As Cool as Ice: Data Mining the RACF Database and RACF Audit Data

Mark Nelson
IBM Corporation
Department BWVA Mail Station P388
2455 South Road
Poughkeepsie, NY 12601

RACF- 2000
Session 126
17 May, 2000
Internet: markan@us.ibm.com

Data Mining Your RACF Data

Trademarks

- **These terms are trademarks of the IBM Corporation in the United States, other countries, or both:**
  - DB2
  - DFSORT
  - IBM
  - OS/390
  - RACF
  - SQL/DS
  - S/390
- **SAS is a trademark of the SAS Institute, Inc.**
- **UNIX is a registered trademark in the United States, other countries, or both and is licensed exclusively through X/Open Company Limited.**

Data Mining Your RACF Data

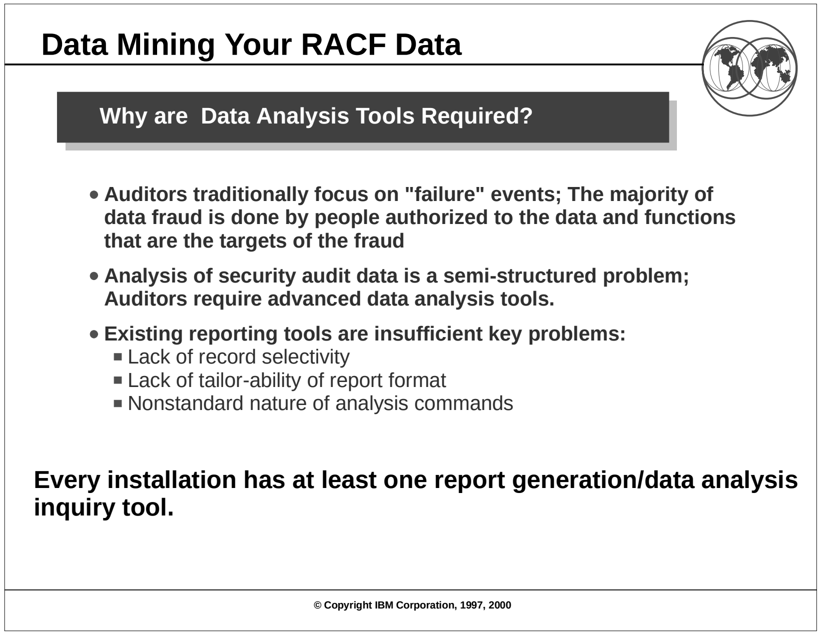
Agenda

- **What is Data Mining and how does it relate to RACF?**
  - What is auditing?
  - Why are analysis tools required?
  - What is data mining?
- **A four step approach:**
  - Understand the data and tools at our disposal
  - Formulating a search
  - Selecting the right tool
  - Refining the search

Data Mining Your RACF Data

What is Auditing?

- **Verification of compliance with the Installation Security Policy, by examining:**
  - Procedures and policies
  - Access rules
  - Physical access
  - User identification
  - Event data
    - **Looking at both successful (allowed) and unsuccessful (denied) events, looking for patterns**
  - Etc.

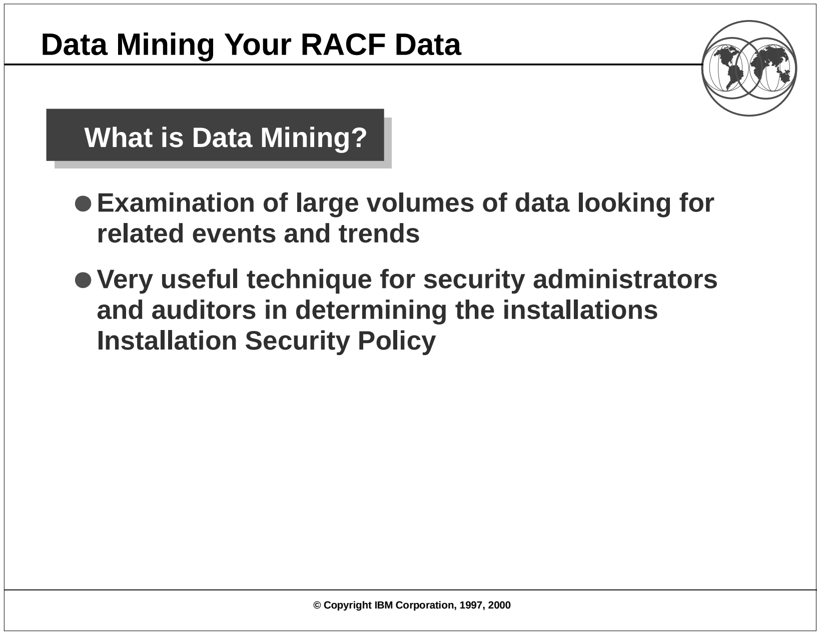


Data Mining Your RACF Data

Why are Data Analysis Tools Required?

- **Auditors traditionally focus on "failure" events; The majority of data fraud is done by people authorized to the data and functions that are the targets of the fraud**
- **Analysis of security audit data is a semi-structured problem; Auditors require advanced data analysis tools.**
- **Existing reporting tools are insufficient key problems:**
  - Lack of record selectivity
  - Lack of tailor-ability of report format
  - Nonstandard nature of analysis commands

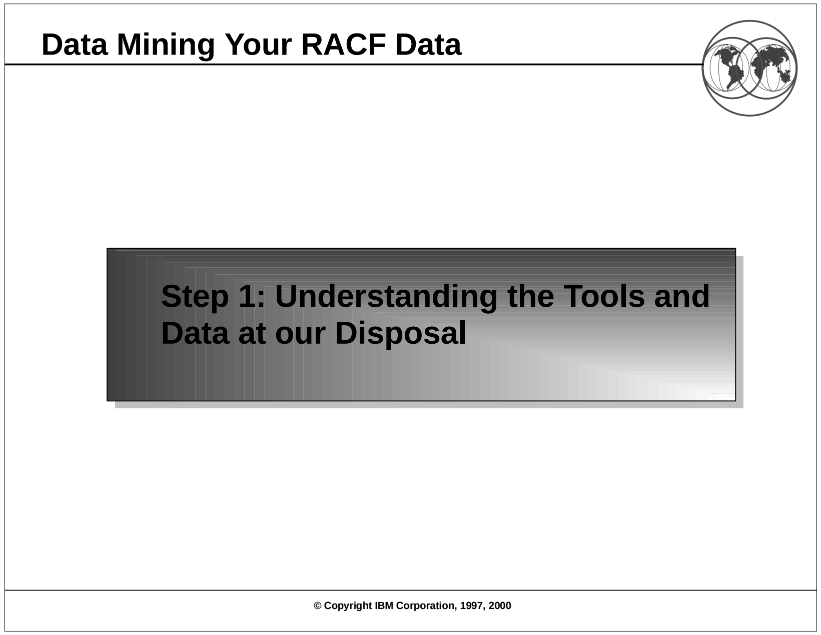
Every installation has at least one report generation/data analysis inquiry tool.



Data Mining Your RACF Data

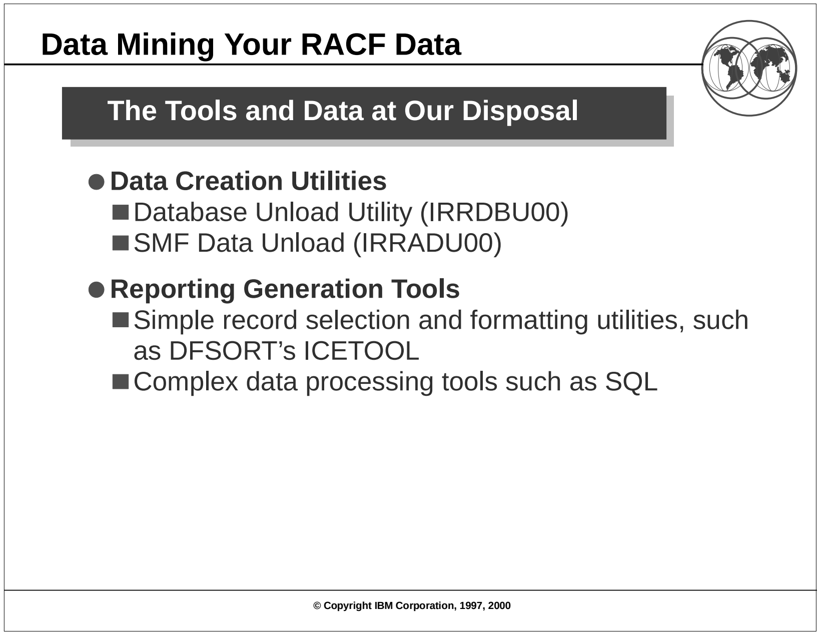
What is Data Mining?

- **Examination of large volumes of data looking for related events and trends**
- **Very useful technique for security administrators and auditors in determining the installations Installation Security Policy**



Data Mining Your RACF Data

Step 1: Understanding the Tools and Data at our Disposal



Data Mining Your RACF Data

The Tools and Data at Our Disposal

- **Data Creation Utilities**
  - Database Unload Utility (IRRDBU00)
  - SMF Data Unload (IRRADU00)
- **Reporting Generation Tools**
  - Simple record selection and formatting utilities, such as DFSORT's ICETOOL
  - Complex data processing tools such as SQL

Data Mining Your RACF Data

What is the RACF Database Unload Utility?

- **Creates a flat, relational representation of the RACF database, suitable for a DBMS load utility**
- **Conventions used in unloading the data:**
  - All fields unloaded, with the exception of encrypted and "reserved for IBM" fields
  - Fields decoded and presented in a readable format
    - **Example: UACC is output as "READ," "UPDATE," "ALTER," or "CONTROL" rather than as a binary field**
  - One record type per segment and per repeat group
    - **Identified by a 4 byte record type**
  - Each record contains a "name" field which identifies the profile being described

Data Mining Your RACF Data

IRRDBU00 Record Formats: Example

- **Records which define user IDs look like:**

Record ID	User ID	Created	Owner	ADSP	Special	Operations	Revoked	GRPACC	PWD INT
0200	MARKN	1997-07-03	SYSADMIN	NO	YES	YES	NO	NO	030 ...
0200	SMITH	1996-04-25	IBMUSER	NO	YES	YES	YES	NO	030 ...
0200	WOLENSKY	1997-03-03	MARKN	NO	NO	NO	NO	NO	030 ...

Data Mining Your RACF Data

IRRDBU00 Invocation

- **If your database is split, can process all parts or each part separately**
- **Uses the enhanced generic naming (EGN) setting and class descriptor table (CDT) from the execution system.**
- Sample JCL

```
//USERX     JOB  Job card. . .
//UNLOAD    EXEC PGM=IRRDBU00,PARM=NOLOCK
//INDD1     DD   DISP=SHR, DSN=SYS1.RACFDB.PART1.COPY
//OUTDD     DD   DISP=SHR, DSN=SYS1.RACFDB.FLATFILE
//SYSPRINT  DD   SYSOUT=*
```

Data Mining Your RACF Data

When to Use IRRDBU00

- **When you want to create tailored reports on your RACF user, group, and access control definitions**
- **When you want to perform a detailed analysis of the contents of the RACF database**
- **When working with an off-loaded copy of the RACF data is OK**

Data Mining Your RACF Data

RACF SMF Data Unload Utility (IRRADU00)

What is the SMF Data Unload Utility?

- **A RACF utility that translates the security relevant audit information into a set of records that can be imported to a relational data base management system, such as SQL/DS, DB2 or SAS.**
  - One record type per event type
  - Processes SMF type 30, 80, 81, and 83 records
- **Primary users are the system auditor and security administrator**
- **Requires READ authority to the SMF data**

Data Mining Your RACF Data

How is the Utility Invoked?

- **Invoked as exits to the SMF Dump Utility (IFASMFDP)**
  - RACF SMF Data Unload modules invoked through the **USER2** and **USER3** exit points
  - IFASMFDP can be used to provide data, time, system ID, and record type selection

```
//USERX    JOB   Job Card...
//UNLOAD   EXEC  PGM=IFASMFDP
//DUMPIN   DD    DISP=SHR,DSN=USER01.SMFDATA
//DUMPOUT  DD    DUMMY
//OUTDD    DD    DISP=SHR,DSN=USER01.SMFDATA.IRRRID00
//SYSPRINT DD    SYSOUT=*
//ADUPRINT DD    SYSOUT=*
//SYSIN    DD    *
     USER2(IRRADU00)  USER3(IRRADU86)
     DATE (99001,99123)
     START (0800)
     END(1700)
     SID(SYS1)
/*
```

Data Mining Your RACF Data

What Does the Utility Produce?

- **Relational representation of the security relevant audit data, suitable for export to a relation data base management system (RDBMS) or browsing**
- **One record type per event code**

ACCESS	Resource access
ADDSD	ADDSD command
ADDUSER	ADDUSER Command
CONNECT	Connect a user to a group
DELRES	Delete resource
DELVOL	Delete volume
DEFINE	Define resource
JOBINIT	Job initiation
RENAMEDS	Rename dataset
.....	Etc.

Data Mining Your RACF Data

What Does the Utility Produce (Continued)?

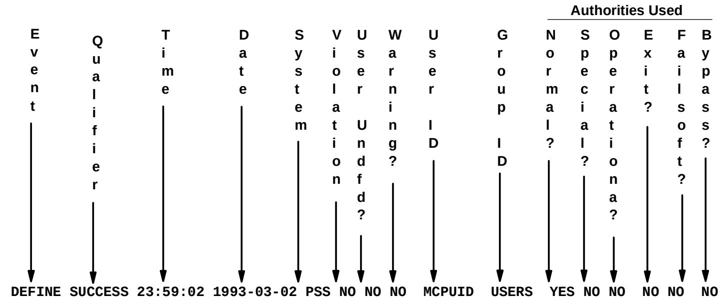
- **All data decoded**
- **Commands translated into command text format**
- **Event code qualifiers decoded into meaningful eight byte values**

INVPSWD	Not valid password
INVTERM	Not valid terminal
NASECL	Not authorized to SECLABEL
NJENAUTH	NJE job not authorized
.....	Etc.

Data Mining Your RACF Data

Record Formats

- All records of a specific event code are identical
- Base portion of all Type 80-based records are identical



```
DEFINE SUCCESS 23:59:02 1993-03-02 PSS NO NO NO  MCPUID  USERS  YES NO NO  NO NO  NO
ACCESS SUCCESS 23:59:03 1993-03-02 PSS NO NO NO  SYSUSER TASKS  NO  NO YES NO NO  NO
DELRES SUCCESS 23:59:04 1993-03-02 PSS NO NO NO  MCPUID  USERS  YES NO NO  NO NO  NO
ACCESS SUCCESS 23:59:04 1993-03-02 PSS NO NO NO  MCPUID  USERS  NO  NO YES NO NO  NO
ACCESS SUCCESS 23:59:05 1993-03-02 PSS NO NO NO  MCPUID  USERS  NO  NO YES NO NO  NO
```

Data Mining Your RACF Data

When to Use IRRADU00

- When you have complex selection criteria
- When you want to create tailored reports
- When you want to look at trends of events

Data Mining Your RACF Data

Step 2: Formulating a Query

Data Mining Your RACF Data

What are you Looking For?

- The first step in formulating a query is to identify what you are looking for
  - "I'm need to find all users who have an extraordinary RACF global authority (SPECIAL, OPERATIONS, or AUDITOR)"
  - "I want to find all failed logon attempts when the user ID changes but the terminal name does not"

Data Mining Your RACF Data

What are you Looking For?

- **Step two is to find the location (record type and name or offset) that helps answer your question.**
  - "I'm looking for all user IDs which have an extraordinary global authority (SPECIAL, OPERATIONS, or AUDITOR)
    - USBD_NAME (column 6-13) in the USER BASIC DATA record (type 0200)
    - USBD_SPECIAL (column 40-43) in USER BASIC DATA
    - USBD_OPER (columns 45-58) in USER BASIC DATA
    - USBD_AUDITOR (columns 386-389) in USER BASIC DATA
    - Find all records where either USBD_SPECIAL, USBD_OPER, or USBD_AUDITOR is set to "YES "

Data Mining Your RACF Data

What are you Looking For?...

- **Another example...**
  - "I want to find all failed logon attempts when the user ID changes but the terminal name does not"
    - INIT_EVENT_TYPE (column 1-8) is "JOBINIT "
    - INIT_EVENT_QUAL (column 10-17) is "INVPSWD"
    - Count the number of times that each terminal (INIT_TERM column 171-178 in JOBINIT record) is used

Data Mining Your RACF Data

Step 3: Selecting the Analysis Tool

Data Mining Your RACF Data

Selecting the Data Analysis Tool

- **The number of separate record types and the complexity of the selection criteria determine what type of data analysis tool is required**
  - When only a single record type is involved and the selection criteria is simple (equality, non-equality, greater than, less than, counting, simple pattern matching) a simple reporting tool such as ICETOOL may be used.
  - When multiple record types or complex selection criteria are required, then a more powerful analysis tool such as a relational data base manager (e.g. DB2, SQL/DS) is used.
- **Most queries are fall into single record/simple criteria**

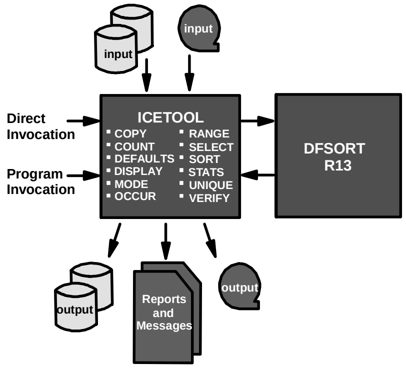
Data Mining Your RACF Data

Using the DFSORT™ ICETOOL Utility

- IBM's DFSORT product contains a simple yet powerful report generation tool, ICETOOL.
- ICETOOL adds an easy-to-use reporting facility to DFSORT'S powerful record selection and ordering capabilities.
- ICETOOL can easily be used with RACF's SMF unload utility (IRRADU00) and database unload utility (IRRDBU00) output.
- 30+ sample reports are shipped in the RACFICE package on the RACF web page (http://www.ibm.com/s390/racf/).

Data Mining Your RACF Data

DFSORT's ICETOOL Utility



Data Mining Your RACF Data

RACFICE and ICETOOL

- All of the RACFICE Reports are created using only 3 of the 15 ICETOOL operators:
  - **SORT/COPY**
    - Record ordering and selection
  - **DISPLAY**
    - Select input fields, create report and column headers, and specify output report format
  - **OCCURS**
    - Counts occurrences of values
    - Can be used to report counts over a specified threshold value

Data Mining Your RACF Data

Selecting Records Using DFSORT

- Records are included in a report using DFSORT's INCLUDE statement:

```
INCLUDE COND=((start,length,type,eval,value,AND|OR,
             start,length,type....)
```

- `start` is the starting position
- `length` is the length of the string being compared
- `type` describes the data type
  - "CH" indicates character
  - "SS" indicates substring
- `eval` is the type of comparison
  - "EQ" is equal
  - "NE" is not equal
  - "LT" is less than
  - "LE" is less than or equal to
  - "GT is greater than
  - "GE is greater than or equal to

Data Mining Your RACF Data

Sample RACFICE Report: SORT Keywords

```
SORT FIELDS=(10,8,CH,A)
INCLUDE COND=((44,1,CH,EQ,C’Y’,OR,
               49,1,CH,EQ,C’Y’,OR,
               390,1,CH,EQ,C’Y’),AND,
               5,4,CH,EQ,C’0200’)
OPTION   VLSHRT
```

Data Mining Your RACF Data

Sample RACFICE Report: ICETOOL Keywords

```
**********************************************************************
* Name: UGLB                                                         *
*                                                                    *
* Find all of the user IDs which have extraordinary RACF privileges, *
* such as SPECIAL, OPERATIONS, and AUDITOR at the global level.      *
**********************************************************************
 SORT    FROM(DBUDATA) TO(TEMP0001) USING(RACF)
 DISPLAY FROM(TEMP0001) LIST(PRINT) -
        PAGE -
        TITLE(’User IDs With Extraordinary Global Authorities’) -
        DATE(YMD/) -
        TIME(12:)  -
        BLANK -
        ON(10,8,CH)  HEADER(’User ID’) -
        ON(79,20,CH) HEADER(’User Name’) -
        ON(44,4,CH)  HEADER(’Special’) -
        ON(49,4,CH)  HEADER(’Operations’) -
        ON(390,4,CH) HEADER(’Auditor’)
```

Data Mining Your RACF Data

Sample RACFICE Report : JCL

```
//MARKNICE JOB ’M.NELSON P385’,NOTIFY=&SYSUID,CLASS=A,
//         REGION=0M,MSGCLASS=H
//*-----------------------------------------------------------------
//UNLOAD    EXEC PGM=IRRDBU00,PARM=NOLOCKINPUT
//SYSPRINT  DD SYSOUT=*
//INDD1     DD DISP=SHR,DSN=RACFDRVR.RACF260
//OUTDD     DD DISP=(NEW,PASS),SPACE=(CYL,(5,1)),UNIT=SYSALLDA,
//          LRECL=5096,RECFM=VB,BLKSIZE=0,DSN=USER01.IRRDBU00
//*-----------------------------------------------------------------
//REPORT    EXEC PGM=ICETOOL
//TOOLMSG   DD DUMMY
//PRINT     DD SYSOUT=*
//DFSMSG    DD DUMMY
//DBUDATA   DD DISP=(SHR,DELETE),DSN=USER01.IRRDBU00
//TEMP0001  DD DISP=(NEW,DELETE),SPACE=(CYL,(5,1,0)),UNIT=SYSALLDA
//TOOLIN    DD *
 <icetool control statements>
//RACFCNTL  DD *
 <sort keywords>
```

Data Mining Your RACF Data

Sample RACFICE Report : Output

```
- 1 -       User IDs With Extraordinary Global Authorities        98/12/29

User ID    User Name              Special   Operations   Auditor
--------   --------------------   -------   ----------   -------
GLBAUDIT   ####################   NO        NO           YES
GLBOPER    ####################   NO        YES          NO
GLBSPEC    ####################   YES       NO           NO
IBMUSER                           YES       YES          YES
MARKN      ####################   YES       YES          YES
SPECUSR    ####################   YES       YES          YES
UAUDR$Y    AUDITOR                NO        NO           YES
UOPER$Y    OPERATIONS             NO        YES          NO
USPEC$Y    SPECIAL                YES       NO           NO
```

Data Mining Your RACF Data

Another Sample ICETOOL Report: Sort Keywords

```
INCLUDE COND=(5,8,CH,EQ,C’JOBINIT’,AND,
              14,8,CH,EQ,C’INVPSWD’)

 OPTION  VLSHRT
```

Data Mining Your RACF Data

Another RACFICE Report: ICETOOL Keywords

```
**********************************************************************
* Name: TRMF                                                         *
*                                                                    *
* Find all of the terminals from which an excessive number of        *
* logons with incorrect passwords have been attempted.               *
*                                                                    *
* The ICETOOL "HIGHER(x)" keyword is used to set the failure         *
* threshold.                                                         *
**********************************************************************
 COPY    FROM(ADUDATA) TO(TEMP0001) USING(RACF)
 OCCURS  FROM(TEMP0001) LIST(PRINT) -
         PAGE -

         TITLE(’TRMF: Terminals with Excessive Incorrect Passwords’)-
         DATE(YMD/) -
         TIME(12:)  -
         BLANK -
         ON(175,8,CH) HEADER(’Terminal ID’) -
         ON(VALCNT)   HEADER(’Number of Incorrect Passwords’) -
         HIGHER(3)
```

Data Mining Your RACF Data

Another RACFICE Report: Output

```
- 1 -        TRMF: Terminals with Excessive Incorrect Passwords        00/03/24

Terminal ID   Number of Incorrect Passwords
-----------   -----------------------------
P4622212                                  7
P4622600                                  9
TERM0001                                  4
```

Data Mining Your RACF Data

Using the "substring" Conditional Test

- **DFSORT release 13 introduced the substring ("SS") comparison test, which indicates that a record is included if the selected value appears anywhere within the specified field**
  ```
  INCLUDE COND=(10,44,CH,SS,"*")
  ```
  - selects any record in which the character "*" appears within columns 10 to 53
- **Consider this example:**
  ```
  INCLUDE COND=(5,4,CH,EQ,C’0500’,AND,
                266,4,CH,EQ,C’NO  ’,AND,
                   (10,249,SS,EQ,C’*’,OR,
                    10,249,SS,EQ,C’%’,OR,
                    10,249,SS,EQ,C’&’))
  ```
  - Which finds all general resource profiles (record type ’0500’) which are not generic (record offset 266 contains ’NO’) but have a generic character in the name (the "SS" operands)

Data Mining Your RACF Data

Using DFSORT Symbols

- DFSORT release 14 introduced the DFSORT SYMBOL, which can be used to replace fields (and constants) in DFSORT and ICETOOL statements with easy-to-read labels
  - **USBD_OPER** could be used as a symbol for **44,1,CH**
- RACFICE contains DFSORT symbols for all of the IRRADU00 and IRRDBU00 fields.
- Using these symbols, you could specify these DFSORT statements:

```
SORT FIELDS=(USBD_NAME,A)
INCLUDE COND=(GRBD_RECORD_TYPE,EQ,C’0500’,AND,
              GRBD_GENERIC,EQ,C’NO  ’,AND,
                 (GRBD_NAME,SS,EQ,C’*’,OR,
                  GRBD_NAME,SS,EQ,C’%’,OR,
                  GRBD_NAME,SS,EQ,C’&’))
```

Data Mining Your RACF Data

A Sample SQL Query

- As an alternative, a relational database manager such as DB2 can be used. RDBMs are most useful for complex selection criteria which span record types.
- Find all of the data set accesses made to data sets whose name begins with "PAYROLL." that were made before 8:00 AM and after 4:59 PM. Ignore all of the requests made by the user OPERBKUP.

```
SELECT
     *
FROM
     USER01.ACCESS
WHERE
     (HOUR(SMF80_TIME_WRITTEN)<8 OR HOUR(SMF80_TIME_WRITTEN)>16)
     AND
     SMF80_EVT_USER_ID^= ’OPERBKUP’
     AND
     ACC_RES_NAME LIKE ’PAYROLL.%’
     ;
```

Data Mining Your RACF Data

Step 4: Refining the Search

Data Mining Your RACF Data

Refining the Search

- **No matter what analysis tool is used, your results fall into one of four categories:**
  - Too much data ("over inclusion")
    - Add additional selection criteria
  - Too little data ("inadvertent exclusion")
    - Remove or lessen criteria
  - The wrong data ("creates confusion")
    - Specify the right criteria
  - The correct data ("right conclusion")
- **Executing against known test data is essential!**

Data Mining Your RACF Data

Additional Material

Data Mining Your RACF Data

What Samples are Shipped With RACF?

- **Sample JCL for:**
  - IRRDBU00
  - IRRADU00
- **Sample SQL create tablespace and create table statements for IRRDBU00 and IRRADU00**
- **DBMS Load Utility control statements for DB2 Load Utility for IRRDBU00 and IRRADU00**
- **Sample queries for IRRADU00 and IRRDBU00 output**
- **30+ ICETOOL reports in ’SYS1.SAMPLIB(IRRICE)’**

Data Mining Your RACF Data

What Reports does RACFICE Contain?

- **Users who have extraordinary global/goup RACF attributes**
- **Discrete data set/general resource profiles which contain generic characters**
- **Users who have more than 20 group connections**
- **Count of user/group/data set/general resource (by class) profiles**
- **User IDs with group privileges above USE**
- **Data set standard and general resources with a UACC of other than NONE**
- **Data set standard and conditional access lists with ID(*) of other than NONE**
- **General resource standard and conditional access lists with ID(*) of other than NONE**
- **Users who have explicit RRSF associations defined**
- **User IDs with an OMVS segment**
- **OS/390 UNIX super users (UID of zero)**
- **OS/390 UNIX UIDs which are used more than once**
- **HLQs with excessive generic profiles**
- **HLQs with excessive fully-qualified generic profiles**
- **User profiles defined in the past 90 days**

Data Mining Your RACF Data

What Reports does RACFICE Contain?...

- **Events associated with a specific user**
- **User IDs with excessive incorrect passwords**
- **Terminals with excessive incorrect passwords**
- **Accesses allowed due to WARNING mode profiles**
- **Accesses allowed because the user has OPERATIONS**
- **Users who are using Automatic Command Direction**
- **Users who are directing command explicitly**
- **User who log on with LOGON BY**
- **RACLINK audit records**
- **Users who are using password synchronization**
- **Access violations**

Data Mining Your RACF Data

Where are These Utilities Documented?

- **RACF Database Unload Utility (IRRDBU00)**
  - RACF Security Administrator’s Guide
  - RACF Macros and Interfaces
- **RACF SMF Data Unload Utility (IRRADU00)**
  - RACF Auditor’s Guide and RACF Macros and Interfaces
- **DFSORT ICETOOL Utility**
  - DFSORT Application Programming Guide

Data Mining Your RACF Data

Agenda

- **What is Data Mining and how does it relate to RACF?**
  - What is auditing?
  - Why are advanced analysis tools required?
  - What is data mining?
- **A four step approach:**
  - Understand the data and tools at our disposal
  - Formulating a search
  - Selecting the right tool
  - Refining the search

Data Mining Your RACF Data

Disclaimer

The information contained in this document is distributed on as "as is" basis, without any warranty either express or implied. The customer is responsible for use of this information and/or implementation of any techniques mentioned. IBM has reviewed the information for accuracy, but there is no guarantee that a customer using the information or techniques will obtain the same or similar results in its own operational environment.

In this document, any references made to an IBM licensed program are not intended to state or imply that only IBM’s licensed program may be used. Functionally equivalent programs that do not infringe IBM’s intellectual property rights may be used instead. Any performance data contained in this document was determined in a controlled environment and therefore, the results which may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

It is possible that this material may contain references to, or information about, IBM products (machines and programs), programming, or services that are not announced in your country. Such references or information must not be construed to mean that IBM intends to announce such IBM Products, programming or services in your country.

IBM retains the title to the copyright in this paper as well as title to the copyright in all underlying works. IBM retains the right to make derivative works and to republish and distribute this paper to whomever it chooses.