DandD Client Server System

Daisuke Yokouchi
Keio Univ. JAPAN

Ritei Shibata
Keio Univ. JAPAN
Table of Contents

• Motivation
  – Needs for enough description of data

• DandD
  – DandD rule
  – DandD instance
  – DandD Client Server System
    • DandDServer
    • DandDBrowser
    • DandDR
    • DandDGenerator

• Home Page
Needs for enough description of data

• Data without enough description can’t say anything to us.
  – Data Storage
    • Nobody can understand the meaning of the data without enough description.
  – Collaboration
    • Many people are involved in data on the way from data collection to analysis.

We propose DandD which is a new environment for data and description.

Ritei Shibata(2004), “InterDatabase and DandD”, COMPSTAT2004
The summary of DandD rule

- Principle
  - Element
    - DataVector (A sequence of numbers) + Attributes
  - Organization
    - Relational
    - Array
  - Background Information
    - Introduction
    - Reference
    - Relatives
    - DataSampling

- Implementation
  - XML (eXtensible Markup Language)
  - UTF-16
A part of DandD Instance

```xml
<Data>
  <Relational Id="r1" Columns="sulfate nitrate rainfall"/>
  ...
</Data>
<DataBody>
  <DataVector Id="sulfate"> 3.5 1.5 ⋅⋅⋅</DataVector>
  <DataVector Id="nitrate"> 1.3 0.8 ⋅⋅⋅</DataVector>
  <DataVector Id="rainfall"> 8.7 3.9 ⋅⋅⋅</DataVector>
  ...
</DataBody>
```
A part of DandD Instance(2)

<DataBody>
<DataVector Id="v1" Access="a1" Protocol="b1"
    Query="c1" PostProcessing="d1"/>

......

External DataVector

</DataBody>

<Appendix>
<Access Id="a1" IP="131.113.65.1" UserId="anonymous"/>
<Protocol Id="b1" Physical="tcp">
    <JDBC DatabaseServer="131.113.65.1" DatabaseName="KobeQuake"/>
</Protocol>
<Query ="c1" Type="SQL">select date from kobequake</Query>
<ScanFormat Id="d1"> %*s,%s,%*s </ScanFormat>

......

</Appendix>

Ritei Shibata(2004), “InterDatabase and DandD”, COMPSTAT2004
By using the mechanism of External DataVector, we can integrate several pieces of data scattered over the Internet into a piece of data.
A piece of new Relational data

InterDatabase

DandD instance

RDB1

RDB2

Data File

DataVector (same length)
Client Program → User

Internet

RDBMS

DandD Instance with Data

DandD Instance Without Data

Data file
Client Server System

• Easiness of Programming Support Software
  – Any language can be used as far as it supports socket handling.

• Flexibility
  – When DandD rule is modified, the client programs works as same as before, in most cases

• Mobility
  – The size of client program can be reduced, so it is installable on low ability machine like PDA or cellular phone.
DandD Server

- **Recieve**
  - DOM (Document Object Model) methods & Original methods
- **Send**
  - Flag + Size + String
- **Implementation**
  - Language
    - Java
  - API for XML documents
    - Xerces for Java2 (Apache Project)
  - Access to databases on the network
    - JDBC (Sun Micro Systems)
  - Interpreter
    - Pnuts (Sun Micro Systems)
DandD Client Programs

- DandDBrowser
  - For browsing a DandD instance
- DandDR
  - For analysis and modeling
- DandDGenerator
  - For creating and editing a DandD instance
DandD Browser

• Purpose
  – To help users to understand what kind of data are described in a DandD instance

• Implementation
  – Java Language
DandD Browser

The Structure of DandD Instance
Continuous rainfall is rather subjectively defined. Duration of the continuous rainfall is not recorded. Observations are for days on which continuous rainfall is observed. From May of 1984 to May of 1988. Observation place "Shobara" is located in a basin surrounded by mountains. Population of "Shobara" is approximately 23000. A recording instrument with the diameter of 197 mm is placed on...
Continuous rainfall is rather subjectively defined. Duration of the continuous rainfall is not recorded. Observation are for days on which continuous rainfall is observed, from May of 1984 to May of 1988. Observation place "Stobara Map" is located in a basin surrounded by mountains with altitude 400-900m, and the altitude is 270m. Population of "Stobara" is approximately 23000 and there is no cause for pollution. The measuring instrument with the diameter 197mm is place on the top of a 10m high building.

Link for an auxiliary material of introduction
The function of the link is implemented by using “Unparsed Entity” equipped in XML originally.
DandDR

• Purpose
  – To support data analysis and modeling
  – An Interface between DandDServer and R

• Implementation
  – C language
  – Iconv (GNU)
> library(dad)
> acid=DandD()
> acid #`print.dad' is called
DandDGenerator

• Purpose
  – Creating and editing DandD instance

• Implementation
  – Java language

• Feature
  – Various templates of DandD instance.
  – Interface to a relational database system
A tree structure of the DandD instance

OUTPUT BUTTON
Concrete description of the DandD instance
DandDGenerator

Database Control Window
Columns of the table

Table in the relational database

As Relational BUTTON
DandDGenerator

SQL Viewer

Text field for inputting SQL query
Homepage

• DandD Project Homepage
  – DandD Support Software
  – Examples of DandD instance
  – DTD

• URL
  http://www.stat.math.keio.ac.jp/DandD/