BIM-based Data Mining System Framework to support an Effectiveness Decision-making for Energy Usage Management of Building Space
- 2015.5.24

*Ph.D, Tae wook, kang. Senior Researcher, KICT
Ph.D, Hyun sang, choi. Research fellow, KICT
Interests – BIM, GIS, Vision, Interoperability
laputa99999@gmail.com www.facebook.com/laputa999

This research was performed through a research subsidy from the 2015 Main Business (Development of Integrated Operation Technology on Construction Information & Spatial Information based on BIM/GIS Interoperation Platform) of the KICT and (11 High-tech G11) Architecture & Urban Development Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.
Introduce

Name - Tae Wook, Kang
Ph.D, Senior Research, ICT Lab, Korea Institute of Construction Institute (Current)
https://sites.google.com/site/bimprinciple/

Specialty – Civil Engineering, Software Engineering

Experiences
• Writing book
  IFMA, BIM for FACILITY MANAGERS, Translator (2014.5)
  Architectural collaborative design, Author (2014.2)
  Civil BIM, Author (2013.11)
  BIM interoperability and platform, Author (2013.1)
  BIM principle, Author (2011.6)

• Research
  BIM on GIS (Part 1 and 2) Research, KICT (Current)
  Point cloud-based architectural MEP object reverse engineering research, KICT(Current)
  BIM-based railway system planning project, MLIT (2013)
  VDC support system development planning Project, KICT (2013)
  World best software BIM modeler and check development, Ministry of Knowledge Economy (2012)

• Career
  BIM division head manager, Hangil IT (2011)
  Adjunct Professor, Chung-Ang University (2010)
Background
Background - BIM on GIS project overview

-This study is the part of BIM on GIS platform development project
-Developing BIM on GIS platform which has interoperability, application including various use-cases, standard/policy platform. Fund 8 M $ for 5 years (2012 – 2016). KICT with private sector (GAIA3D, Seokyoung Systems…)

### Research and Development

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>Use-cases and scenario ex – O&amp;M</td>
<td>Develop it ex – FM, BEMS ...</td>
<td>Validate it</td>
<td>Improve it</td>
</tr>
<tr>
<td>SW Platform</td>
<td>Interoperability Support Platform Design between BIM and GIS</td>
<td>Develop it Open API. Linkage. Mapping Rule</td>
<td>Validate it using the pilot project</td>
<td>Improve it</td>
</tr>
<tr>
<td>Standard Model</td>
<td>IFC4 analyze and Infra-BIM IFC design</td>
<td>Infra IFC draft version, Converting Tool</td>
<td>Infra-BIM IFC completion</td>
<td>Infra-BIM IFC Standardization</td>
</tr>
<tr>
<td>Policy system</td>
<td>IPDish design</td>
<td>IPD guideline development</td>
<td>IPDish validation</td>
<td>Collaboration System like IPD Suggestions for Public Sector</td>
</tr>
</tbody>
</table>

Commercialization and Standardization

R&D Results Validation and Feedback

Survey the technology and the solution
Background - BIM on GIS issue survey results

GIS-BIM Object Visualization
- GIS-BIM visualization
- IFC Parsing / Converter
- LandXML Converter
- CityGML mapping
- LOD generator
- Spatial Indexing

GIS-BIM Data Integration
- Query Language
- BPD / ETL
- Mapping Ruleset
- LOD Mapping
- Coordination Mapping
- Topology Linkage
- CityGML-BIM Extension
- Sensor / Actuator Node Management
- Data View Definition
- Open API with Security

GIS-BIM Data Mining
- Data Quality Validation
- Data Mining
- Data Analysis (ex – Trend, Decision Tree, Classification etc)
- GIS-BIM Data Cube (OLAP. Online analytical processing)

Context-based Application
- Model View Definition for Application - MBP Layers (Model + Business Logic + Presentation Layer)
- Context-based App
- Semantic Query and Analysis
Objective

BIM-based Data Mining system framework for supporting building **Space Energy Management (B-DM4SEM)**. The proposed **framework** considers functional variability and extensibility.

Which spaces and users of those spaces exceed the specific reference amount of annual energy usage:

```
SELECT SpaceObject(*) FROM Alignment, Building WHERE Buffer(Alignment, Building, '1km') AND Building.Pset('EM').EnergyUsage < '500 kWh'
```
Objective

Which spaces of those spaces exceed the specific amount of annual energy usage?

Energy manager

... Facility manager

... Actors

Decision making support

Energy usage data

BIM data

Data transform, binding, integrity validation, service extensibility

B-DM 4SEM

B-DM 4SFM

B-DM 4AM

Plugins

B-DM concept framework

IFC, Revit ...

Excel, Xml ...

Xml, Text ...

Sensor data

Energy usage data

BIM data
Motivation – How to obtain the needed data from BIM on GIS

In previous study related to BIM on GIS platform development (part 2), the difficulty in finding and identifying the needed information depending on the context related to specific use-case were founded.

The needed data depending on the use-case should be extracted from BIM database by using data mining to support the decision making.
Motivation – How to extract the information from energy data related to spatial object
BIM-based Data Mining System and prototype
Identification of framework structure

Before developing BIM-based DM, to identify considerations for ensuring the support functions such as data transform, binding, integrity, extensibility is needed.

Data / Service perspective

Data
- Flexible and simple forms of Integrated Data Model (IDB)
- Data model transformation (DMT)
- Data model validation rule (DVR)
- Data binding method considering BIM object (DBM)

Service
- Service extensibility (SEB)
- Service management method considering reusability (SRU)
- Problem description method method for implementing services (PDM)
Identification of the framework structure

The B-DM framework structure is identified based on the considerations:

- Service management method considering reusability
- Service extensibility
- Problem description method for implementing services
- Data model transformation
- Data binding method considering BIM object
- Data model validation rule
Identification of the framework structure

The B-DM framework structure is identified based on the considerations:

- BIM-based Integrated Database
- BIM-based Data Mining
- Operator
- BIM-based Data Mining Script
- BIM-based Service Container

Service #1
Service #2
Service #n

Actors manage energy of building space
Decision making support

Plug in

B-DM

<table>
<thead>
<tr>
<th>BIM-based Service Container</th>
<th>Service Plug-ins</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIM-based Data Mining Script</td>
<td>Data Mining Function Library</td>
</tr>
<tr>
<td>BIM-based Data Mining Operator</td>
<td>BIM Server</td>
</tr>
<tr>
<td>BIM-based Data Model Validation Rule set</td>
<td>BIM Object Data Binder</td>
</tr>
<tr>
<td>BIM-based Integrated Database</td>
<td>BIM-based ETL</td>
</tr>
</tbody>
</table>

Block diagram of B-DM4SEM framework

IFC, Revit ...
Excel, Xml ...
Xml, Text ...

BIM data
Energy usage data
Sensor data
Integrated Data Model (IDM) for B-DM

**IDM Schema model**

- **DomainSchema**
  - `<<PK>> + DM_ID
  - + Name: string
  - + Description: string

- **ObjectContainer**
  - `<<PK>> + OC_ID`
  - + Name: string

- **Relationship**
  - `<<PK>> + RS_ID`
  - `<<FK>> + RT_ID`
  - `<<FK>> + OT_ID_LEFT`
  - `<<FK>> + OT_ID_RIGHT`
  - + Name: string

- **ObjectType**
  - `<<PK>> + OT_ID`
  - `<<FK>> + DM_ID`
  - + Name: string
  - + Description: string
  - + Validate(Object): bool
  - + Create(): Object

- **Object**
  - `<<PK>> + OB_ID`
  - `<<FK>> + OT_ID`
  - `<<FK>> + OC_ID`
  - + Name: string

- **PSetDictionary**
  - `<<PK>> + PD_ID`
  - `<<FK>> + PARENT_PD_ID`
  - `<<FK>> + OT_ID_LEFT`
  - `<<FK>> + OT_ID_RIGHT`
  - + Name: string
  - + Validate(PSet): bool
  - + Create(): PSet

- **PSet**
  - `<<PK>> + PS_ID`
  - `<<FK>> + OB_ID`
  - + Name: string

- **PropertyType**
  - `<<PK>> + PT_ID`
  - `<<FK>> + PD_ID`
  - + Name: string
  - + Description: string
  - + Type: PT_ENUM_TYPE
  - + Format: string
  - + Validate(Property): bool
  - + Create(): Property

- **Property**
  - `<<PK>> + P_ID`
  - `<<FK>> + PS_ID`
  - + Name: string
  - + Value: string

**IDM object model**

- **RelationshipType**
  - `<<PK>> + RT_ID`
  - + Name: string
  - + Description: string
  - + Type: RT_ENUM_TYPE
  - + CustomType: string
  - `<<FK>> + OT_ID_LEFT`
  - `<<FK>> + OT_ID_RIGHT`
  - + Validate(Relationship): bool
  - + Create(): Object

- **PT_ENUM_TYPE**
  - `<<enumeration>>`
  - + Generalization
  - + Association
  - + Custom

- **DomainSchema**
  - `<<PK>> + DM_ID`
  - `<<FK>> + PARENT_DM_ID`
  - + Name: string
  - + Description: string

- **ObjectType**
  - `<<PK>> + OT_ID`
  - `<<FK>> + DM_ID`
  - + Name: string
  - + Description: string
  - + Validate(Object): bool
  - + Create(): Object

- **Object**
  - `<<PK>> + OB_ID`
  - `<<FK>> + OT_ID`
  - `<<FK>> + OC_ID`
  - + Name: string

- **PSet**
  - `<<PK>> + PS_ID`
  - `<<FK>> + OB_ID`
  - + Name: string

- **Property**
  - `<<PK>> + P_ID`
  - `<<FK>> + PS_ID`
  - + Name: string
  - + Value: string
Process definition concept of B-DM

In example, ODB, ETL and DVR can be defined as rule set.

Example of B-DM:

**B-ODB**
- B-DATA-BIND-RuleSET
- B-DATA-BIND-MAPPING
- B-DATA-BIND-LOGGER

**B-ETL**
- B-DATA-EXTRACTOR
- WORKFLOW
- B-DATA-TRANSFORMER

**B-DVR**
- B-DATAMODEL-RuleSET
- B-DATAMODEL-VALIDATOR
- B-DATAMODEL-LOGGER

Operational concepts and correlations of DMT, DBM, and DVR in B-DM4SEM
Process definition concept of B-DM

The heterogeneous EM data from the external legacy system

Subcontractor Perspective

View and Use the data

Facility = B199
Storey = S05
Code = B199.S05
Name = Main research building

Floor finish history =
At 2005.3.2, Tile#024 replacement, 0.5 year
At 2006.7.5, Tile#024 replacement, 0.5 year
At 2006.10.9, Tile#099 replacement, 0.5 year
At 2007.12.5, Tile#099 replacement, 0.5 year
Definition of operators for B-DM

The operator definition is divided into CRUDE in consideration of MECE conceptually. Here, CRUDE and the "*" mark have the following definitions.

- C, create; R, read query; U, update query; D, delete; E, execute; and *, plural indicator of an object

<table>
<thead>
<tr>
<th>Operator</th>
<th>Namespace</th>
<th>Operator-Supported Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRUDE</td>
<td>B-ODB-OP</td>
<td>Data Binding Rule (BRU), Data Binding Logger (BLG)</td>
</tr>
<tr>
<td></td>
<td>B-DVR-OP</td>
<td>Data Model Validation Rule (MRU), Data Model Validator (MVA), Data Model Validation Logger (MVL)</td>
</tr>
<tr>
<td></td>
<td>B-DM-OP</td>
<td>Classification Model (DMCF), Prediction Model (DMPD), Clustering Model (DMCL), Association Model (DMAS)</td>
</tr>
<tr>
<td></td>
<td>B-ETL-OP</td>
<td>Data Extractor (TDE), Data Transformer (TDT), Data Loader (TDL)</td>
</tr>
<tr>
<td>CRUD</td>
<td>B-DMS-OP</td>
<td>DomainSchema (BDS), ObjectType (BOT), RelationshipType (BRT), PSetDictionary (BPD), PropertyType (BPT), ObjectContainer (BOC), Object (BOJ), Relationship (BRS), PSet (BPS), Property (BPR)</td>
</tr>
</tbody>
</table>
1. Which spaces exceed the annual energy usage criteria (EUC)?

2. When was the first time this annual EUC was exceeded, and what is the space and the annual energy consumption trend?

The EUC value for use case 1 was 1,000 kWh, and that for use case 2 was 5,000 kWh.
Case study – test data set

KICT energy usage data set from 2013 to 2014 year which has 32850 records about 45 spaces
Prototype development for case study

Block diagram of B-DM4SEM framework

- Actors manage energy of building space
- Decision making support
- Plug in

<table>
<thead>
<tr>
<th>B-DM</th>
<th>Service Plug-ins</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIM-based Service Container</td>
<td>Service Plug-ins</td>
</tr>
<tr>
<td>BIM-based Data Mining Script</td>
<td>Data Mining Function Library</td>
</tr>
<tr>
<td>BIM-based Data Mining Operator</td>
<td>BIM Server</td>
</tr>
<tr>
<td>BIM-based Data Model Validation Rule set</td>
<td>BIM Object Data Binder</td>
</tr>
<tr>
<td>BIM-based Integrated Database</td>
<td>BIM-based ETL</td>
</tr>
</tbody>
</table>

- **Python**
- **NumPy**
- **SciPy**
- **B-ETL**
- **Talend**

- **Python engine**
- **IDB (Integrated Database)**

- BIM data
- Energy usage data
- Sensor data
- IFC, Revit ...
- Excel, Xml ...
- Xml, Text ...

Actors

Service #1

Service #2

Service #n

Manage energy of building space

Decision making support

Plug in
class BPAL:
    def GetMiningData(self):
        # Query the data set from IDB
        miningBaseTable = B_IDB.SelectDataTable("SELECT * FROM BIMOBJECT, ENERGY_OBJECT, ENERGYUSAGE_RECORD WHERE (BIMOBJECT.PK = ENERGY_OBJECT.FK AND ENERGY_OBJECT.PK = ENERGYUSAGE_RECORD.FK) AND SUM(ENERGYUSAGE_RECORD.USAGE, ...) ")

        # Obtain DataTable
        miningHeader, miningBaseData = self.DataTableToList(miningBaseTable)

        # Execute data mining prediction model(DMPD)
        npData = B_DMPD.LinearRegressionModel(miningHeader, miningBaseData)

        # Display the data mining results
        miningData = self.NumPyArrayToList(npData)
        miningData = self.ListToDataTable(miningHeader, miningData)

        UI.AddChart(drawGrid, drawChart, miningData, axisXLabel, axisYLabel, axisXColumn, axisYColumns)

BPAL()
Case study: use-case #1

Which spaces exceed over 1,000 kWh (annual EUC)?

1. Executing use case -1
2. Searching & executing a service
3. Executing data mining model for service
4. Obtaining information for service
5. Rendering data mining results with spatial information
Identifying space object in BIM with data mining results.

When was the first time this annual EUC (5,000 kWh) was exceeded, and what is the space and the annual energy consumption trend?
Case study: prototype

Which spaces exceed over 1,000 kWh (annual EUC)?
# Experiment result

## Results of the work performance time (min) for each case (manual)

<table>
<thead>
<tr>
<th>Test method division and phase</th>
<th>Work performance time</th>
<th>Related data processing time</th>
<th>Work performance/Report generation</th>
<th>Total work time A</th>
<th>Work time difference D = A – (B + C)</th>
<th>Improved performance ratio R = A / (B + C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Existing work method (manually)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td>A (1000)</td>
<td>827.7</td>
<td>620.1</td>
<td>578.6</td>
<td>537.1</td>
<td>495.5</td>
</tr>
<tr>
<td>Collect</td>
<td></td>
<td>185.6</td>
<td>167.4</td>
<td>160.2</td>
<td>161.3</td>
<td>154.1</td>
</tr>
<tr>
<td>Integration</td>
<td></td>
<td>1200.8</td>
<td>996.0</td>
<td>955.1</td>
<td>914.1</td>
<td>873.2</td>
</tr>
<tr>
<td>Work performance/Report generation</td>
<td></td>
<td>405.0</td>
<td>336.3</td>
<td>322.5</td>
<td>308.8</td>
<td>295.0</td>
</tr>
<tr>
<td><strong>B-DM4SEM method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work performance/Report generation</td>
<td>B</td>
<td>12.0</td>
<td>10.5</td>
<td>11.0</td>
<td>11.5</td>
<td>11.0</td>
</tr>
<tr>
<td>Work time difference D = A – (B + C)</td>
<td></td>
<td>2491.5</td>
<td>1993.7</td>
<td>1889.8</td>
<td>1794.2</td>
<td>1691.3</td>
</tr>
<tr>
<td>Improved performance ratio R = A / (B + C)</td>
<td></td>
<td>20.5</td>
<td>16.8</td>
<td>15.9</td>
<td>15.1</td>
<td>14.4</td>
</tr>
<tr>
<td><strong>B-DM4SEM data integration rule development time</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Related data integration rule development time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DB extraction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
</tr>
<tr>
<td>BIM integration time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>65.0</td>
</tr>
<tr>
<td>B-DM script development time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>47.5</td>
</tr>
<tr>
<td>Total work time C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>115.5</td>
</tr>
</tbody>
</table>
Experiment result

B-DM had an improved performance of up to 14.4 to 20.5 times

The maximum annual energy reference usage was divided into the following phases:
A (5000 kWh), B (6000 kWh), C (7000 kWh), D (8000 kWh), and E (9000 kWh)

Comparison of work performance time for each case
Conclusion
Conclusion & Future research

Conclusion

• B-DM4SEM framework concept was proposed to support effective decision-making depending on the use-case perspective such as building-space energy usage management.

• The proposed framework considers functional variability and extensibility.

• To implement B-DM, B-DM4SEM prototype system was developed and some use-cases was conducted effectively.

• The proposed method makes it easier to re-define a data-mining model when the use cases change as compared to existing methods.
  • Because defining the parameter values or logic applied in the proposed method can be achieved based on XML or scripts, improvements and modifications of the model can be easily conducted.

Future research

• B-DM framework development in more detail

• Effectiveness deduction of B-DM in detail

• N-screen & Web-based prototype development
Thanks for your interest

Tae Wook, Kang (www.facebook.com/laputa999)

This research was performed through a research subsidy from the 2015 Main Business (Development of Integrated Operation Technology on Construction Information & Spatial Information based on BIM/GIS Interoperation Platform) of the KICT and (11 High-tech G11) Architecture & Urban Development Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.