Data Stream Mining & Processing

Proceedings of
IEEE Third International Conference on
Data Stream Mining & Processing

August 21-25, 2020
Lviv, Ukraine

IEEE

Ukrainian Catholic University

Manhattan College

DSMP

Proceedings of the

2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)

Organized by

IEEE Ukraine Section

IEEE Ukraine Section (Kharkiv) SP/AP/C/EMC/COM Societies Joint Chapter

IEEE Ukraine Section (West) AP/ED/MTT/CPMT/SSC Societies Joint Chapter

IEEE Ukraine Section IM/CIS Societies Joint Chapter

Ukrainian Catholic University

Manhattan College

Kharkiv National University of Radio Electronics

Lviv, Ukraine
August 21-25, 2020

Additional copies may be ordered from:

IEEE Conference Operations
445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331 USA

DSMP'2020 Organizing Committee
E-mail: dsmp.conference@gmail.com

IEEE Catalog Number: CFP20J13-USB

ISBN: 978-1-7281-3213-6

DSMP'2020 Conference Committee

Honorary Chairpersons

Yuriy Rashkevych, Ukraine

Yevgeniy Bodyanskiy, Ukraine

Igor Aizenberg, USA

General Chairs

Dmytro Peleshko, Ukraine
Olena Vynokurova, Ukraine
Yaroslav Prytula, Ukraine

Technical Program Committee Chair

Dmytro Peleshko, Ukraine

Publication and Finance Chair

Olena Vynokurova, Ukraine

Local Organizing Committee Chair

Tatiana Sviridova, Ukraine

Technical Program Committee

Ankur Agrawal, USA
Svitlana Antoshchuk, Ukraine
Sergii Babichev, Czech Republic
Jayaram Balasubramaniam, India
Oleksandr Baiev, Ukraine
Oleg Berezkiy, Ukraine
Petro Bidyuk, Ukraine
Sergii Bogomolov, Australian
Vilalii Boyun, Ukraine
Gennadii Churyumov, Ukraine
Ibraim Didmanidze, Georgia
Kai Du, China
Mykola Dyvak, Ukraine
Oleksandr Dumin, Ukraine
Andrey Fisunenko, Ukraine
Mohammed Gabsi, France
Eduard Petlenkov, Estonia
Shao Cheng Qu, China
Taras Panchenko, Ukraine
Nataliya Pankratova, Ukraine
Bohdan Pavlyshenko, Ukraine
Ievgen Pichkalov, Ukraine
Kashifuddin Qazi, USA
Olga Radyvonenko, Ukraine
Ali Rekik, Tunisia
Yurii Romanyshyn, Ukraine
Bohdan Rusyn, Ukraine
Anatolii Sachenko, Ukraine
Galina Setlak, Poland
Nataliya Sharonova, Ukraine
Igor Shelevytsky, Ukraine
Aleksandr Slipchenko, Netherlands

Partners

Gold partner

GeoGuard
www.geoguard.com

Partners

Lviv City Council
http://city-adm.lviv.ua/

Lviv Convetion Bureau
http://www.lvivconvention.com.ua/en/

Welcome Letter

Dear Colleagues,

We would like to personally encourage each of you to join us at IEEE Second International Scientific Conference Data Stream Mining and Processing (DSMP'2020), which is held in Lviv, UKRAINE, 21-25 August, 2020. Our main goal is not only to provide an opportunity for networking and learning recent scientific achievements but also a chance to be involved in real time panel discussions with IT representatives to review and discuss their practical outcomes on real projects.

The DSMP is organized by IEEE Ukraine Section, IEEE Ukraine Section (Kharkiv) SP/AP/C/EMC/COM Societies Joint Chapter, IEEE Ukraine Section (West) AP/ED/MTT/CPMT/SSC Societies Joint Chapter, IEEE Ukraine Section IM/CIS Societies Joint Chapter, Ukrainian Catholic University, Manhattan College and Kharkiv National University of Radio Electronics.

Agenda of the DSMP'2020 is very rich. This year we have nominated a 88 number of accepted papers coming from about 12 countries which makes DSMP a truly international high impact conference. Major highlights of DSMP'2020 are its keynotes speakers. This conference proved to be extremely important given the fruitful dialog and a chance to exchange ideas and sharing valuable hands-on experience.

This year program is based on the following topics: Hybrid Systems of Computational Intelligence, Machine Vision and Pattern Recognition, Dynamic Data Mining & Data Stream Mining, Big Data & Data Science Using Intelligent Approaches and also panel with participation of IT Companies.

We are proud of the fact that DSMP proceedings have been included into the IEEE Xplore Digital Library as well as other Abstracting and Indexing (A&I) databases (Scopus, Web of Science and etc.). High quality of the DSMP program would not be possible without the contribution of authors, keynote speakers, organizers, students, 53 reviewers who devoted a lot of enthusiasm and hard work to prepare papers, presentations, organization infrastructure and carefully review all submissions. We are very grateful for their efforts.

We would like to thank each of your for attending our conference and bringing your expertise to our gathering.

We would like to express our gratitude to our partners and sponsors for being so generous and sponsoring our conference.

We wish all participants an excellent conference, fruitful discussions and pleasant stay in Lviv and Conference venue.

Sincerely

Yuriy Rashkevych Yevgeniy Bodyanskiy Igor Aizenberg

[2] M. Anjay and S. Anjuli, "Assessment of Exit Requirements for Fire Safety of Commercial Buildings, Kathmandu, Nepal", International Journal of Emerging Technologies and Innovative Research, vol.4, Issue 10, 2017, pp. 248-255. DOI: 10.1717/JETIR.17074

[3] Building Department, Code of Practice for Fire Safety in Buildings, 2011, Available at: [www.bd.gov.hk/english/documents/code/fs_code2011.pdf].

[4] OSHA, Available at: [https://www.osha.gov/OshDoc/data_General_Facts/emergency-exit-routes-factsheet.pdf].

[5] G. Ma, S. Tan and S. Shang, "The Evaluation of Building Fire Emergency Response Capability Based on the CMM", International Journal of Environmental Research and Public Health, 16(11), 2019, 1962. DOI: 10.3390/ijerph16111962.

[6] A. Rego, L. Garcia, S. Sendra and J. Lloret, "Software Defined Network-based control system for an efficient traffic management for emergency situations in smart cities", Future Generation Computer Systems, 88, 2018, pp. 243-253. DOI: 10.1016/j.future.2018.05.054

[7] S. Granda and T. Ferreira, "Assessing Vulnerability and Fire Risk in Old Urban Areas: Application to the Historical Centre of Guimar√£es", Fire Technology, 55, 2019, pp. 105-127. DOI: 10.1007/s10694-018-0778-z.

[8] N. Kwok, C. Bratiotis, M. Luu, N. Lauster, K. Kysow and S. Woody, "Examining the Role of Fire Prevention on Hoarding Response Teams: Vancouver Fire and Rescue Services as a Case Study", Fire Technology, 2017. DOI: 10.1007/s10694-017-0672-0.

[9] J.C. Cheng, Y. Tan, Y. Song, Z. Mei, V.J. Gan and X. Wang, "Developing an evacuation evaluation model for offshore oil and gas platforms using BIM and agent-based model", Autom. Constr., 89, 2018,pp. 214-224. DOI: 10.1016/j.autcon.2018.02.011.

[10] R. Ratushnyi, P. Khmel, A. Tryhuba, E. Martyn and O. Prydatko, "Substantiating the effectiveness of projects for the construction of dual systems of fire suppression", Eastern-European Journal of Enterprise Technologies, 4, 2019, pp. 46-53. DOI: 10.15587/1729-4061.2019.175275.

[11] O. Smotr, N. Burak, Yu. Borzov and S. Ljaskovska, "Implementation of Information Technologies in the organization of Forest Fire Suppression Process", in Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, pp. 157-161, August 21-25, 2018. DOI: 10.1109/DSMP.2018.8478416

[12] I. Malets, O. Prydatko, V. Popovych and A. Dominik, "Interactive Computer Simulators in Rescuer Training and Research of their Optimal Use Indicator", in Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, pp. 558-562, August 21-25, 2018. DOI: 10.1109/DSMP.2018.8478486

IEEE Third International Conference on Data Stream Mining & Processing
August 21-25, 2020, Lviv, Ukraine

Increasing the Animation Study Management Services Functioning Efficiency

Olga Smotr
Department of Project Management, Information Technologies and Telecommunications
Lviv State University of Life Safety
Lviv, Ukraine
olgasmotr@gmail.com

Solomija Ljaskovska
Department of designing and operation of machines
Lviv Polytechnic National University
Lviv, Ukraine
solomiam@gmail.com

Igor Malets
Department of Project Management, Information Technologies and Telecommunications
Lviv State University of Life Safety
Lviv, Ukraine
igor.malets@gmail.com

Oksana Karabyn
Department of of Applied Mathematics and Mechanics
Lviv State University of Life Safety
Lviv, Ukraine
karabynoks@gmail.com

***Abstract*—the research deals with the issues of animation of the animation studio services market, and the advantages and disadvantages of their functioning, both for the ordinary user and from the developer. It made clear that it is appropriate to provide information to the client about two areas: pricing policy and content of services. The services automation methods within animation studios are researched using modern methods of Data Mining.**

It is proposed to build the work logic using data mining methods, Decision Trees methods, for the animation studio management services effective functioning. It is determined that in case of client priority search criteria choice is related to pricing policy, it makes sense to organize the operation of the animation studio management service based on the backtracking algorithm. In the case of client's search criteria choice priority are content-related topics of services, apply algorithms for constructing decision-making tree.

Keywords — information technology, Data Mining, Decision Trees, animation studio, entertainment

I. INTRODUCTION

Recently, there has been a rapid development of the entertainment industry worldwide. This is the entry into a post-industrial stage of development most the developed countries in the world, a characteristic feature is the significant growth of services and service industries. Despite the fact that the entertainment industry is one of the youngest sectors of the socio-cultural sphere, it accounts for about 6 % of the world's capital [1].

In today's conditions of informatization and computerization of the society [2] and the rapid increase in demand for services of the entertainment industry [3-4]. Business development in this area requires new approaches to information processing and decision-making.

The most notable change is in the travel companies that actively involve organizations for process activities advanced digital technology [5-9]. These include global computer reservation / reservation systems, integrated communications networks, multimedia systems, smart cards, management information systems, and more.

However, they are mostly dedicated to addressing the digitization of tourism animation services and, unfortunately, very little attention has been paid to digitizing the children's animation studio segment. Modern animation is the activity of developing and implementing special leisure programs [10]. Although the demand for services in this particular segment of the entertainment industry is huge. With a data web - site for parents Britain Netmum, parents spend on children's holiday from $ 200 to several thousands of dollars, and in general, the market value of children's parties is over $ 1.5 billion [11]. Similarly, in Ukraine, on average, the organization of children's holidays costs about 5 thousand UAH [4].

II. PROBLEM STATEMENT

Of course, today most animation studio, event -agency etc. have their own web services, where the highlight information about their services and show potential customers their own successful experiences conducted their activities, etc. However, this is not enough. The client (animation customer) needs the ability to quickly and conveniently view the offerings of different animation studios with the ability to compare their services, pricing, quality of service and the ability to provide these services in the right location and time range, etc.

Analyzing typical for the market of Ukraine, web services animation studio "Children's Planet" [12] "Igroland" [13], "Papashon" [14], event - agency "Empire holidays" [14] it is difficult not to notice, that the data the services are not convenient and do not meet the above requirements.

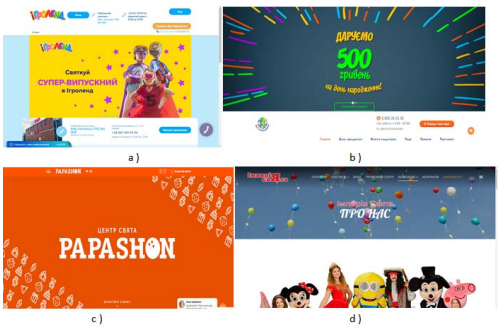

Fig. 1. View of the main pages of services of children's animation studios: a) children's animation studio "Igroland"; b) children's animation studio "Children's Planet"; c) the PAPASHON entertainment complex; d) "Empire of the Holiday" event-agency

Obviously, all four main pages of web services (Fig. 1) are not radically different and do not carry information regarding the choice of possible services. They do not have an intuitive or adapted interface. After pressing certain buttons and selecting menus, which, by the way, are not easy to find, the information becomes not much more. So, for example, to find the list of services of the children's animation studio "Children's Planet", you must use the page scrolling, on the website of the children's animation studio " Igroland " you need to find the "Menu" button, and click on the "Details" button on the main page of the entertainment complex PAPASHON. However, even after such manipulation of information comes not much more. None of these services is immediately accessible to the complex of services that can be obtained based on a certain budget. Also, none of these services has the opportunity to introduce, for example, topics that are interesting for the child to get information about the range of services that the company can provide and the price range of these services, it is difficult to find differentiation of services by age categories and so on.

In addition to imperfect client-side work, these services are not functional enough to meet the needs of animation studio staff. Here are some of the main unsolved problems:

- does not have a cumulative database of customer preferences and needs;
- there is no service of forming comparative characteristics (cost, duration in time, age restrictions and number of participants, etc.) of different offers;
- when ordering a client an offer to the administrator comes just a message about this event, which is not processed by the service itself, respectively, with a large influx of people who want to order the same service, the service will not help the administrator with this problem;
- there is no dynamic selection of the set of possible offers for the client according to the criteria he / she has set;
- there is no dynamic selection of possible offers for the client, taking into account time and quantitative parameters, the possibilities of the studios themselves, etc.

III. THE METHOD OF IMPROVING THE WORK OF CHILDREN'S ANIMATION STUDIO SERVICES WITH THE USE OF DATA MINING METHODS IS PROPOSED

For fast and qualitative sampling of data, modern methods of analysis of intellectual data must be applied by certain criteria. We believe that it is most appropriate for the organization of the effective operation of a service for managing the work of animation studios to use the Decision Trees methods, such as to enable, step by step , based on the client's answers, to form from the existing set of services, the set of the most client-friendly solutions.

To construct Decision Trees most widely used in practice, algorithms binary search tree (BST) ; returning search algorithm (backtracking algorithm) and algorithms decision tree.

In our opinion, to sample a variety of possible services, it is most advisable to use the backtracking algorithm when choosing a client as a priority, the direction of pricing.

The operation of this algorithm can be interpreted as a process of bypassing a tree. Each peak corresponds to it a sequence $(x_1, \ldots, x_i)$, with peaks that correspond to sequences of the form $(x_1, \ldots, x_i, y)$, sons of the summit. The root of the tree corresponds to an empty sequence.

This tree is being traversed by searching deep. In addition, a predicate P is specified on all tree vertices. If $P(v) = False$, then the subtree vertices with root at vertex v are not considered, and the volume of the bust decreases. The predicate $P(v)$ acquires the value False when it becomes clear that the sequence $(x_1, \ldots, x_i)$, corresponding to vertex v cannot in any way be added to the complete solution. To apply this method, the solution of the problem must look like a finite sequence of elements $(x_1, \ldots, x_n)$.

In our case, the elements $x_1, \ldots, x_n$ of this sequence are the cost of the services that can be provided to the client within the budget that he or she determines (predicate P). This is exactly what a client wants, based on a specific budget.

Initially, the client is offered a set of services that can be provided within a given budget, considering a possible set of services, the client is able to remove / add certain of them, then the budget for other services increases / decreases. The algorithm returns the customer to the previous selection step, with the other set of service elements until the desired set of services within the client's budget is formed.

Description of the backtracking algorithm

After entering in the search window the "budget" of the amount that the client focuses on, the logic of the system employs a backtracking algorithm , which allows to form from the existing set of services, the set of services that can be provided within the specified budget.

Let us illustrate the algorithm for generating multiple offers (search with returns) in case the client chooses as a priority, the direction of the pricing policy, that is, according to the criterion "budget", in a specific example.

Suppose that table 1 sets the sets of services $\{p_1, \ldots, p_n\}$.

TABLE I. SET OF EXISTING SERVICES

Code	Service	Cost	Duration
p1	Costume photoshoot	1000 UAH	1 hour
p2	Rolledrome	600 UAH	1 hour
p3	Birthday greetings with cake	550 UAH	30 minutes
p4	Weaving of African pigtails	500 UAH	1 hour
p5	Soap bubbles show	500 UAH	1 hour
p6	Trampoline arena	120 UAH	1 hour
p7	Aqua makeup	90 UAH	30 minutes
p8	Trampoline «Treasure Island»	80 UAH	30 minutes
p9	Maze	60 UAH	unlimited

Each service is matched by its cost. You need to find a subset of which the cost of the elements does not exceed the client's criterion "budget" (CB). If the sum is so large that the addition of any new number exceeds CB then we go back and change the last addition of the sum. Let $CB = 1$ *thousand UAH*. Figure 2. illustrates part of the backtracking algorithm for the problem of finding a subset of a set of services $\{p_1, \ldots, p_n\}$ with the criterion "budget" equal to 1 thousand UAH.

In general, when a subset of the set of services $\{p_1, \ldots, p_n\}$, given in Table 1. from the criterion "budget" is equal to 1 thousand UAH. The client may be offered 6 possible service options. Formed variants of subsets of possible services are given in Tables 2-7, respectively. And it should be noted that these options are provided, with information on the possible duration of the organized holiday.

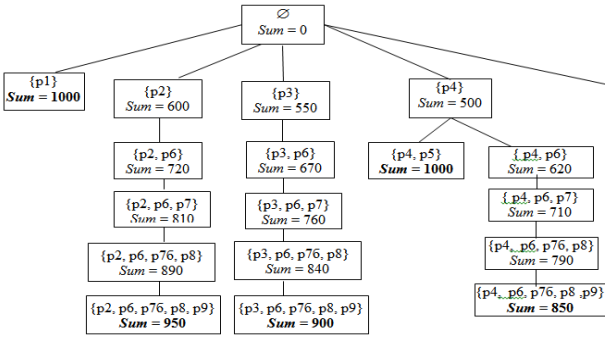

Fig. 2. Part of the backtracking algorithm for problem of finding the subset of a set of services $\{p_1, \ldots, p_n\}$ with the criterion "budget" equal to 1 thousand UAH.

TABLE II. SET OF OFFERED SERVICES

Code	Service	Cost	Duration
p1	Costume photoshoot	1000 UAH	1 hour
	Sum	**1000 UAH**	**1 hour**

TABLE III. SET OF OFFERED SERVICES

Code	Service	Cost	Duration
p2	Rolledrome	600 UAH	1 hour
p6	Trampoline arena	120 UAH	1 hour
p7	Aqua makeup	90 UAH	30 minutes
p8	Trampoline «Treasure Island»	80 UAH	30 minutes
p9	Maze	60 UAH	unlimited
	Sum	**950 UAH**	**unlimited**

TABLE IV. SET OF OFFERED SERVICES

Code	Service	Cost	Duration
p3	Birthday greetings with cake	550 UAH	30 minutes
p6	Trampoline arena	120 UAH	1 hour
p7	Aqua makeup	90 UAH	30 minutes
p8	Trampoline «Treasure Island»	80 UAH	30 minutes
p9	Maze	60 UAH	unlimited
	Sum	**870 UAH**	**unlimited**

TABLE V. SET OF OFFERED SERVICES

Code	Service	Cost	Duration
p4	Weaving of African pigtails	500 UAH	1 hour
p5	Soap bubbles show	500 UAH	1 hour
	Sum	**1000 UAH**	**2 hour**

TABLE VI. SET OF OFFERED SERVICES

Code	Service	Cost	Duration
p4	Weaving of African pigtails	500 UAH	1 hour
p6	Trampoline arena	120 UAH	1 hour
p7	Aqua makeup	90 UAH	30 minutes
p8	Trampoline «Treasure Island»	80 UAH	30 minutes
p9	Maze	60 UAH	unlimited
	Sum	**820 UAH**	**unlimited**

TABLE VII. SET OF OFFERED SERVICES

Code	Service	Cost	Duration
p5	Soap bubbles show	500 UAH	1 hour
p6	Trampoline arena	120 UAH	1 hour
p7	Aqua makeup	90 UAH	30 minutes
p8	Trampoline «Treasure Island»	80 UAH	30 minutes
p9	Maze	60 UAH	unlimited
	Sum	**820 UAH**	**unlimited**

Once a client is selected, a subset of the solutions that are most appropriate for him / her, he / she can make changes by removing / adding certain services (this will result in a certain budget change) and make an order.

If the client chooses as a priority, the content area of services, the most appropriate is to use the algorithm of building a decision tree. In this case, we need to solve the typical classification problem.

After all, when we consider a client's request in terms of the content topics of animation studio services , we need to make decisions about the set of existing objects (animation studio services), assigning them to certain thematic classes, that is, providing these objects with classification features. So we need to solve a typical classification problem whose set of conditional attributes ***A*** will be made up by the client's requirements. The set ***W*** is an active animation studio service; the set ***d*** is a decision attribute - two elements {"good luck", "bad luck"}.

Description of the algorithm for building a decision tree

After selecting a search box of your request priority area "subject", the customer will be asked to answer a few questions that are conditional attributes and accordingly help shape due to the algorithm for constructing Decision Tree , the set of possible proposals for a client under given his attributes.

Let's use one of the algorithms for building a Decision Tree, namely the algorithm ID3 (Iterative Dichotomizer-3 algorithm) [15]. To illustrate the algorithm for generating set of proposals using the algorithm for constructing Decision Tree to a specific example.

Let Table 8 provide information on options for a children's holiday. We construct Decision Tree for given Table 8.

TABLE VIII. SET OF OFFERED SERVICES

Version	Age	Gender	Number of participants	Subject	Type of entertainment	Duration of entertainment	Budget (UAH)	Holiday
1.	up to 5 years	boy	up to 4	not thematic	are active	unlimited		Yes
2.	up to 5 years	boy	up to 4	ninja	science	30 minutes – 1 hour	5000-1000	No
3.	5-10 years	boy	4-8	ninja	science	30 minutes – 1 hour	5000-1000	Yes
4.	up to 5 years	boy	more than 8	Peppa	art	до 2 hour	unlimited	No
5.	up to 5 years	girl	more than 8	Peppa	art	unlimited	5000-1000	Yes
6.	5-10 years	girl	4-8	ninja	team entertainment	unlimited	5000-1000	No
7.	5-10 years	girl	4-8	Lady bug	team entertainment	unlimited	5000-1000	Yes
8.	5-10 years	boy	more than 8	Lady bug	team entertainment	unlimited	5000-1000	No
9.	5-10 years	boy	more than 8	pirates	team entertainment	unlimited	up to 500	No
10.	more than 10 years	girl	more than 8	Peppa	art	до 2 hour	5000-1000	No
11.	more than 10 years	girl	up to 4	Peppa	art	до 2 hour	5000-1000	Yes

12.	more than 10 years	girl	more than 8	Peppa	are active	unlimited	5000-1000	No
13.	more than 10 years	girl	more than 8	not thematic	beauty and fashion	до 2 hour	unlimited	Yes
14.	more than 10 years	boy	more than 8	not thematic	beauty and fashion	до 2 hour	unlimited	No

A holiday is a decision-making attribute. The set of all conditional attributes $\boldsymbol{A}$ = {"age", "gender", "number of participants", "subject matter", "type of entertainment", "duration of entertainment", "budget"} corresponds to the root node. Select the attribute "age" and mark it the root vertex. The set of values of this attribute consists of three elements: up to 5 years, 5-10 years, more than 10 years. Put the root vertex in correspondence with three edges, each of which is attributed to the value of the attribute "age". Set examples will be divided into three subsets that correspond to the values of the attribute "age"; these subsets correspond to each of the vertices 2, 3, 4 of the tree shown in Figs. 3 We remove the attribute "age" from the set A and get the set $\boldsymbol{A}$ = {"gender", "number of participants", "subject matter", "type of entertainment", "duration of entertainment", "budget"}.

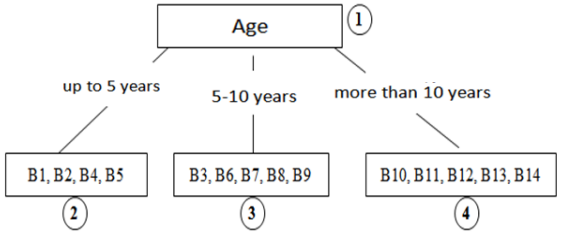

Fig. 3. The first step of the ID3 algorithm (removing the "age" attribute).

Consider the vertex number 3. It is matched by the subset of examples {B3, B7} that have the value of the decision attribute "*yes*" and the subset of examples {B6, B8, B9} that have the value of the attribute of the decision "*no*". We select the following attribute from the set $\boldsymbol{A}$; let it "gender". Denote by vertex 3, construct two edges with the values of this attribute, and divide the set of examples in vertex 3 into two subsets, in each of which the values of gender are the same.

Consider the vertex number 6. It corresponds to the subset {B3}, which has the value of the decision attribute "yes", and the subset of examples {B8, B9}, which have the value of the decision attribute "no". We select the following attribute from the set $\boldsymbol{A}$; let it be the "number of participants". Denote by vertex 6, construct three edges with values of this attribute and divide the set of examples in vertex 3 into two subsets, in each of which the values of the number of participants are the same.

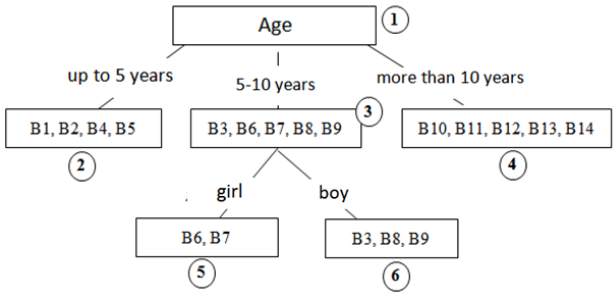

Fig. 4. The second step of the ID3 algorithm (removing the "gender" attribute)

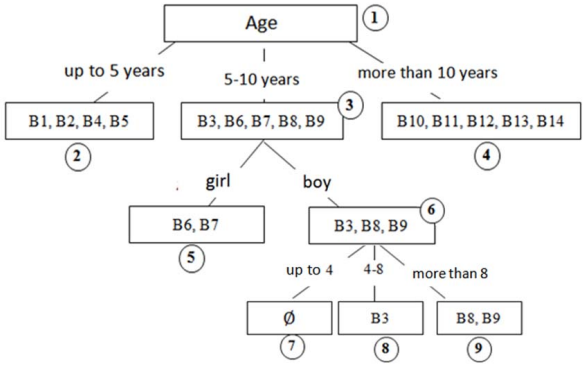

Fig. 5. The third step of the ID3 algorithm (removing the "number of participants" attribute)

In Fig. 5 in verse 7 we have an empty set, which indicates that under such criteria given by the client, we will not be able to offer him anything, that is, a holiday cannot be organized, so we will mark this vertex "no" and it will become a leaf. In verse 9, examples B8 and B9 have the same attributes of the game attribute - no. Therefore, we denote this vertex by "no" and it will become a leaf. Similarly, we denote vertex 8 as "yes" and it will also become a leaf

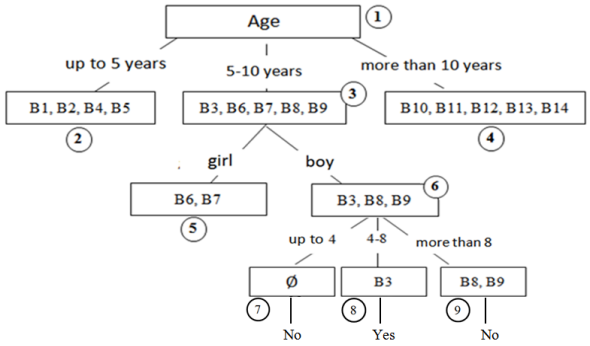

Fig. 6. The fourth step of the ID3 algorithm.

IV. CONSIDERATION OF THE DEVELOPED FUNCTIONAL OF THE SERVICE

The main page (see Fig. 7) of the service for managing the work of children's animation studios, includes buttons: "Offers" - for viewing, all available offers, "Order" - go to the window where you can place an order, "Register" - to register for service, "Login" - to enter the service and also 3 offers that can be accessed by clicking on the "Details" button.

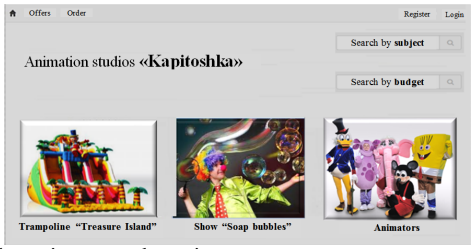

Fig. 7. The main page web service

After the user, when logging in to the site, chooses a budget search option and enters the amount for which he wants to organize a holiday, he will be offered as many services as possible within the selected budget, as shown in the figure. 8.

Fig. 8. A list of suggested holiday options

If someone at the entrance to the site will select the search option offers the direction of "subjects" and introduce subjects on which it wants to organize a holiday, he will be asked to respond to several questions, such as: the age of the child; to become a child; budget, the number of children on holiday, he will be offered as many services within the chosen subject. One of the variants of the proposals when choosing the theme "Active entertainment" and given attributes: "age" - 11 years; "gender" - boy ", budget - 1000 UAH, shown in the figure. 9.

Having determined the optimum number of services, the user can order the selected offers.

V. CONCLUSIONS

To organize the effective functioning of animation studio management services, taking into account the specifics of their work, the logic of query processing systems must be built using modern methods of Data Mining, namely the "Decision Tree" methods. In order to provide quality services, in a user-friendly format, it is most appropriate to provide information to the client in the context of two areas: price policy and content topics of services. If the client chooses as a priority , the direction - pricing policy, it is advisable to use the backtracking algorithm when choosing the client, as the priority , direction - the content of services - algorithms for building a decision tree.

Fig. 9. A list of the proposed "Active Entertainment" holiday option

REFERENCES

[1] Why tourism? World Tourism Organisation UNWTO, access mode: [http://www2.unwto.org/ content/why-tourism].

[2] Top 10 Digital Trends in 2018 year, Idea Digital Agency. 2018. 19 марта. [URL: https://ideadigital.agency/10-golovnih-digital-trendiv-u-2018-rotsi/].

[3] The Business of Theme Parks (Part II): How Much Do They Cost? And Earn? Retrieved from : [http://www.theparkdb.com/blog/the-business-of-theme-parks-part-ii-how-much-do-they-cost-and-earn/].

[4] Official web-page of State Statistics Service of Ukraine, 2019. Retrieved from : [http://www.ukrstat.gov.ua/].

[5] H. Gelter, "Digital Tourism – An analysis of digital trends in tourism and customer digital mobile behavior for the Visit Arctic Europe project", Summary. Interreg Nord, 2017.

[6] TheParkDatabase, World Amusement Parks statistics. Retrieved from: [http://www.theparkdb.com].

[7] S. Babushko, "Digitizing the Activity of Travel Agencies", Topical issues of tourism and tourism practice: Collection of materials of the scientific-practical conference on the 25th anniversary of the Institute of Tourism of the Federation of Trade Unions of Ukraine, Kiev: APSVT, pp. 13-15, April 18, 2019.

[8] O.I Artemenko, V.V Pasichnyk and V.V Egorova. "Information technologies in the field of tourism. Analysis of research applications and results" Bulletin of the National university "Lviv Polytechnic". Information systems and networks, № 814, 2015, pp. 3-22.

[9] N.I. Dumich and O.O. Smotr "Construction composite web application animation studio", in the Third All-Ukrainian Scientific and Practical Conference of Young Scientists, Students and Cadets "Information Protection in Information and Communication Systems", Lviv, pp. 268-269, November 28, 2019.

[10] S.M. Kylymystyj, "Classification of animation activities", International Bulletin: cultural studies, philology, musicology, Issue II (5), 2015, pp. 77-83.

[11] The Adult Price for a Children's Holiday: How Much Do Children's Parties Have? BBC News. Retrieved from: [https://www.bbc.com/ukrainian/entertainment/2015/04/150407children_party_dt].

[12] The official website of the children's animation studio "Children's Planet". Retrieved from: [https://dityacha-planeta.com/].

[13] The official website of the children's animation studio "Igroland". Retrieved from: [http://igroland.com.ua/ua].

[14] The official website of PAPASHON entertainment complex. Retrieved from: [https://papashon.com/].

[15] The official website of the event-agency "Empire of the Holiday" Retrieved from: [http://isvyata.org/].

T. Corman, Algorithms: Construction and Analysis, 3rd ed, Trans. with English, M.: LLC "I. D. Williams", 2013.

IEEE Third International Conference on Data Stream Mining & Processing
August 21-25, 2020, Lviv, Ukraine

Method of Ontology Use in OODA

Vasyl Lytvyn, Dmytro Dosyn, Victoria Vysotska
Information Systems and Networks Department
Lviv Polytechnic National University
Lviv, Ukraine
Vasyl.V.Lytvyn@lpnu.ua

Andrii Hryhorovych
Informatics and Information Systems Department
Drohobych Ivan Franko State Pedagogical University
Drohobych, Ukraine
a.hryhorovych@gmail.com

***Abstract*—The article examines the behavior of an intellectual agent in a competitive environment. An OODA loop is selected to model the behavior. The interaction of the stages of the OODA loop (observation, orientation, decision-making, action) with the ontology of the tasks and the domain within which this agent functions is considered.**

Keywords—OODA loop; observation; orientation; analysis; synthesis; decision making; action, ontology.

I. INTRODUCTION

According to the ideas of John Boyd and his followers, any activity in a competitive environment (for example, in the military field) with some degree of approximation can represented as an OODA cybernetic model [1]. This model involves repeated repetition of a loop of actions, consisting of four consecutive interacting processes (Fig. 1): orientation, decision, action. This model has successfully used to model activities and decision-making in business, politics, sociology, etc., in those areas where there is a competing party. According to Boyd's theory, every person or organization has its own decision-making and activity loop in solving tasks.

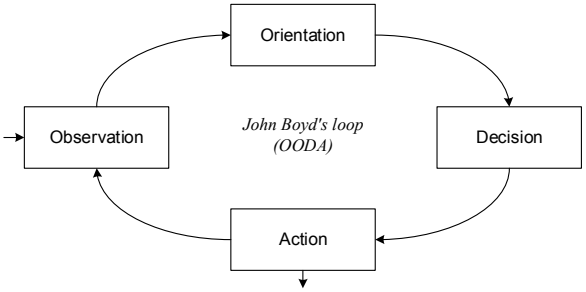

Fig. 1. OODA loop processes

Develop an approach to improve the performance of a management entity in a competitive environment. The purpose of the study is to develop and validate a method for modeling the OODA loop using the knowledge base ontologies of the environment in which the control object operates.

II. ANALYSIS OF RECENT RESEARCH AND PUBLICATIONS

According to the literature on the theory of J. Boyd, summarized in [1], the purpose of the loop stages and their functional purpose is as follows. Observation is the process of gathering the information needed to make a decision in a particular case. The necessary information can obtained from both external and internal sources. Internal sources of information are understood as feedback elements of the loop. As external sensors and other channels of information are used. In order for the observation to be scientific, it must:

- Systematic, not accidental;
- Carried out consistently and systematically;
- Sufficiently broad information was provided about the phenomenon being monitored (as many facts as possible should be operated on);
- provided accurate recording of observation results.

Data can collected:

- Mechanically. Mechanical registration of data is that the source of information, i.e. «event» or «phenomenon», manifests itself as a change of some physical state, and this new state is recorded mechanically.
- Expert observation, followed by the recovery of memory results, which is referred to as «recording».
- Through experimental research, the peculiarity of which is that the phenomenon (the subject of the study) is studied under different conditions and circumstances; the use of this method of research contributes to a deep and very accurate study of a certain psychological patterns.

Orientation is the most responsible and cognitively complex stage in the entire OODA cycle. The orientation stage consists of two sub-stages: destruction and creation. Destruction involves breaking the situation into smaller elemental parts that are easier to understand. The decision-maker or organization will seek to decompose the task to such a level until the newly created components of the task are close to the standard or typical situations for which the decision-maker (ODA) has an action plan. Being acquainted with these elementary typical sub-tasks is achieved through training, coaching, experience and coaching. ODA identifies the current situation, that is, assigns it to a specific sub-task, and applies a pre-prepared action plan for that sub-task. Then these constituent elemental subplots are combined into a general plan of action, which corresponds to the sub-stage of «creation». If there are no plans from which to choose a solution, then the process remains at the orientation stage and further decomposition of the task is carried out. Failure to develop a plan with a real chance of success can further grind to the last cycle. Orientation uses methods of analysis and synthesis, which are closely linked. They are designed to process information obtained through the application of research methods. Analysis is the study of the qualities, properties and characteristics of the object under study by its conditional separation into separate components. In turn, the

synthesis is to summarize information about individual components and to form a set of information data on the object of study as a whole. The results of the analysis and synthesis process are the basis for making various forecasts for the near and far future. Prediction can done by calculation and extrapolation methods.

At the stage of decomposition of the system is carried out:

- Defining and decomposing the overall purpose of the study and the main function of the system as a limitation of the trajectory in the space of states of the system or in the field of acceptable situations. Most often, decomposition is performed by building a goal tree and a function tree:

- Isolation of the system from the environment (division into «system» and «non-system»);
- Description of influential factors;
- Description of development trends;
- Description of the system as a «black box»;
- Functional (by function), component (by type of elements), structural (by type of relations between elements) decomposition of the system.

Decision-making is the third stage of the OODA cycle. If, at this stage, the ODA has been able to form only one real plan, then the decision is made whether to implement the plan or not. If several alternative options are formed, the ODA at this stage selects the best one for further implementation. Choosing the best plan can made on the criterion of value for money. With a time limit, the plan that meets the requirements of fast reliability is preferable. The following methods are used to make a decision:

- The cost-effectiveness method takes into account three stages: construction of the model of efficiency, construction of the model of cost, synthesis of cost and efficiency;
- Methods of theory and practice of reliability are based on the application of the apparatus of probability theory and random processes, mathematical statistics and modeling.

Action is the final stage of the cycle, which involves the practical implementation of the chosen course of action or plan. There are two main ways to achieve competitive advantage in pursuing different types of professional activity. The first way is to quantify your action cycles faster. This enables us to be the first to make decisions and make competitors react to our actions. The second way is to improve the quality of the decisions made, that is, to make decisions that are more relevant to the situation than the competitors are. Improving the quality of one's own decisions can achieved in various ways, which include the use of modern formal mathematical methods, the use of automated control systems, decision support systems, and expert systems. If the latter are used, the modern approach to their construction is used as the nucleus of the ontology knowledge bases [2]. Therefore, the task of developing ontology usage methods in the OODA loop arises.

III. BASIC MATERIAL

The distinguishing feature of the OODA cycle from other cyclical models is that in any situation, it is always assumed that there is a competitive side. In Fig. 2 lists three controls that are in some of their initial states and have their own states of purpose. The fact that these objects operate in a competitive environment, during the passage of the OODA loop, to analyze the states in which the competitors are and their actions, as shown in the figure by the appropriate arrows. It is proposed to use the OODA loop system to use an intelligent system whose core knowledge base is an ontology consisting of domain ontology and task ontology that may arise in this environment. In our opinion, the content of the ontology directly affects the 2nd and 3rd stages of the cycle, and the structure and content of the ontology depends on the 1st and 2nd stages (Fig. 3). Let's take a closer look at each stage of the OODA loop as it interacts with the domain ontology and the tasks that arise in this area. The observation phase enables the ontology development process to analyzed and analyzed to select the relevant information that is required in the next stages of the OODA loop. Formally, an ontology consists of terms (concepts, concepts) organized into a taxonomy, their definitions and attributes, as well as the associated axioms and rules of inference. Therefore, the model of ontology O is understood as three types [2]

$$O = \langle C, R, F \rangle, \tag{1}$$

where C is a finite set of concepts (concepts, terms) of software, R is finite set of relations between concepts (concepts, terms) of a given software, F is interpretation of concepts and relations (axioms). Axioms impose semantic constraints on the system of concepts and relations [3].

In order to build an ontology that adequately describes the semantic software model, it is necessary, first, to solve the tasks of obtaining knowledge from different sources to identify many concepts and establish a hierarchy on that set. Since much of the information is contained in natural-language texts (NLTs), it is promising to gain knowledge of textual information as well as intelligent processing of specially selected NLT collections. One of the most effective approaches to complete ontology is its automated teaching of natural language texts. Automated filling can implemented by analyzing text documents using a knowledge processor (Fig. 4). This approach is discussed in more detail in a monograph [4]. In the given scheme, the task of the linguistic processor is to perform its lexical, grammatical, syntactic and semantic analysis. As a result, the ontology is replenished with concepts, TLDs (subject - action - object) and cause and effect relationships between TLDs. The other part of the important relationship between concepts and their properties is established by the ontology processor, which builds an ontological structure for each concept obtained after analyzing the text.

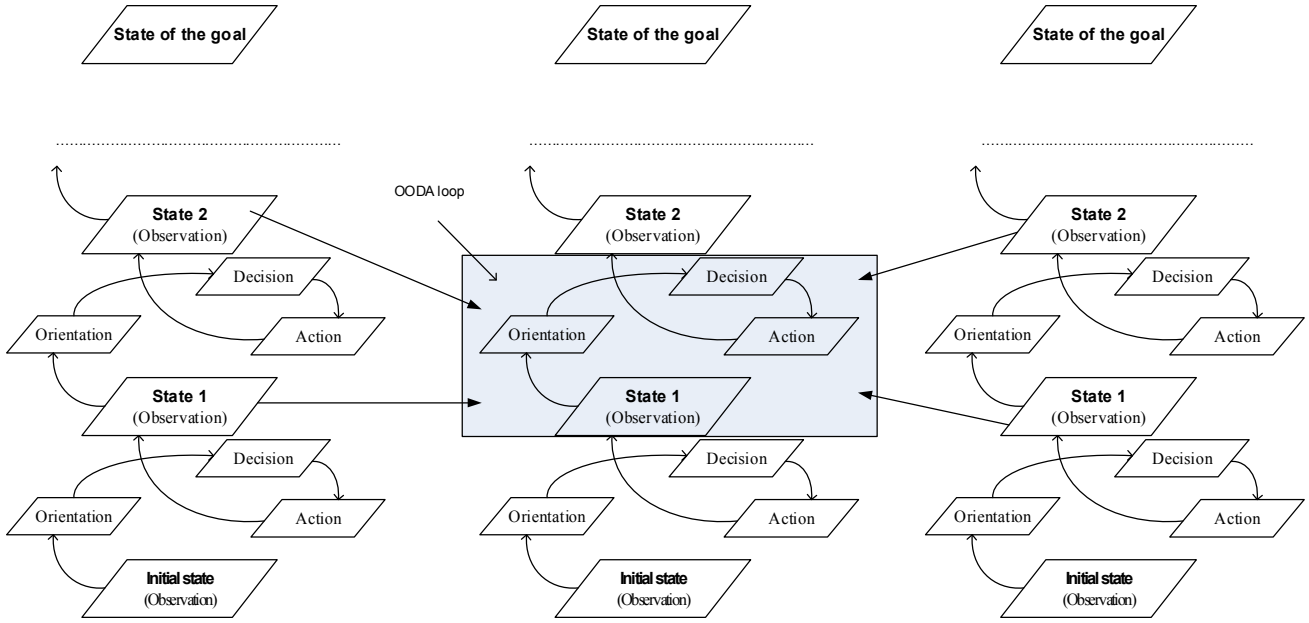

Fig. 2. Operation of facilities in a competitive environment

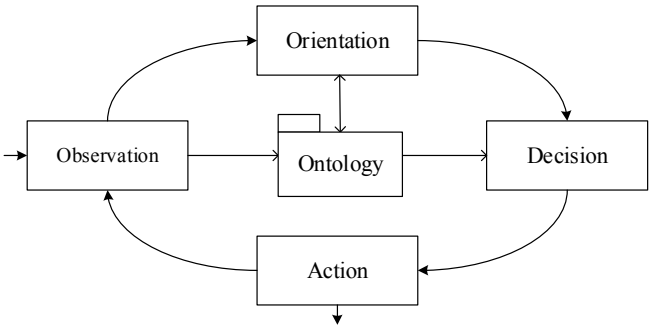

Fig. 3. Using ontologies in the OODA loop

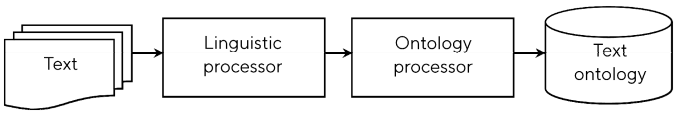

Fig. 4. Structural and functional diagram of the knowledge processor

The work of the ontology processor is supported by the corresponding database, the main components of which are: first, a set of rules, and second, a universal logical database MModWN [5] type WordNet [6]. The Knowledge Processor is used in a system of automated learning from text documents, which in turn is used to solve the task of semantic search in full-text databases. Among the systems developed in Ukraine, it should noted that the development of the staff of the Department of Mathematical Informatics at Taras Shevchenko National University of Kyiv is a system of text processing in natural language [7]. The system is designed to solve tasks such as analysis and synthesis of texts in natural language, automated generation of abstract text, automated indexing (definition of the subject) of the text. Our system is described in detail in [4, 8-12]. The main advantage of our approach is to build an Intelligent Agent (IA) that determines the value of the messages that are proposed to added to the ontology depending on the management plan chosen. The peculiarities of the functioning of a specialized IA are determined by its interest - a vector of estimates of the desirability of possible states of the agent. To describe the agent's interest, by which he distinguishes the states of the surrounding world and positions himself in it, a utility function is used, which is a numerical evaluation of his desirability for the agent [13-18]. Utilities are combined with action probabilities to determine the expected utility of each action. Suppose $U(S)$ the usefulness of the condition S from the point of view of the agent making the decision to perform some action A. An arbitrary undetermined action can cause a result state $Result_i(A)$ where the index i runs over all possible results. Before taking action A, the agent evaluates the likelihood of each of the possible outcomes $P(Result_i(A) | Do(A), E)$, where E is the set of parameters available to the agent of its state, and $Do(A)$ is the statement according to which in the current state the action is performed. Thus, it is possible to calculate the conditional utility of the action based on known status parameters:

$$EU(A \mid E) = \sum P(Result_i(A) \mid Do(A), E) \cdot U(Result_i(A)) \quad (2)$$

If the Maximum Expected Utility (MEU) guides a rational intellectual agent, it is forced to choose an action that

maximizes the expected utility for the agent. This is how a mechanism for motivating the behavior of a rational intellectual agent works, regardless of its scope. In the case of an information retrieval agent Maximum Expected Utility − *MEU*, its interest may be determined by evaluating the novelty of the messages received, which requires the use of methods of intellectual analysis of natural-language texts. We believe that the text is constructed as a message. The message structure is oriented to the perception of another agent, so it consists of two parts (Fig. 5): the ascertaining part, by which the addressee evaluates the relevance of the message (1) and defines its context (2), and the constructive part - potentially new to the reader in in this context (3).

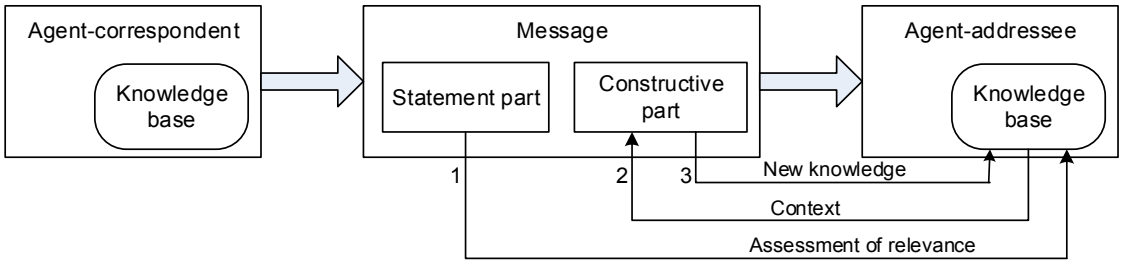

Fig. 5. Recognition of new knowledge in a message to fill their knowledge base

If the new knowledge is not a complete algorithm, but only a single fact or rule that clarifies the algorithms already known to the agent, changing their utility functions serves to evaluate the novelty of the fact (rule) and their importance to the agent. This approach is described in more detail in a monograph [9]. The representation of knowledge in the form of an ontology implies that any possible generalization, that is, a complex, complex concept, is always explicitly articulated, named, and appears as a separate concept in the knowledge base. At the orientation stage, a plan of action is drawn up. In order to reach the target state of IA, a plan to achieve this state with all possible alternatives must first be developed [10]. Planning is based on decomposition. The ZP planning task has three components: the set of states S , the set of actions A , the set of target states $Goal$ (goal states):

$$\text{ZP} = \langle St, A, Goal \rangle . \tag{3}$$

Further, we will assume that the goal state is the same. If there are several states of goal, then the goal can written as a disjunction of these states. Then achieving this state is a solution to some sub-task, so assuming the uniqueness of the goal state is normal. The action A , in turn, has four parts: action name, parameter list, prerequisite, and result. The plan itself is defined as a four-element tuple - <Multiple Actions, Multiple Organization Constraints, Multiple Causal Relationships, Multiple Prerequisites Multiple> [11]. The condition $S(i)$ is given as a set of facts with appropriate probabilistic estimates. The action is presented as a mapping from state $S(i)$ to state $S(j)$ with a corresponding probability p_{ij} , that is $S(i) \to S(j)$, a probability p_{ij} . The IA must be able to evaluate the condition in order to select the necessary actions. It is easier to do this with states in which he was already. It is more difficult to estimate future states. Heuristic functions or metacognition are used to evaluate them. So first, let's look at the assessment of the completed states, then the actions and finally their combination, leading to a new (future) state. Let $v(S(i))$ is an assessment of the condition $S(i)$. We will use software ontology to evaluate states that already had IA. The goal state $Goal$ is determined by the need for some set of attributes to reach certain values $z(x, Goal)\ \forall x \in X$. Each state $S(i)$ is defined by its many traits that acquire value $z(y, S(i))\ \forall y \in Y_i$. To evaluate a condition $S(i)$, it is necessary to map ψ the plurality of features and their status values $S(i)$ to the plurality of features and status values $Goal$. Obviously, the BR, namely the optional module of the Semantic Web Rule Language (SWRL) ontologies, should use such a mapping.

$$\psi : Y_i \xrightarrow{\hat{O}} X . \tag{4}$$

Then an assessment of the condition $v(S(i))$ is calculated

$$v(St(i)) = d(S(i), Goal) = \sum_{x \in X_W} \varphi(z(\psi(y), S(i)), z(x, Goal)) . \tag{5}$$

where X_W is many essential features. For example, to determine the weight of the elements of ontology [4, 12]. φ is a certain metric that depends on the specifics of the software [4, 12]. In our research for choosing IA actions, we will rely on the agent's rationality as an effort to minimize the cost of resources to reach the end state. Therefore, we assume that each action a_{ij} is uniquely determined by the cost of resources g_{ij}^k (cost of transition from state to state), where $k = 1, 2, ..., n_i$ - the number of alternatives α_k for making the transition a_{ij} . Therefore, in the future, we denote the action by three indexes: a_{ij}^k transition from state $S(i)$ to state $S(j)$, using the alternative α_k [13].

Because the lower the score, the better, the action estimate is directly proportional to the cost of resources, that is:

$$v(a_{ij}^k) = E \cdot g_{ij}^k , \tag{6}$$

where E is a scalar value that reduces the measurement of an action estimate to a single dimension with a state estimate. Choosing the best plan is at the decision stage. In general, the decision to choose an action based on an alternative is made according to some relationship between condition and action:

$$o_i\left(a_{ij}^k\right)=\delta\left(v\left(a_{ij}^k\right),v\left(S\left(j\right)\right)\right). \quad (7)$$

In particular, this relationship can be linear:

$$o\left(a_{ij}^k\right)=\omega v\left(a_{ij}^k\right)+\left(1-\omega\right)v\left(S\left(j\right)\right). \quad (8)$$

where $\omega\in\left[0,1\right]$ is part of the alternative action that the IA gives when making the decision, the other share belongs to the state in which it will pass.

After evaluating actions and states, the path selection task is reduced to the task of asynchronous dynamic programming [14]. We get the following model of transitions between states:

$$S\left(j\right)=a\left(S\left(i\right),o_i\right) \quad (9)$$

with the optimization criterion

$$S\left(j\right)=a\left(S\left(i\right),o_i\right) \quad (9)$$

$$\Theta\left(S\left(0\right),\vec{o}\right)\Rightarrow\min(\max). \quad (10)$$

Task (9) - (10) is a dynamic programming task. Using methods suitable for the solution of such tasks, we find the solution in the form of a path from transition to the initial state, that is, the path of the plan. In our further works, it is planned to complicate model (9) - (10) taking into account the states of competitors and to use the developed mathematical support for the modeling of military actions.

IV. CONCLUSIONS

This paper deals with the modeling of the behavior of an intelligent agent that functions in a competitive environment. An OODA loop is selected to model the behavior. To improve the efficiency of the OODA cycle, it is proposed to use the ontology of the domain within which the intelligent agent operates and the ontology of the tasks that arise in this field. The influence of the stages of the OODA loop on the ontology content and vice versa - the ontology content on the course of the stages is determined. The interaction between the ontology and the stages of the OODA loop is suggested. The task of planning the activities of an intellectual agent in a competitive environment is reduced to the task of dynamic programming.

[1] V. Lytvyn, V. Vysotska, D. Dosyn, O. Lozynska, and O. Oborska, "Methods of Building Intelligent Decision Support Systems Based on Adaptive Ontology," IEEE 2nd International Conference on Data Stream Mining and Processing, DSMP, 2018, pp. 145-150.

[2] T. Gruber, "A translation approach to portable ontologies," Knowledge Acquisition, 5 (2),1 993, pp. 199-220.

[3] S. Babichev, "An Evaluation of the Information Technology of Gene Expression Profiles Processing Stability for Different Levels of Noise Components," Data, vol. 3 (4), 2018, art. no. 48.

[4] S. Babichev, B. Durnyak, I. Pikh, and V. Senkivskyy, "An Evaluation of the Objective Clustering Inductive Technology Effectiveness Implemented Using Density-Based and Agglomerative Hierarchical Clustering Algorithms," Advances in Intelligent Systems and Computing, vol. 1020, 2020, pp. 532-553.

[5] S. Sachenko, S. Rippa, and Y. Krupka, "Pre-Conditions of Ontological Approaches Application for Knowledge Management in Accounting," International Workshop on Antelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2009, pp. 605-608.

[6] G.A. Miller, "WORDNET: A lexical database for English," Communications of ACM, 11, 1995, pp. 39-41.

[7] M. Davydov, and O. Lozynska, "Mathematical method of translation into Ukrainian sign language based on ontologies," Advances in Intelligent Systems and Computing, vol. 871, 2018, pp. 89-100.

[8] V. Lytvyn, V. Vysotska, D. Dosyn, and Y. Burov, "Method for ontology content and structure optimization, provided by a weighted conceptual graph," Webology, vol. 15(2), 2018, pp. 66-85.

[9] V. Lytvyn, Y. Burov, P. Kravets, V. Vysotska, A. Demchuk, A. Berko, Y. Ryshkovets, S. Shcherbak, and O. Naum, "Methods and Models of Intellectual Processing of Texts for Building Ontologies of Software for Medical Terms Identification in Content Classification," CEUR Workshop Proceedings, Vol-2362, 2019, pp. 354-368.

[10] Y. Burov, V. Vysotska, and P. Kravets: "Ontological approach to plot analysis and modeling," CEUR Workshop Proceedings, Vol-2362, 2019, pp. 22-31.

[11] I. Khomytska, V. Teslyuk, A. Holovatyy, and O. Morushko, "Development of methods, models, and means for the author attribution of a text," Eastern-European Journal of Enterprise Technologies, vol. 3(2-93), 2018, pp. 41-46.

[12] N. Antonyuk, L. Chyrun, V Andrunyk., A. Vasevych, S. Chyrun, A. Gozhyj, I. Kalinina, and Y. Borzov, "Medical News Aggregation and Ranking of Taking into Account the User Needs," CEUR Workshop Proceedings, Vol-2362, 2019, pp. 369-382.

[13] L. Chyrun, L. Chyrun, Y. Kis, and L. Rybak, "Automated Information System for Connection to the Access Point with Encryption WPA2 Enterprise," Lecture Notes in Computational Intelligence and Decision Making, 1020, 2020, pp. 389-404.

[14] Y. Kis, L. Chyrun, T Tsymbaliak., and L. Chyrun, "Development of System for Managers Relationship Management with Customers," Lecture Notes in Computational Intelligence and Decision Making, 1020, 2020, pp. 405-421.

[15] L Chyrun., A. Kowalska-Styczen, Y. Burov, A. Berko, A. Vasevych, I. Pelekh, and Y. Ryshkovets, "Heterogeneous Data with Agreed Content Aggregation System Development," CEUR Workshop Proceedings, Vol-2386, 2019, pp. 35-54.

[16] L. Chyrun, Y. Burov, B. Rusyn, L. Pohreliuk, O. Oleshek, A. Gozhyj, I. Bobyk, "Web Resource Changes Monitoring System Development," CEUR Workshop Proceedings, Vol-2386, 2019, pp. 255-273.

[17] A. Gozhyj, L. Chyrun, A. Kowalska-Styczen, and O. Lozynska, "Uniform Method of Operative Content Management in Web Systems," CEUR Workshop Proceedings, Vol-2136, 2018, pp. 62-77.

[18] L. Chyrun, A. Gozhyj, I. Yevseyeva, D. Dosyn, V. Tyhonov, and M. Zakharchuk, "Web Content Monitoring System Development," CEUR Workshop Proceedings, Vol-2362, 2019, pp. 126-142.

IEEE Third International Conference on Data Stream Mining & Processing
August 21-25, 2020, Lviv, Ukraine

Information Resources Analysis System of Dynamic Integration Semi-Structured Data in a Web Environment

Andrii Berko, Irina Pelekh
Information Systems and Networks Department
Lviv Polytechnic National University
Lviv, Ukraine
Andrii.Y.Berko@lpnu.ua

Lyubomyr Chyrun, Ivan Dyyak
Programming Department
Ivan Franko National University of Lviv
Lviv, Ukraine
Lyubomyr.Chyrun@lnu.edu.ua

***Abstract*— The article describes the procedure for analyzing information resources of the system of dynamic integration of weakly structured data in the web-environment to determine the common features of information resources and identify links between them. Model-based shared resource definitions and language describing the rules of access to resources, examines the process of creating an object adapter defining the common features of the information resources and identify relationships between them using the rules of the method of "black box". The process of determining the structural-dynamic model of the domain Mashup-application is described.**

Keywords—information resource; dynamic integration; Mashups-application; object adapter; "black box" method; agent-oriented approach

I. INTRODUCTION

The growing number of resources and services available on the Internet and the opportunities that Web 2.0 offers are pushing end users to evolve from passive consumers of information into information producers who are able to access and manipulate existing information resources in order to generate new content. In terms of human-machine interaction, this requires new communication paradigms that should allow people to access content, move it to personal interactive workspaces where they can integrate it, and, if necessary, create new content using existing content. The solution of this problem requires the use of methods and means of dynamic data integration. The data integration system frees users from the need to know which data sources, other than integrated, they use, what properties of these sources and how to access them [1-4]. Today, due to the rapid development of the Internet, a huge number of information resources of various nature and content has accumulated and is constantly growing [5-9]. This raises the problem of their operational, dynamic integration, while highlighting the content of information and preserving the semantics of data [10-13]. The complexity and nature of the methods used to solve this problem depends significantly on the level of integration that needs to provided, the properties of individual data sources and the totality of sources as a whole and the necessary methods of integration [14-19]. Due to the rapid growth of Ukrainian and world information resources, methods of determining the common features of information resources and dynamic integration of information are in the focus of researchers of federated environments and are of scientific and practical interest to developers of modern distributed intelligent information systems, but numerous problems still remain unresolved. Therefore, at present especially urgent task to develop new approaches, technologies, architectural solutions and flexible tools for analyzing the information resources of the system of dynamic integration semi-structured data in a web environment for defining common information resources and identify relationships between them.

II. ANALYSIS OF RECENT RESEARCH AND PUBLICATIONS

Speaking about the dynamic integration of data in the web-environment, the research and development of data integration systems, working using the technology of Mash-Up [4, 20-23]. Mashups are an application development approach that allows users to combine data from multiple sources into a single integration tool. [4, 24-26] Some Meshes can created by a simple combination of JavaScript code with XML, which makes it possible to build a new innovative web service. Other, larger Meshes, which are the basis of relevant web sites, use the technology of services such as Google Maps and an address database, linking them together and showing project information on a map. Unlike web service compositions, where the focus is only on combining business services, the Mashup framework is expanding, which has more functionality and allows for the integration of heterogeneous resources, such as data services. Programs created with the help of Mash-Up technology are called Mashups or Mashup applications. Research shows that there is a high level of interest within Mashup. In addition, with the rapid development of IT-technologies, the needs for dynamic integration of heterogeneous data in the web-environment become especially urgent. Mashups technology opens up new and wide opportunities for data transmission to consumers. However, the fact remains unchanged that the user of Mashups should know, at least, how to write program code using programming languages (for example, Java Script, XML / HTML), how to use various web APIs. In order to solve this problem, there is a considerable amount of effort that is put into developing tools that are designed to support users with little programming knowledge to develop Mashups applications [27-32].

Since the content, nature and structure of a Mashup application cannot fully known in advance, because the Mashup is formed as a response to a user request. Therefore, it is advisable to choose such a method of constructing the

978-1-7281-3214-3/20/$31.00 ©2020 IEEE