

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теоретических основ
компьютерной безопасности и
криптографии

Восстановление баз данных Microsoft SQL Server

АВТОРЕФЕРАТ

дипломной работы

студентки 6 курса 631 группы
специальности 10.05.01 Компьютерная безопасность
факультета компьютерных наук и информационных технологий
Тихоновой Марии Игоревны

Научный руководитель

доцент, канд. юрид. наук

А.В. Гортинский

22.01.2024 г.

Заведующий кафедрой

д. ф.-м. н., доцент

М. Б. Абросимов

22.01.2024 г.

Саратов 2024

ВВЕДЕНИЕ

Microsoft SQL Server является одной из самых надежных систем хранения данных, которой доверяют крупнейшие предприятия и организации. В современном мире важнейшие корпоративные системы нуждаются в высоком уровне доступности своих данных. Надежность систем обеспечивается с помощью различных технологий. К ним можно отнести организацию регулярного резервного копирования, механизмы кластеризации, репликации, использование RAID-массивов для повышения отказоустойчивости систем.

Несмотря на то, что, как правило, системы надежны, всегда существуют различные риски, в результате которых данные могут стать недоступными. К ним относятся повреждение оборудования, заражение вирусами, ошибки в программном обеспечении, человеческий фактор.

В критических случаях, когда невозможно воспользоваться стандартными методами восстановления или технологиями высокой доступности, используются специальные программы для восстановления баз данных. Они позволяют восстановить максимально возможное число объектов базы данных и их содержимого.

В настоящий момент существует множество программ для восстановления файлов баз данных при их физическом повреждении. Большинство из них выполняют лишь предварительный просмотр, а восстановление данных предусмотрено в платных версиях программ. К таким инструментам относятся SysTools SQL Recovery и Recovery Toolbox For SQL Server v.2.4.21.0, которые будут рассмотрены в работе.

Целью данной работы является написание программы, позволяющей восстанавливать и просматривать структуру базы данных и ее объекты, а также данные, которые удалось восстановить в случае критических повреждений файла базы данных. Повреждения могут включать в себя системные области, а также области, в которых хранятся данные. При этом стандартными средствами просмотреть, восстановить структуру базы данных невозможно.

Дипломная работа состоит из введения, 2 разделов, заключения, списка использованных источников и 9 приложений. Общий объем работы – 91 страница, из них 43 страницы – основное содержание, включая 39 рисунков и 1 таблицу, список использованных источников из 20 наименований.

КРАТКОЕ СОДЕРЖАНИЕ

1 Теоретическая часть

1.1 Исследование проблемы восстановления баз данных

Повреждения файла базы данных могут возникнуть в любой области файла. В данной работе будут рассмотрены критические случаи повреждения, захватывающие блоки данных в области системных таблиц, пользовательских таблиц, специальных страниц, хранящих метаданные о файле базы данных.

Существуют программы, позволяющие просматривать структуру базы данных MS SQL Server, пользовательские данные, а также восстанавливать информацию из поврежденных файлов баз данных.

1.1.1 MS SQL Server Management studio

MS SQL Server Management studio – стандартный инструмент для работы с базами данных от Microsoft.

В случае рассматриваемых повреждений восстановление файла базы данных открыть его невозможно.

1.1.2 SYSTools SQL Recovery

Инструмент для восстановления баз данных SysTools SQL Recovery используется для восстановления поврежденных mdf-файлов.

В случае повреждения страницы заголовка файла программа восстанавливает все пользовательские объекты и данные, но не может восстановить имя базы данных.

При попытке восстановления файла с повреждением области системной таблицы sys.syscolpars, программа не может найти информацию о пользовательских таблицах.

При повреждении области пользовательской таблицы dbo.DimPublisher программа восстановила все пользовательские объекты, при этом данные таблицы dbo.DimPublisher восстановлены частично.

1.1.3 Recovery Toolbox For SQL Server v.2.4.21.0

Программа Recovery Toolbox For SQL Server v.2.4.21.0 при повреждении страницы, содержащей информацию о заголовке файла, позволяет просмотреть список восстановленных пользовательских объектов, а также их содержимое.

При повреждении страниц, содержащих информацию о системной таблице sys.syscolpars, программа восстановила описание триггеров, представлений и хранимых процедур, но структуру пользовательских таблиц и данные восстановить не удалось.

При повреждении страницы, содержащей данные пользовательской таблицы dbo.DimPublisher, программа восстановила все пользовательские объекты. Данные таблицы dbo.DimPublisher были частично восстановлены.

Таким образом, в случае физического повреждения файла базы данных невозможно использовать программу MS SQL Server Management Studio для восстановления. Инструменты SYSTools SQL Recovery и Recovery Toolbox For SQL Server v.2.4.21.0 показывают одинаковую эффективность.

1.2 Организация хранения данных в Microsoft SQL Server

1.2.1 Структура файла базы данных

В MS SQL Server данные могут храниться в одном или в нескольких файлах. Основной файл данных содержит все метаданные базы данных, которые могут быть просмотрены, используя системные представления.

Файл данных поделен на страницы по 8 килобайт. MS SQL Server поддерживает несколько типов страниц, таких как страницы данных, индексные страницы, страницы заголовка файла, загрузочные страницы и другие.

Страницы в базе данных непрерывно нумеруются, начиная с нуля, и на них можно ссылаться, указав идентификатор файла и номер страницы.

1.2.2 Виды системных страниц в MS SQL Server

- Страница с номером 0 всегда является заголовком файла;
- Страница с номером 1 – страница свободного места;

- Страница с номером 2 содержит информацию о выделенных экстентах;
- Страница с номером 3 хранит информацию о том, являются ли экстенты смешанными;
- Страница с номером 6 содержит схему разностных изменений;
- Страница с номером 9 является загрузочной страницей файла базы данных;
- Страницы Index Allocation Map содержат сведения об экстентах, используемых таблицей или индексом для единицы распределения.

1.2.3 Структура страницы

Каждая страница mdf-файла имеет определенную структуру (рисунок 1):

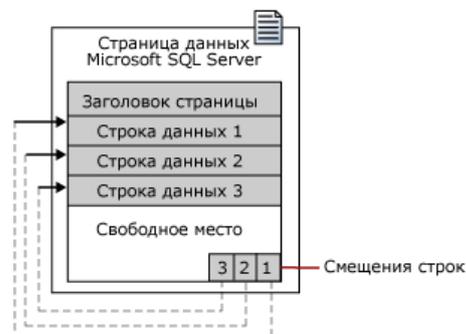


Рисунок 1 – Структура страницы mdf-файла.

Заголовок страницы имеет размер 96 байт. После заголовка страницы находится область, в которой хранятся фактические данные. За ней следует свободное пространство. В конце расположен массив слотов, который представляет собой блок двухбайтовых записей, указывающих смещение, с которого начинаются соответствующие строки данных на странице.

1.2.4 Структура строки

Строка состоит из: двух битов состояния (2 байта), длины данных фиксированной длины (2 байта), данных фиксированной длины (N-байт), количества колонок (2 байта), битовой карты для столбцов, содержащих значение NULL (Количество колонок, деленное на 8 байт), количества колонок с данными переменной длины (2 байта), массива со смещением столбцов

(Количество столбцов переменной длины, умноженное на 2 байта), данных переменной длины (M-байт), тега управления версиями (14 байт).

1.2.5 Виды резервного копирования и восстановления данных

Физическое резервное копирование представляет собой периодическое сохранение информации базы данных на основе создания физических копий файлов.

Логическое резервное копирование подразумевает создание текстовых файлов, содержащих команды SQL, с помощью которых можно восстановить объекты базы данных по состоянию на определенный момент времени.

1.2.6 Особенности хранения некоторых типов данных

1.2.6.1 Хранение данных типа DATE в Microsoft SQL Server

Данные типа DATE хранятся как 3 байта и содержат количество дней, прошедших с 0001-01-01.

1.2.6.2 Хранение данных типа DATETIME в Microsoft SQL Server

Данные типа DATETIME используют для хранения 8 байт. Первые 4 байта хранят информацию о дате, вторые 4 байта хранят информацию о времени. Точкой начала отсчета для извлечения даты является 1900-01-01.

1.2.6.3 Хранение данных типа DECIMAL в Microsoft SQL Server

Количество байт, необходимых для хранения чисел с фиксированной точкой, зависит от точности числа.

1.2.6.4 Хранение данных типа FLOAT в Microsoft SQL Server

Числа с плавающей точкой хранят максимально возможное приближение к числу. Для хранения данных этого типа используется 8 байт.

2 Практическая часть

2.1 Описание работы программы. Общий алгоритм

Программа CourseWork.exe представляет собой приложение с графическим интерфейсом для просмотра структуры пользовательских объектов, данных таблиц, извлеченных из поврежденного файла базы данных. Также она позволяет выполнить логическое восстановление базы данных. Программа работает с повреждениями, характеризующимися потерей значительного количества содержимого.

2.2 Возможности программы

Приложение позволяет просматривать структуру объектов пользователей и восстановленные данные в случае, если файл базы данных имеет повреждения: области данных таблицы пользователя, области данных системных таблиц, системных областей файла базы данных.

Программа предназначена для работы с базами данных версии MS SQL Server 2017.

2.3 Интерфейс

Пользователь может выбрать файл базы данных для восстановления, версию MS SQL Server. Далее производится обработка файла, и результирующая информация выводится на экран (рисунок 2).

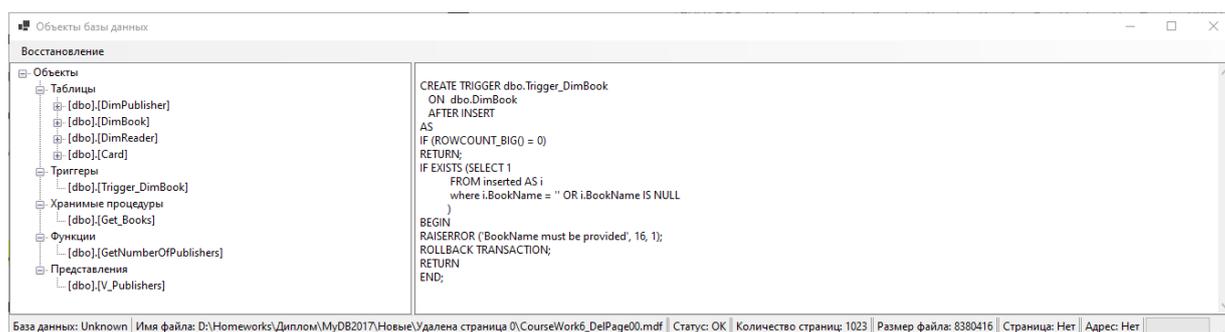


Рисунок 2 – Просмотр структуры объектов.

Меню «Восстановление» позволяет создать текстовый файл «Script.txt», содержащий набор SQL-команд, которые могут быть использованы для логического восстановления объектов базы данных.

2.4 Пример использования программы

Для проведения эксперимента создадим базу данных CourseWork6. Создадим в ней таблицы `dbo.DimPublisher`, `dbo.DimBook`, `dbo.DimReader` и `dbo.Card` с полями, имеющими различные типы данных. Заполним таблицы данными. Далее создадим триггер `dbo.Trigger_DimBook`, хранимую процедуру `dbo.Get_Books`, представление `dbo.V_Publishers` и функцию `dbo.GetNumberOfPublishers()`.

2.4.1 Повреждение в области системной таблицы `sys.syssschobjs`

В случае удаления страниц с номерами 0110h–0117h (272–279), принадлежащих системной таблице `sys.syssschobjs`, программа восстановила все пользовательские объекты и данные. Также был создан текстовый файл «Script.txt», с помощью которого можно восстановить объекты базы данных, используя MS SQL Server Management Studio.

2.4.2 Повреждение в области системной таблицы `sys.syscolpars`

В случае удаления страниц с номерами 00A8h–00ACh (168–172), принадлежащих системной таблице `sys.syscolpars`, программа корректно восстановила название базы данных, названия таблиц, содержимое триггеров, хранимых процедур, функций и представлений. В рабочей папке программы был создан текстовый файл «Script.txt».

2.4.3 Повреждение в области системной таблицы `sys.sysclsobjs`

В случае удаления страницы с номером 0057h (87), принадлежащей системной таблице `sys.sysclsobjs`, программа также корректно восстановила все объекты базы данных, включая данные. В рабочей папке программы был создан текстовый файл «Script.txt».

2.4.4 Повреждение в области системной таблицы `sys.sysobjvalues`

В случае удаления страниц с номерами 00E0h–00E6h (224–230), принадлежащих системной таблице `sys.sysobjvalues`, программа восстановила все объекты базы данных, включая данные. В рабочей папке программы был создан текстовый файл «Script.txt».

2.4.5 Повреждение в области системной таблицы sys.sysallocunits

В случае удаления страниц с номерами 0014h, 008Fh (20, 143), принадлежащих системной таблице sys.sysallocunits, программа восстановила все объекты базы данных, включая данные. В данном случае также есть возможность логического восстановления файла.

2.4.6 Повреждение в области системной таблицы sys.sysrowsets

В случае удаления страниц с номерами 0011h, 00A1h и 00A5h (17, 161, 165), принадлежащих системной таблице sys.sysrowsets, программа корректно восстановила название базы данных, структуру таблиц, содержимое триггеров, хранимых процедур, функций и представлений. Данные восстановить не удалось. Текстовый файл «Script.txt» не содержит команды для вставки данных.

2.4.7 Повреждение в области системной таблицы sys.sysscalartypes

В случае удаления страницы с номером 004Bh (75), принадлежащей системной таблице sys.sysscalartypes, программа восстановила название базы данных, содержимое триггеров, хранимых процедур, функций и представлений. Структуру таблиц и сами данные восстановить не удалось. Текстовый файл «Script.txt» не содержит команды для создания таблиц и вставки данных.

2.4.8 Повреждение страницы, содержащей заголовки файла

В случае удаления страницы с номером 0000h (0), содержащей заголовки файла, программа восстановила все объекты базы данных. В рабочей папке программы был создан текстовый файл «Script.txt».

2.4.9 Повреждение страницы, содержащей информацию о свободном месте

В случае удаления страницы с номером 0001h (1), содержащей информацию о свободном месте, программа восстановила все объекты базы данных. В данном случае возможно логическое восстановление с помощью файла «Script.txt».

2.4.10 Повреждение страницы, содержащей информацию о выделенных экстентах

В случае удаления страницы с номером 0002h (2), содержащей информацию о выделенных экстентах, программа восстановила все пользовательские объекты. В рабочей папке программы был создан текстовый файл «Script.txt».

2.4.11 Повреждение страницы, содержащей информацию о том, являются ли экстенты смешанными

В случае удаления страницы с номером 0003h (3), содержащей информацию о том, являются ли экстенты смешанными, программа восстановила все объекты базы данных. В рабочей папке программы был создан текстовый файл «Script.txt». Название базы данных не восстановлено.

2.4.12 Повреждение страницы, содержащей схему разностных изменений

В случае удаления страницы с номером 0006h (6), содержащей схему разностных изменений, программа восстановила все объекты базы данных. В рабочей папке программы был создан текстовый файл «Script.txt». Название базы данных восстановить не удалось.

2.4.13 Повреждение загрузочной страницы

В случае удаления загрузочной страницы базы данных, имеющую номер 0009h (9), программа восстановила все объекты базы данных. В рабочей папке программы был создан текстовый файл «Script.txt». Название базы данных восстановить не удалось.

2.4.14 Повреждение в области пользовательской таблицы dbo.DimPublisher

В случае удаления страницы с номером 00E8h (232), содержащей данные пользовательской таблицы dbo.DimPublisher, программа восстановила содержимое пользовательских объектов. Данные таблицы dbo.DimPublisher были восстановлены частично. В рабочей папке программы был создан текстовый файл «Script.txt».

ЗАКЛЮЧЕНИЕ

Результатом работы является программа для просмотра и логического восстановления структуры базы данных из физически поврежденного файла.

Рассмотрены следующие виды повреждений:

- 1) Повреждена область системных данных, хранящая сведения об объектах базы данных (sys.sysschobjs, sys.syscolpars, sys.sysclsobjs, sys.sysallocunits, sys.sysrowsets, sys.sysscalartypes);
- 2) Повреждены системные области файла базы данных, содержащие метаданные о базе данных;
- 3) Повреждены системные области файла базы данных, содержащие информацию о файле базы данных;
- 4) Повреждена область пользовательских данных.

При повреждении, описанном в первом случае, не всегда удается полностью восстановить объекты баз данных и сами данные, так как перечисленные системные таблицы содержат метаданные о пользовательских объектах.

При повреждении, описанном во втором случае, пользовательские объекты можно всегда восстановить, однако, некоторая информация о базе данных может быть частично утеряна.

При повреждении, описанном в третьем случае, информация о базе данных и о пользовательских объектах может быть всегда восстановлена.

При повреждении, описанном в четвертом случае, данные, содержащиеся в пользовательских таблицах, не всегда удастся восстановить в полном объеме.

Разработанное приложение рекомендуется использовать при наличии физических повреждений файлов баз данных.