Local Ranking Model of Exhibits for Semantic Services in Smart Museum

Oksana B. Petrina
Petrozavodsk State University (PetrSU)
Petrozavodsk, Russia
{petrina}@cs.karelia.ru

***Abstract*—This work introduces the local ranking model of exhibits in semantic network in smart museum. In the museum environment there is a need to find the most interesting exhibits for the visitor. Semantic description about exhibits is represented as a semantic network formed according to ontological model. The ontological model was developed earlier and is provides structural rules for creating the semantic network. Semantic services are based on the selection of objects from the semantic network. The problem of selection is reduced to ranking. Work considers the problem of ranking the available objects by their proximity in the user profile context. This work propose a solution based on probabilistic approach to distance comparison between subsets of Finite Set. Probabilistic approach was used to distance comparison between the semantic description of an museum exhibit and user profile. This approach allows to describe the characteristics of the semantic network for performance evaluation of local ranking.**

I. Ranking for semantic services of smart museum

The concept model of smart museum environment was proposed in [1], [2]. The semantic layer is responsible to construct over this semantic network the following advanced services for smart museum: Visit service, Exhibition service, and Enrichment service. A semantic network is created on top of the available knowledge corpus of descriptions collected from the above information sources. A semantic network is a directed graph consisting of nodes (vertices) representing domain objects and links (edges) representing semantic relations [1].

The ontological model is used for effective service construction, in particular for smart museum services [3]. The model describes shared information, including available descriptions of museum exhibits and facts of everyday life history [4]. The model provides structural rules for creating the required semantic network in the RDF triplestore (maintained by the SIB). Semantic services are based on the selection of objects from the semantic network. The problem of selection is reduced to ranking. The connection structure of the semantic network is the base for the algorithms of exhibits and descriptions ranking. Local ranking assumes that the matching of two objects based on the semantic description of these objects, that is, the matching of the properties values of the classes instances. We consider the following types of local ranking [5].

1. The rank calculates for each exhibit based on the user profile. The rank of the exhibit directly depends on the user profile: the more the similarity of the semantic description of an exhibit to the user profile, the higher an exhibit rank is. Two semantic network objects participate in the ranking: the user profile c_u is defined by the class Profile, the exhibit m is defined by the class ManMadeThing. The rank $r_{c_u,m}$ is the result of semantic matching between the set of property values of the class Profile instance and of the set of property values of the class ManMadeThing instance. As a result, for a specific user forms a list of ranks of exhibits. Such ranging is used from Visit service for construction of the personal program. Visit service needs to find a small group of tightly interrelated exhibits to form a given thematic exposition. Exhibits in this thematic exposition are ranked according to the user profile. Visit service provides the opportunity to present personal program, associated information, where exhibits are semantically matched with the user.

2. The rank calculates for each exhibit relatively to other exhibits. This semantic ranking applied that reflects the semantic connectivity of any museum artifact with each another. The more coincidence of properties from the exhibits is, the higher the rank between them. The rank $r_{m_1 m_2}$ is the result of semantic matching between two the set of property values of the class ManMadeThing instance. Such ranging is used from Exhibition service when a visitor views the current exhibits. Exhibition service can visualize descriptions of the recommended exhibits on the surrounding screens or on the personal mobile device when a visitor comes to the exhibit and studies associated information.

That is, the more the rank r between two objects the more similar they are. The proximity of two objects can also be expressed by distance. In the next section, we consider a probabilistic approach to determining the distance between two objects.

II. Probabilistic approach to distance comparison between the set of properties

Paper [6] consider the problem of ranking the available objects by their proximity in the user profile context used probabilistic approach to distance comparison between subsets of Finite Set [7]. Let us consider this task for a case the smart museum, where the semantic description of an museum exhibit is as the available object.

The local ranking model can be represented as follows. Parameter c_u is the user profile, M is the set of museum exhibits ($m \in M$), T_m^i is the set of properties of the class ManMadeThing, T_t^i is the set of properties of the class Profile.

Then the distance $\rho(c_u, m)$ between the user profile c_u and the museum exhibits $m \in M$ is the minimal (or average) distance between all of the sets of properties T_t^i for the $t \in c_u$ and of the set of properties T_m^i:

$$\rho^i(c_u, m) = \min_{t \in c_u} \rho(T_t^i, T_m^i). \quad (1)$$

The distance $\rho(c_u, m)$ between the user context c_u and exhibit $m \in M$ is the minimal (or average) distance within the exhibit of all distances $\rho^i(c_u, m)$. In accordance with the value $\rho(c_u, m)$ the objects $m \notin m_0 \cup c_u$ can be ranked.

To implement this approach the distances $\rho^i(c_u, s)$ are made comparable to each other. Formally, given $U = u_1, u_2, ..., u_n$, let X and Y be two random subsets of U.

$U = u_1, u_2, ..., u_n$ is the finite set of all properties in ontology;

X, Y are two random subsets of U. For example, X is the set of properties of Profile class instance and Y is the set of properties of ManMadeThing class instance. Or X and Y are set of properties of two random ManMadeThing class instance.

We represent the subsets X and Y as two binary vectors x and y, which are constructed as follows: $x_i = 1$ if only $u_i \in X$, otherwise $x_i = 0$ (similarly for y and Y).

We denote $p_i, i = 1, ..., n$ the probability of u_i appearance in the subset. Then we can carry out a random experiment, which consists of n independent tests. Each test can have one of the four possible outcomes $A\alpha\beta^i = x_i = \alpha, y_i = \beta$, where $\alpha, \beta \in 0, 1$, i is the number of a test. Let $I(A)$ is an indicator of a random event $A; a, b, c, d$ are overall quantities of the outcomes $A_{11}^i, A_{10}^i, A_{01}^i, A_{00}^i$ correspondingly. Note that $a = X \cup Y$, $b = X \setminus Y$, $c = Y \setminus X$, $d = U(X \cup Y)$, where $|X|$ is the quantity of elements in X. Then $I(A_{11}^i) + I(A_{10}^i) + I(A_{01}^i) + I(A_{00}^i) = 1$, $a + b + c + d = n$.

The set of random events can be introduced as follows:

$$B : \{(A_{11}^1, A_{10}^1, ..., A_{00}^n) : \sum_{i=1}^{n} I(A_{11}^i) = a, \sum_{i=1}^{n} I(A_{10}^i) = b, \sum_{i=1}^{n} I(A_{01}^i) = c, \sum_{i=1}^{n} I(A_{00}^i) = d\} \quad (2)$$

Usually different distances between the sets are described as functions of a, b, c, d, i.e., $\rho(X, Y) = h(a, b, c, d)$.

The function of distribution of the random value $\rho(X, Y)$ is described as follows:

$$F(t) = P(\rho(X, Y) < t) = \sum_{(a,b,c,d) \in C} \sum_{(A_{11}^1, ..., A_{00}^n)} \prod_{i=1}^{n} p_i^{2I(A_{11}^i)} \times (1 - p_i)^{2I(A_{00}^i)} (p_i(1 - p_i))^{I(A_{01}^i) + I(A_{10}^i)}, \quad (3)$$

where $C = \{(a, b, c, d) \in Z^4 : a, b, c, d \geq 0, a + b + c + d = n, h(a, b, c, d) > t\}$.

At this point, the value $F(\rho)$ can be taken as a new (probabilistic) distance between the subsets X and Y, where ρ is the previously determined value of distance between X and Y. This probabilistic distance assumes values between 0 and 1. If the value of the probabilistic distance is close to 1, then there are not many pairs of the sets X and Y having a distance value between them that exceeds or is equal to ρ. Similarly, if the value is close to 0, then the value less than ρ can seldom be stated.

This approach seems to be effective, if in the semantic description two exhibit two sets of properties are used instead of subsets. Besides the possibility to compare the distances between the sets of properties, the probabilistic distance provides one more advantage. Evaluation of distance is in terms long/short before the whole set of the distances ρ has been obtained. In the mobile environments, this opportunity is very useful, since it reduces the amount of computation by means of ranging object in accordance with the preferences and interests of a user.

This approach can be useful for the following cases. 1. The experimental use for simulation. Analyzing the performance of the software solution can be performed on simulated data. To bring the situation closer to the real one, the smart space data is generated according to a known distribution function. On such data, you can check the performance and other characteristics of the services. 2. The analytic use to evaluate specific information metrics. For a real data sample, you can calculate the distribution function for a particular metric. The metric is chosen according to the task being investigated for a particular service. For example, in [6] the metric "Category" was chosen. The distribution function will have a form that allows us to characterize the properties of information. 3. The use for experimental detection of analytical properties. This is possible to obtain the experimental distribution function for real data. This show that the proposed type of ranking corresponds to the theoretical one. Thus, the analytic properties of the distribution function are confirmed experimentally.

III. Case study: local ranking model in smart museum

In the local ranking model the distance is the reciprocal of the rank $r_m^{c_u}$ between the user profile c_u and the exhibit $m \in M$:

$$\rho(c_u, m) = 1/r_{c_u m}, \quad (4)$$

The rank is the result of a semantic matching the set of property values of the Profile class instance and of the set of property values of the ManMadeThing class instance:

$$r_{c_u m} = \sum_{j=1}^{J} f(k_j^{c_u}, k_j^{m_i}), \quad (5)$$

where J - quantity of matching, $k_j^{c_u}$ - the value of j property k in user profile c_u, $k_j^{c_u}$ - the value of j property k in semantic description of the exhibit m, f - the binary function of semantic matching between two properties:

$f(k_j^{c_u}, k_j^{m_i}) = 1$, if $k_j^{c_u} = k_j^{m_i}$, else 0.

Similarly (4), you can calculate the distance between two exhibit $m \in M$: $\rho(m_1, m_2) = 1/r_{m_1 m_2}$

Let us consider an example of calculation of a semantic distance between a museum visitor profile and the exhibit descriptions. A user profile will be used from [5]. The user profile includes the data about personal education. We calculate the distance between the user profile and exhibits. The distance

exhibit	semantic matching with the user profile	distance
MIU2281 Photo of members of the student construction brigades «Goliard» 	– date **1976** – refer to department **the Department of History and Philology** – has been made in place **Padanyi** – represents possible friends: 1) S. Verigin (has Person Education: **1973-1978 the Department of History and Philology in PetrSU**) 2) T. Yuldashev (has Person Education: **1973-1978 the Department of History and Philology in PetrSU**) 3) V. Efimov (has Person Education: **1974-1979 the Department of History and Philology in PetrSU**)	0,17
MIU2155 Photo of members of the student construction brigades «Goliard» 	– date **1976** – refer to department **the Department of History and Philology** – has been made in place **Padanyi** – represents possible friends: 1) V.Birin (has PersonEducation: **1974-1979 the Department of History and Philology in PetrSU**) 2) N. Pochtovalov (has Person Education: **1974-1979 the Department of History and Philology in PetrSU**)	0,20
MIU873 Uniform of G. Chumakov 	– date **1976** – refer to department **the Department of History and Philology** – owner is a possible friend: G. Chumakov (has Person Education: **1973-1978 the Department of History and Philology in PetrSU**)	0,33
MIU2297 Photo of G. Chumakov 	– date **1976** – refer to department **the Department of History and Philology** – represents a possible friend: G. Chumakov (has Person Education: **1973-1978 the Department of History and Philology in PetrSU**)	0,33

Fig. 1. Distance between user profile and description of exhibit

between user profile and some relevant exhibits are shown in Fig. 1. The more the similarity of exhibit information to the user profile, the higher an exhibit rank is. Distance is the opposite of rank. In fact, there are more than four relevant exhibits, but for convenience, the Fig. 1 shows exhibits with the shortest distance. The more the user profile and semantic description of exhibit are matched, the shorter the distance between them.

This experiment was conducted based on the semantic network of the History museum of Petrozavodsk State University. Semantic network correspond to physical and digital exhibits, associated historical events, persons, and other objects. The ontological model in [4] provides structural rules for creating the required semantic network in the RDF triplestore. Table I shows the basic metrics of the semantic network of the History museum of Petrozavodsk State University.

TABLE I. SEMANTIC NETWORK METRICS OF SMART MUSEUM

Metric	**Value**
Objects of semantic network	
Total Objects number	729
- exhibits number	297
- person number	60
Metrics of ontological model	
Classes number	33
Property number	112
Metrics of RDF-triplestore	
Triples number in triplestore	70226

IV. CONCLUSION

This work present the mathematical local ranking model of exhibits in semantic network of smart museum. The semantic network is creating according to structural rules of ontological model. Probabilistic approach was used to distance comparison between the set of properties of ontological model classes. This approach seems to be effective, if in the semantic description two exhibit two sets of properties are used instead of subsets. This model is useful, since it reduces the amount of computation by means of ranking object in accordance with the preferences and interests of a user. Probabilistic approach allows to describe the characteristics of the semantic network for performance evaluation of local ranking.

ACKNOWLEDGMENT

The research study is supported by the Ministry of Education and Science of Russia within project # 2.5124.2017/8.9 of the basic part of state research assignment for 2017–2019. The work is implemented within the Government Program of Flagship University Development for Petrozavodsk State University in 2017–2021.

REFERENCES

[1] D. Korzun, A. Varfolomeyev, S. Yalovitsyna, and V. Volokhova, “Semantic infrastructure of a smart museum: toward making cultural heritage knowledge usable and creatable by visitors and professionals,” *Personal and Ubiquitous Computing*, pp. 1–10, Nov. 2016.

[2] S. Marchenkov, A. Vdovenko, O. Petrina, and D. Korzun, “A smart space-based design of semantic layer for advancing museum information services,” in *Proc. 19th Conf. Open Innovations Association FRUCT*, S. Balandin and T. Tyutina, Eds., Nov 2016, pp. 159–166.

[3] A. Smirnov, N. Shilov, and A. Kashevnik, “Ontology-based mobile smart museums service, approach for small & medium museums,” in *Proc. of the 4th Intern. Conf. on Advances in Future Internet (AFIN 2012)*. IARIA, August 2012, pp. 48–54.

[4] O. B. Petrina, D. G. Korzun, V. V. Volokhova, S. E. Yalovitsyna, and A. G. Varfolomeyev, “Semantic approach to opening museum collections of everyday life history for services in Internet of Things environments,” *International Journal of Embedded and Real-Time Communication Systems (IJERTCS). Special Issue on Big Data Analytics and Intelligent Environments in Internet of Things*, vol. 8, no. 1, pp. 31–44, 2017.

[5] S. Marchenkov, A. Vdovenko, O. Petrina, and D. Korzun, “Smart museum of everyday life history in petrozavodsk state university: Software design and implementation of the semantic layer,” in *Proc. 21st Conf. Open Innovations Association FRUCT*, Nov 2017, pp. 224–230.

[6] A. Varfolomeyev, D. Korzun, A. Ivanovs, and O. Petrina, “Smart personal assistant for historical tourism,” in *Recent Advances in Environmental Sciences and Financial Development. Proc. 2nd Int’l Conf. on Environment, Energy, Ecosystems and Development (EEEAD 2014)*, C. Arapatsakos, M. Razeghi, and V. Gekas, Eds., Nov. 2014, pp. 9–15.

[7] A. Varfolomeyev, P. Kirikov, and A. Rogov, “Probabilistic approach to distance comparison between subsets of finite set,” *Uchenye zapiski Petrozavodskogo gosudarstvennogo universiteta - Proceedings of Petrozavodsk State Univeristy*, vol. 8, Dec. 2010, pp. 83–88.