Computing and Informatics, Vol. 40, 2021, 1197–1216, doi: 10.31577/cai_2021_6_1197

APPLICATION OF THE FUZZY MODEL THEORY FOR MODELING QA-SYSTEMS

Gulnara Yakhyaeva, Anastosia Karmanova, Aleksei Ershov

Department of Information Technologies
Novosibirsk State University
Pirogova st. 1, 630090 Novosibirsk, Russian Federation
e-mail: gul_nara@mail.ru, {anast.karmy, alaershov}@gmail.com

Abstract. The work is devoted to the description of the question-answer system QA-RiskPanel, which provides means of determining and forecasting the risks related to computer attacks. The QA-RiskPanel system uses a constantly updated database of previous computer attacks as a source of knowledge. We thus guarantee the most up-to-date risk prediction. The ontological approach to the formalization of the object domain allows the analysis of risks at various levels of specification/generalization. In this paper we provide a model-theoretic formalization of the Knowledge Base of the described object domain. Then we describe the classification of question types, which are probabilistic in this system. Finally we present algorithms for finding the answers to all question types of our classification.

Keywords: Question-answering system, knowledge base, theory of the fuzzy models, generalized fuzzy model, information security, computer attacks, case of the computer attack

Mathematics Subject Classification 2010: 68-T35

1 INTRODUCTION

Nowadays it is hard to overestimate the value of effective risk management in the field of information security [1]. Big companies often report great financial losses as consequences of hacker attacks and exposure of valuable information. When a threat is present, it is essential to know the symptoms of the attack, potential losses, possible solutions and any other information that may help to quickly take the right countermeasures and solve the problem.

To solve this problem, a knowledge base system for information risk management was developed at Novosibirsk State University [2]. The system allows the information security administrator to explicitly work with cases of computer attacks and take measures to prevent attacks and protect computer systems from damage. The system also assists the administrator in stopping destructive actions at the first stages and promptly handle its consequences in order to minimize the damage.

The software system RiskPanel is modular, which allows plugging in of new modules in case of need. For instance, core modules were complemented by the early attack recognition module [3], based on JSM-method [4] and formal concept analysis [5].

This work describes the question answering system QA-RiskPanel, developed as a part of the software system RiskPanel. The Knowledge Base of the described system is based on the case model approach. The user of the QA-RiskPanel system can ask probability questions, directed towards computer attack prediction and studying of related risks.

2 RELATED WORK

Currently, there are two main domains of question answering research: information retrieval systems and intelligence support systems [6].

Question answering systems, developed in the context of information retrieval approach, aim to find text fragments on the Internet that answer the question of the user [7]. Conventionally, question-answering systems using this approach are divided into two types: open domain and restricted domain systems [8, 9].

Open domain systems are *general*, because their purpose is to answer questions from any object domain. Similarly, restricted domain systems are *specialized* and developed to answer only questions from some specific object domain.

However, both kinds of these systems use natural language processing and text mining methods. The only difference is the degree of the ontology development.

Open domain question answering systems use general ontologies of the natural language (e.g. WordNet). Therefore, such issues of the natural language processing as synonymy, lexical ambiguity, polymorphism etc. have to be addressed in these systems [10, 7]. As a rule, there is a limited class of questions that modern question answering systems with open domain can formalize for further processing. Such questions are called factographic. Factographic questions can be divided into several subclasses: questions about people, time, toponyms, lists of some things, definitions etc. [11].

Question answering systems with restricted domain use dedicated ontologies. On the one hand, it helps to solve the problem of ambiguity of the natural language words. On the other hand, such systems can find answers to more specific questions of the object domain.

Studies of artificial intelligence led us to the development of knowledge-based question answering systems, which use various Knowledge Bases as data sources. Obviously, such systems have a restricted object domain. It is important to note that knowledge-based systems can only find answers in the information contained in the Knowledge Base, and therefore are less flexible in construction of questions.

However, the main advantage of this approach lies in the conceptual model of the object domain, represented by the structure of the Knowledge Base. This model allows advanced methods of structured information processing to be used, such as logical deduction, analogical reasoning etc. This, in its turn, shifts the goals of development of such systems. Usually, they focus not on searching and localizing the requested information, but on revelation of hidden rules and patterns, analysis of critical situations and description of risks in the given object domain.

For instance, the system L & C [12], developed for medical tasks, solves the problem of integrating expert medical knowledge with personal information about patients. The Demner-Fushman's question answering system [13] is based on the application of statistical methods to clinical medicine. The *Diagnostic Panel* system [14] designed for "spinal deformity and degenerative diseases of the spine" object domain and is based on the methods of statistical processing of the data retrieved from the medical documents written in natural language. System [15] is devoted to methods for identifying payment plans and services by mobile operators which are the best for the given subscribers. The WEBCOOP system [16] uses logical deduction to generate answers in the field of tourism.

3 THE KNOWLEDGE BASE OF COMPUTER SECURITY

3.1 The Structure of the Knowledge Base

The object domain Δ = *"Computer attacks"* is formalized in terms of the Fuzzy Model Theory [17] with use of the Description Logic methods [18]. The first step is to introduce the set $\mathbf{P}$ of atomic concepts of the object domain Δ. All atomic concepts are divided into six subsets:

$$\begin{aligned}
&\mathbf{P}_1: \text{"Symptoms"};\\
&\mathbf{P}_2: \text{"Threats"};\\
&\mathbf{P}_3: \text{"Vulnerabilities"};\\
&\mathbf{P}_4: \text{"Consequences"};\\
&\mathbf{P}_5: \text{"Losses"};\\
&\mathbf{P}_6: \text{"Countermeasures"}.
\end{aligned}$$

The set of all atomic concepts $\mathbf{P}$ is formed from the content of the National Vulnerability Database (the agency NIST[1]). NVD is the U.S. government repository of standards based vulnerability management data. It is a regularly updated database of security vulnerabilities. The vulnerability description contains information about the version of software that has the vulnerability, ways of exploiting the vulnerability, possible countermeasures, etc.

Each subset $\mathbf{P}_i$ of the atomic concepts is hierarchically ordered set. The ontology of the object domain Δ (let us call it *TBox*, which is traditional in Description Logic) includes the set $\mathbf{P}$ and the set of all specialization axioms which represent the hierarchical structure of the set $\mathbf{P}$.

The set of all concepts *CON* of the object domain Δ is built according to the syntax of the Description Logic. Each concept $\varphi \in CON$ is a Boolean combination of the atomic concepts from $\mathbf{P}$.

TBox is the first component of the Knowledge Base of the object domain Δ. Description of the computer attack cases is the second component of the Knowledge Base. The information about computer attacks for QA-RiskPanel mainly comes from NIST and MITRE[2] databases. Each attack e is characterized by the presence/absence of traits from classes $\mathbf{P}_i$. Therefore, the atomic concept at this stage is regarded as a unary predicate, i.e. $P(x) \in \mathbf{P}$.

Further, we supplement the knowledge about truth of the atomic concepts for various computer attack cases through the use of axioms from *TBox*. Let us introduce the following notation:

$$ABox = \{P(e) \mid \text{atomic concept } P(x) \text{ is true on the case } e\}.$$

In what follows the pair $\mathbf{KB} = \langle TBox, ABox \rangle$ will be called Knowledge Base of the object domain Δ. The Knowledge Base will be expanded with the appearance of new concepts of the given object domain (e.g. new threats, vulnerabilities etc.) or new cases of computer attacks. However, the structure of the Knowledge Base remains the same.

3.2 Theoretical-Modal Formalization of the Knowledge Base

We use a Case Model and a Fuzzy Model of the object domain Δ for statistical processing of data [17]. We construct these models on the basis of the class of interpretations of the Knowledge Base $\mathbf{KB}$.

Let us consider the finite set of computer attacks $E = \{e_1, \ldots, e_n\}$ which were previously used to describe *ABox* and the class $\mathbf{P}$ of unary predicates describing *TBox*.

[1] `http://www.nist.gov/`

[2] `http://www.mitre.org/`

Definition 1. The algebraic system $\mathcal{A}_E = \langle E, P \rangle$ is called *Interpretation* of the Knowledge Base **KB** if $\mathcal{A}_E \models ABox$ (i.e. $\mathcal{A}_E \models \varphi(e_i)$ for each sentence $\varphi(e_i) \in ABox$).

Definition 2. Ordered triple $Case(\mathcal{A}_E) = \langle \{a\}, P, \tau \rangle$ is called *Case Model* generated by the Interpretation $\mathcal{A}_E = \langle E, P \rangle$ if for each concept $\varphi(x) \in CON$ we have $\tau(\varphi(a)) = \{e \in E | \mathcal{A}_E \models \varphi(e)\}$.

The Case Model associates each concept with the set of computer attack cases that have this concept. Notice, that model $Case(\mathcal{A}_E)$ is a Boolean-valued model. This model associates each sentence of the signature $P \cup \{c_a\}$ with an element of the Boolean algebra $\rho(E)$ [17].

Most methods of statistical data processing use an objective and/or subjective probability concept. The objective probability is the relative frequency of actual occurrences of some event within a total set of observations. Another interpretation of the objective probability is the relation of the amount of 'positive' observations to the total amount of observations. The subjective probability is a degree of confidence of an expert or a group of experts that some particular event will happen [19].

The proposed approach uses concept of the Fuzzy Model to describe objective probability.

Definition 3. The ordered triple $Fuz(\mathcal{A}_E) = \langle \{a\}, P, \mu \rangle$ is called *Fuzzy Model* of the object domain Δ generated by the Interpretation $\mathcal{A}_E = \langle E, P \rangle$ if for each concept $\varphi(x) \in CON$ the following is true:

$$\mu(\varphi(a)) = \frac{\|\{e \in E \mid \mathcal{A}_E \models \varphi(e)\}\|}{\|E\|}.$$

The truth value of sentence (concept) is a number from the interval [0,1] on the Fuzzy Model. This number represents the objective probability that the concept is true on the potential computer attack. The works [17, 20] contain more detailed description of the properties of the Case Models and the Fuzzy Models.

It is worth mentioning that the information from the Internet is often incomplete. Therefore, there is a class of various Interpretations of the Knowledge Base **KB**. Let us denote this class by $\mathcal{I}_E$, i.e. $\mathcal{I}_E = \{\mathcal{A}_E = \langle E, P \rangle \mid \mathcal{A}_E \models ABox\}$.

Definition 4. An ordered triple $Fuz(\mathcal{I}_E) = \langle \{a\}, P, \xi_E \rangle$ is called *Generalized Fuzzy Model* generated by the class of Interpretations $\mathcal{I}_E$ if for each concept $\varphi(x) \in CON$ we have

$$\xi_E(\varphi(a)) = \{\mu(\varphi(a)) \mid Fuz(\mathcal{A}_E) = \langle \{a\}, \sigma_\Delta, \mu \rangle \text{ and } \mathcal{A}_E \in \mathcal{I}_E\}.$$

Therefore, truth values of sentences are subsets of rational numbers from the interval $[0, 1]$ on the Generalized Fuzzy Model. Furthermore truth values of sentences are the intervals defined on the set

$$Q^n = \left\{0, \frac{1}{n}, \ldots, \frac{n-1}{n}, 1\right\}$$

where $n = \|E\|$ is the number of cases in the Knowledge Base [21].

In strictly mathematical sense the Generalized Fuzzy Model $Fuz(\mathcal{I}_E)$ is not the interval model. However if $n \to \infty$ then truth values of sentences will be approximated to the intervals on the set $[0,1] \cap Q$. Therefore we can consider the truth values on the model $Fuz(\mathcal{I}_E)$ as intervals of rational numbers in practice. Based upon that, let us denote

$$Fuz(\mathcal{I}_E) \models_{[\alpha,\beta]} \varphi(a)$$

where $\alpha = \inf(\xi_E(\varphi(a)))$ and $\beta = \sup(\xi_E(\varphi(a)))$. In the special case when $\alpha = \beta$, we denote $Fuz(\mathcal{I}_E) \models_\alpha \varphi(a)$. We say that $\varphi(a)$ is true on the model $Fuz(\mathcal{I}_E)$ if $Fuz(\mathcal{I}_E) \models_1 \varphi(a)$ and false if $Fuz(\mathcal{I}_E) \models_0 \varphi(a)$.

Consider the following subsets of the set of cases E:

$$T(E,\varphi) = \{e \in E \mid \forall \mathcal{A}_E \in \mathcal{I}_E : \mathcal{A}_E \models \varphi(e)\}, \tag{1}$$

$$F(E,\varphi) = \{e \in E \mid \forall \mathcal{A}_E \in \mathcal{I}_E : \mathcal{A}_E \not\models \varphi(e)\}, \tag{2}$$

$$N(E,\varphi) = E \backslash (T(E,\varphi) \cup F(E,\varphi)). \tag{3}$$

Let $Fuz(\mathcal{I}_E) \models_{[\beta_1,\beta_2]} \varphi(a)$. Then, by the Definitions 3 and 4 we have

$$\|T(E,\varphi)\| = \beta_1 \|E\|, \tag{4}$$

$$\|N(E,\varphi)\| = (\beta_2 - \beta_1)\|E\|, \tag{5}$$

$$\|F(E,\varphi)\| = (1 - \beta_2)\|E\|. \tag{6}$$

Notice that $\inf(\xi_E(\varphi(a))) = \sup(\xi_E(\varphi(a)))$ if and only if $N(E,\varphi) = \emptyset$. Furthermore

$$Fuz(\mathcal{I}_E) \models_1 \varphi(a) \Leftrightarrow \|E\| = \|T(E,\varphi)\|,$$

$$Fuz(\mathcal{I}_E) \models_0 \varphi(a) \Leftrightarrow \|E\| = \|F(E,\varphi)\|.$$

3.3 Optimal Restrictions of the Knowledge Base

To find answers to conditional questions (see Section 4.4) we need to restrict the Knowledge Base **KB**, leaving as many cases of computer attacks from the set E as necessary to make the Knowledge Base satisfy the given condition. Moreover, we try to exclude only the minimum necessary amount of cases from the Knowledge Base, so that it remains as big as possible.

Definition 5. Let $E_1 \subseteq E(E_1 \neq \emptyset)$ and the model $\mathcal{A}_E = \langle E, P \rangle$ be the Interpretation of the Knowledge Base **KB**. Then

1. Model $\mathcal{A}_{E_1}$ is called **Restriction** of the Interpretation $\mathcal{A}_E$ on the set of cases E_1 (denoted $\mathcal{A}_{E_1} \subseteq \mathcal{A}_E$).
2. Class of models $\mathcal{I}_{E_1} = \{\mathcal{A}_{E_1} = \langle E_1, P \rangle \mid \exists \mathcal{A}_E \in \mathcal{I}_E : \mathcal{A}_{E_1} \subseteq \mathcal{A}_E\}$ is called the *Restriction* of the class $\mathcal{I}_E$ to the set E_1.
3. Generalized Fuzzy Model $Fuz(\mathcal{I}_{E_1})$ generated by the class $\mathcal{I}_{E_1}$ is called the *Restriction* of the model $Fuz(\mathcal{I}_E)$ and denoted $Fuz(\mathcal{I}_{E_1}) \preceq Fuz(\mathcal{I}_E)$.

Let $\rho(Fuz(\mathcal{I}_E))$ denote the set of all Restrictions of the model $Fuz(\mathcal{I}_E)$. The signature P of the Generalized Fuzzy Model $Fuz(\mathcal{I}_E)$ is purely predicative. Hence each model $Fuz(\mathcal{I}_{E_1}) \preceq Fuz(\mathcal{I}_E)$ is uniquely defined by the set of cases $E_1 \subseteq E$. It is easy to prove that partially ordered set $\langle \rho(Fuz(\mathcal{I}_E)), \preceq \rangle$ is a Boolean lattice.

Let $\varphi(x) \in CON$ and $\alpha \in [0,1]$. Consider the following subsets of the set $\rho(Fuz(\mathcal{I}_E))$:

$$M(\varphi \geq \alpha) = \{Fuz(\mathcal{I}_{E_1}) \preceq Fuz(\mathcal{I}_E) \mid \xi_{E_1}(\varphi(a)) \subseteq [\alpha, 1]\},$$
$$M(\varphi \leq \alpha) = \{Fuz(\mathcal{I}_{E_1}) \preceq Fuz(\mathcal{I}_E) \mid \xi_{E_1}(\varphi(a)) \subseteq [0, \alpha]\}.$$

Note that

$$M(\varphi \geq \alpha) = \emptyset \Leftrightarrow \alpha \neq 0 \text{ and } Fuz(\mathcal{I}_E) \models_0 \varphi(a),$$
$$M(\varphi \leq \alpha) = \emptyset \Leftrightarrow \alpha \neq 1 \text{ and } Fuz(\mathcal{I}_E) \models_1 \varphi(a).$$

Let $Fuz(\mathcal{I}_E) \models_{[\beta_1,\beta_2]} \varphi(a)$. If $\alpha \leq \beta_1$ then $Fuz(\mathcal{I}_E) \in M(\varphi \geq \alpha)$. Consequently this model is the largest in the ordered set $\langle M(\varphi \geq \alpha), \preceq \rangle$.

On the other hand if $\alpha = 1$ then $M(\varphi \geq \alpha) = \{Fuz(\mathcal{I}_{E_1}) | E_1 \subseteq T(E, \varphi)\}$. Hence the model $Fuz(\mathcal{I}_{T(E,\varphi)})$ is the largest in the ordered set $\langle M(\varphi \geq \alpha), \preceq \rangle$.

By analogy if $\alpha \geq \beta_2$ then model $Fuz(\mathcal{I}_E)$ is the largest in the ordered set $\langle M(\varphi \leq \alpha), \preceq \rangle$. And if $\alpha = 0$ then the model $Fuz(\mathcal{I}_{F(E,\varphi)})$ is the largest in the ordered set $\langle M(\varphi \leq \alpha), \preceq \rangle$.

In other cases, if the sets $M(\varphi \geq \alpha)$ and $M(\varphi \leq \alpha)$ are not empty then partially ordered sets $\langle M(\varphi \geq \alpha) \preceq \rangle$ and $\langle M(\varphi \leq \alpha), \preceq \rangle$ have more than one maximum element.

Proposition 1. Let $E_1 \subseteq E(E_1 \neq \emptyset)$.

1. If model $Fuz(\mathcal{I}_{E_1})$ is the maximum in the ordered set $\langle M(\varphi \geq \alpha); \leq \rangle$ then $T(E_1, \varphi) = T(E, \varphi)$.
2. If model $Fuz(\mathcal{I}_{E_1})$ is the maximum in the ordered set $\langle M(\varphi \leq \alpha); \leq \rangle$ then $F(E_1, \varphi) = F(E, \varphi)$.

Proof. Let us prove the item 1. The item 2. is proven the same way.

Consider the model $Fuz(\mathcal{I}_{E_1}) \in \rho(Fuz(\mathcal{I}_E))$ such that $Fuz(\mathcal{I}_{E_1}) \in M(\varphi \geq \alpha)$ and $T(E,\varphi) \setminus T(E_1,\varphi) \neq \emptyset$. Let us prove that the model $Fuz(\mathcal{I}_{E_1})$ is not maximum in the set $\langle M(\varphi \leq \alpha); \preceq\rangle$.

If $Fuz(\mathcal{I}_{E_1}) \in M(\varphi \geq \alpha)$ then $\alpha \leq \frac{\|T(E_1,\varphi)\|}{\|E_1\|}$. And if $T(E,\varphi) \setminus T(E_1,\varphi) \neq \emptyset$ then there is at least one case e such that $e \in T(E,\varphi) \setminus T(E_1,\varphi)$.

Consider the model $Fuz(\mathcal{I}_{E_1 \cup \{e\}})$. Obviously $Fuz(\mathcal{I}_{E_1}) \preceq Fuz(\mathcal{I}_{E_1 \cup \{e\}})$.

From the equalities

$$\|T(E_1 \cup \{e\},\varphi)\| = \|T(E_1,\varphi)\| + 1 \text{ and } \|E_1 \cup \{e\}\| = \|E_1\| + 1$$

it follows that

$$\alpha \leq \frac{\|T(E_1,\varphi)\|}{\|E_1\|} \leq \frac{\|T(E_1 \cup \{e\},\varphi)\|}{\|E_1 \cup \{e\}\|}.$$

This means that $Fuz(\mathcal{I}_{E_1 \cup \{e\}}) \in M(\varphi \geq \alpha)$, i.e. the model $Fuz(\mathcal{I}_{E_1})$ is not the maximum in the set $\langle M(\varphi \geq \alpha); \preceq\rangle$. $\square$

Definition 6. Let $\varphi(x) \in CON$ and $\alpha \in [0,1]$. Model $Fuz(\mathcal{I}_{E_1})$ is called $(\varphi \geq \alpha)$-**Optimal Restriction** ($(\varphi \leq \alpha)$-**Optimal Restriction**) of the model $Fuz(\mathcal{I}_E)$, if it satisfies the following conditions:

1. Model $Fuz(\mathcal{I}_{E_1})$ is maximum in the partially ordered set $\langle M(\varphi \geq \alpha), \preceq\rangle(\langle M(\varphi \leq \alpha), \preceq\rangle)$;
2. If model $Fuz(\mathcal{I}_{E_2})$ is maximum in $\langle M(\varphi \geq \alpha), \preceq\rangle$ ($\langle M(\varphi \leq \alpha), \preceq\rangle$) then condition $\xi_{E_2}(\varphi(a)) \subseteq \xi_{E_1}(\varphi(a))$ is true.

Theorem 1. Consider $\varphi(x) \in CON$ and $\alpha \in [0,1]$. Let $Fuz(\mathcal{I}_E) \models_{[\beta_1,\beta_2]} \varphi(a)$.

1. If model $Fuz(\mathcal{I}_{E_1})$ is $(\varphi \geq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$ then

$$\|E_1\| = \min\left\{\lfloor \frac{\beta_1}{\alpha}\|E\|\rfloor, \|E\|\right\};$$
$$\inf(\xi_{E_1}(\varphi(a))) = k\beta_1;$$
$$\sup(\xi_{E_1}(\varphi(a))) = \min\{1, k\beta_2\}.$$

2. And if model $Fuz(\mathcal{I}_{E_1})$ is $(\varphi \leq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$ then

$$\|E_1\| = \min\left\{\lfloor \frac{1-\beta_2}{1-\alpha}\|E\|\rfloor; \|E\|\right\};$$
$$\inf(\xi_{E_1}(\varphi(a))) = \max\{0; 1 - k(1-\beta_1)\};$$
$$\sup(\xi_{E_1}(\varphi(a))) = 1 - k(1-\beta_2)]),$$

where $\lfloor x\rfloor$ is the integer part of number x and $k = \frac{\|E\|}{\|E_1\|}$.

Proof. Let us proof the item 1. The item 2. is proven the same way.

If $\alpha \leq \beta_1$ then model $Fuz(\mathcal{I}_E)$ is its own $(\varphi \geq \alpha)$-Optimal Restriction. Consider a case when $\alpha > \beta_1$.

Model $Fuz(\mathcal{I}_{E_1})$ is $(\varphi \geq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$. Then (by Proposition 1) we have $T(E_1, \varphi) = T(E, \varphi)$. Furthermore, by virtue of the fact that $Fuz(\mathcal{I}_E) \models_{[\beta_1, \beta_2]} \varphi(a)$ we get

$$\|T(E_1, \varphi)\| = \|T(E, \varphi)\| = \beta_1 \|E\|.$$

As $Fuz(\mathcal{I}_{E_1}) \in M(\varphi \geq \alpha)$ then $\alpha \leq \frac{\|T(E_1,\varphi)\|}{\|E_1\|}$. Consequently, we get the following inequality:

$$\|E_1\| \leq \frac{\|T(E_1, \varphi)\|}{\alpha} = \frac{\beta_1}{\alpha}\|E\|.$$

Also, by virtue of the fact that the model $Fuz(\mathcal{I}_{E_1})$ is maximum in the partially ordered set $\langle M(\varphi \geq \alpha), \preceq \rangle$ we have

$$\|E_1\| \leq \frac{\beta_1}{\alpha}\|E\| \ \text{ and } \ \|E_1\| - \|T(E, \varphi)\| \leq \|F(E, \varphi)\| + \|N(E, \varphi)\|.$$

Let us show that the number $\lfloor \frac{\beta_1}{\alpha}\|E\| \rfloor$ satisfies these conditions. Obviously, that number $\lfloor \frac{\beta_1}{\alpha}\|E\| \rfloor$ is the biggest number that satisfies the first inequality. Accept

$$\lfloor \frac{\beta_1}{\alpha}\|E\| \rfloor - \|T(E, \varphi)\| > \|F(E, \varphi)\| + \|N(E, \varphi)\|.$$

Then we get

$$\frac{\beta_1}{\alpha}\|E\| - \beta_1\|E\| > (1 - \beta_2)\|E\| + (\beta_2 - \beta_1)\|E\|.$$

If we divide this inequation in $\|E\|$, we get $\frac{\beta_1 - \alpha\beta_1}{\alpha} > 1 - \beta_1$. Consequently, $\beta_1 - \alpha\beta_1 > \alpha - \alpha\beta_1$. Therefore, $\alpha < \beta_1$, i.e. we have a contradiction.

Next, from $\|T(E_1, \varphi)\| = \beta_1\|E\|$ follows that

$$\inf(\xi_{E_1}(\varphi(a))) = \frac{\|T(E_1, \varphi)\|}{\|E_1\|} = \frac{\beta_1\|E\|}{\|E_1\|} = k\beta_1.$$

On the other side, since the model $Fuz(E_1)$ is $(\varphi \geq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$ then

$$\begin{aligned}\sup(\xi_{E_1}(\varphi(a))) &= \min\left\{1, \frac{\|T(E_1, \varphi)\| + \|N(E, \varphi)\|}{\|E_1\|}\right\} \\ &= \min\left\{1, \frac{\beta_2\|E\|}{\|E_1\|}\} = \min\{1, k\beta_2\right\}.\end{aligned}$$

□

Consequence 1. Consider $\varphi(x) \in CON$ and $\alpha \in [0,1]$. Let $Fuz(\mathcal{I}_E) \models_{[\beta_1,\beta_2]} \varphi(a)$, $\beta_1 < \alpha$, model $Fuz(\mathcal{I}_{E_1})$ is $(\varphi \geq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$ and $k = \frac{\|E\|}{\|E_1\|}$. Then

1. if $\alpha \geq \frac{\beta_1}{\beta_2}$, then $Fuz(\mathcal{I}_{E_1}) \models_{[k\beta_1;1]} \varphi(a)$;
2. if $\alpha < \frac{\beta_1}{\beta_2}$, then $Fuz(\mathcal{I}_{E_1}) \models_{[k\beta_1;k\beta_2]} \varphi(a)$.

Proof. By Theorem 1, $\sup(\xi_{E_1}(\varphi(a))) = \min\{1, k\beta_2\}$. Consider the case when $\alpha \geq \frac{\beta_1}{\beta_2}$. Then

$$\|E_1\| = \lfloor \frac{\beta_1}{\alpha} \|E\| \rfloor \leq \lfloor \beta_2 \|E\| \rfloor \leq \beta_2 \|E\|,$$

i.e. $\|E_1\| \leq \beta_2\|E\|$. Then $k = \frac{\|E\|}{\|E_1\|} \geq \frac{1}{\beta_2}$. Consequently, $k\beta_2 \geq 1$. Therefore, we get that $\sup(\xi_{E_1}(\varphi(a))) = 1$.

On the other side, if $\alpha < \frac{\beta_1}{\beta_2}$ then $k < \frac{1}{\beta_2}$ So we get $\sup(\xi_{E_1}(\varphi(a))) = k\beta_2$. □

Consequence 2. Consider $\varphi(x) \in CON$ and $\alpha \in [0,1]$. Let $Fuz(\mathcal{I}_E) \models_{[\beta_1,\beta_2]} \varphi(a)$, $\alpha < \beta_2$, model $Fuz(\mathcal{I}_{E_1})$ is $(\varphi \leq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$ and $k = \frac{\|E\|}{\|E_1\|}$. Then

1. if $\alpha \leq 1 - \frac{1-\beta_2}{1-\beta_1}$ then $Fuz(\mathcal{I}_{E_1}) \models_{[0;1-k(1-\beta_2)]} \varphi(a)$;
2. if $\alpha > 1 - \frac{1-\beta_2}{1-\beta_1}$ then $Fuz(\mathcal{I}_{E_1}) \models_{[1-k(1-\beta_1);1-k(1-\beta_2)]} \varphi(a)$.

Proof is analogous to the proof of the Consequence 1.

4 FORMALIZATION AND CLASSIFICATION OF QUESTION TYPES

Question-answering systems are aimed satisfying a user's informational need. However, analysis, understanding, and satisfying this need is a complicated task even for a human, not to speak about software systems. That is why the formalization of question-answering relationships is an important task. Obviously, the quality and the effectiveness of interaction between a user and a QA-system will depend on the quality of the formalizing process.

A set of all questions, that a user can ask in natural language form, is infinite even under a narrow subject domain, and consequently, the formalization of this set seems to be impossible. Accordingly, we can only describe some of the typical question templates. The selection of the formalization applies some restrictions on those question types that can be processed by a question-answering system. In the current work, the Knowledge Base of a question-answering system is formalized as an algebraic system $Fuz(\mathcal{I}_E)$. It will let us make a formalization and a classification in the spirit of the erotetic logic, which foundational ideas can be found in the work [22].

Traditionally, in erotetics two types of questions are considered: "whether-questions" and "what-questions". As our judgements are of a probabilistic nature, we will consider one more question type – "the probabilistic questions".

4.1 "Whether-Questions"

Questions of "whether" type are aimed to explore the truth value of a judgement. Often a question of this type starts with *"Whether true or not that..."* As an answer, we expect to get *"yes"* or *"no"*. From a theoretical-modal point of view, the "whether-question" is formalized as "query": whether the current judgement is true or not within the given algebraic system.

In the QA-RiskPanel system "whether-questions" are modified. Because, on the one hand, the aim of the system is a definition of the probability value of some risks occuring and, on the other hand, the system base of knowledge is formed as Generalized Fuzzy Model $Fuz(\mathcal{I}_E)$. Consequently, while formulating a "whether-question", we will point out not only a judgement itself but also probability characteristics of this judgement. Let us take an example:

> Whether true or not that the probability of unreliable password use in a computer attack is less than 0.3?
>
> Whether true or not that the probability of SQL-injection use in a computer attack is equal to 0.8?

4.2 Probabilistic Questions

A probabilistic question is a demand to estimate the probability of some judgement. For an example, a user's informational need is, that he wonders how often a vulnerability *"unreliable password"* is used in computer attacks. In this case the question will look like this:

> What is the probability, that an unreliable password will be used in an attack?

The answer will be an interval of rational numbers from $[0; 1]$.

In a more general case, if we consider the interpretation of questions on models with truth-functions of a different nature (for instance, Boolean-valued or case models [23]), then we will get evaluative questions. The general scheme is the following:

> What is the truth value, that $\langle$judgement$\rangle$?

In case of interpretation on classical model evaluative question will be equal to "whether-question".

4.3 "What-Questions"

In erotetics, "what-question" is interpreted as evaluating the value of the question subject. Each "what-question" is considered as a set (possibly, infinite) of different "whether-questions". The answer is a list of subjects for which the answer on "whether-question" is true.

In our approach every "what-question" will be split into a finite amount of "whether-questions". Let us give an example of "what-question":

> What hidden attacks can appear with probability of more than 0.8?

An algorithm to find an answer to "whether-question" should be understood as a procedure, that consists of three steps:

1. detection of a subset C of all atomic concepts set P;
2. estimation of truth values of all concepts from C;
3. detection of subset $C' \subseteq C$, so that truth values of its elements satisfy the probability characteristics from the question.

The answer to the question of this type is a list of concepts from C'.

4.4 Conditional Questions

Conditional questions are questions that contain a condition and require an answer only in the case when this condition is true. Conditional questions play an important role in formalization of question-answering systems. The fact is that the real world questions are intended to mean, that the questioner has some preliminary knowledge about a question domain, and disregarding this knowledge can result in reduction of pertinence of the answer, that has been elicited by the question-answering system. As a result, a user will receive an answer that does not satisfy his need, and we should mention that the main aim of QA systems is to produce relevant answers that match the user's need.

In the developed system there is an opportunity to add a condition to a question of every type. Consequently, a user can ask conditional "whether-questions", conditional probability questions, conditional "what-questions", and conditional modal questions (see Section 4.5). If we do not add a condition to a question, let us call this question *unconditional*.

Let us consider an example of the conditional question:

> "If there is no network encryption, what is the probability of information disclosure in an attack?"

This question may be understood as a "query", and the purpose is to find a conditional probability of the event *"there is an information disclosure"*, upon the condition, that an event *"there is no network encryption"* is veracious. From the theoretic-model point of view, in order to get an answer to a question like this we should find out the truth value within the Generalized Fuzzy Model that is ($\varphi = 1$)-optimal submodel of the model $Fuz(\mathcal{I}_E)$, where $\varphi =$ *"there is no network encrypting"* (see Section 3.3).

Notice, that in our case the judgements are of a probability nature, and then, conditions may be held with some value of probability. Let us give an example of a question of this type:

> If there is no network encryption in at least 30 % of cases, what is the probability of information disclosure in attack?

To retrieve an answer we should find ($\varphi \geq 0.3$)-Optimal Restriction, on which we will evaluate the truth value of the judgement *"there is an information disclosure"*. Note, that in the general case, we will get a class of Optimal Restrictions with the target condition.

4.5 Modal Question

Computer attacks can be divided into two classes: single-stage attacks and multi-stage attacks [24]. In a single-stage attack an intruder uses a vulnerability directly in order to accomplish his final purpose. In a multistage attack an intruder can use an existing vulnerability to open a new one, that will be used to perform another attack.

For analysis of multistep attacks in the QA-RiskPanel system we construct attack graphs. A set of graph nodes is a set of single-stage attacks $E = \{e_1, \ldots, e_n\}$. An edge between two attacks built if one attack creates the conditions that are necessary for performing another attack. In such a way, a multistage attack turns out to be a path in a directed graph that has every next node reachable from the previous one by a directed edge single transition.

Thus, we have the ability to get statistical information about multistage attacks while adding different modalities to a question of any type. Currently, in the QA-RiskPanel system the opportunity to ask modal questions of two types – the *"possibly"* questions and *"soon or late"* questions – is implemented. Let us provide an example of the unconditional probability question of the *"possibly"* type:

> What is the possibility that as the result of a multistage attack, there will be database contamination?

And the conditional probability question of *"soon or late"* type:

> If the possibility of buffer overflow is not less than 0.8, what is the possibility that in the result of a multistage attack there will be a database contamination?

For the purposes of formalization of questions and retrieval of answers about multistage attacks, we will use modal logic methods and model checking algorithms.

5 ALGORITHMS FOR RETRIEVING ANSWERS

5.1 Unconditional Questions

An algorithm to find an answer to unconditional "whether-question" and an algorithm to find an answer to an unconditional probability question are very similar to each other. In both cases we need to find the truth value of a sentence $\varphi(a)$ (that was generated by a concept $\varphi \in CON$) on the model $Fuz(\mathcal{I}_E)$. This truth value is the answer to the probability question. An answer to a "whether-question" is the result of the comparison of the current truth value with the restriction which was given in the question.

"What-question" sets an infinite set of sentences $\varphi_1(a), \ldots, \varphi_n(a)$; for each of them we need to evaluate truth values on the model $Fuz(\mathcal{I}_E)$. An answer to this question is a list of sentences with truth values that satisfy requirements from the question.

Thus, the core of the algorithm for answering any of the unconditional questions is the procedure of searching the truth value of the given sentence on the Fuzzy Model.

Note that the truth value of the sentence $\varphi(a)$ on the Generalized Fuzzy Model $Fuz(\mathcal{I}_E)$ should be built as a union of truth values of this sentence on all Interpretations from the class $\mathcal{I}_E$. Each interpretation $\mathcal{A}_\Delta \in \mathcal{I}_E$ is a finite model, and its signature consists of a finite number of monadic predicates. That is why the procedure of calculation of the truth value of the sentence on the model $\mathcal{A}_\Delta$ is solvable and can be solved by propositional logic methods.

Class $\mathcal{I}_E$ is also finite. However, its cardinality exponentially depends on the degree of uncertainty in the Knowledge Base **KB**. Accepting that Knowledge Base **KB** consists of 10 000 fully described precedents of computer attacks and in only 10 cases there is no information about virus usage in these computer attacks. This inconspicuous uncertainty generates 2^{10} different Interpretations. Obviously, complete enumeration of all Interpretations is too much time-consuming and could not be implemented in software.

In [21] there is a consideration of an algorithm that evaluates the truth value of a quantifier-free sentence on the generalized Fuzzy Model, the signature of which consists of a finite number of monadic predicates. The current algorithm is based on the idea of decomposition of the Generalized Fuzzy Model into the direct product of generalized precedents and has polynomial complexity.

5.2 Modal Questions

Let us describe theoretical-modal formalization of the questions of *"possibly"* and *"soon or late"* types. Let us introduce a binary relation R on set E: attacks e_1 and e_2 are in relation R if the consequences of the attack e_1 open vulnerabilities that are typical for the attack e_2. In the current work the transitive closure R^T of the relation R lies in the area of our interest.

Let us complete the terminology *TBox* of Knowledge Base **KB** with a role-based concept R^T. Let us introduce a designation: $TBox' = TBox \cup \{R^T\}$. This will result in an extension of the set of truthful atomic concepts *ABox* to *ABox'*-set and, consequently, the extension of the Knowledge Base $\mathbf{KB}'$ and the generalized Fuzzy Model $Fuz(\mathcal{I}_E)'$ that formalizes this Base.

Let $\varphi \in CON$, i.e. it does not contain role-based concept R^T in its signature. Then, according to the description logics syntax, *"probably"* questions will be formalized with the help of concept $\varphi_\diamond = \varphi \vee \exists R^T.\varphi$ and *"soon or late"* questions will be formalized with the help of concept $\varphi_\Box = \varphi \vee (\forall R^T.(\exists R^T.\varphi))$. Answers to these questions will depend on the truth values of relevant judgements on the Generalized Fuzzy Model $Fuz(\mathcal{I}_E)'$.

For realization of judgements $\varphi_\diamond$ and $\varphi_\Box$ truth values estimation we consider graph $G = \langle E, R^T\rangle$. We will make postfix traversal of the graph G. In the result of graph traversal, we will assign one of three values, TRUE, UNKNOWN or FALSE, to every node of graph according to the current rule:

$$e \in T(E, \varphi_{\diamond/\Box}) \Rightarrow e := \text{TRUE};$$

$$e \in F(E, \varphi_{\diamond/\Box}) \Rightarrow e := \text{FALSE};$$

$$e \in N(E, \varphi_{\diamond/\Box}) \Rightarrow e := \text{UNKNOWN}.$$

This marking of the graph allows us to estimate inf and *sup* of truth values of the judgements. In work [25] the description of FuzGLEMP and FuzGLEMN algorithms is given; these algorithms estimate the truth values of judgements of $\varphi_\diamond$-type and $\varphi_\Box$-type on the Generalized Fuzzy Model $Fuz(\mathcal{I}_E)'$. An algorithm of creating the graph from [26] was taken as the basis and it was modified for work with the incomplete knowledge.

5.3 Conditional Questions

Currently, the QA-RiskPanel system has answering algorithms for two types of conditional questions:

If $\varphi \geq \alpha$ then ⟨unconditional modal/nonmodal question⟩?

If $\varphi \leq \alpha$ then ⟨unconditional modal/nonmodal question⟩?

Therefore, the answering algorithm for a conditional question reduces to finding $(\varphi \geq \alpha)$-Optimal Restriction or $(\varphi \leq \alpha)$-Optimal Rectriction $Fuz(\mathcal{I}_{E_1})$, where $E_1 \subseteq E$. Then we run one of the algorithms from Sections 5.1 and 5.2 on the model $Fuz(\mathcal{I}_{E_1})$.

Let $Fuz(\mathcal{I}_E) \models_{[\beta_1,\beta_2]} \varphi(a)$. If $\alpha \leq \beta_1$ or $\alpha = 1$ then there is only one $(\varphi \geq \alpha)$-Optimal Restriction, and if $\alpha \geq \beta_2$ or $\alpha = 0$ then there is only one $(\varphi \leq \alpha)$-Optimal Restriction of the model $Fuz(\mathcal{I}_E)$. This Optimal Restriction is used to run the algorithm of finding the truth value of the sentence $\psi(a)$, which formalizes the corresponding unconditional question.

In other cases there is a class $K_{\varphi\geq\alpha}$ of the $(\varphi \geq \alpha)$-Optimal Restrictions and a class $K_{\varphi\leq\alpha}$ of the $(\varphi \leq \alpha)$-Optimal Restrictions. Our goal is to find such models $\mathit{Fuz}(\mathcal{I}_{E_{min}}), \mathit{Fuz}(\mathcal{I}_{E_{max}}) \in K_{\varphi\geq\alpha}$ (or $\mathit{Fuz}(\mathcal{I}_{E_{min}}), \mathit{Fuz}(\mathcal{I}_{E_{max}}) \in K_{\varphi\leq\alpha}$) that the following conditions are met:

$$\inf(\xi_{E_{min}}(\psi(a))) = \min\{\inf \xi_{E'}(\psi(a)) \mid \mathit{Fuz}(\mathcal{I}'_E) \in K_{\varphi\geq\alpha/\varphi\leq\alpha}\};$$
$$\sup(\xi_{E_{max}}(\psi(a))) = \max\{\sup \xi_{E'}(\psi(a)) \mid \mathit{Fuz}(\mathcal{I}'_E) \in K_{\varphi\geq\alpha/\varphi\leq\alpha}\}.$$

Consider the algorithm of choosing $\mathit{Fuz}(\mathcal{I}_{E_{min}})$ and $\mathit{Fuz}(\mathcal{I}_{E_{max}})$ for the class $K_{\varphi\geq\alpha}$. In this case the condition $\beta_1 < \alpha < 1$ is true. According to the Consequence 1, there are two cases.

Case 1. The condition $\alpha \geq \frac{\beta_1}{\beta_2}$ is met. Then for each model $\mathit{Fuz}(\mathcal{I}'_E) \in K_{\varphi\geq\alpha}$ we have $E' = T(E,\varphi) \cup N$, where $N \subseteq N(E,\varphi)$. Therefore, the range of truth values of the sentence $\psi(a)$ on models from the class $K_{\varphi\geq\alpha}$ depends only on the choice of cases from the set $N(E,\varphi)$. Let us divide the set $N(E,\varphi)$ into three subsets (according to formulas from (1)):

$$T(N(E,\varphi),\psi), F(N(E,\varphi),\psi), N(N(E,\varphi),\psi).$$

To construct the model $\mathit{Fuz}(\mathcal{I}_{E_{min}})$ we, in the first place, choose cases from the set $F(N(E,\varphi),\psi)$. After that if $\|F(N(E,\varphi),\psi)\| < \|N\|$ we choose cases from the set $N(N(E,\varphi),\psi)$. And if $\|F(N(E,\varphi),\psi)\| + \|N(N(E,\varphi),\psi)\| < \|N\|$ the final step is to choose cases from the set $T(N(E,\varphi),\psi)$.

To construct the model $\mathit{Fuz}(\mathcal{I}_{E_{max}})$ the priority of choice is the following:

$$T(N(E,\varphi),\psi), N(N(E,\varphi),\psi), F(N(E,\varphi),\psi).$$

Case 2. The condition $\alpha < \frac{\beta_1}{\beta_2}$ is met. Then for each model $\mathit{Fuz}(\mathcal{I}'_E) \in K_{\varphi\geq\alpha}$ we have $E' = T(E,\varphi) \cup F$, where $F \subseteq F(E,\varphi)$. Then (similar to the case 1) we divide the set $F(E,\varphi)$ into three subsets:

$$T(F(E,\varphi),\psi), F(F(E,\varphi),\psi), N(F(E,\varphi),\psi).$$

To construct the model $\mathit{Fuz}(\mathcal{I}_{E_{min}})$ the priority of choice is the following:

$$F(F(E,\varphi),\psi), N(F(E,\varphi),\psi), N(F(E,\varphi),\psi).$$

To construct the model $\mathit{Fuz}(\mathcal{I}_{E_{max}})$ the priority of choice is the following:

$$T(F(E,\varphi),\psi), N(F(E,\varphi),\psi), F(F(E,\varphi),\psi).$$

The algorithm of choosing $Fuz(\mathcal{I}_{E_1})$ and $Fuz(\mathcal{I}_{E_2})$ for the class $K_{\varphi\leq\alpha}$, according to the Consequence 2, also has two cases: $\alpha \leq 1 - \frac{1-\beta_2}{1-\beta_1}$ and $\alpha > 1 - \frac{1-\beta_2}{1-\beta_1}$.

6 CONCLUSION

This article describes the mathematical formalization and algorithms used in the QA-RiskPanel, the question-answering system with restricted domain. The system implements the case-based methodology of object domain modelling, and allows the user to ask probability questions about the definition and prediction of computer attack risks.

The Knowledge Base of the QA-RiskPanel system contains a set of computer attack cases. These cases are used to estimate the probability of various statements related to the security of corporate information systems.

Currently, the question-answering system QA-RiskPanel contains modules for processing question of three types: unconditional, conditional and modal. The first module is designed to handle situations when there is no information about the ongoing computer attack. The second module handles the cases when there is some probability information about the attack. The purpose of the third module is to provide information about multi-step attacks.

Question patterns and answering algorithms are developed for each module. All algorithms are based on the Fuzzy Model Theory and have polynomial complexity.

REFERENCES

[1] Vasenin, V. A.: Towards Creating an International System for Monitoring and Analysis of the Information Space for the Purposes of Prevention and Termination of Military-Political Cyber Conflicts. Informasionnye Tehnologii, 2012, No. 9, pp. 2–10 (in Russian).

[2] Yakhyaeva, G. E.: Application of Boolean Valued and Fuzzy Model Theory for Knowledge Base Development. Proceedings of the International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON 2019), 2019, pp. 868–871, doi: 10.1109/SIBIRCON48586.2019.8958245.

[3] Yakhyaeva, G.—Yasinskaya, O.: Application of Case-Based Methodology for Early Diagnosis of Computer Attacks. Journal of Computing and Information Technology, Vol. 22, 2014, No. 3, pp. 145–150, doi: 10.2498/cit.1002388.

[4] Anshakov, O. M.—Gergely, T.: Cognitive Reasoning: A Formal Approach. Springer, 2010, doi: 10.1007/978-3-540-68875-4.

[5] Ganter, B.—Stumme, G.—Wille, R.: Formal Concept Analysis: Foundations and Applications. Springer, 2005, doi: 10.1007/978-3-540-31881-1.

[6] Mollá, D.—Vicedo, J. L.: Question Answering in Restricted Domains: An Overview. Computational Linguistics, Vol. 33, 2007, No. 1, pp. 41–61, doi: 10.1162/coli.2007.33.1.41.

[7] Zhao, Q.—He, Y.—Jiang, C.—Wang, P.—Qi, M.—Li, M.: Integration of Link and Semantic Relations for Information Recommendation. Computing and Informatics, Vol. 35, 2016, No. 1, pp. 30–54.

[8] Allam, A. M. N.—Haggag, M. H.: The Question Answering Systems: A Survey. International Journal of Research and Reviews in Information Sciences, Vol. 2, 2012, No. 3, pp. 211–221.

[9] Dwivedi, S. K.—Singh, V.: Research and Reviews in Question Answering System. Procedia Technology, Vol. 10, 2013, pp. 417–442, doi: 10.1016/j.protcy.2013.12.378.

[10] Liu, Y.—Li, M.—Khan, M.—Qi, M.: A MapReduce Based Distributed LSI for Scalable Information Retrieval. Computing and Informatics, Vol. 33, 2014, No. 2, pp. 259–280.

[11] Li, X.—Roth, D.: Learning Question Classifiers. Proceedings of the 19th International Conference on Computational Linguistics (COLING 2002), 2002, pp. 556–562.

[12] Ceusters, W.—Smith, B.—Van Mol, M.: Using Ontology in Query Answering Systems: Scenarios, Requirements and Challenges. Proceedings of the 2nd CoLogNET-Elsnet Symposium, Amsterdam, 2003, pp. 5–15.

[13] Demner-Fushman, D.—Lin, J.: Answering Clinical Questions with Knowledge-Based and Statistical Techniques. Computational Linguistics, Vol. 33, 2007, No. 1, pp. 63–103, doi: 10.1162/coli.2007.33.1.63.

[14] Palchunov, D.—Yakhyaeva, G.—Yasinskaya, O.: Software System for the Diagnosis of the Spine Diseases Using Case-Based Reasoning. Proceedings of the 2015 International Conference on Biomedical Engineering and Computational Technologies (SIBIRCON 2015), Novosibirsk, 2015, pp. 150–155, doi: 10.1109/SIBIRCON.2015.7361884.

[15] Palchunov, D.—Yakhyaeva, G.—Dolgusheva, E.: Conceptual Methods for Identifying Needs of Mobile Network Subscribers. In: Huchard, M., Kuznetsov, S. (Eds.): Proceedings of the Thirteenth International Conference on Concept Lattices and Their Applications (CLA 2016). CEUR Workshop Proceedings, Vol. 1624, 2016, pp. 147–160.

[16] Benamara, F.: Cooperative Question Answering in Restricted Domains: The WEBCOOP Experiments. Proceedings of the Conference on Question Answering in Restricted Domains, Barcelona, Spain, ACL, 2004, pp. 31–38.

[17] Palchunov, D. E.—Yakhyaeva, G. E.: Fuzzy Logics and Fuzzy Model Theory. Algebra and Logic, Vol. 54, 2015, No. 1, pp. 74–80, doi: 10.1007/s10469-015-9326-9.

[18] Baader, F.—Calvanese, D.—McGuinness, D. L.—Nardi, D.—Patel-Schneider, P. F.: The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press, Cambridge, 2002.

[19] Gillies, D.: Philosophical Theories of Probability. Routledge, 2000, 240 pp.

[20] Yakhyaeva, G.: Logic of Fuzzifications. Proceedings of the 4th Indian International Conference on Artificial Intelligence (IICAI 2009), Tumkur, Karnataka, India, 2009, pp. 222–239.

[21] Yakhyaeva, G.—Yasinskaya, O.: An Algorithm to Compare Computer-Security Knowledge from Different Sources. Proceedings of the 17^{th} International Conference on Enterprise Information Systems (ICEIS 2015), Vol. 2, Barcelona, Spain, 2015, pp. 565–572, doi: 10.5220/0005347205650572.

[22] Belnap, N. D.—Steel, T. B.: The Logic of Questions and Answers. London, 1976.

[23] Yakhyaeva, G.: Fuzzy Model Truth Values. 6^{th} Conference on Applied Mathematics (APLIMAT 2007), Bratislava, Slovakia, 2007, pp. 423–431.

[24] Alhomidi, M.—Reed, M.: Attack Graph-Based Risk Assessment and Optimization Approach. International Journal of Network Security and Its Applications, Vol. 6, 2014, No. 3, pp. 31–43, doi: 10.5121/ijnsa.2014.6303.

[25] Yakhyaeva, G.—Ershov, A.: Knowledge Base System for Risk Analysis of the Multi-Step Computer Attacks. Proceedings of the 18^{th} International Conference on Enterprise Information Systems (ICEIS 2016), Vol. 2, Rome, Italy, 2016, pp. 143–150.

[26] Blackburn, P.—Van Benthem, J.—Wolter, F. (Eds.): Handbook of Modal Logic. Elsevier, Amsterdam, Studies in Logic and Practical Reasoning, Vol. 3, 2007.

Gulnara Yakhyaeva is Associate Professor and Deputy Head of the General Informatics Department of the Faculty of Information Technologies at Novosibirsk State University, Russia. Her main research topics are explainable AI methods and algorithms and fuzzy model theory. She is author of about 70 refereed publications.

Anastosia Karmanova is Master of informatics, Project Manager at the Tauconsult Company, Russia. Her main research topics are business analysis and domains modeling.

Aleksei Ershov is Master of informatics, Lead Software Developer at the Whoosh LLC, Russia. His main research topic is development of intelligent information systems.