见。举 1 个简单的例子,假设在 A 表中某条记录的值是字符"'",程序将其保存到变量 a 中,程序中执行以下 SQL 字符串 select * from B where X ='"+a+"'"时(+为连接符),就会导致执行错误,达到攻击系统的目的。

2　SQL 注入防御相关工作

根据对 SQL 注入过程的分析,SQL 注入攻击的本质是数据的可执行性。要阻止 SQL 注入入侵,关键是要避免生成含可执行命令或错误提示,详细来说做到以下 4 点可以防止 SQL 注入:

(1)避免产生和显示错误信息;

(2)避免输入数据被插入可执行命令;

(3)避免数据库系统对操作系统关键信息的操作;

(4)避免数据库中的数据被用于二阶注入。

针对以上几条,很多文献提出了一些应对方法。有的文献提出了禁止非法输入数据和建立黑名单[4,5]的方法;有的文献提出输入数据转义和预处理[4]之类的方法;还有从开发插件等角度[5]考虑解决 SQL 注入问题的方案;SQL 参数化[1,7]被认为是有效防止一阶注入的方法。用关联分析等数据挖掘方法[4]检测注入攻击也是有效的处理方式。对于管理员,常用的办法是加强权限管理[3]和删除一些可控制操作系统的命令如 SQL Server 的 xp_cmdshell 命令以尽量减少损失。对于二阶攻击,目前还没有有效的方法。

3　通过编码数据库内容防止注入攻击

3.1　基本方法和评价标准

注入攻击的本质就是使本来应是数据的输入 I 变成可执行的输入 I_e。设 A 是所有可能作为输入的字符,A_i 是输入中可能出现的所有字符的集合,E 是导致字符串可执行的所有字符的集合,目前所有直接阻止注入攻击的方法都是使输入满足式 1:

$$A_i \cap E = \phi \quad (1)$$

易证只要满足式 1,就可以彻底避免注入攻击,但是如果因此导致 $A_i \neq A$,就会出现副作用。数据编码的方式就是根据式 1 在防止注入攻击的同时,经过编码使得 $A_i = A$ 也成立。

定义　一个编码系统可表示为一个四元组 $S(A_i, A_o, f_E, f_D)$。其中 A_i 是输入字母表,A_o 是输出字母表,f_E 是编码函数,f_D 是解码函数。

一个好的反注入编码函数应符合以下要求:

(1)对任意字符串 s,有 $f_D(f_E(s)) = s$

(2)$A_o \cap E = \phi$

(3)$A_i = A$

(4)对于任意字符 c_1 和 c_2,有 $(c_1 > c_2) \rightarrow (f_E(c_1) > f_E(c_2))$

(5)对于常用字符串函数 f 和字符串 s,有 $\exists g \mid f_D(g(f_E(s))) = f(s)$

条件(1)和(2)是正确编码和反注入攻击的基本要求,条件(3)保证可以输入任何数据,条件(4)和(5)保证数据库的可操作性。

3.2　ASCII – 16 进制编码

16 进制 ASCII 编码(即用 2 位 16 进制数表示一位字符)就是一种良好的编码方法,其编码函数见式 2。

$$f_E(c) = \mathrm{String}(\mathrm{Hex}(\mathrm{ASCII}(c))) \quad (2)$$

该函数将每个字符的 ASCII 码转换为 2 位 16 进制字符串 ,其解码函数 f_D 对每个 2 位字符进行 16 进制和 ASCII 解码,可以准确还原原数据,符合条件(1)。

该编码方法输出字母表如式 3 所示:

$$Ao = \{1,2,3,4,5,6,7,8,9,0,a,b,c,d,e,f\} \quad (3)$$

对于 SQL 语言和该字母表,该编码方法符合条件(2);由于可以编码任意字符,该编码方法符合条件(3);由于 ASCII 码编码顺序与字符串偏序一致,16 进制转换也保持了原来的次序,易证此方法符合条件(4);由于字符和编码的一一对应性,条件(5)也可被满足。字符串连接可以直接执行,字符串匹配则需要将 2 位编码作为一个单位,也容易做到,唯一的问题是由于编码的加长,查询的性能不如短的字符串。下面我们将进行一系列实验确认编码对性能的影响及实践中要注意的问题。

3.3　实验和结果

图 3 是一个编码转换和逆转换的例子,并记录了转换所用的时间。

图 3　编码解码及消耗的时间

Fig. 3　Encoding and decoding result and time consumed

从图 3 可以看出,即使原始字符串中含有非法字符,仍然可以正确地编码和解码,编码和解码消耗的时间也非常短。为进一步了解长编码对数据库操作的影响,我们设计了以下几个测试:

(1)对 100 条编码和未编码数据进行 100 次选择操作;

(2)对 100 条编码和未编码数据进行 100 次投影操作;

(3)插入 100 条编码和未编码数据(只有一个字段);

(4)对两张有 100 条数据的表用编码和未编码数据进行 10 次自然连接。

实验结果如图4所示,可见编码长度对数据

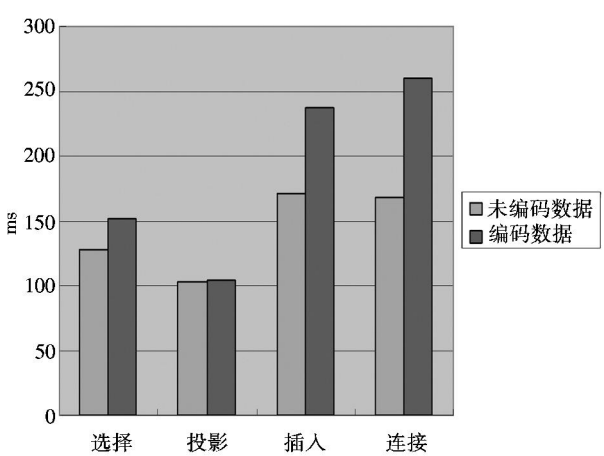

图 4　编码对数据库各种操作的影响

Fig. 4　Affects to database operations of the encoding

库的选择、插入、连接操作都有一定影响,但对于小规模数据影响不明显。对于大规模数据的连接操作,需要慎重考虑编码后的性能(外键长度不可太长)。

4 讨论

各种文献提到的防止SQL注入方法有黑名单法、参数化法、正则替换法、驱动程序和插件法以及加强权限管理。表1从防治效果、副作用、通用性、易用性和系统开销4方面比较编码法和其他各种方法。

上面的方法中,多数算法都只是对输入的控制,不能防治二阶注入。只有正则替换法和16进制编码法可用于数据库编码,但正则替换法无法实现编码系统的要求(4)和(5),不能对编码的数据库进行直接操作,因此可操作性不佳,而16进制编码法却有较好操作性。

表1 各种SQL注入防治方法的比较

Table 1 Comparasion of SQL Inj ction Prevention Methods

方法	防治效果	副作用	通用性	易用性	系统开销
黑名单法	较好	错误阻止正常输入和用户访问	好	容易	小
参数化法	可防全部一阶注入	无	部分数据库支持	较难	较小
正则替换法	可防大部分注入	无	较好	较难	较大
驱动程序和插件法	较好	无	一般	较难	较大
加强权限管理	有限	过度限制,影响功能	好	容易	很小
16进制编码法	彻底防治	几乎没有	好	较易	较大

综合看来,16进制编码法是一种良好的SQL注入防止方法,是唯一可彻底解决二阶注入的方案。由于16进制编码法可以防止二阶注入,对被攻击过的数据库实施转码也是可实施的方案,但由于16进制编码较长,在性能上有一定限制,将来可以考虑选择更短的编码系统以获得更佳的性能。

5 结论

对数据库中从外部输入的字段进行16进制编码并将这种形式存入数据库中可以彻底解决SQL注入问题。作为对被攻击过系统的补救措施,16进制编码也能完全防御二阶注入。但是出于对性能的考虑,不建议系统使用很长的外键。

参考文献:

[1] Justin Clarke. SQL注入攻击与防御[M]. 北京:清华大学出版社,2013:10.
[2] tangxs. 盘点2011年6月黑客攻击事件[J]. 网络与信息,2011(8):66-68.
[3] 余志高,周国祥. Web应用中SQL注入攻击研究[J]. 信息安全与通信保密,2010(4):81-83.
[4] 张博. SQL注入攻击与检测技术研究[J]. 信息安全与通信保密,2010(5):90-92.
[5] 杨小丽,袁丁. 防SQL注入攻击的数据库驱动设计与实现[J]. 计算机工程与设计,2010(12):2 691-2 694.
[6] 陈小兵,张汉煜,骆力明,等. SQL注入攻击及其防范检测技术研究[J]. 计算机工程与应用,2007,43(11):150-152.
[7] KHIN SHAR L, KUAN TAN H B. Defeating SQL injection[J]. Computer, 2013,46(3):69-77.

An Encoding Method to Solve the SQL Injecting Problem Thoroughly

FU Xixu, GONG Xizhang
(Shanghai Ocean University, Shanghai 201306, China)

Abstract: The essence of SQL injection attack is inserting executable codes in character fields. Encoding the characters into hexadecimal ASCII codes can prohibit the insertion of executable codes so as to prevent SQL injection thoroughly. On the other hand, all properties of stings are remained in this encoding method. Encoded string can be used as foreign keys as well as used in comparing and ordering.

Keywords: information security; SQL injection; ASCII encoding; hexadecimal number system

(责任编辑:张英健)

一种通用的彻底解决 SQL注入漏洞的编码方法

作者： 付熙徐， 龚希章， FU Xixu， GONG Xizhang
作者单位： 上海海洋大学现教中心,上海,201306
刊名： 盐城工学院学报（自然科学版）
英文刊名： Journal of Yancheng Institute of Technology (Natural Science Edition)
年，卷(期)： 2015(1)