Development of a technology of designing intelligent information systems for the estimation of social objects

Natalia Zenkova
Alexander Arzamastsev
Klaus G. Troitzsch

Nr. 1/2011

Arbeitsberichte aus dem Fachbereich Informatik

The "Arbeitsberichte aus dem Fachbereich Informatik" comprise preliminary results which will usually be revised for subsequent publication. Critical comments are appreciated by the authors. All rights reserved. No part of this report may be reproduced by any means or translated.

Arbeitsberichte des Fachbereichs Informatik

ISSN (Print): 1864-0346
ISSN (Online): 1864-0850

Herausgeber / Edited by:

Der Dekan:
Prof. Dr. Zöbel

Die Professoren des Fachbereichs:
Prof. Dr. Bátori, Prof. Dr. Burkhardt, Prof. Dr. Diller, Prof. Dr. Ebert, Prof. Dr. Furbach, Prof. Dr. Grimm, Prof. Dr. Hampe, Prof. Dr. Harbusch, JProf. Dr. Kilian, Prof. Dr. Lämmel, Prof. Dr. Lautenbach, Prof. Dr. Müller, Prof. Dr. Oppermann, Prof. Dr. Paulus, Prof. Dr. Priese, Prof. Dr. Rosendahl, Prof. Dr. Schubert, Prof. Dr. Staab, Prof. Dr. Steigner, Prof. Dr. Sure, Prof. Dr. Troitzsch, Prof. Dr. von Kortzfleisch, Prof. Dr. Walsh, Prof. Dr. Wimmer, Prof. Dr. Zöbel

Kontaktdaten der Verfasser

Klaus G. Troitzsch, Natalia Zenkova, Alexander Arzamastsev
Institut für Wirtschafts- und Verwaltungsinformatik
Fachbereich Informatik
Universität Koblenz-Landau
Universitätstraße 1
D-56070 Koblenz
EMail: kgt@uni-koblenz.de, natulin@mail.ru, arz_sci@mail.ru
Development of a technology of designing intelligent information systems for the estimation of social objects

Natalia Zenkova, Alexander Arzamastsev, Klaus G. Troitzsch

The estimation of various social objects is necessary in different fields of social life, science, education, etc. This estimation is usually used for forecasting, for evaluating of different properties and for other goals in complex man-machine systems.

At present this estimation is possible by means of computer and mathematical simulation methods which is connected with significant difficulties, such as:

- time-distributed process of receiving information about the object;
- determination of a corresponding mathematical device and structure identification of the mathematical model;
- approximation of the mathematical model to real data, generalization and parametric identification of the mathematical model;
- identification of the structure of the links of the real social object.

The solution of these problems is impossible without a special intellectual information system which combines different processes and allows predicting the behaviour of such an object. However, most existing information systems lead to the solution of only one special problem. From this point of view the development of a more general technology of designing such systems is very important.

The technology of intellectual information system development for estimation and forecasting the professional ability of respondents in the sphere of education can be a concrete example of such a technology. Job orientation is necessary and topical in present economic conditions. It helps to solve the problem of expediency of investments to a certain sphere of education. Scientifically validated combined diagnostic methods of job orientation are necessary to carry out professional selection in higher education establishments. The requirements of a modern society are growing, with the earlier developed techniques being unable to correspond to them sufficiently.

All these techniques lack an opportunity to account all necessary professional and personal characteristics. Therefore, it is necessary to use a system of various tests.

Thus, the development of new methods of job orientation for entrants is necessary. The information model of the process of job orientation is necessary for this purpose.

Therefore, it would be desirable to have an information system capable of giving recommendations concerning the choice of a trade on the basis of complex personal characteristics of entrants.

Concept of building intellectual information systems

A concept of building intellectual information systems includes the following main components: an information part and a module which works with artificial neural networks (ANN).

The purpose of the information part consists in the accumulation, storage and provision of the information about the objects, and also maintenance of the interface for the end user.
The purpose of the module working with ANN is the construction of the artificial neural network and its training on the sets of input variables of the objects with their corresponding output conditions. An ANN is capable of classifying new information and, on its basis, of being trained further. The result of the work of this component is the intellectual model of the object.

**Information part of the intellectual information system (IIS)**

The information part of the system represents a set of PHP-scripts co-operating with a MySQL database. This component functions according to the scheme presented in Fig. 1.

![Fig. 1. The scheme of functioning of the information part of the system](image1)

Figure 2 contains a more detailed description of the processes in the information part of the IIS.

![Fig. 2. The mechanism of functioning of the information part of the IIS.](image2)
The links presented in Fig. 2 have the following meaning:

1. The user sends the request to the HTTP-server, containing some data. Any Internet browser (Internet Explorer, Mozilla Firefox, Opera, etc.) can be used for connecting to the HTTP-server as the client of the system.
2. HTTP-server translates PHP-scripts, acting in accordance with the user’s request.
3. PHP-scripts refer to a database with request.
4. Request results are sent to the PHP-scripts to process.
5. The result of processing is a generated HTML-page.
6. The HTTP-server sends the user the response which is a generated HTML-page which is displayed in the user’s Web-browser.

This component supports three levels of access — the user, an expert and an administrator. A specific set of authorities and functional capabilities corresponds to each level:

a. administrator-level authorities control the user management;
b. the expert level is empowered to manage the object;
c. the user level allows to enter information about the parameters of the object and to get a result on the basis of entered data.

The specific problem under consideration in a given subject area is called an object in the intellectual information system.

The object of the system is defined by a set of input variables with the corresponding output conditions. Creating an object is possible for the expert and associated with the determination of such characteristics as:

- the number of input parameters;
- input parameters that will be used for later analysis;
- the proposed output conditions.

The accumulation of information about the object can be done in two ways.

**First (with user and expert):** In this case, the user is registered in the system and introduces the variable values of the object. These data are stored in the database. Then, the expert determines the appropriate output condition of the object by analysing a set of input data that was previously entered by the user. Thus, the knowledge is stored in the database and forms the base of the knowledge about the object. The information which is stored in the knowledge base generates the classification defined in accordance with the output conditions of the object.

**Second (without user):** In this case, the expert, already having sets of input data about the object, can load them into the system. The set can already be analysed or evaluated by the expert subsequently.

**Intellectual subsystem of the intellectual information system**

The intellectual subsystem is intended to implement the processes of constructing, training, modification and operation of ANN models. ANN models are autonomous systems of knowledge representation in the IIS (Fig. 3).

This subsystem includes the following basic components:
1. block of data monitoring, which provides transfer of input and output data; the monitoring block tracks the incoming flow of new information into the database;
2. converter, preparing the training set for the network;
3. constructor of the ANN, providing automated creation and training of the intellectual model (ANN) of the object.
4. The mechanism of interaction between the components of the integrated system is presented in Fig. 4.

Fig. 3. The scheme of functioning of the intellectual part of the system

Fig. 4. Structural diagram of the intellectual information system
This intellectual information system includes the basic distinctive components, such as: web-interface, the database used for long-term storage of the technical information, the intellectual core including the developing ANN model based system of the experts' knowledge representation. The intellectual core of the information system will consist of ANN models, i.e. models constructed on the basis of artificial neural networks.

Figure 5 shows the functioning of the intellectual information system (IIS).

![Diagram of IIS functioning](image)

The top part of Fig. 5 presents the dependence of the primary information accumulation over time (in arbitrary time units), $R$ is the number of records in the database received from the terminals. The bottom part of fig. 5 presents the process of training the ANN model; $E$ is the reduced mean-square error of the ANN model, and $L$ shows the levels of training of the ANN model.

\[
E(t) = E^*(1 - e^{-kt}) + E_0 e^{-kt},
\]
where

- \( E(t), E_0 \) and \( E^* \) are the values of the reduced mean-square error, its initial value and the level at which it will be fixed at the end of each training cycle of the ANN model;
- \( k \) is the parameter of the specific training rate which depends on \( n \), the capacity size of the training set, using the method of residual minimisation;
- \( t \) is the time in arbitrary units.

The learning process of the ANN model (the intellectual core of the IIS) depends on the number and quality of data in the training set.

It is necessary to have a lot of empirical data for the identification of the complex structure of the object. While new data come in, the expert will complicate the structure of the ANN model to achieve an acceptable level of adequacy. Therefore, this is a multistage training process. When level \( L_5 \) (an acceptable level of the model training) is reached, then the expert will finish the training of the model and switch the IIS into the operation mode.

The expert has the following rights:

- to determine the acceptable level of the error of the model,
- to complete the process of training the model,
- to determine the output conditions of the object corresponding to the input data (if the training set does not contain output parameters).

The mathematical formalization of the processes of information accumulation and building the developing intellectual core on its base was described earlier in [1]. It allows solving the following problems:

- the estimation of the primary period of information accumulation, which is necessary for building the ANN model of the research object;
- the determination of the primary structure of the ANN model on the basis of input characteristics of the research object;
- the estimation of the necessary number of training cycles of the intellectual core by the expert.

The technology of the design of intellectual information systems on the basis of an ANN model and with a distributed input of the data consists of the following steps:

- study of the research object (for developing IIS); determination of the input variables and the possible output states of the object;
- accumulation of primary information about the object in the database through a distributed system of data collection;
- generation of the primary structure of the ANN model;
- planning of the step-by-step process of the developing ANN model; training of the ANN model;
- development of the ANN model on the basis of new empirical data coming in; this step needs the control of an expert who can determine when the formation of the knowledge representation will be finished;
• operation of the intellectual information system (without expert).

**Application of the technology of intellectual information system design for the estimation of social objects**

The social object can be represented by investigating results of students of Tambov State University named after G.R. Derzhavin (TSU). For the construction of the intellectual core of the information system we can use the empirical data received earlier by the staff of the Computer and Mathematical Simulation Department at TSU. These empirical data come from questioning high school students in Tambov [2].

These students answered questions of a questionnaire which consisted of the following sections: statistical information; progress; professional potential; public activity; creativity; leisure; health; family.

The results of the questionnaire will be both input and output variables (which correspond to database fields) on the basis of which the training of the artificial neural network will be done.

For the artificial neural network we can use the software which we developed during our previous activity [3]. This program is a universal neural network simulator using an original algorithm of self-organization of neural network structure which allows achieving good flexibility and adaptability of the model to empirical data. The algorithm is described in [3].

This technology of intellectual information system design can be used for simulation of other various social, medical and other objects as well.

The advantages of this technology are: universality for a variety of subject areas, multi-user access to the same object of research, developing an intellectual core on the basis of ANN models, and support of the networking technology of interaction.

**Acknowledgments**

This work was supported by the DAAD program “Michail Lomonosov” (B) (grant # A0974042, 2010/2011).

**References**


Bisher erschienen

Arbeitsberichte aus dem Fachbereich Informatik
(http://www.uni-koblenz-landau.de/koblenz/fb4/publications/Reports/arbeitsberichte)

Klaus G. Troitzsch, Natalia Zenkova, Alexander Arzamastsev, Development of a technology
designing intelligent information systems for the estimation of social objects,
Arbeitsberichte aus dem Fachbereich Informatik 1/2011

Kurt Lautenbach, A Petri Net Approach for Propagating Probabilities and Mass Functions,
Arbeitsberichte aus dem Fachbereich Informatik 13/2010

Claudia Schon, Linkless Normal Form for ALC Concepts, Arbeitsberichte aus dem
Fachbereich Informatik 12/2010

Alexander Hug, Informatik hautnah erleben, Arbeitsberichte aus dem Fachbereich Informatik
11/2010

Marc Santos, Harald F.O. von Kortzfleisch, Shared Annotation Model – Ein Datenmodell für
kollaborative Annotationen, Arbeitsberichte aus dem Fachbereich Informatik 10/2010

Gerd Grüner, Steffen Staab, Categorization and Recognition of Ontology Refactoring Pattern,
Arbeitsberichte aus dem Fachbereich Informatik 9/2010

Daniel Eißing, Ansgar Scherp, Carsten Saathoff, Integration of Existing Multimedia Metadata
Formats and Metadata Standards in the M3O, Arbeitsberichte aus dem Fachbereich Informatik
8/2010

Stefan Scheglmann, Ansgar Scherp, Steffen Staab, Model-driven Generation of APIs for
OWL-based Ontologies, Arbeitsberichte aus dem Fachbereich Informatik 7/2010

Daniel Schmeiß, Ansgar Scherp, Steffen Staab, Integrated Mobile Visualization and
Interaction of Events and POIs, Arbeitsberichte aus dem Fachbereich Informatik 6/2010

Rüdiger Grimm, Daniel Pähler, E-Mail-Forensik – IP-Adressen und ihre Zuordnung zu
Internet-Teilnehmern und ihren Standorten, Arbeitsberichte aus dem Fachbereich Informatik
5/2010

Christoph Ringelstein, Steffen Staab, PAPEL: Syntax and Semantics for Provenance-Aware
Policy Definition, Arbeitsberichte aus dem Fachbereich Informatik 4/2010

Nadine Lindermann, Sylvia Valcárcel, Harald F.O. von Kortzfleisch, Ein Stufenmodell für
kollaborative offene Innovationsprozesse in Netzwerken kleiner und mittlerer Unternehmen
mit Web 2.0, Arbeitsberichte aus dem Fachbereich Informatik 3/2010

Maria Wimmer, Dagmar Lück-Schneider, Uwe Brinkhoff, Erich Schweighofer, Siegfried
Kaiser, Andreas Wieber, Fachtagung Verwaltungsinformatik FTVI Fachtagung
Rechtsinformatik FTRI 2010, Arbeitsberichte aus dem Fachbereich Informatik 2/2010

Max Braun, Ansgar Scherp, Steffen Staab, Collaborative Creation of Semantic Points of
Interest as Linked Data on the Mobile Phone, Arbeitsberichte aus dem Fachbereich Informatik
1/2010

Marc Santos, Einsatz von „Shared In-situ Problem Solving“ Annotationen in kollaborativen
Lern- und Arbeitsszenarien, Arbeitsberichte aus dem Fachbereich Informatik 20/2009
Carsten Saathoff, Ansgar Scherp, Unlocking the Semantics of Multimedia Presentations in the Web with the Multimedia Metadata Ontology, Arbeitsberichte aus dem Fachbereich Informatik 19/2009

Christoph Kahle, Mario Schaarschmidt, Harald F.O. von Kortzfleisch, Open Innovation: Kundenintegration am Beispiel von IPTV, Arbeitsberichte aus dem Fachbereich Informatik 18/2009


Andreas Fuhr, Tassilo Horn, Andreas Winter, Model-Driven Software Migration Extending SOMA, Arbeitsberichte aus dem Fachbereich Informatik 16/2009

Eckhard Großmann, Sascha Strauß, Tassilo Horn, Volker Riediger, Abbildung von grUML nach XSD soamig, Arbeitsberichte aus dem Fachbereich Informatik 15/2009

Kerstin Falkowski, Jürgen Ebert, The STOR Component System Interim Report, Arbeitsberichte aus dem Fachbereich Informatik 14/2009

Sebastian Magnus, Markus Maron, An Empirical Study to Evaluate the Location of Advertisement Panels by Using a Mobile Marketing Tool, Arbeitsberichte aus dem Fachbereich Informatik 13/2009

Sebastian Magnus, Markus Maron, Konzept einer Public Key Infrastruktur in iCity, Arbeitsberichte aus dem Fachbereich Informatik 12/2009

Sebastian Magnus, Markus Maron, A Public Key Infrastructure in Ambient Information and Transaction Systems, Arbeitsberichte aus dem Fachbereich Informatik 11/2009


Hannes Schwarz, Jürgen Ebert, Andreas Winter, Graph-based Traceability – A Comprehensive Approach. Arbeitsberichte aus dem Fachbereich Informatik 4/2009


Christoph Wernhard: Tableaux Between Proving, Projection and Compilation, Arbeitsberichte aus dem Fachbereich Informatik 18/2007

Ulrich Furbach, Claudia Obermaier: Knowledge Compilation for Description Logics, Arbeitsberichte aus dem Fachbereich Informatik 17/2007


Rüdiger Grimm, Anastasia Meletiadou: Rollenbasierte Zugriffskontrolle (RBAC) im Gesundheitswesen, Arbeitsberichte aus dem Fachbereich Informatik 15/2007


Ulrich Furbach, Markus Maron, Kevin Read: Location based Informationsystems, Arbeitsberichte aus dem Fachbereich Informatik, 11/2007


