Provide Financial Risk Management as Service

Tie Liu[1], Feng Li[1], Hao Zhang[1], Xiangyang He[1], Rongzeng Cao[1], Yongming Wei[2]

[1]IBM Research - China
Beijing, P.R.China, 100193
{liultie,lfeng,zzhangh,hexyang,caorongz}@cn.ibm.com

[2]IBM Software Lab Service
Beijing, P.R.China, 100193
weiyongm@cn.ibm.com

***Abstract*—We present an innovative flexible computation support service for financial risk management. Commercial bank faces a higher requirement on quantitative risk management from Basel II accord. In practice, quantitative risk computation needs both the internal and external data, and the risk calculation models and corresponding parameters should be updated for a periods of time. To deal with these challenges, we propose a service oriented architecture based system, where flexible computation support services are provided for financial risk management. We present three main services in details. First, risk calculation service provides mathematical risk models and parameter learning functions. Second, business logic service adopts rule engine technology to integrate expert knowledge and mathematical model. Third, data management service provides the function for collection and management of internal and external data. These services collaborate to provide flexible quantitative risk management functions for commercial banks with a lower cost compared with the effort to develop their own risk management methodologies and systems.**

Keywords–Quantitative risk management; rule engine

I. Introduction

Financial risk management is a critical problem for commercial banks especially when the finance systems become highly complicated. There is a long history for risk management of banks, which includes risk identification, warning, quantification and integration. Quantitative risk management is becoming an important orientation for commercial banks [6]. Especially after Basel II accord was delivered, banks are paying more attentions to measure and manage risk quantitatively, including credit risk, operational risk and market risk. However, the implementation of quantitative risk management is not an easy work for banks, especially for small city and countryside banks. For example, in China, there are hundreds of banks and most banks are small and median sized banks. These banks face more challenges in quantitative risk management compared with station owned banks. The guideline from Basel II accord points out a path to quantitative risk management, but these small and median sized banks face huge challenges to follow Basel II accord. In this paper, we present a flexible computation support service for quantitative risk management, which can help these banks to improve their risk management capability.

A. Related work

One important aspect is the risk calculation model. There are many studies about quantitative model for risk computation [6], especially for credit risk [3]. CreditMetrics proposed by JP Morgan models the probability of moving from one credit quality to another within a given time horizon. KMV [7] uses the asset value model to compute the probability of default (PD) and another method CreditRisk+ also focuses on default. The linear probability models are used to compute the default probability such as Logit and Probit models [1]. Other statistical methods are also used for quantitative risk computation [4], such as Bayesian network, neural network [2] and so on. The calculation of PD, loss given default(LGD), expected loss(EL) and unexpected loss(UL) are very fundamental for quantitative risk management. However, most methodologies are designed with supported constraints, and the parameters should be tuned for different kinds of clients.

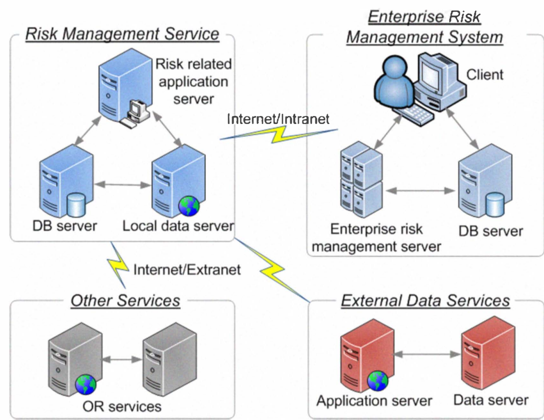

Fig. 1: The overall vision of risk management service.

Another important aspect is the data. Usually, a good risk calculation model should be tuned on the internal and external data. Most statistical models will work well with given training data, but the trained models face challenges when they are applied in another country or region. This is why that there is not a general quantitative model for all banks.

B. Our approach

We propose a flexible computation support service for quantitative risk management, where flexible methodologies and data management are both provided as services. The computation support services are provided with a service oriented architecture (SOA) based system. The main computation support services include: risk computation service provides mathematical risk models and parameter learning functions; business logic service adopts a rule engine to integrate expert knowledge and mathematical models; data management service includes the collection and management of internal and external data. Our approach provides a flexible risk management service for commercial banks, and the overall vision of our approach is shown in Fig.1.

The remainder of the paper is organized as follows. The problem is defined in section II. The risk computation services are presented in details in section III. The SOA architecture and the implementation of our approach is introduced in section IV. A practical case is studied in section V and the conclusion follows in section VI.

II. Problem Description

Quantitative risk management is a critical problem for commercial banks. In Basel II accord, the methodologies of quantitative risk calculation should be selected depending on the practical situations of banks. The implementation of Internal Ratings-Based (IRB) approaches for credit risk includes calculations of the probability of default (PD), loss given default (LGD), the exposure at default (EAD), and effective maturity (M). The methodologies can be selected. For example, the linear model of PD calculation can be as:

$$PD = \sum\nolimits_i a_i x_i, \tag{1}$$

where x_i is the factors which influence the probability of default and a_i is the parameter. PD can be also calculated with a nonlinear model, and an example of Logit model is:

$$PD = \frac{1}{1 + e^{-y}}, \quad y = a_0 + \sum\nolimits_{i=1}^{n} a_i x_i, \tag{2}$$

where x_i are the factors which influence the probability of default. For these two models, we must select the effective factors and assign the corresponding parameters by learning.

Besides the flexibility of mathematical models, the training data are also a critical factor for risk computation especially for statistical methods. For banks using either one of the IRB approaches for credit risk or the Advanced Measurement Approaches (AMA) for operational risk, the historical data are important. As stated in Basel II accord [9], commercial banks are recommended to collect the data for at least three years. However, commercial banks in developing country, such as China, there is not enough data accumulations. Furthermore, banks also need to pay for the collection of external data.

Commercial banks, especially for small and median sized banks, face challenges to implement quantitative risk management when they develop mathematical methodologies and collect enough data for the modeling training. We conclude these challenges as the following three aspects.

First, there is not a general mathematical method or model which can be used for all situations. Although many models [6] are proposed for risk computation, there are constraints or suppositions on these models mostly and many suppositions are not always appropriate for emerging market. In another aspect, commercial banks accumulate their own experiences for risk management, and it is indicated that these experiences are also important.

Second, quantitative risk calculation models must be customized when they are applied for a special commercial banks which may be very different because of countries, regions, business types, customer distributions and other factors. For example, we should select models for some specific group of customers and the corresponding parameters should be tuned on the historical data of this group of customers.

Third, a large scale of historical internal and external data are necessary to develop effective models. In the emerging market, historical internal data from banks were not collected well. This situation brings more dependence on external data. However, the collection and management of a large scale of external data is expensive and the update of external data is also costly.

To deal with these challenges when small and median sized banks implement quantitative risk management following Basel II accord, can we provide quantitative risk management services with a comparative lower cost for these small and median sized banks? What is the main component and how they work together? We will try to give the answers in this paper.

III. Quantitative Risk Management Service

In this section, we first give a brief introduction to quantitative risk management services, then we present three most important computation support services.

A. Brief introduction to quantitative risk management service

As shown in Fig.1, risk management services are provided to collaborate with the enterprise risk management system from bank to implement the risk calculation and collect the external data. The risk management service are categorized into different service components and each component is implemented and provided by a kind of service. A powerful rule engine is designed to integrate these service components, while commercial banks can select their own service and customize their necessary methods or models and the corresponding parameters. Furthermore, commercial banks can customize their business logic. For example, different business operations and risk categories should be considered. The data management service, especially about external data which will vary frequently, is also important. These data are not necessary for some specific applications but commercial banks usually have to pay for all the data package.

Our risk management services provide the maximal flexibility with a lower cost for banks. In the following, we will introduce three important risk computation support services: risk calculation service, business logic service and data management service.

B. Risk calculation service

There are different mathematical models following Basel II accord for the calculation of credit risk, operational risk and market risk. Risk calculation service provides the different risk calculation models, the general optimization and inference models and the learning technologies. These models can be selected by customers who access our services, and the corresponding parameters can be learned using internal data. Fig. 2 illustrates the risk calculation service components and the collaboration with the rule engine from business logic service. The risk calculation models is composed with the basic mathematical models and optimization models, and they collaborate with the learning algorithms through the business logic service. In other words, these models are called by some task templates which are described by business logics. The function of feature selection is also provided in the learning algorithm and it can be also customized. We will introduce several mathematical methods as the examples in the following.

1) Linear model for credit risk calculation: Linear classification technologies are integrated and provided as a service which can be used for risk measurement. For example, Logic and Probit methods are both included for quantitative credit risk measurement, as in equation (1) and (2). The most effective features can be identified with Adaboost algorithm for credit risk calculation and measurement, and the parameters are learned from the training data with gradient descent method.

2) Bayesian network for operational risk calculation: Bayesian network is provided as a service for risk factor analysis and risk measurement. The core idea is to model the priors and the likelihood functions. For operational risk, the failed events E_j which generates the loss directly are identified and modeled as hidden variables by experts or statistical technologies. The likelihood functions $p(C_i|E_j)$ describe the potential distribution of interdependency between failed events E_j and losses C_i. Similarly, the potential risk factors R_k are also identified and modeled to induce the failed events directly with another likelihood function $p(E_j|R_k)$. Then the overall potential loss from operational risk can be calculated as follows:

$$p(C_i) = \sum\nolimits_j p(E_j)p(C_i|E_j),$$

$$p(E_j) = \sum\nolimits_k p(R_k)p(E_j|R_k), \qquad (3)$$

The prior distribution of risk factors $p(R_k)$ and failed events $p(E_j)$, the likelihood functions $p(C_i|E_j), p(E_j|R_k)$ can be inferred through the learning technologies or expert assignment.

3) Model combination by business logic: We provide many more methods and models, such as neural network, which will not be introduced in details. These models are categorized by given tasks. For examples, we provide the credit risk calculation models such as KMV, Logit, Probit, etc. Beside these models, the knowledge different industries are also collected and provided for the usage of these models. As shown in Fig.2, the knowledge are integrated in the model selection and parameter assignment. For the customer who has not the capability to collect a big database and train the models by themselves, we provide the default selection of methods and models.

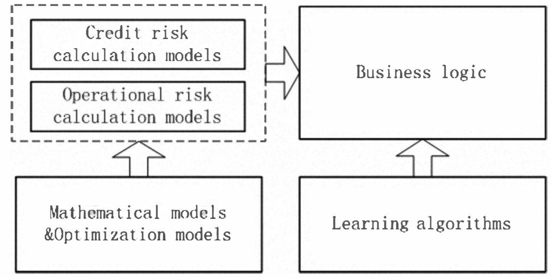

Fig. 2: Risk calculation service.

Risk calculation service provides mathematical models and algorithms for risk measurement, but how to use these models and algorithms to settle a specific task from the customer? We will introduce the business logic service which answers this question in the following subsection.

C. Business logic service

How to use the provided models and algorithms is related with the specific task and the experiences of customers. Experiences are also very important for risk computation and management. For example, the loans from a specific industry may deteriorate during a period of time, and the banks should adopt a specific model to compute the credit risk from this industry. We can utilize the production rule to describe this selection:

$$\text{If} \quad Customer.industry = IT \ ,$$

$$\text{Then} \quad model.selection = Logit, \qquad (4)$$

where we can edit this rule and execute the new algorithm without the compiling process. With this method ,business logic service provides the maximal flexibility to response quickly for such requirement.

Figure 3 gives an example where the methodologies for credit risk computation are written with business rules: the customer is first judged as individual or corporation; the corporation customer will be judged wether belongs on-list corporation and then industry categories will be considered; finally a specific credit risk calculation model is selected a specific group of customers. The business logic can be expressed by the production rules as follows:

If *Customer.type = Corporation,*
Then
. If *Customer.industry = IT ,*
. Then *model.selection = Logit,*
. If *Customer.industry = Energy ,*
. Then *model.selection = Probit,*
. If ......,

Business logic service will explain these rules and call the models or methodologies introduced in above subsection.

Another important function of business logic service is to integrate the experiences for risk computation and management. For example, experts observed that some loans from a specific

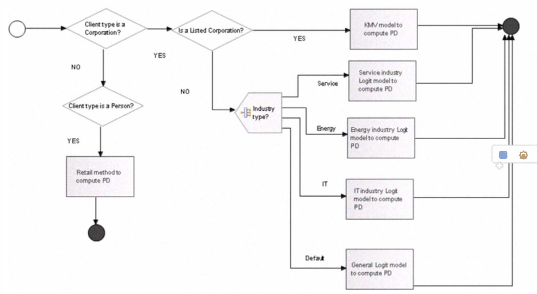

Fig. 3: A PD calculation example from business logic service.

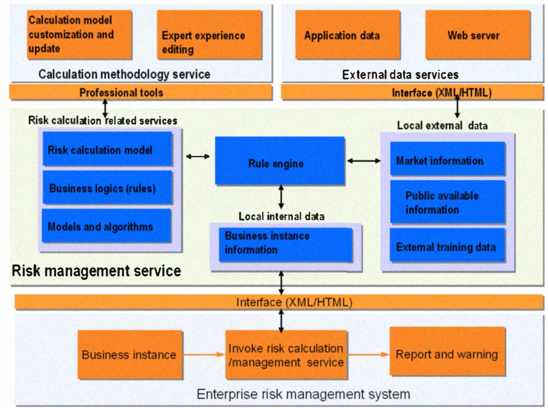

Fig. 4: The architecture view of risk management service.

corporation default in the past three months. Before they investigate the potential reasons clearly, they may consider to give a negative mark or a warning for the credit risk measurement of individuals from the same corporation temporarily. These expert knowledge can abstracted as rules and integrated in the credit risk computation method. Of course, there are a large number of knowledge for risk computation and management which can not be depicted by a precise mathematical model, and these experiences or knowledge can be integrated using business logic service.

The core of business logic service is a powerful rule engine, and risk calculation experiences and model customizations are separated and represented by rules with the reduced cost of designing, developing, and delivering the risk calculation module. To be concluded, business logic service provides the capability to describe the risk calculation methodologies. Expert experiences and model customizations are integrated to provide risk calculation function. Model selection and parameter learning can be implemented by the learning technologies. Therefore, the training data are very important. In the following subsection, we will introduce data management service which provide the base for the learning based models and methodologies.

D. Data management service

Data manage service provides two basic functions: internal data extraction; external data collection, update and management. High quality data are very important for quantitative risk calculation, especially for some models which should be trained. Further, some external data are high valued assets. Therefore, data collection, process and maintenance are very important for quantitative risk management.

Internal data is very important for model selection and parameter training, and we provide the Basel II risk data model and the standard interface for data preparation. For the bank whose data are not integrated and standardized, data extraction and transformation provided by data management service are necessary. The customer can write their own ETL commands to transform their data to the standard interface by our services. Furthermore, our approaches provides Business Solution Template(BST) which defines the standardized data interfaces to IBM banking data warehouse(BDW).

External data collection and management is another important function for data management service. There are two kinds of data: public available data and paid industry data. In our service, public available data, such as stock market data, are collected and updated periodically. These data can be provided as services directly, or the analyzed results can be provided. For paid industry data, the bank only needs to pay the service cost.

To be concluded, data management service tries to liberate commercial banks from data collection and management with high convenience and low expenditure.

Besides these three important services, there are also many other functions which are also implemented as services. We will not introduce all the services here.

IV. SOA-BASED IMPLEMENTATION

A. Quantitative risk management service architecture

The architecture of risk management service and the interfaces with other services are illustrated in Fig. 4. The enterprise risk management system can be constructed based on our risk management services, where these services are invoked to implement the risk calculation task. Then the results will be sent to the reporting and warning system. The risk management services are provided by a much flexible manner. For example, data management service is composed with local internal data service and local external data service. Another example is that business logic service can be accessed by a web based rule editor where the customer can write their task with business rule languages, and they can also specify their methodologies and integrate their experiences or expert knowledge conveniently. These rules are stored as XML format and exchanged together with the internal data.

As shown in Fig. 4, our approach provides multiple categories of services which can be accessed by different roles. For example, administrators can also customize and update some methodologies for specific risk calculation tasks through the professional services, and these knowledge will be kept by business rules and can be accessed by the customers.

The risk management system from customers, such as banks or enterprises will invoke risk management services through

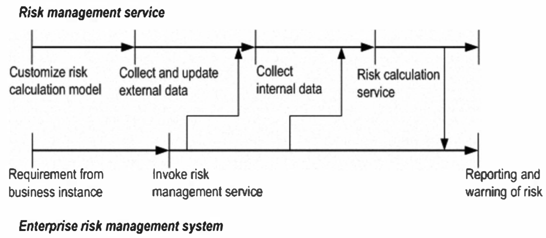

Fig. 5: The usage process of risk management service.

intranet or internet. Then the customer can define their task using business logic service where risk computation service will be called. The internal data from customer and the edited task description written in business rules are both sent to risk management service server. When the requirement arrives, data management service will prepare the local internal data complying with the methodologies from business rules. The rule engine will be in charge of the final integration and processes of business rules, risk models, statistical methods, internal and external data for risk computation. The whole usage process of risk management service for the customer is illustrated in Fig. 5.

B. Model driven approach

We implement risk management services using the Eclipse platform while Eclipse Modeling Framework (EMF) and Graphical Editing Framework (GMF) provide many helps to implement the visual interfaces. The model driven approach [13] uses abstract models to formulate the target application and refine design models at different levels step by step through model transformation techniques until the implementation of the whole system. In the implementations, these EMF models can be utilized for the service implementation, and the algorithms from risk calculation models can be separated.

V. A Practical Case Study

Quantitative credit risk computation and management is a critical problem for banks. In this section, we will use a case implementation for credit risk computation to illustrate in details how the risk management services can be utilized to resolve a real industrial problem.

A. Problem description

There are many studies about credit risk computation [10], and most studies focus on the computation models. Here, we try to integrate the general mathematical models with expert knowledge from experiences with a special rule engine, where we can provide a customized service for quantitative risk computation. Therefore, the most important issues here are: first, how to extract the flexible parts from credit risk models; second, how to provide the most general models where the customization can be well done; third, how to provide a mechanism to help customers transform their experienced knowledge to business rules. We will address these issues in the following subsections, and the computation of probability of default is illustrated how to implement and use our risk management services.

B. Implementations of credit risk calculation services

We implement the most general credit risk models, including Z-score, ZETA, Logit, Probit, KMV, CreditMetics and so on. The features and parameters of these models can be written with business rules, and the models can be called with the maximal flexibility. Each model has the limitation and can not be used directly for all cases. For example, Logit model need the selection of financial or economic feature, and the parameters must be learned using the training data. We use Adaboost [11] to implement a feature selection service, and linear regression for parameter learning of linear classification. We also design a Bayesian network and the corresponding leaning algorithm for PD computation, where the inputs come from the financial data of the client and the macro economic environmental index.

Business logic describe the flexible credit risk computation methodologies. We use business rules to describe the flexible methods and expert knowledge from experiences. There are many papers on how to represent the business logics with efficient rules [5] [12] [8]. We use production rule here which is usually represented as "*if* condition(C), *then* action(A)". Business instances include the financial data and other attributes such industry type. A hybrid rule engine system will process the business instance through evaluating the conditions of rules and executing the actions if the conditions are satisfied. A web-based rule editor is provided to edit the calculation methodologies. With the help of EMF and GMF, different drag-drop components are designed to speed up the design of methodologies and parameters.

In China, there is not a large number of internal data of default events and the information about the clients of banks are not well collected. To train the models, we use the public available data to provide the default parameters, such as the data from listed companies of Chinese [1]. With these financial data, we use a simplified method to get the training data: the corporations tagged with "ST" are supposed as the negative samples with "default". Other corporations are seemed as normal. With this simplified method, we can train our models such Logit, Probit and so on. We must point out that the default parameters can be only referred for banks and they can train their own parameters of credit risk models with their own data. Another important function of data management service is to update and manage external data periodically, and the business solution template from IBM banking data warehouse is used to design these data model.

C. Application of credit risk calculation services

There are several steps when we use credit risk calculation service for credit risk management:

[1] The data from listed companies of Shenzhen stock market: http://table.finance.yahoo.com/table.csv?s=000001.sz, and Shanghai stock market: http://table.finance.yahoo.com/table.csv?s=600000.ss. The financial data are from: http://stockdata.stock.hexun.com/2009_zxcwzb_600037.shtml(in Chinese).

1) Select services: we can access the risk management services through internet or intranet, where we can browse and select the provided services. For example, for credit risk computation, you can find that these services provide the computation of PD, LGD and so on. Suppose we select PD computation as the selected service.
2) Select methodologies: after you select the PD services, the system will send you a list of methodologies for your options, and the expert knowledge related with each method are also listed for your reference. Suppose that you select Logit model for your PD computation, and the recommended financial features and parameters are also provided. For example, the financial features may be: current assets, earnings before interest and taxes, retained earnings, short term debts and long term debts.
3) Training: if you have your own collected data, you can train your selected models with these data. The standard format is defined to promise that your input data are well used. Learning based feature selection can be also done if you are interested.
4) Business logic editing: usually you can get a default business logic when you have selected the models, where the models are called in the business logic. You can also modify these logics through a graphical interface and write business rules by yourself.
5) Customize your service: after you have finished the above steps (you can also use the default selections without training and business logic editing), you can customize your methodology as a service and store them under your account. At the next time, you can also access, use and modify this methodology.
6) Risk computation for a business instance: after you customize your service, you can compute the default probability using this service for a given business instance.
7) Historical data management service: you can browse your historical records from risk management service, and the original data are also kept. An interesting feature we plan to implement in the future is to provide the interactions that the customers can evaluate the compute results. This is also a method to collect training data, where the evaluations are thought as the ground-truth and we can provide a better service based on your business types and your evaluations.
8) Reporting service: risk management services also integrate the functions from COGNOS@ to provide reporting service. Besides the standard reporting format, customers can also select or define their own reporting format. Fig.6 gives an example of reporting after the computations from other related services are finished.

Partial services are only accessed by administrator, such as the mathematical model management, external data update, expert knowledge input and so on. We will not dive into details on all these services.

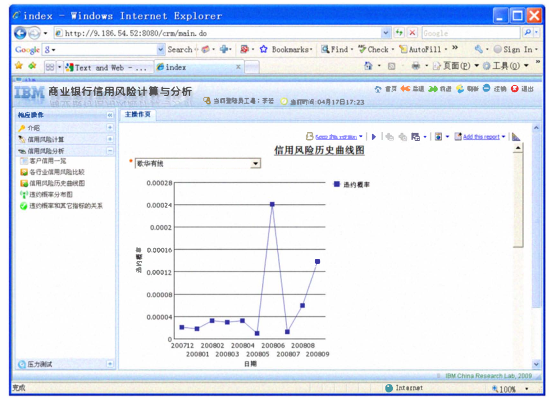

Fig. 6: A application case using risk management service.

VI. Conclusion

We propose a novel and flexible computation support service for quantitative risk management. We present in details about three main component services: risk calculation service, business logic service and data management service. A powerful rule engine is leveraged to integrate models, experiences methods, expert knowledge, internal and external data. The SOA-based implementation provides the maximal flexibility of risk management applications. These risk management services can be provided to the banks with much lower cost compared with that they develop and maintain a quantitative risk management system by themselves. The tradeoff between flexibility and complexity should be considered in the following work.

References

[1] John H Aldrich and Forrest D Nelson. *Linear Probability, Logit, and Probit Models*. SAGE, 2006.
[2] E. Angelini, G. Tollo, and A. Roli. A neural network for credit risk evaluation. *Prepring submitted to Elsevier Science*, 2007.
[3] M. Crouhy, D. Galai, and R. Mark. A comparative analysis of current credit risk models. 24:59–117, 2000.
[4] J. Galindo and P. Tamayo. Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications. *Computational Economics*, 15:107–143, 2004.
[5] Business Rules Group. Defining business rules what are they really? In *http://www.businessrulesgroup.org/first-paper/BRGwhatisBR-3ed.pdf*, 2000.
[6] Alexander J. McNeil, Rudiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton and Oxford, 2005.
[7] R. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 28:449–470, 1974.
[8] M.Muehlen, M.Indulska, and G.Kamp. Business process and business rule modeling: A representational analysis. In *IEEE EDOC Conference Workshop*, pages 189–196, 2007.
[9] Basel Committee on Banking Supervision. The new basel capital accord, 2004.
[10] Anthony Saunders and Linda Allen. *Credit risk measurement: new approaches to value at risk and other paradigms*. John Wiley and Sons, 2002.
[11] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE CVPR*, 2001.
[12] G. Wagner. Rule modeling and markup. *Reasoning Web*, 3564:251–274, 2005.
[13] Wenjun Yin, Xinxin Bai, Minghua Zhu, Ming Xie, and Jin Dong. ifao: Spatial decision support services for facility network transformation. In *IEEE SCC*, 2008.