工业和信息化“十三五”人才培养规划教材

信息安全技术类

Database Security Technology

数据库安全技术

◎ 贺桂英 周杰 王旅 主编　◎ 焦冬艳 郭玲 副主编

- 全书共分 8 章，书中每个**知识点**都有相应的**实例说明**，帮助读者**理解**和**消化**所学的内容
- 全书采用由**浅**入**深**的**递进式**讲解思路，力求每个内容的介绍从**简单**到**复杂**，一步一个实例说明，使读者不厌倦、有激情、想学习
- 数据库安全技术**内容分散**、**理论性强**，实际掌握**应用技术**更加重要

工信出版集团　人民邮电出版社 POSTS & TELECOM PRESS

工业和信息化“十三五”人才培养规划教材
信息安全技术类

Database Security Technology

数据库安全技术

◎ 贺桂英 周杰 王旅 主编　◎ 焦冬艳 郭玲 副主编

人 民 邮 电 出 版 社
北 京

图书在版编目（CIP）数据

数据库安全技术 / 贺桂英，周杰，王旅主编. — 北京 : 人民邮电出版社，2018.3
工业和信息化“十三五”人才培养规划教材. 信息安全技术类
ISBN 978-7-115-44230-7

Ⅰ. ①数… Ⅱ. ①贺… ②周… ③王… Ⅲ. ①关系数据库系统－安全管理 Ⅳ. ①TP311.138

中国版本图书馆CIP数据核字(2017)第321740号

内 容 提 要

本书共 8 章，重点介绍与数据库安全相关的理论和技术，主要内容包括数据库安全基础、数据库安全层次、SQL 和 Web 应用基础、SQL 注入与防范、数据库访问控制、数据库备份与恢复、数据加密与审核、大数据与安全。

本书适合作为高等院校信息安全、信息管理、大数据等相关专业的教材，也可作为对数据库安全感兴趣的读者的自学教材。

◆ 主　　编　贺桂英　周　杰　王　旅
副 主 编　焦冬艳　郭　玲
责任编辑　范博涛
责任印制　马振武

◆ 人民邮电出版社出版发行　　北京市丰台区成寿寺路 11 号
邮编　100164　　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
固安县铭成印刷有限公司印刷

◆ 开本：787×1092　1/16
印张：12.75　　2018 年 3 月第 1 版
字数：314 千字　　2018 年 3 月河北第 1 次印刷

定价：39.80 元

读者服务热线：(010)81055256　印装质量热线：(010)81055316
反盗版热线：(010)81055315
广告经营许可证：京东工商广登字 20170147 号

前言 FOREWORD

随着网络技术的飞速发展，信息安全问题备受关注。数据库系统中存储了大量数据且一般都是集中存放的，这些数据通过网络为许多最终用户所共享，因此其安全性问题日益突出。数据库管理系统本身能够实现数据库安全性控制的方法和技术有很多，包括用户标识和鉴别、存取控制、视图机制、审计和数据加密等；还有基于 Web 应用的 SQL 注入防范技术等。学习数据库安全技术旨在保护数据库以防止非法用户使用数据库，从而造成数据泄露、更改或破坏等。

本书是为应用型院校信息安全及其相关专业编写的一本数据库安全技术实用教材。全书共分 8 章。第 1 章介绍数据库安全的基础知识；第 2 章介绍数据库安全层次，即网络安全、服务器安全和数据库安全；第 3 章介绍 SQL 和 Web 应用基础知识；第 4 章介绍 SQL 注入原理和相关的防御技术；第 5 章介绍数据库管理系统中的访问控制的相关知识和应用设置；第 6 章介绍数据库备份与恢复的相关知识和技术；第 7 章介绍如何对数据库中的数据进行加密和审核；第 8 章介绍大数据应用以及相关的安全技术。本书的编写注重实践操作技能的培养，同时兼顾理论知识，通过讲授和实操两条主线来安排课程内容，旨在使读者通过实例和实操内容讲解来掌握数据库管理系统本身的安全技术和 SQL 注入防范技术。

本书的主要特色如下。

① 内容全面，实例丰富。书中每个知识点都有相应的实例说明，帮助读者理解和消化所学的内容。

② 递进式的讲解思路。全书采用由浅入深的递进式讲解思路，力求每个内容的介绍从简单到复杂，一步一个实例说明，使读者不厌倦、有激情、想学习。

③ 注重技术应用。数据库安全技术比较分散，理论性强，实际掌握应用技术更加重要。因此，本书编制了大量操作实例，并配有详细的操作说明，帮助读者掌握数据库安全相关技术。

本书第 1 章、第 3 章、第 8 章由贺桂英教授编写，第 4 章、第 5 章、第 6 章、第 7 章和附录部分由周杰老师编写，第 2 章由王旅老师编写，焦冬艳和郭玲老师参与了教材部分章节的编写，全书由贺桂英教授审阅定稿。在本书的编写过程中，我们得到了蓝盾信息安全技术股份有限公司田文春博士和深信服科技股份有限公司广东区安全专家袁小辉的大力支持，在此向对本书编写提供帮助的老师和工程技术人员表示衷心的感谢！

由于编者水平有限，书中疏漏之处在所难免，敬请相关专家和广大读者批评指正。

作者

2017 年 11 月于广州

目 录 CONTENTS

Chapter 1

第1章 数据库安全基础

随着信息技术的发展，特别是移动互联网的飞速发展，基于网络的分布式信息系统已经在各个行业，特别是电子商务、政府办公、企业事务管理等领域广泛应用。数据库作为信息承载的主体，是信息管理系统的核心和基础，数据的机密性、完整性、可用性、隐私性面临着严重的挑战和风险。保护数据不泄露或不被窃取是数据库管理员和应用系统开发者的重要工作。

本章介绍了数据库相关概念及发展、SQL Server 数据库、数据库面临的主要威胁来源及安全对策。

1.1 数据库相关概念及发展

数据库（Database，DB）是存储数据的一个集合，一个数据库系统（Database System，DBS）中可以有多个数据库，数据库管理系统（Database Management System，DBMS）是管理数据库的软件，数据库管理员（Database Administrator，DBA）则是使用数据库管理系统管理数据库的人。下面我们对数据库相关概念进行详细介绍。

1.1.1 数据

数据是指对客观事物进行记录并可以鉴别的符号，是对客观事物的性质、状态以及相互关系等进行记载的物理符号或这些物理符号的组合，是可识别的、抽象的符号。简而言之，数据是符号的集合，是对事物特性的描述。这里的“符号”不仅仅指文字、字母、数字和其他特殊符号，还包括图形、图像、声音等多媒体的表示。例如，“0、1、2”“阴、雨、下降、气温”“学生的档案记录、货物的运输情况”等都是数据。

人们通过获得、识别自然界和社会的不同信息来区别不同事物，得以认识和改造世界。信息是加载于数据之上，对数据进行具有含义的解释。数据和信息是不可分离的，信息依赖数据来表达，数据则生动具体地表达出信息。数据是符号，是物理性的，信息是对数据进行加工处理之后所得到的并对决策产生影响的数据，是逻辑性和观念性的；数据是信息的表现形式，信息是数据有意义的表示。数据是信息的表达、载体，信息是数据的内涵，是形与质的关系。数据本身没有意义，数据只有对实体行为产生影响时才成为信息。

例如，在学生基本信息表中，如果单独看待学号、姓名、性别、身份证号、专业、班级，它们就是数据，如果将这些数据共同组合起来看待，则就是学生基本信息。

1.1.2 数据表及其设计规范

数据表（或称表，Data Table）是数据库最重要的组成部分之一。数据库只是一个框架，数据表才是其实质内容。如教务管理系统中，教务管理数据库包括：学生基本信息表、班级表、课程表、专业表、成绩表和毕业表等，这些表用来管理学生入学到毕业期间产生的数据，这些数据表通过一定的规则相互关联，相互作用，共同构成、管理学生学籍信息。

数据表是以行列的形式组织及展现数据的，跟 Excel 表格一样，都要有一个表头（字段），但数据表中存储着若干相互关联的数据，同一列的数据属性相同，同一行的数据不能重复。图 1-1 为教务管理系统中学生信息表部分数据，为了保密，学号、姓名和身份证号被部分隐藏。

	学号	姓名	性别	民族	报读专业层次	专业名称	报读类型	入学学期	身份证号
1	1715000120000**	黄冠...	男	汉族	本科	信息安全	学历	2017春季	441900********5359
2	1716000120000**	王俊...	男	汉族	本科	信息安全	学历	2017春季	441827********7930
3	1716000120000**	杜志...	男	汉族	本科	信息安全	学历	2017春季	442000********1594
4	1716000120000**	王政...	男	汉族	本科	信息安全	学历	2017春季	442000********2055
5	1716000120000**	唐梓...	男	汉族	本科	信息安全	学历	2017春季	442000********2331
6	1716000120000**	关树...	男	汉族	本科	信息安全	学历	2017春季	442000********7657
7	1716000120000**	李立...	男	汉族	本科	信息安全	学历	2017春季	442000********6671
8	1716000920000**	卢艺...	男	汉族	本科	信息安全	课程	2017春季	442000********0231
9	1716000920000**	郭子...	男	汉族	本科	信息安全	课程	2017春季	442000********7657
10	1716000920000**	黄悦...	男	汉族	本科	信息安全	课程	2017春季	440506********003X

图1-1 教务管理系统中学生信息表部分数据

数据库的设计范式是数据库设计所需要满足的规范，满足规范的数据库是简洁的、结构明晰的、无数据冗余的，同时，不会造成操作（插入 INSERT、删除 DELETE 和更新 UPDATE）异常。

数据表的建立需要符合一定的要求，就是至少要满足数据库第三范式，否则与 Excel 表没有区别，发挥不出数据库的优势。

- 第一范式（1NF）：强调的是列的原子性，即列不能再分成其他几列，也就是说一列只代表一个属性，不能代表多个属性。在学生信息表（学号、姓名、性别、电话）中，“电话”字段一般有固定电话和手机，因此，不符合 1NF，需要将“电话”字段拆分，得到学生信息表（学号、姓名、性别、固定电话、手机）。1NF 比较容易判断。
- 第二范式（2NF）：首先满足 1NF，其次表必须有一个主键，且没有包含在主键中的列必须完全依赖于主键，而不能只依赖于主键的一部分。

例如，在成绩表（学号、课程号、课程名称、成绩、学分）中，主键应该是学号和课程号，成绩完全依赖学号和课程号，但课程名称、学分只依赖课程号，因此成绩表不符合 2NF，必须将成绩表拆分为成绩表（学号、课程号、成绩）和课程表（课程号、课程名称、学分）才符合 2NF。

不符合 2NF 的设计容易产生冗余数据和操作异常，例如课程名称和学分，如果同一门课程有 N 个学生选修，则课程名称和学分就要重复 $N-1$ 次。其次，更新、插入和删除都会产生异常，如果想添加一个没有学生选修的课程，这是根本行不通的。删除和更新都会操作多条记录，否则就会出现数据不一致的情况。

- 第三范式（3NF）：首先满足 2NF，其次非主键列必须直接依赖于主键，不能存在传递依赖，即不能存在非主键列 A 依赖于非主键列 B，非主键列 B 依赖于主键的情况。

在学生基本信息表（学号、姓名、年龄、专业名称、毕业学分）中，主键是学号，因此在表中存在学号决定专业名称，专业名称决定毕业学分的传递依赖，因此不符合 3NF；可以将其拆分为学生基本信息表（学号、姓名、年龄、专业名称）和专业信息表（专业名称、毕业学分）两个表就符合 3NF。

不符合 3NF 的设计同样容易产生冗余数据和操作异常。

1NF 保证字段不可拆分，2NF 消除非主属性对主键的部分函数依赖，3NF 消除非主属性对主键的传递函数依赖。表的建立满足 3NF 就满足需求，此外，还可以升级到巴斯-科德范式（Boyce-Codd Normal Form，BCNF），这样就消除了主属性对主键的传递函数依赖。

1.1.3　数据库

数据库是按照数据结构来组织、存储和管理数据的仓库。它出现于 20 世纪 60 年代，伴随着信息技术的发展，在 20 世纪 90 年代后期得到迅速发展，进入 21 世纪，数据库理论日臻完善，管理和存储功能也日益智能化。

数据库是以一定方式存储在一起、能为多个用户共享、具有尽可能小的冗余度、与应用程序彼此独立的数据集合。通俗地说，数据库是一个存储数据的仓库，这些数据是按照一定的数学模型组织起来的，具有较小冗余度和较高的数据独立性，能够与其他用户共享数据，是有组织、有管理的数据集合。

教务管理系统数据库中包括学生基本信息表、班级表、课程表、专业表、成绩表和毕业表等众多数据表，这些表对象及其视图、存储过程等对象共同组成了一个数据库。

1.1.4 数据库管理系统

数据库管理系统是一种操作和管理数据库的大型软件，用于建立、使用和维护数据库，可对数据库进行统一的管理和控制，以保证数据库的安全性和完整性。用户可以通过数据库管理系统访问数据库，数据库管理员也可以通过它进行数据库的维护工作。

数据库管理系统提供数据定义语言（Data Definition Language，DDL）、数据操作语言（Data Manipulation Language，DML）、数据控制语言（Data Control Language，DCL）和数据查询语言（Data Query Language，DQL），供用户定义数据库的模式结构与权限约束，实现对数据的创建、删除、修改、查询及对用户的授权等操作。

- 数据定义语言，用于建立、修改数据库的结构，定义数据库的三级模式结构、两级映像以及完整性约束和保密限制等约束。
- 数据操作语言，为用户提供 UPDATE、INSERT 和 DELETE 功能，完成对数据库的更新、插入和删除操作。
- 数据控制语言，用来授予或回收访问数据库的某种特权，并控制数据库操作事务发生的时间及效果，对数据库实行监视等功能，如 GRANT、ROLLBACK 和 COMMIT。
- 数据查询语言，基本结构是由 SELECT 子句、FROM 子句、WHERE 子句组成的查询块。

1.1.5 数据库的发展

数据模型是数据库的核心和基础，决定着数据在数据库中的存储策略，数据库技术的发展阶段是以数据模型的发展演变为主要标志，主要分为三个阶段：第一代是层次、网状数据库系统，第二代是关系数据库系统，第三代是面向对象数据库系统。

1. 层次数据库

层次模型是最早出现的数据模型，它是以树状（层次）结构来表示实体类型及实体间联系的数据模型。现实世界中，许多实体之间的联系本来就呈现出一种很自然的层次结构，如家族关系、行政结构等。层次模型是用树状结构表示实体与实体之间的联系，树中的每一个节点代表一个记录类型，树状结构表示实体类型之间的联系，记录之间的联系通过指针实现，查询效率高。层次模型的限制条件是：① 有且只有一个节点，无父节点，此节点代表树的根；② 其他节点有且只有一个父节点，是树的枝。

采用层次数据模型的数据库称为层次数据库系统，典型代表是 IBM 公司 1968 年推出的 IMS（Information Management System），这是一个大型的商用数据库管理系统，曾经得到广泛的应用。

2. 网状数据库

在现实世界中，事物之间的联系非常复杂，并非都是层次关系的，因此利用层次结构来表示非树状结构就非常不直接，为了实现非树状结构的表示，因此出现了网状模型。网状模型可以有效地解决非树状结构的表示。

网状模型允许一个以上的节点无双亲以及一个节点可以有多于一个的双亲。网状模型能够表示比层次模型更具有普遍性的结构，不受层次模型两个限制的制约，可以直接地去描述现实世界，而层次模型只是它的一个特例。

与层次模型相同的是，网状模型中也是以记录为数据的存储单位，一个记录包含若干数据项，

该数据项可以是多值的、复合的数据。每个记录有一个唯一能够标识它的内部标识符，称为码（Database Key，DBK），它是在记录存入数据库时由数据库管理系统自动赋予。码可以看作记录的逻辑地址，可作为记录的替身，或用于寻找记录。网状数据库是导航式（Navigation）数据库，用户在操作数据库时，不但要说明做什么，还要说明怎么做。例如在查找语句中，不但要说明查找的对象，而且要规定存取路径。

1964 年美国通用电气公司 Bachman 等人开发了第一个网状数据库管理系统 IDS（Integrated Data Store），奠定了网状数据库的基础。在 20 世纪 70 年代，曾经出现过大量的网状数据库管理系统产品，比较著名的有 Cullinet 软件公司的 IDMS、Honeywell 公司的 IDSII、Univac 公司的 DMS1100、HP 公司的 IMAGE 等。

网状模型对层次结构和非层次结构的事物都能够进行比较自然的模拟，在关系数据库之前，网状数据库管理系统比层次数据库管理系统应用得更普遍，在数据库发展史上，网状数据库曾经占有重要的地位。

3. 关系数据库

1970 年，IBM 的研究员 E.F.Codd 博士发表了论文《大型共享数据库的关系模型》，文章提出了关系模型的概念，后来陆续发表了多篇文章，奠定了关系数据库的理论基础。

关系模型是用二维表的形式表示实体与实体间的联系的数据模型，关系模型是当前的主流数据模型，它的出现使层次模型和网状模型逐渐退出了数据库的历史舞台。关系数据模型提供了关系操作的特点和功能要求，但对数据库管理系统的语言没有具体的语法要求，对关系数据库的操作是高度非过程化的，用户不需要指出特殊的存取路径，路径的选择由数据库管理系统优化机制来完成。Codd 在 20 世纪 70 年代初期的论文中论述了范式理论和衡量关系系统的 12 条标准，用数学理论奠定了关系数据库的理论基础。Codd 博士也以其对关系数据库的卓越贡献获得了 1981 年 ACM（Association for Computer Machinery）图灵奖。

关系模型有着严格的数学基础，是以集合论中的关系概念为基础发展起来的，无论实体还是实体间的联系均由单一的结构类型——关系来表示，简单清晰，便于理解和使用。在实际的关系数据库中，关系也称为表，它由表名、行和列组成。表的每一行代表一个元组，每一列称为一个属性，一个二维表就是一个关系，数据则看成是二维表中的元素，操作的对象和结果都是二维表。关系数据库是由若干个表组成的。

关系模型与层次模型、网状模型的本质区别在于数据描述的一致性，模型概念单一，描述实体的数据本身能够自然地反映它们之间的联系，而层次模型和网状模型使用指针来存储和体现联系。尽管关系数据库出现得比层次数据库和网状数据库晚，但它以完备的理论基础、简单的模型、说明性的查询语言和便于使用等特点得到了最广泛的应用。

目前关系数据库是市场上的主流，著名产品有甲骨文公司的 Oracle 数据库，Microsoft 公司的 SQL Server 和 Access 数据库，此外还有 MySQL、Sybase、Informix、Visual FoxPro 等。

4. 面向对象数据库

面向对象是一种认识方法学，也是一种新的程序设计方法学。把面向对象模型和数据库技术结合起来可以使数据库系统的分析、设计与人们对客观世界的认识最为相近。面向对象数据库系统是为了满足新的数据库应用需要而产生的新一代数据库系统。

面向对象模型具有以下优点。

① 易维护。采用面向对象思想设计的结构，可读性高，由于继承的存在，即使改变需求，

维护也只在局部模块，所以维护起来非常方便，成本也较低。

② 质量高。在设计时，可重用现有的、且在以前的项目中已被测试过的类，使系统满足业务需求并具有较高的质量。

③ 效率高。在软件开发时，根据设计的需要对现实世界的事物进行抽象，从而产生类。使用这样的方法解决问题，接近于日常生活和自然的思考方式，势必提高软件开发的效率和质量。

④ 易扩展。由于面向对象模型具有继承、封装、多态的特性，基于此设计出的系统结构具有高内聚、低耦合的特点，故面向对象数据库系统更灵活、更容易扩展，而且成本较低。

人工智能（Artificial Intelligence，AI）应用的需求（如专家系统）也推动了面向对象数据库的发展，专家系统常常需要处理各种复杂的数据类型。与关系数据库不同，面向对象数据库不因数据类型的增加而降低处理效率。由于这些应用需求，20 世纪 80 年代已开始出现一些面向对象数据库的商品和许多正在研究的面向对象数据库。多数的面向对象数据库被用于基本设计的学科和工程应用领域。

面向对象数据库研究的另一个进展是在现有关系数据库中加入许多纯面向对象数据库的功能。在商业应用中对关系模型的面向对象扩展着重于性能优化，即处理各种环境对象的物理表示的优化和增加 SQL 模型以赋予面向对象特征。如 Versant、UNISQL、O2 等， 它们均具有关系数据库的基本功能，采用类似于 SQL 的语言，用户很容易掌握。

1.1.6 主流数据库管理系统

目前，商品化的数据库产品主要以关系型数据库为主，技术也比较成熟。SQL Server、Oracle、MySQL、DB2 是当前数据库管理系统市场中四大主流产品，市场占有率很高。

1. 微软公司的数据库产品

微软公司除了 SQL Server 这个数据库产品外，还有一个桌面级的产品——Microsoft Access，它是 Office 的一个组件。Access 是一个小型的桌面数据库，应用简单，操作容易，主要用于少量数据的处理，在早期的网站和小型公司网站中，都采用了 Access 数据库。Access 的发展是随着 Office 版本发展的，有 Access 2000、Access 2003、Access 2007、Access 2010、Access 2013 和 Access 2016。

SQL Server 数据库是一个企业级的产品，正版的软件是收费产品，能够支持海量数据的存取，满足企业对快速响应、数据安全等要求。SQL Server 的版本也不断更新，比较成熟的是 SQL Server 2000，在当时非常流行。随着版本不断升级、功能不断增加，出现了 SQL Server 2005、SQL Server 2008、SQL Server 2012、SQL Server 2014、SQL Server 2016。

SQL Server 在事务处理、数据挖掘、负载均衡等方面功能强大，使数据库应用系统的开发、设计变得快捷方便，同时 SQL Server 在数据库市场占有相当高的份额。

本书主要以 SQL Server 2008 为基础来讲解数据库安全技术，大家如果有兴趣可以了解更高版本的功能。

2. 甲骨文公司的数据库产品

甲骨文（Oracle）公司的 Oracle 数据库应用非常广泛，与微软公司的数据库产品相比，其操作难度会大一些，对数据库管理人员要求较高。Oracle 数据库作为一个成熟的数据库产品，适用于大型数据库系统，稳定性高。

Oracle 公司旗下的另一个产品 MySQL 也是一个关系数据库管理系统，应用非常广泛，特别是在基于 Linux 系统的 Web 应用方面，MySQL 通常都是最佳的后台数据库（Linux 作为操作系统，Apache 或 Nginx 作为 Web 服务器，MySQL 作为数据库，PHP、Perl、Python 作为服务器端脚本解释器）。

3. IBM 公司的数据库产品

DB2 是 IBM 公司推出的一个重量级数据库产品，主要应用于金融领域等超大型应用系统，具有较好的可伸缩性，可支持从大型机到单用户环境，应用于所有常见的服务器操作系统平台下。

对用户来说，如何选择数据库管理系统呢？可以从构造数据库的难易程度、程序开发的难易程度、对分布式应用的支持、并行处理、可移植性和可扩展性、数据完整性、并发控制、容错能力、安全控制、支持多种文字处理能力、数据恢复的能力、成本等方面进行综合考虑，选择一个最适合自己的数据库管理系统。

1.1.7　数据库安全

数据库安全包含两层含义：第一层是指系统运行安全，系统运行安全通常受到的威胁主要指一些网络不法分子通过互联网、局域网等入侵电脑，使系统无法正常启动，或超负荷让电脑运行大量算法，并关闭 CPU 风扇，使 CPU 过热烧坏等破坏性活动；第二层是指系统信息安全，系统信息安全通常受到的威胁主要有攻击者入侵数据库，并盗取想要的资料。数据库系统的安全特性主要是针对数据而言的，包括数据独立性、数据安全性、数据完整性、并发控制、故障恢复等几个方面。

根据一些权威机构的数据泄露调查分析报告，以及对已经发生的信息安全事件进行技术分析，总结出信息泄露呈现出的两个趋势。

① 通过 B/S（Browser/Server，浏览器/服务器）模式应用，以 Web 服务器为跳板，窃取数据库中的数据，非常典型的攻击就是 SQL 注入攻击，主要原因是应用和数据库直接访问协议而没有任何控制。

② 数据泄露常常发生在内部，大量的运营维护人员直接接触敏感数据，导致以防外为主的网络安全失去了用武之地。

数据库安全必须在信息安全防护体系中处于被保护的核心位置，不易受到外部攻击者攻击，同时数据库自身应该具备强大的安全措施，能够抵御并发现入侵者。为了保证数据库安全，应该进行事前诊断、事中控制和事后分析三步操作。

① 事前诊断。利用数据库漏洞扫描系统扫描数据库，给出数据库的安全评估结果，暴露当前数据库系统的安全问题。利用专业的安全软件扫描应用系统，发现应用漏洞，及时堵住；模拟攻击者攻击，对数据库进行探测性分析，重点检查用户权限是否越权等，并收集应用系统漏洞和数据库的漏洞；检查敏感数据是否加密，危险的扩展存储过程是否禁用，端口是否安全，访问协议是否安全等。事前诊断越充分，越有利于系统安全。

② 事中控制。及时关闭数据库服务器，切断攻击者与数据库的联系。尽管会面临一定的损失，但总比数据丢失造成的危害小得多。

③ 事后分析。采用数据库审计功能，对数据库访问日志进行分析，及时发现可疑操作和可疑的数据，及时利用数据库备份进行数据恢复。

1.2 SQL Server 数据库

在 SQL Server 中，数据库是表、索引、存储过程和视图等数据库对象的集合，是数据库管理系统的核心内容。数据库的数据分别存储在不同的对象中，而这些对象有些是用户在操作时能够看得到的，如表、索引、存储过程和视图等，这就是数据库的逻辑存储结构。但至于这些对象是如何存放在磁盘中的，作为用户我们不需要关心，只有数据库管理员才能处理相应的物理实现，这些就是数据库的物理存储结构。

1.2.1 物理存储结构

1. 数据库文件

在物理存储方面，SQL Server 数据库至少具有两个操作系统文件：数据文件和日志文件。数据文件是用于存放数据库数据和数据库对象的文件，包括数据和对象，例如表、索引、存储过程和视图。数据文件分为主要数据文件和次要数据文件。日志文件包括恢复数据库中的所有事务所需的信息。为了便于分配和管理，可以将数据文件集合起来，放到文件组中。

- 主要数据文件：包括数据库的启动信息，并指向数据库中的其他文件。用户数据和对象可存储在此文件中，也可以存储在次要数据文件中。每个数据库有一个主要数据文件。主要数据文件的扩展名是.mdf。
- 次要数据文件：是可选的，由用户定义并存储用户数据。次要数据文件通过将每个文件放在不同的磁盘驱动器上，从而将数据分散到多个磁盘上。另外，如果数据库超过了单个 Windows 文件大小限制，可以使用次要数据文件，这样数据库就能继续增长。次要数据文件的建议扩展名是.ndf。
- 事务日志文件：保存用于恢复数据库的日志信息，每个数据库必须至少有一个日志文件。事务日志文件的建议扩展名是.ldf。

默认情况下，数据文件和日志文件被放在同一个驱动器上的同一个路径下。这是处理单磁盘系统采用的方法。但是，在生产环境中，这可能不是最佳的方法，建议将数据文件和日志文件放在不同的磁盘上，从而保护数据库的安全。

2. 文件组

每个数据库都有一个主要文件组，包括主要数据文件和未放入其他文件组的所有次要文件。用户可以创建自定义的文件组，用于将数据文件集合起来，以便于管理、数据分配和放置。

例如，可以分别在 3 个磁盘驱动器上创建 3 个文件 Data1.ndf、Data2.ndf 和 Data3.ndf，然后将它们分配给文件组 filegroup1。然后，可以明确地在文件组 filegroup1 上创建一个表。对表中数据的查询将分散到 3 个磁盘上，从而提高了性能。通过使用在 RAID（Redundant Array of Independent Disk，独立磁盘冗余阵列）条带集上创建的单个文件也能获得同样的性能提高。但是，文件和文件组能够轻松地在新磁盘上添加新文件。

- 主要文件组：包括主要文件的文件组。所有系统表都被分配到主要文件组中。
- 自定义文件组：用户首次创建数据库或以后修改数据库时明确创建的任何文件组。

系统有一个默认文件组，如果在数据库中创建对象时没有指定对象所属的文件组，对象将被分配给默认文件组。不管何时，只能将一个文件组指定为默认文件组。默认文件组中的空间必须

足够大，能够容纳未分配给其他文件组的所有新对象。在 SQL Server 中，PRIMARY 文件组是默认文件组，除非使用 ALTER DATABASE 语句进行了更改，否则系统对象和表仍然分配给 PRIMARY 文件组，而不是新的默认文件组。

1.2.2　逻辑存储结构

从数据库应用和管理角度看，SQL Server 数据库分为系统数据库和用户数据库两类。用户数据库存放的是与用户业务有关的数据，其中的数据是由用户来维护的。安装 SQL Server 时会自动安装 master、msdb、model、resource 和 tempdb。系统数据库是由 SQL Server 数据库管理系统自动维护，这些数据库用于存放维护系统正常运行的信息，通常不需要进行管理，只是了解就行。

- master 数据库：记录 SQL Server 实例的所有系统级信息。
- msdb 数据库：用于 SQL Server 代理计划警报和作业。
- model 数据库：用作 SQL Server 实例中创建的所有数据库的模板。对 model 数据库进行的修改（如数据库大小、排序规则、恢复模式和其他数据库选项）将应用于以后创建的所有数据库。
- tempdb 数据库：一个工作空间，用于保存临时对象或中间结果集。
- resource 数据库：一个只读数据库，包括了 SQL Server 2008 中的系统对象。系统对象在物理上保留在 resource 数据库中，但在逻辑上显示在每个数据库的 sys 架构中。

系统数据库存储着整个 SQL Server 的系统级信息，对系统来说非常重要，也是攻击者获得数据库信息的直接来源。在后续章节中，还会给大家介绍一些非常重要的系统函数或方法来查询系统级信息。

1.2.3　数据表

数据表是 SQL Server 存储数据的基本单元，用来存储数据和操作的逻辑结构，包括行和列。行是组织数据的单位，每一行表示唯一的一条记录；列主要描述数据的属性，每一列表示记录的一个属性，而且同一个表中的列名必须唯一。

SQL Server 数据表分为三类：系统表、用户自定义表（用户表）和临时表，其中视图也是一种特殊的表。

- 系统表（视图）：由系统自动创建并维护。master 数据库的数据表存储的是所有与 SQL Server 有关的信息，包括所有的登录账号、数据库的初始信息等。从 SQL Server 2012 开始，这些信息存储到 resource 数据库中了，并以系统视图的形式显示在每个数据库的 sys 架构中，而 master 数据库里只存储系统级信息，因此大家在 master 数据库中仅看到少数几个数据表。其次，msdb 数据库的数据表存储的信息大多与备份、恢复以及作业调度等相关。model 数据库中所有数据库的模板信息是以系统视图的形式存在的。
- 用户自定义表：也称为用户表，是用户根据自己的业务需要在 SQL Server 中建立的表，用户可以根据权限创建、修改、删除用户表。
- 临时表：临时表存储在 tempdb 数据库中，而不是存储在用户数据库中。临时表的创建和使用与用户表一样，只不过命名上临时表以“#”开始。例如，创建一个临时表：Create Table #ZhuanYe(ZhuanYeId int,ZhuanYeMingCheng varchar(20))；或清除一个临时表：Drop Table

#ZhuanYe。临时表一般用于数据处理，处理完数据后临时表失去意义，删除即可。

1.2.4 数据完整性

数据完整性是保证数据的正确性和相容性，防止不合语义或不正确的数据进入数据库，是否具备完整性关系到数据库能否真实地反映现实世界。数据库安全首先是保证数据的安全，如果被非法者写入冗余数据，会造成操作数据混乱，就会对正确的数据造成干扰。因此，在建立数据库的数据表时，首先要建立一定的规则，保证数据库的完整性。例如：学生的性别必须是男或女，不能为其他的任何值。课程成绩的范围是 0～100 分，学生学号是固定长度的一系列数字等，其实每个字段都有特殊的意义，也就有了特殊的取值范围。

数据库的完整性有 4 类，分别是实体完整性、域完整性、参照完整性和用户定义完整性。

- 实体完整性：表中有一个主键，其值不能为空或重复，且能唯一地标识对应的记录。实体完整性又称为行完整性，通过 PRIMARY KEY 约束、UNIQUE 约束、索引或 IDENTITY 属性等可以实现数据的实体完整性。例如，学生基本信息表中，学号为主键，每一个学号列能唯一地标识该学生对应的行记录信息，通过学号列建立主键约束实现学生基本信息表的实体完整性。
- 域完整性：列数据输入的有效性，又称为列完整性，通过 CHECK 约束、DEFAULT 约束、NOT NULL 约束、数据类型和规则等实现域完整性。CHECK 约束通过显示输入到列中的值来实现域完整性。例如，学生信息表中的性别列，取值只能是男或女，不能为其他值；在成绩表中，课程成绩的范围是 0～100 分，低于 0 分或高于 100 均为非法数值。
- 参照完整性：保证主表中的数据与从表中的数据一致，又称为引用完整性，在 SQL Server 中，通过定义主键与外键之间的对应关系实现参照完整性，参照完整性确保键值在所有表中一致。

主键：主表中能唯一标识每个数据行的一个或多个列。

外键：主表中的主键是从表中的一个字段，也就是说从表中的一个或多个列的组合是主表的主键。

在学生信息表中，学号是主键，在成绩表中学号和课程号的组合是主键，这时候，成绩表中的学号就是学生信息表的外键。

参照完整性就是要保证成绩表中的学号必须能够在学生信息表中存在。

- 用户自定义完整性：不属于其他任何完整性类别的特定业务规则，所有完整性类别都支持用户自定义完整性，包括 CREATE TABLE 中所有的列级约束和表级约束、存储过程和触发器。

1.2.5 数据约束

数据库的约束主要是保证数据库的完整性，常用的约束有 UNIQUE 约束、CHECK 约束、PRIMARY KEY 约束、FOREIGN KEY 约束、NOT NULL 约束。

- PRIMARY KEY 约束：主键约束，主键列的值必须是非空且不重复的。
- UNIQUE 约束：用在非主键列，保证非主键列的值是唯一、不重复的。但允许有一个空值 NULL。
- FOREIGN KEY 约束：从表中的一个或多个列的组合是主表的主键，要保证从表中的字段必须能够在主表中存在。
- NOT NULL 约束：不允许为空值，也就是说必须输入字符，包括空格字符。