

An Integrated Policy System (IPS) for Supporting Policy Making

Maria MOISE¹, Victor POPA²

¹ Romanian American University,
1B Expozitiei Blvd., Bucharest, 012101, Romania,
maria.moise@rau.ro

² ICI Bucharest,
8-10 Averescu Blvd., Bucharest, Romania,
vpopa@ici.ro

Abstract: This paper describes how policy design process is enabled by multichannel social computing, policy topic extraction, semantic analysis, opinions formation, simulation of agent based models and simulation of global models. An Integrated Policy System for policy design process includes: 1) extracting opinions, 2) summarising opinions, 3) generating scenarios, 4) simulating the scenarios for selecting and publishing the best ones. Opinions can be extracted from text analysis or using model for opinion formation on governmental decisions.

Keywords: Individual models, global models, e-Participation, opinions formation, Fuzzy Cognitive Maps (FCMs), Adaptive Neuro Fuzzy Inference Systems (ANFIS).

1. Introduction

Modeling the policy making process is a complex task. The policy making process can be modeled at individual and global level. At individual level, public bodies use agent-based simulation, where each agent represents an individual entity (citizen, company, stakeholder, house holder). The agent behavior is observed and taken into account in the policy making process. At global level, the model takes into account objectives, constraints, impacts on environment, economy and society. The integration of individual and global levels is a big challenge.

An Integrated Policy System (IPS) is a multi-disciplinary approach focusing on integration of global models, individual models and e-Participation model of decision process. The IPS uses simulation techniques, social networks, automatic text analysis and opinion formation methods to facilitate integrated e-governance. The citizens are involved in each step of the policy design process and informed about the impact of envisaged decisions.

The policy design process is enabled by multichannel social computing, policy topic extraction, semantic analysis, opinions formation, simulation of agent based models, simulation of global models. An IPS for policy design process includes:

- extracting opinions;
- summarising opinions;

- generating scenarios, simulating the scenarios for selecting the best scenarios, publishing the best scenarios.

Opinions can be extracted from text analysis or using model for opinion formation on a governmental decisions.

The model for opinion formation on a governmental decisions considers a group of agents (individuals) among whom a process of opinion formation takes place. The opinions of an agent will be influenced by opinions of others. These influences are modeled by weights which the agents put on opinions of other agents.

The IPS approach integrates the individual models, global models and e-Participation models using Machine-Learning component.

The Machine Learning component support:

- learning FCM weights from data;
- learning agent rules from data;
- learning rules for clustering and classification of documents.

Learning weights from data for FCM model uses Adaptive Network Based Fuzzy Inference Systems (ANFIS) and Least Square Error (LSE) method. Combining LSE and ANFIS will be done in three steps:

- decomposing FCM concept domains in regions using ANFIS approach;
- computing FCM weights for each region using LSE method;

- aggregating FCM weights for all regions using ANFIS approach.

The ANFIS approach learns the rules and membership functions from data. The ANFIS is an adaptive network of nodes and directional links. Associated with the network is a learning rule - for example, back propagation rule. The network is learning a relationship between inputs and output.

The majority of policy models relay on agent-based simulation [Gill, 2010] where agents represents parties involved in decision-making process. In addition to agent-based simulation models (individual level models), policy planning needs a global perspective that faces problem at a global level and should tightly interact with individual level. There are two methods concerning integration of global level models with individual level models:

- global level models are used for taking decisions and individual level models are used to understand the impact of taken decisions;
- global level models cooperate with individual level models to take decisions and to assess the impact of decisions.

The second approach is studied in several works as:

- Future Policy Modelling (FUPOL), it is a FP7 project (piloting in Europe and China) focusing on: land use models, housing models, energy models, immigration models, tourism models etc. The project uses opinion mining, FCM and agent-based approaches. FUPOL does not provide integration solutions;
- ePolicy project, it is a FP7 project, based on game theory to integrate optimization models with simulation models;
- [Gavanelli 2012] uses Benders decomposition – he decomposes the policy making problem in two components:
 - o Master component that defines objectives and constraints;
 - o Detail component that defines the strategic action plan to achieve objectives. The detail component communicates with an agent based simulator for generating training data. objectives. The detail component communicates with an agent based simulator for generating training data.

His learning method is based on Regression model.

- [Deng 2007] integrates simulation with optimization (optimization aids simulation for choosing optimal parameters).

The IPS approach extends above research works by:

- using causal models and opinions formation models [Akiyoshi, 2009], [Moise, Popa, 2014];
- adding new learning methods based on combining the optimization method LSE with adaptive fuzzy neural networks (ANFIS);
- learning methods aid both global and local models.

Compared with ePolicy project, IPS adds a new e-participation model for opinion formation, based on agent simulation approach.

2. IPS Architecture

The proposed architecture (Figure 1) to satisfy IPS goal includes the components:

- Global Modeling;
- Machine Learning;
- Local (individual) Modeling.

The Global Modeling takes into account the causal relations between concepts, and the Individual Modeling is based on agent behavior. Machine-Learning component is used to learn functions and rules from data.

Global modeling includes causal models based on Fuzzy Cognitive Maps (FCMs).

FCMs are signed digraphs with feedback. They are capable of modeling scenarios described in terms of concepts (activity, event, state) and their cause-effect relationships. FCM nodes represent concepts, and edges represent causal links among the concepts. Given an initial state of a system, represented by a set of values of its constituent concepts, a FCM can simulate its evolution over time to predict its future behavior.

Thus, a FCM model includes the following components:

- a. Nodes: to represent concepts C_i , $i = 1, \dots, N$, where N is the total number of concepts. Each concept/node indicates a characteristic or key factor of the system such as, actions, event and states. Each

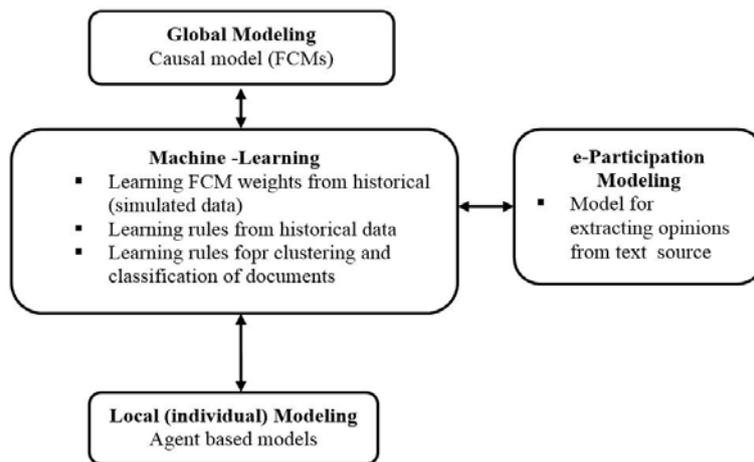


Figure 1. System architecture

concept/node it is characterized by a value. The concepts/nodes are interconnected through weighted arcs, which signify the relations among them;

- b. Edges/arcs: to represent causal links among the concepts;
- c. Weights: to represent how much one node influence another. The weight $w_{ij} \in [-1, 1]$, is analogous to the strength of the causal link between two concepts C_i and C_j .

The positive sign of w_{ij} indicates a direct relation between the two concepts that means a positive causality, the negative sign of w_{ij} indicates an indirect relation between the two concepts that expresses negative causality, and $w_{ij} = 0$ expresses no relation.

Computation procedure starts at time $t=0$ and computes values of dependent nodes using values of causal nodes. Thus, if nodes $C_1, C_2 \dots C_p$ influence node C_j then value of node C_j at time $(k+1)$ is computed from values of nodes $C_1, C_2 \dots C_i$ at time (k) using formula bellow:

$$C_j(k+1) = \sum C_i w_{ij}.$$

Individual Modeling includes agent based models for opinion formation on governmental decisions and agent based models for green park design.

The model for opinion formation on a governmental decisions considers a group of agents (individuals) among whom a process of opinion formation takes place. The opinions of an agent will be influenced by opinions of others. These influences are modeled by weights which the agents put on opinions of other agents. The process of forming of opinion of any agent takes an average over opinions of

other agents .This process is repeatedly until a consensus between agents is attained. Thus, if $x_i(t)$ is opinion of agent i at time t , and a_{ij} is the influence of agent j on agent i ($a_{i1} + a_{i2} + \dots + a_{in} = 1$), then the opinion of agent i at time $t+1$ is described by $x_i(t+1) = a_{i1}x_1(t) + a_{i2}x_2(t) + \dots + a_{in}x_n(t)$ or equivalently, by $x(t+1) = Ax(t)$.

So, the agent i adjust his opinion in period $t+1$ taking a weighted average of opinions of other agents. The computation is made until is reached a consensus, polarization or a opinion fragmentation.

Machine Learning component provides:

- Learning FCM weights from data using ANFIS and LSE methods;
- Learning agent rules from historical (simulated) data using ANFIS an Subtractive Clustering methods;
- Learning the classification and clustering of text documents.

The architecture used in learning process is illustrated in Figure 2.

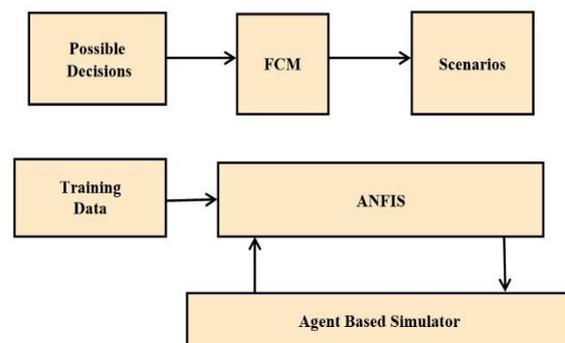


Figure 2. Machine Learning Architecture

2.1 Learning agent rules from historical (simulated) data

There are several methods for extracting fuzzy rules from input/output data for both function approximation and pattern classification applications. For function approximation, where the output data correspond to a continuous valued variable, the extracted rules express continuous-valued relationships (e.g., “if input is small then output is big”). For pattern classification, where the output data correspond to class assignments, the extracted rules have discrete-valued consequents (e.g., “if input is small then output is class 1”). In both cases, fuzzy rules provide a powerful framework for capturing and, perhaps more importantly, explaining the input/output data behaviour.

Several methods for extracting fuzzy rules for function approximation have used clustering to determine the number of rules and initial rule parameters. Each cluster essentially identifies a region in the data space that contains a sufficient mass of data to support the existence of a fuzzy input/output relationship.

Clustering of numerical data forms the basis of many modelling and pattern classification algorithms. The purpose of clustering is to find natural groupings of data in a large data set, thus revealing patterns in the data that can provide a concise representation of the data behaviour. Clustering algorithms typically require the user to specify the number of cluster centers and their initial locations. The Fuzzy CMeans algorithm and Kohonen’s Self-Organizing Map, the *mountain* (subtractive) method is well-known examples of such clustering algorithms. In our project we use *Adaptive Network Based Fuzzy Inference Systems (ANFIS)* to extract rules from data,

because the ANFIS approach is generally more accurate than other approaches.

The ANFIS approach learns the rules and membership functions from data. ANFIS is an *adaptive network*. An adaptive network is a network of nodes and directional links. Associated with the network is a learning rule - for example back propagation. It’s called adaptive because some, or all, of the nodes have parameters which affect the output of the node. These networks are learning a relationship between inputs and outputs. The basic learning rule for ANFIS is based on gradient descent and chain rule. The combination of a fuzzy model with a training of the outputs has the advantage that the model can be understood by the expert and adapt to the measured values. This leads to a better fit of the model.

The ANFIS architecture is shown in Figure 3. The circular nodes represent nodes that are fixed whereas the square nodes are nodes that have parameters to be learnt.

Using ANFIS approach the component generates rules of form:

If Age is Gauss(a,b) and Sex=M then Cycling-Area is f(Age, Sex)

2.2 Learning FCM weights from data

In manual development of FCM models all experts are asked to determine the relevant factors in a brain storm meeting. They discuss about main characteristics of the system, number and kind of concepts, which should be contained in the FCM. Then, they determine the structure and the interconnections of the network using fuzzy conditional statements. Each expert may draw his own individual

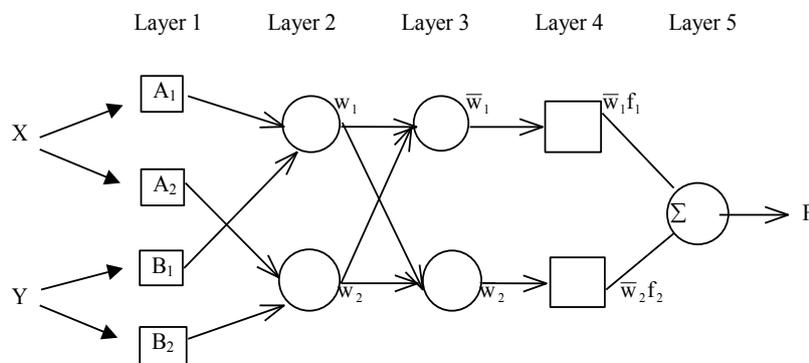


Figure 3. ANFIS Architecture (Sugeno type)

FCM, which can differ from the others. In order to deal with these diagrams, the assigned weights by each expert can be considered and a new FCM will be constructed by all the experts. Thus, this FCM will represent the knowledge and experience of all related experts FCMs can produce by expert manually or generate by other source of information computationally. Experts developed a FCM or a mental model manually based on their knowledge in related area. At first, they identify key domain issues or concepts. Secondly, they identify the causal relationships among these concepts and thirdly, they estimate causal relationships strengths.

Although the manual procedures for developing FCM are well-established, they have several drawbacks. They require expert knowledge, which has to be supported by knowledge of FCM methodology. Since the number of possible connections among concepts increases quadratically with the increase of the number of concepts, expressing complex systems that consists of large number of nodes is often very difficult or even impossible to perform by a human. This may result in simplifications, which eventually translate into inaccuracy or incorrectness. The development process often requires many iterations and simulations before a suitable model is established. In case of group development, the quality of the final model can be improved by varying impact of a given expert model on the final model based on credibility of a particular expert. However, it requires additional parameters, such as credibility coefficients of each individual expert, which complicates the FCM development task.

Manual methods for development FCM models have also a major disadvantage of relying on

human knowledge. It is very difficult to assess the model's accuracy in an unbiased way. Moreover, even if there are available historical data to justify the model's quality, obtaining appropriate model that mimics the data is a laborious task, which requires drawing and simulating successive models.

In many cases developing a FCM manually becomes very difficult and experts. Intervention could not resolve the problem. When the experts were not able to express their expertise or even there is no expert in the related area, therefore there is a gap.

For these reasons, the development of computational methods for learning FCM is necessary. For this purpose, not only the edge or casual relations between nodes, but also the strength on each edge must be achieved based on historical data. In this way, the expert knowledge is substituted by the knowledge, which produced from historical data by means and new computational procedures. Many algorithms for learning FCM model structure have been recently proposed. In general, the following categories of these algorithms are used:

- Fuzzy decision trees;
- Association rules algorithms;
- Algorithms for classification □ based on distance;
- Hebbian algorithms;
- System Dynamics Simulation;
- Genetic algorithms and neural network.

ITP approach for learning FCM weighs from data combines LSE and ANFIS formalisms in 3 steps:

Step1: Decompose each concept domain in regions using ANFIS formalism.

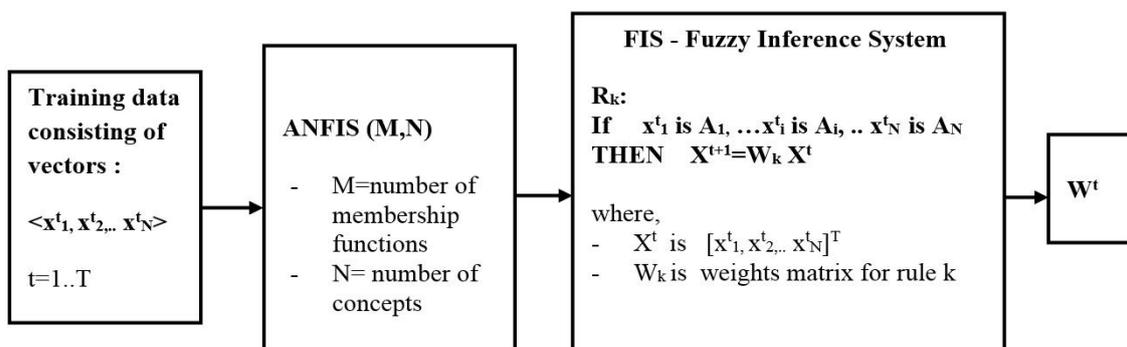


Figure 4. FCM weights Learning

Step2: For each combination of regions A_1, A_2, \dots, A_N generate the fuzzy rule:

**Rk: If x^t_1 is A_1, \dots, x^t_i is A_i, \dots, x^t_N is A_N
THEN $X^{t+1} = W_k X^t$**

where, matrix W_k is computed using LSE method.

Step3: For each time $t = 1 \dots T$ compute $W(t) = ANFIS(X^t)$, where $W(t)$ is weights matrix at time t .

2.3 eParticipation

The component supports the citizens participation in eGovernment issues via the availability of an interactive framework that mines the citizens opinions about governmental decisions. The model analyses the Web postings and detects opinions and topics to which the opinions refer. The syntactic parser identifies pairs (adjective, noun) using the steps:

Step1:

Eliminating from citizen's posted text the non-text elements: images, graphical representations of text (including symbols), frames, scripts and video etc.;

Step2:

Identifying the word tokens in the text and annotating every token with an appropriate grammatical category;

Step3:

Extracting the nouns appearing in the text and their referring adjectives;

Step 4:

Obtaining the semantic of the adjectives using Frame Net Ontology.;

Step 5: Opinion Analysis

The large number of potential contribution needs to be aggregated and summarized. This data aggregation step is crucial for e-governance since it expresses the voice of the citizen. The text will be analysed whether a document or a section thereof is subjective or objective and whether the opinion expressed is positive or negative using a mixed-initiative framework which combines human annotations and automatic annotation based on machine learning/natural language processing techniques. The strength of the expressed opinion will also be determined automatically.

Step 6: Opinion Summarisation

Opinion summarization aims at giving the overall sentiment of a large number of opinion resources at various granularities. The presentation and visualization is and still to a large extent unexplored and new research field. Results of the Opinion extraction as well as the underlying raw data will be classified and linked with the topic.

3. Conclusions

The IPS aims policy makers at building scenarios and simulation of the scenarios in order to assessing impacts of scenarios. The IPS builds scenarios taking into account the citizens opinions. The ITP simulates the scenarios using: 1) global models (Fuzzy Cognitive Maps) and 2) individual models (agent based models). Global models take into account causal dependencies between concepts, and individual models takes into account the behavioral rules of agents. For integration of global models, individual models and eParticipation models, IPS uses a machine learning component. Machine learning component uses historical data (simulated data) to learn function dependencies and rules in order to enrich the global and individual models.

REFERENCES

1. ROBINSON, D. T., D. G. BROWN, D. C. PARKER, P. SCHREINEMACHERS, M. A. JANSSEN, M. HUIGENS, H. WITTMER, N. GOTTS, P. PROMBUROM, E. IRWIN, T. BERGER, F. BATZWEILER, C. BARNAUD, **Comparison of Empirical Methods for Building Agent-based Models in Land Use Science**. Journal of Land Use Science, vol. 2(1), 2007.
2. HEYDEBRECK, P., M. KLOFSTEN, L. KRÜGER, **F2C-An Innovative Approach to Use Fuzzy Cognitive Maps (FCM) for the Valuation of High-Technology Ventures**, International Business Information Management Association, IBIMA vol. 2011, 2011, Article ID 483.
3. GAVANELLI, M., F. RIGUZZI, M. MILANO, D. SOTTARA, A. CANGINI, P. CAGNOLI, **An Application of Fuzzy Logic to Strategic Environmental Assessment**. In Pirrone, R. and Sorbello,

- F., editors, AI*IA, volume 6934 of LNCS. Springer, 2011.
4. GAVANELLI, M., F. RIGUZZI, M. MILANO, P. CAGNOLI, **Constraint and Optimization Techniques for Supporting Policy Making**. In Computational Intelligent Data Analysis for Sustainable Development, chapter 16. Taylor & Francis 2012.
 5. DENG, G., **Simulation-based Optimization**. PhD thesis, University of Wisconsin – Madison, 2007.
 6. JAFFAR, J, M. MAHER, **Constraint Logic Programming**, Journal of Logic Programming volume 19/20, 1994, pp. 503-581.
 7. MATTHEWS, R. N. GILBERT, A. ROACH, G. POLHILL, N. GOTTS, **Agent-based Land-use Models: A Review of Applications**. Landscape Ecology, volume 22(10), 2007.
 8. GILBERT, N., **Computational Social Science**. SAGE, 2010.
 9. BRECK, E., C. YEJIN, C. CLAIRE, **Identifying Expressions of Opinion in Context**, in Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-2007). 2007.
 10. CARVALHO, P., L. SARMENTO, J. TEIXEIRA, M. J. SILVA. **Liars and Saviors in a Sentiment Annotated Corpus of Comments to Political Debates**, in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: short-papers. 2011
 11. FUPOL, <http://www.fupol.eu/>
 12. e-Policy, <http://www.epolicy-project.eu/>
 13. AKIYOSHI, M., K. KIMURA, H. OISO, N. KOMODA, **Analysis Support System of Open-ended Questionnaires Based on Atypical and Typical Opinions Classification**, Studies in Informatics and Control ,Vol. 18, No. 3, 2009.
 14. MOISE, M. V. POPA **Methods and Experiments Regarding Learning Fuzzy Cognitive Maps**, Journal of Information Systems & Operations Management, Vol. 8, No 1, May 2014.

