Anna Fiedukowicz[1]
Agata Pillich-Kolipińska
Robert Olszewski

Department of Cartography
Warsaw University of Technology

[1] *This work has been supported by the European Union in the framework of European Social Fund through the Warsaw University of Technology Development Programme.*

Utilisation of computational intelligence for simplification of linear objects using extended WEA algorithm

Abstract

One of the most important challenges of the temporary cartography is the automation of cartographic modelling process – and how to compromise it with the accuracy of said process. The paper describes the author proposal of exploiting the computational intelligence methods, such as Artificial Neural Networks enhanced with fuzzy logic, to modify the multi-parameter WEA line-simplification algorithm. The knowledge database, created with such methods, would allow substituting the fixed weights used by (Zhou, Jones, 2004) in WEA algorithm, thus making the whole process more flexible and suitable for different data (and bringing it closer to automation).

The paper briefly describes the Fuzzy Inference Systems and the Artificial Neural Networks and presents some of the first results achieved by utilising those methods in the process of generalisation of geographic information – by modifying the original WEA algorithm. Proper formalisation of the knowledge base is of the key importance in the process, but can be highly subjective.

Introduction

Generalisation of geographic information may be implemented in many ways, which differ with respect to the assumed methodology, as well as to the level of automation of the modelling process. Therefore, the answer to the question concerning the possible compromise between the cartographic modelling process and its automation, is the important challenge of the contemporary cartography.

Following the authors' opinion, it is possible to find the compromise solution using the, so-called, computational intelligence (CI), and, in particular machine learning (ML) and the data mining (DM). This solution comprises a specific "transfer" of subjective, cartographic knowledge to a digital tool, which will automatically specify the adequate method and its parameters, in the process of spatial data modelling for a given level of details (LoD).

The objective of performed research was to develop the methodology of linear objects simplification with the use of the well known WEA (Weighted Effective Area) algorithm (Zhou, Jones, 2004) and the selected computational intelligence methods: artificial neural networks and fuzzy inference systems. The authors proposed to modify the multi-parameter WEA algorithm, substituting the fixed weights with a knowledge base, integrated with a computational engine, which utilises the fuzzy logic and/or artificial neural networks. This approach allows developing the knowledge base of cartographic generalisation methods, basing on two different approaches:

- the approach which utilises machine learning methods (**implicit methods**) consisting of aquiring examples of correct solutions,

- the approach which defines open, although purposefully "fuzzy" generalisation rules (**explicit methods**).

Artificial neural networks are the examples of the first approach - the *NEURO* method, and fuzzy inference systems are the examples of the second approach - the *FUZZY* method. Utilisation of both approaches allows comparing obtained results. The proposed solution allows selecting the most appropriate generalisation operators and determining their parameters.

The inference engine, in both approaches, is the knowledge base, which utilises computational intelligence methods. Depending on the knowledge acquisition ways and types of its formal representation in the knowledge base, the computational system is based either on artificial neural networks (the implicit knowledge specified as a set of examples of correct solutions) or on fuzzy inference (the explicit knowledge specified in the form of open, fuzzy rules, which utilises the, so-called, linguistic variables and membership functions) The duality of the proposed solution not only allows diversifying the way of acquisition and representation of cartographic knowledge required for automation of the generalisation process, but also comparing obtained results objectively.

Computational Intelligence

Following the IEEE Computational Intelligence Society, **computational intelligence** deals with „the theory, designing and utilisation of biologically inspired computational methods, with particular respect to neural networks, genetic algorithms, evolution programming, fuzzy inference systems and hybrid systems" (http://ieee-nns.org/). The computational intelligence, which is characterised by the iterative approach to computational tasks, is formed by artificial neural networks, fuzzy inference systems, evolution algorithms etc. Unlike the conventional meaning of the artificial intelligence, the computational intelligence does not utilise algorithms based on symbolic representation of knowledge.

The group of computational intelligence algorithms, among others, consists of fuzzy inference systems (FIS), artificial neural networks, rough sets and decision trees. The following parts of this section refer to artificial neural networks and fuzzy inference systems, as they were the methods applied by the authors.

Fuzzy Inference Systems (FIS)

The theory of fuzzy sets, proposed by Zadeh (1965, 1973) assumes that the following components are used for description of the system operations:

- the linguistic variables (e.g. *big, small, about half, enough, rather important*),
- the fuzzy conditional clauses, which express relations between linguistic variables in the form of IF-THEN rules, e.g. *if the number of inhabitants is big, then the city is important,*
- the complex inference rules, allowing for inducing resulting values based on the knowledge of the primary variable (e.g. *if dinner served in a restaurant is tasty)* and relations between variables (e.g. *if dinner is tasty and the waiter is nice, the tip will be high).*

The vast majority of cases when the fuzzy logic is applied, are connected with the widely considered process of control. The fuzzy rules calculus is applied, in which relations are expressed in the form of IF-THEN rules with predecessors and consequents, which contain linguistic variables. This allows developing a system which automates operations, similarly to the human intuition. Thus, the starting point in fuzzy controlling is the solution proposed by an expert in a given sector, in the form of fuzzy rules. In this sense, it is the descriptive and not the prescriptive solution (Zadeh, 1997). The important element of the fuzzy inference system is the database of rules. This database contains conditional clauses, which determine causal relations between input and resulting variables. The objective of the inference process, implemented in the FIS system is to find the value of the resulting variable, induced by the current values of initial variables and the database of rules.

Artificial Neural Networks (ANN)

Utilisation of neural networks has become more popular as a result of their usefulness for modelling and calculations. Advantages of the ANN may be specified as follows (Tadeusiewicz, 1998; Patterson, D., 1996):

- the ANN allows creating complex, non-linear models in a relatively simple way and makes them "learn" at the same time with the use of presented examples. This method does not require to assume, a priori, any presupposition concerning the shape and the non-linearity level of the created regression function,
- neural networks are characterised by the high level of resistance to errors. Information used for learning may be incomplete or erroneous. The neural network, which has been correctly taught, may effectively filter noises and perform calculations basing on resultant tendencies and trends,
- the important advantage of utilising neural networks is the simplicity of their use. ANN operate as the "black box": question - answer. In practice, models required by the user are constructed by the neural networks themselves, since they learn using the specified examples. The process of learning substitutes programming. As a result, computational tasks may be solved without the knowledge of algorithms, having the test set, containing "questions" and "answers" only.
- the appropriately prepared (taught) ANN is characterised by the ability to generalise, i.e. the ability to generalise the acquired knowledge.

Many types of neural networks exist – they differ by the structure (Fig. 1) and rules of operations. The mostly applied **regression networks** include the, so-called, multilayer perceptrons (MLP) and networks of radial basis functions (RBF), as well as generalized regression neural network (GRNN).

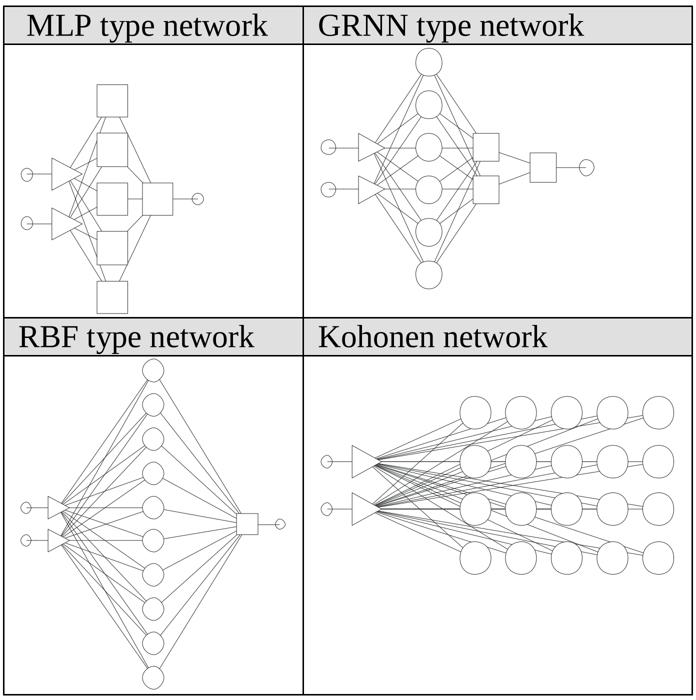


Fig. 1. Architecture of MLP, GRNN, RBF and Kohonen networks.

Utilisation of computational intelligence algorithms in cartography

Both fuzzy inference systems and artificial neural networks may be applied for wide analyses and spatial data mining (Olszewski, 2009). Each of these approaches has its advantages. Artificial neural networks are characterised by the natural ability to generalise knowledge acquired as a result of analyses of specified examples; however, it is difficult to interpret the process leading to results generated by the appropriately trained ANN. In order to simplify interpreting the process of extracting those rules, which openly describe operations of the modelled system, is often applied. It is the approach similar to defining IF-THEN relations in the fuzzy inference method. The approach of this type is specified as the neuro-fuzzy modelling. Extracted symbolic rules allow neglecting unimportant factors and extracting the basic components of a model being constructed (Gopal, Liu, Woodcock, 2001). However, optimisation of obtained solutions is important. It is particularly visible in rule-based systems. Optimisation of several rules, which create the computational engine of the system is much easier than optimisation of operations in a multi-layer, complex and specially trained artificial neural network.

Research

During their research, the authors developed the knowledge bases, using two (implicit and explicit) methods for the system of generalisation, which support operations of the WEA algorithm. This algorithm, being the extension of a standard, one-parameter EA procedure (Visvalingam-Whyatt, 1993), uses several parameters for the needs of evaluation of the importance of particular intermediate points (vertices) (Zhou, Jones, 2004) – Fig. 2:

- the elementary triangle area,
- the triangle's flatness (calculated on several ways),
- the triangle's skewness,
- the convexity.

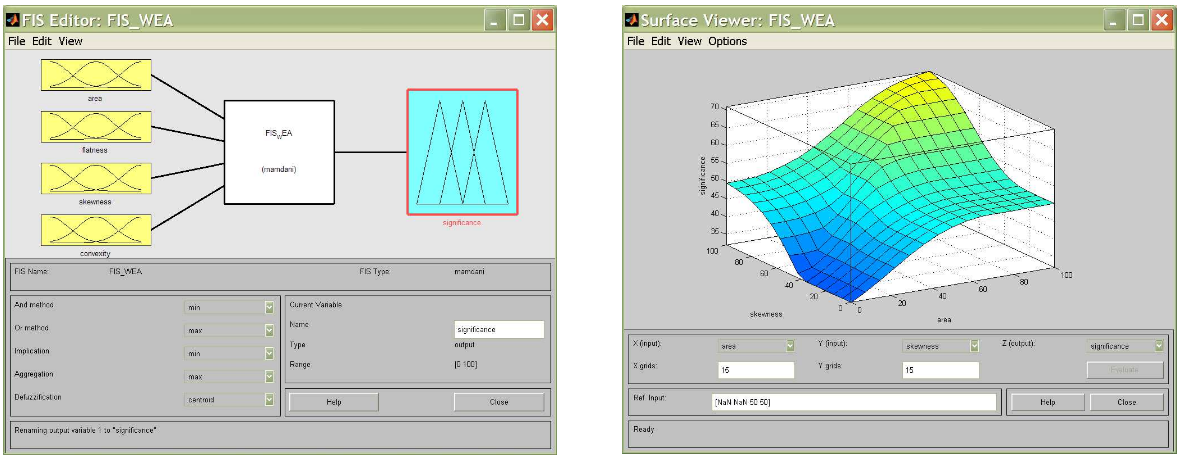


Fig. 2. Fuzzy Inference System.

In the original approach, Zhou and Jones (2004) propose to express the resulting weight of particular characteristic points of a line object using the formula:

$$WEA = W_{\text{Flat}} * W_{\text{Skew}} * W_{\text{Convex}} * EA$$

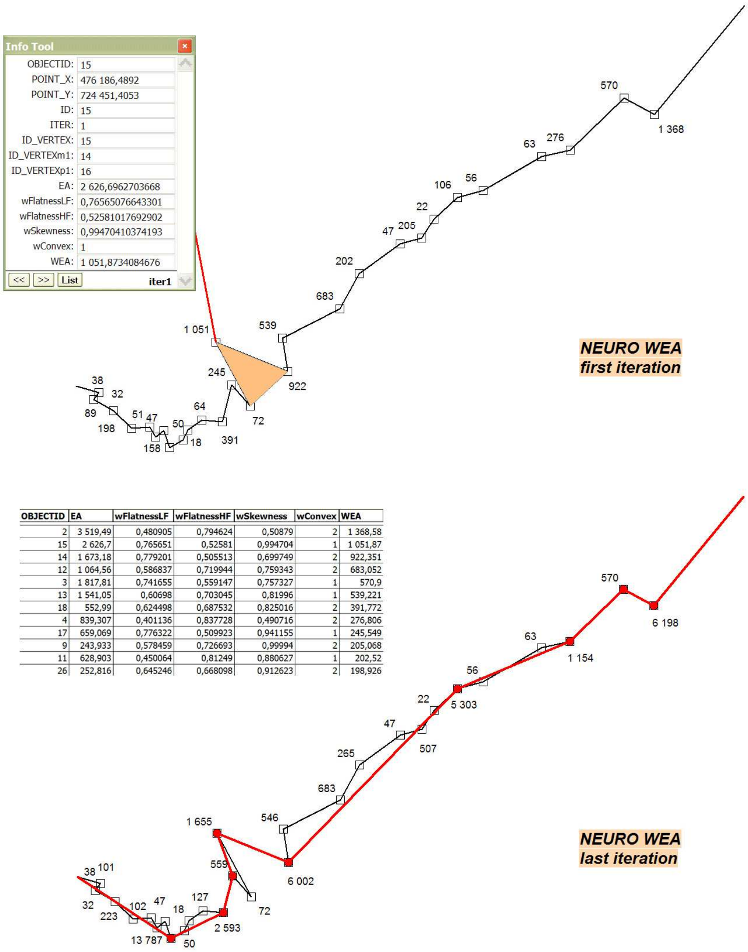
Research works performed by the authors of the presented work were based on the assumption that - instead of assuming "fixed" values of particular parameters and determining

WEA by multiplying them, it is possible to develop the cartographic knowledge base of the generalisation system. This base, after required "transfer" of the cartographic knowledge to the system, will allow implementing the expert system, which will determine the importance of particular points creating the linear object.

Data from the geodetic and cartographic resources were used for research - the river network in Poland and the coastline of Norway. Having 1:10 000, 1:50 000, 1:250 000 and 1:1000 000 scale maps, the authors:

1. selected "teaching" data for the system - the characteristic river and a fragment of the coastline and their digital cartographic representation at several scales,
2. manually specified weights of particular points utilised as a teaching set for the artificial neural network,
3. defined openly fuzzy rules concerning the importance of particular points (e.g. if a triangle is flat and its area is small, then this point is unimportant) for the fuzzy inference system (Fig. 2).

The NEURO and FUZZY knowledge bases, developed this way were utilised for generalisation of water streams (Fig. 3) and sections of a coastline other than examples used for teaching. The obtained results were compared both, with the results of manual generalisation applied on analogue maps at the scales of 1: 10K – 1: 1M, with the results of the "conventional" WEA algorithm, as well as other line simplification algorithms: Douglas-Peucker (first described in (Douglas, Peucker, 1973)) and Wang (Fig. 4).



OBJECTID	EA	wFlatnessLF	wFlatnessHF	wSkewness	wConvex	WEA
2	3 519,49	0,480905	0,794624	0,50879	2	1 368,58
15	2 626,7	0,765651	0,52581	0,994704	1	1 051,87
14	1 673,18	0,779201	0,505513	0,699749	2	922,351
12	1 064,56	0,586837	0,719944	0,759343	2	683,052
3	1 817,81	0,741655	0,559147	0,757327	1	570,9
13	1 541,05	0,60698	0,703045	0,81996	1	539,221
18	552,99	0,624498	0,687532	0,825016	2	391,772
4	839,307	0,401136	0,837728	0,490716	2	276,806
17	659,069	0,776322	0,509923	0,941155	1	245,549
9	243,933	0,578459	0,726693	0,99994	2	205,068
11	628,903	0,450064	0,81249	0,880627	1	202,52
26	252,816	0,645246	0,668098	0,912623	2	198,926

Fig. 3. Test data simplification.

Conclusions

Performed research proved the high usefulness of the computational intelligence and construction of the knowledge base in the geographic information generalisation process. The obtained results are strongly influenced not only by selection of teaching examples (their amount and quality), but also by the complexity of the neural network, its architecture (MLP, RBF or GRNN), the number of neurones in hidden layers, applied algorithms, time of teaching the ANN, division of data into a teaching and validation sets, the level of smoothing of radial functions, the nature of source data etc. Determination of these parameters, being of the key importance for formalisation of the system knowledge base, is - to the wide extent - the subjective process, which may be automated only at the low level.

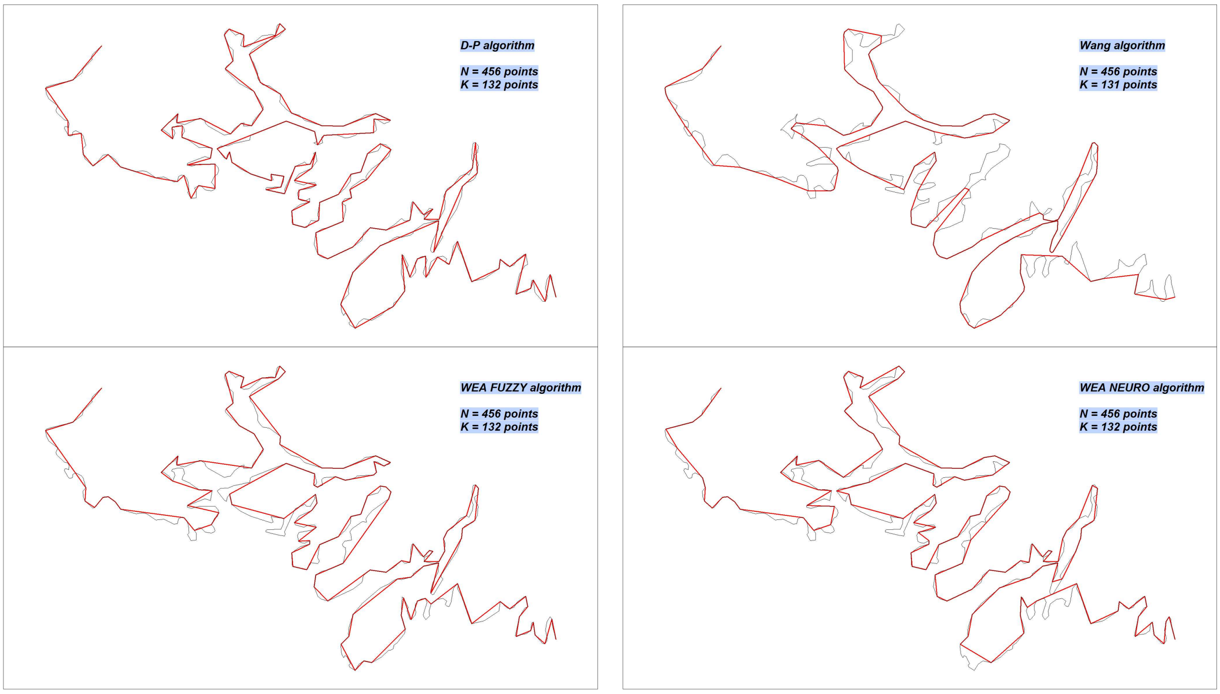


Fig. 4. The results of simplification.

Similar conclusions result from the analysis of research works performed with the use of fuzzy inference systems. The obtained results depend not only of the nature of source data, its accuracy and geometric complexity, but also on the type of the applied knowledge base of the system, the way of concluding, the number and the nature of fuzzy rules, determination of the membership function, utilised linguistic variables etc. However, similarly to the case of utilisation of ANN, methodical aspects of obtained results should be stressed. Performed research works proved that using the correctly created knowledge base and fuzzy decision rules, correct cartographic results may be obtained. However, it is difficult to point to the "optimal" set of coefficients and parameters of the method. They depend on the nature of source data, scale and destination of the results, as well as on the cartographer's knowledge and expertise.

Concluding the performed research it may be stated (without losing the generality of concluding) that obtained results of generalisation of linear objects, performed with the use of one simplification algorithm (WEA), will contribute to development of the complex methodology of geographic information generalisation by computational intelligence methods, applied for many, various operators and algorithms. The proposed approach also integrates the apparently incompatible features of the geographic information generalisation process - it allows automating the process, at the same time preserving its subjectivity.

Bibliography

1. Douglas D. H., Peucker T. K., 1973. *Algorithms for the reduction of the number of points required to represent a digitized line or its caricature*, The Canadian Cartographer, 10(2)
2. Gopal, S., Liu, W., and C. Woodcock, 2001. *Visualization based on the fuzzy ARTMAP neural network for mining remotely sensed data*, in: Geographic Data Mining and Knowledge Discovery, H.J. Miller and J. Han (eds), Taylor and Francis, p. 315-336.
3. Olszewski R., 2009. Cartographic *modelling of terrain relief using computational intelligence methods*, Proceedings of ICA Conference, Santiago de Chile
4. Patterson, D., 1996. *Artificial Neural Networks*, Singapore, Prentice Hall
5. Tadeusiewicz R., 1998. *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza PLJ
6. Visvalingam, M. and Whyatt, J.D., 1993. *Line generalisation by repeated elimination of points*, Cartographic Journal, 30(1)
7. Zhou S., Jones Ch. B., 2004. *Shape-aware line generalisation with Weighted Effective Area*, Proceedings of SDH'04
8. Zadeh L.A., 1965. *Fuzzy sets*, Information and Control 8, s. 338-353
9. Zadeh L.A., 1973. *Outline of a new approach to the analysis of complex systems and decision processes*, IEEE Transactions on systems, Man and Cybernetics SMC-2, s. 28-44
10. Zadeh L.A., 1997. Foreword to: Hoffmann F., 1997. *Entwurf von Fuzzy-Reglern mit Genetischen Algorithmen*, DUV Informatik, XI, Deutscher Universitätsverlag