

УДК 004.65:004.91

Применение генератора отчетов для преобразования таблиц из формата текстового файла в формат электронной таблицы

А. Б. Бетелин¹¹НИЦ «Курчатовский институт» – НИИСИ, Москва, Россия, ab@niisi.msk.ru

Аннотация. В статье приводится описание методики преобразования таблиц из формата текстового файла в формат электронной таблицы при помощи генератора отчетов АСУ НИИСИ.

Ключевые слова: преобразование текстовых таблиц, электронная таблица, табличный процессор, генератор отчетов

1. Введение

В учетно-управленческой деятельности организации иногда возникает потребность преобразовать файл, содержащий таблицу с некоторыми данными, из формата простого текстового файла (plain text) в формат табличного процессора типа OpenOffice Calc [1] или Microsoft Excel [2].

Такая ситуация может сложиться, к примеру, когда в организации применяются "древние" программы, ориентированные исключительно на алфавитно-цифровое (а/ц) взаимодействие с пользователем. И если подобная программа порождает данные в виде таблицы, то в ее оформлении используются только алфавитно-цифровые знаки из присутствующих на клавиатуре а/ц терминала и еще, возможно, символы псевдографики.

Дело может осложняться необходимостью провести дополнительную переработку таблицы, например: удалить ряд колонок, добавить несколько новых, вычислив значения полей в этих колонках, изменить заголовки колонок и всей таблицы. Вдобавок, с целью повышения удобства визуального восприятия результата имеет смысл добавить внешнее оформление (например, цвет фона для ячеек), задать гарнитуры и кегли шрифтов, скорректировать ширину колонок и т.п. Т.е. кроме изменения формата еще и выполнить существенное преобразование структуры таблицы и содержащейся в ней информации.

Если подобное действие происходит достаточно редко или вообще однократно, то проще всего проделать его вручную средствами самого табличного процессора: загрузить в него

требуемый текстовый файл, отредактировать и сохранить в нужном формате.

Если же это действие регулярное, а содержимое обрабатываемых файлов однотипно, то имеет смысл автоматизировать этот процесс и осуществлять его в пакетном режиме, используя возможности командной строки.

Покажем, как такая задача может быть решена с применением генератора отчетов, разработанного в НИИСИ РАН (ныне НИЦ «Курчатовский институт» – НИИСИ - НИИСИ) для АСУ НИИСИ [3].

2. Пошаговое описание методики

Предлагаемая методика предполагает получение преобразованной таблицы в виде файла, порождаемого генератором отчетов АСУ НИИСИ (он подробно описан в [4]).

Для того, чтобы воспользоваться этим генератором отчетов, нужно:

а) сформировать входной поток информации для генератора;

б) создать шаблон отчета в формате ODS.

Входной поток необходимо получить из исходного текстового файла с таблицей, которую требуется подвергнуть преобразованиям.

Шаблон может быть изготовлен при помощи табличного процессора OpenOffice Calc или аналогичного ему. В шаблоне задаются внешнее оформление отчета и переменные, которые впоследствии будут заменены на конкретные значения из входного потока.

Поскольку мы предполагаем воспользоваться возможностями, предоставляемыми АСУ НИИСИ, для

формирования входного потока будет удобно загрузить исходный текстовый файл с таблицей в базу данных (БД), применяемую в АСУ НИИСИ, в ней средствами СУБД провести все необходимые преобразования таблицы и затем изготовить входной поток в виде CSV-файла.

АСУ НИИСИ функционирует в среде ОС Linux [5], поэтому мы будем пользоваться в том числе средствами командного интерпретатора bash [6].

Опишем по шагам последовательность действий для решения нашей задачи.

Будем считать, что программа генератора отчетов размещена в каталоге `/bin`, а все действия мы проводим в специально созданном для этого рабочем каталоге с именем `/wrk`. При этом предполагаем, что у нас имеются права на запись в этот каталог.

При помощи любого подходящего текстового редактора создадим специализированный командный файл на диалекте языка SQL для сервера СУБД MySQL [7], который эксплуатируется в АСУ НИИСИ. В этом файле запрограммируем загрузку исходного файла в БД, все необходимые преобразования данных исходной таблицы, а также команды подготовки входного потока информации для генератора отчетов. Назовем этот файл `tab.sql` и разместим его в каталоге `/wrk`.

Шаблон, содержащий образец внешнего вида отчета, также разместим в каталоге `/wrk` в файле с именем `tab.ods`. Сделаем каталог `/wrk` текущим (здесь и далее приводятся команды для командного интерпретатора bash):

```
cd /wrk
```

Изготовим входной поток в виде CSV-файла для генератора отчетов:

```
mysql -hlocalhost -u<username> \  
-p<userpasswd> -B <DBname> < tab.sql >  
tab.csv
```

Здесь:

<username> - имя пользователя СУБД MySQL;
<userpasswd> - пароль пользователя СУБД MySQL;
<DBname> - имя базы данных.

Далее необходимо выделить и переработать содержательную часть шаблона отчета - файл `content.xml`. Для этого выполним следующее.

В каталоге `/wrk` создадим временный каталог с именем `tmp`:

```
mkdir /wrk/tmp
```

Скопируем наш шаблон в этот каталог и там разархивируем его при помощи утилиты `unzip`:

```
cp /wrk/tab.ods /wrk/tmp  
cd /wrk/tmp && unzip tab.ods
```

После этого в каталоге `/wrk/tmp` должны появиться файлы из zip-архива шаблона, и среди них - файл `content.xml`, который необходимо модифицировать в процессе изготовления отчета. Для этого запустим генератор отчетов (программу `xmlfilter.pl` - см. [4]) следующим образом:

```
cat /wrk/tab.csv | /bin/xmlfilter.pl \  
--in /wrk/tmp/content.xml \  
--out /wrk/tmp/content.xml
```

Мы вправе указать один и тот же файл в качестве входного и выходного для генератора отчетов, т.к. сначала генератор полностью считывает входной файл и только после этого формирует выходной файл.

В результате в каталоге `/wrk/tmp` появится переработанный файл `content.xml` для готового отчета. Остальные файлы изменениям не подвергаются. Осталось упаковать все эти файлы в ODS-документ, который заменит шаблон отчета в каталоге `/wrk/tmp`:

```
cd /wrk/tmp && zip -r tab.ods *
```

Поместим полученный отчет в наш рабочий каталог `/wrk` под именем `tab_res.ods` и удалим уже ненужный временный каталог `/wrk/tmp`:

```
cp /wrk/tmp/tab.ods /wrk/tab_res.ods  
cd /wrk && rm -rf /wrk/tmp
```

В конечном итоге в каталоге `/wrk` образуется готовый отчет в виде ODS-документа, который содержит таблицу, подвергшуюся всем требуемым преобразованиям.

Если же имеется необходимость получить файл в формате электронной таблицы MS Excel, то можно сконвертировать наш ODS-документ в этот формат средствами пакета OpenOffice, пользуясь языком OpenOffice Basic.

Запустим OpenOffice в диалоговом режиме, затем последовательно откроем пункты меню "Сервис" -> "Макросы" -> "Управление макросами" -> "OpenOffice.org Basic". В открывшемся окне в каталоге объектов выберем "Мои макросы" -> "Standard" -> "Module1". Нажав кнопку "Редактировать", в модуль "Module1" согласно рекомендациям [8,9] добавим одну процедуру и одну функцию (в OpenOffice Basic и те и другие называются макросами):

Код:

```
sub To_MS_Excel ( File_name )
  File_URL = ConvertToURL ( File_name )
  Obj_doc = StarDesktop.loadComponentFromURL _
( File_URL, "_blank", 0, Array ( MakePropertyValue _
( "Hidden", True ), ) )
  if not isNull ( Obj_doc )
  then
    select case LCase ( Right ( File_name, 3 ) )
      case "ods"
        Flt_name = "Microsoft Excel 97"
        File_ext = "xls"
      case else
        File_ext = "000"
    end select
    File_name = Left ( File_name, _
Len ( File_name ) - 4 ) + "." + File_ext
    File_URL = ConvertToURL ( File_name )
    Obj_doc.storeToURL ( File_URL, Array _
( MakePropertyValue ( "FilterName", Flt_name ), ) )
    Obj_doc.close ( True )
  endif
end sub

function MakePropertyValue _
( optional Prop_name as String, optional Prop_value ) _
as com.sun.star.beans.PropertyValue
  dim ob_Struct as new _
com.sun.star.beans.PropertyValue
  if not IsMissing ( Prop_name )
  then
    ob_Struct.Name = Prop_name
  endif
  if not IsMissing ( Prop_value )
  then
    ob_Struct.Value = Prop_value
  endif
  MakePropertyValue () = ob_Struct
end function
```

Процедура To_MS_Excel выясняет расширение входного файла и если оно "ods", то организует конвертацию этого файла в формат MS Excel. Вспомогательная функция MakePropertyValue необходима для работы механизмов пакета OpenOffice. Она задает соответствие между указанным свойством (property) и его значением, и возвращает пару «свойство-значение» в виде структуры. Эта функция заранее не определена, поэтому ее нужно добавить самостоятельно.

Теперь можно вызвать OpenOffice в режиме командной строки, указав абсолютное имя входного файла в качестве аргумента процедуры To_MS_Excel:

```
soffice -headless -norestore -invisible \
-nofirststartwizard \
"macro:///Standard.Module1.To_MS_Excel \
(/wrk/tab_res.ods)"
```

Если приведенную выше пошаговую последовательность действий оформить в виде скрипта для командного интерпретатора bash, то мы получим средство автоматического преобразования таблиц из формата текстового файла в формат электронной таблицы MS Excel в пакетном режиме.

3. Дополнительные соображения

Функциональность процедуры To_MS_Excel может быть расширена, с тем чтобы она выполняла конвертацию и других типов файлов, обрабатываемых программами пакета OpenOffice - ODP и ODT. Добавим в процедуру To_MS_Excel необходимые строки кода (см. [8,9]) и для соответствия выполняемым операциям переименуем ее в To_MS_Office:

Код:

```
sub To_MS_Office ( File_name )
  File_URL = ConvertToURL ( File_name )
  Obj_doc = StarDesktop.loadComponentFromURL _
( File_URL, "_blank", 0, Array ( MakePropertyValue _
( "Hidden", True ), ) )
  if not isNull ( Obj_doc )
  then
    select case LCase ( Right ( File_name, 3 ) )
      case "odt"
        Flt_name = "Microsoft Word 97"
        File_ext = "doc"
      case "ods"
        Flt_name = "Microsoft Excel 97"
        File_ext = "xls"
      case "odp"
        Flt_name = "Microsoft PowerPoint 97"
        File_ext = "ppt"
      case else
        File_ext = "000"
    end select
    File_name = Left ( File_name, _
Len ( File_name ) - 4 ) + "." + File_ext
    File_URL = ConvertToURL ( File_name )
    Obj_doc.storeToURL ( File_URL, Array _
( MakePropertyValue ( "FilterName", Flt_name ), ) )
    Obj_doc.close ( True )
  endif
end sub
```

Теперь процедура To_MS_Office сможет обеспечить конвертацию файлов текстовых

документов, электронных таблиц и презентаций из форматов пакета OpenOffice в соответствующие форматы пакета Microsoft Office [10].

Также следует отметить, что требуемая модификация исходной таблицы может быть реализована не только с помощью средств, предоставляемых применяемым диалектом языка SQL. Для этого можно воспользоваться любым подходящим для этих целей языком программирования, например, Perl [11]. Или же, в зависимости от объема и сложности преобразований, провести обработку исходного файла (частично или полностью) при помощи утилит командной строки, предназначенных для операций с текстом, таких как `awk`, `sed`, `cut` [12] и т.п. Здесь также возможно использование вспомогательных программ обработки текста

`tes`, `tbe`, `pfn` и др., входящих в состав программного комплекса на основе редактора текстов RK [4,13]. Конечным результатом применения перечисленных средств должно стать формирование входного потока информации для генератора отчетов.

4. Заключение

Приведенная выше методика хорошо вписывается в сценарии автоматической обработки файлов. Она с успехом применяется в НИИСИ для регулярного преобразования текстовых файлов, содержащих таблицы с данными, в формат табличного процессора Microsoft Excel.

Using a report generator to convert tables from a text file format to a spreadsheet format

A. B. Betelin

Abstract. The article describes a technique for converting tables from a text file format to an electronic spreadsheet format using the report generator of the Automated Control System of SRISA.

Keywords: text tables converting, spreadsheet, spreadsheet processor, report generator

Литература

1. OpenOffice.org [Электронный ресурс] // Wikipedia. — URL: <https://en.wikipedia.org/wiki/OpenOffice.org> (дата обращения: 25.06.2026).
2. Microsoft Excel [Электронный ресурс] // Wikipedia. — URL: https://en.wikipedia.org/wiki/Microsoft_Excel (дата обращения: 25.06.2026).
3. Бетелин, А. Б. Методы создания распределённого комплекса информационных систем для автоматизации управленческого и бухгалтерского учёта / А. Б. Бетелин, И. Б. Егорычев, А. А. Прилипко, Г. А. Прилипко, С. Г. Романюк, Д. В. Самборский // Труды НИИСИ РАН. — 2018. — Т. 8, № 4. — С. 175–180.
4. Бетелин, А. Б. Простые генераторы отчётов для учётно-управленческих информационных систем / А. Б. Бетелин, И. Б. Егорычев, А. А. Прилипко, Г. А. Прилипко, С. Г. Романюк, Д. В. Самборский // Труды НИИСИ РАН. — 2023. — Т. 13, № 1–2. — С. 10–18.
5. Linux [Электронный ресурс] // Wikipedia. — URL: <https://en.wikipedia.org/wiki/Linux> (дата обращения: 25.06.2026).
6. Bash (Unix shell) [Электронный ресурс] // Wikipedia. — URL: [https://en.wikipedia.org/wiki/Bash_\(Unix_shell\)](https://en.wikipedia.org/wiki/Bash_(Unix_shell)) (дата обращения: 25.06.2026).
7. MySQL [Электронный ресурс] // Wikipedia. — URL: <https://en.wikipedia.org/wiki/MySQL> (дата обращения: 25.06.2026).
8. DuCharme, Bob. Moving to OpenOffice: Batch Converting Legacy Documents / Bob DuCharme. — 11.01.2006. — URL: <https://www.xml.com/pub/a/2006/01/11/from-microsoft-to-openoffice.html> (дата обращения: 25.06.2026).
9. Массовое преобразование файлов формата Microsoft Office в открытый формат Open Document из командной строки [Электронный ресурс]. — URL: <https://ulspo.ru/index.php/instructions/how-to/131-msf2odf-macros.html> (дата обращения: 25.06.2026).

10. Microsoft Office [Электронный ресурс] // Wikipedia. — URL: https://en.wikipedia.org/wiki/Microsoft_Office (дата обращения: 25.06.2026).
11. Perl [Электронный ресурс] // Wikipedia. — URL: <https://en.wikipedia.org/wiki/Perl> (дата обращения: 25.06.2026).
12. Тейнсли, Д. Linux и UNIX: программирование в shell. Руководство разработчика / Д. Тейнсли; пер. с англ. — Киев : Издательская группа BHV, 2001. — 464 с.
13. Прилипко, Г. А. Исследование и разработка интерактивных языково-ориентированных систем редактирования и визуализации программ: автореф. дис. ... канд. физ.-мат. наук / Г. А. Прилипко ; МГУ им. М. В. Ломоносова. — Москва, 1986. — 23 с.

Поступила в редакцию / Received: 19.02.2026.

Поступила после рецензирования / Revised: 26.06.2026.

Принята к печати / Accepted: 26.06.2026.