MATHEMATICAL MODELING AND COMPUTING, Vol. 8, No. 3, pp. 560–572 (2021)

Forecasting economic result of business logic improvements using Game Theory for modeling user scenarios

Kuzmin O. Ye., Stanasiuk N. S., Berdnik D. A., Gaiduchok O. V.

Lviv Polytechnic National University,
12 S. Bandera Str., 79013, Lviv, Ukraine

(Received 7 July 2021; Revised 15 September 2021; Accepted 24 September 2021)

A new approach to user behavior modeling based on Game Theory was proposed. It was developed to consider initial intensity, a strategy applied, a profit gained, and resources utilized as inalienable attributes of users' behavior. The approach covers various aspects of users' motivation and rational actions, not only a statistical image of a pool's summary. Additionally, the given model is strongly connected to profit and loss parameters by operating with profit and utilized resources as parts of model inputs. The proposed model can enable efficient modeling aimed to validate an economic result of existing interfaces and assume results of new ones.

Keywords: *user behavior, game theory, strategy, positive and negative scenarios, economic result.*

2010 MSC: 00A71, 91A35 **DOI:** 10.23939/mmc2021.03.560

1. Introduction

A. Problem Definition. Many aspects of modern economics related to the use of software are strongly dependent on user behavior. Massive websites, providing services to millions of people daily, or software systems operating in a specific area are controlled and evaluated by human beings. For example, users' behavior plays a significant role in the economic results of electronic trades and information processes related to exchanges. This concept is well known and has been developing within the last 40 years. Various business areas are already identified as user-behavior-dependent; these are mostly multi-user systems providing commercial services to individuals. Many modern online marketing tasks are strongly related to identifying, clustering, and modeling users' behavior. Users' behavior is also a significant influencer of business processes involving very few actors. The latest trends of trade surveillance software market are tracking and predicting traders' behavior. The value of misuse or abuse is exceptionally high.

Another driver affecting the user modeling landscape is trial and modeling costs. The majority of modern tools demand trial groups, checking models, assumptions, or changes. Such groups should be of a statistically significant size to deliver value. Typically running such tests is costing and sometimes leads to loss of conversions if real users were involved and the assumption was wrong or less effective. In more general perception, such areas are looking for tools to analyze users' behavior with minimal costs and minimal involvement of real users.

B. Relevance of research. The modern commercial landscape is pretty much competitive. It forces market players to shorten their expenses and maximize income. Meanwhile, income maximization demands observing narrow clusters, detailed customer profiles, and complex strategies. In most cases, all such circumstances lead to the necessity to get deeper into an understanding of customer groups and their behavior.

Another critical driver of relevance for this research is coming from the international trading domain. Stock exchanges are highly dependent on the behavior of individuals performing trading. Only FCA issued fines to individuals with a total amount of more than 75 000 000 GBP only during the year 2019. By MAR statistics, new attempts of market abuse are detected daily.

All listed above are building a solid convenience in developing new user behavior modeling methods to meet current requirements.

C. Research objective. This research aims to elaborate on the user behavior modeling approach, able to meet modern business requirements listed in the Problem definition and requirements applied to behavior models already. In addition, it should propose bases of the holistic method, considering users' actions as aim-based sequences rather than random actions.

D. Review of latest researches. Doctor's Kobsa apologetic work [1] identifies core milestones of user behavior modeling development. Moreover, it highlights the environment formed as a recognizable brunch of modern computer science and the period when it came to using 'user behavior' as a core abstraction to manage. Kobsa formulates three basic requirements applicable for any user modeling system and assumes that commercial use will produce a more significant number of needs of a lower level. The initial conditions are stated as:

- Generality, including domain independence.
- Expressiveness (to be able to express as many types of assumptions about the user as possible).
- Strong Inferential Capabilities (to express clear reasoning).

Later, following new demands, the evolution of this list leads to a final, multifactor set of requirements for a model to be used in contemporary business:

- Comparisons of different users' selective actions
- Import of external user-related information
- Privacy support

A detailed overview of behavioral analysis approaches is provided in [2], but not the paper itself. The references provided give details about the method's efficiency and use in a different application. Thus, [3] highlights the importance of the Bayesian approach in the user modeling for achieving results correlated on each step iteratively. Consequently, such a model may and most probably will be accurate in situations where users' observations on each step may significantly impact their behavior. Users' actions are not random, and the precise sequence of actions may indicate a cluster of users [4]. The research [5], based on Markov Models, proves that user interface significantly influences users' behavior. A detailed review of the group of works applying different approaches in clickstream analysis and brief evaluation was proposed in [6]. In [7] a comprehensive probabilistic model for the impact of time on users' actions and its value was built. More insights from users' behavior are highlighted in [8, 9]:

- Time is an essential parameter for the determination of users' behavior. It can define the moment of the event and the duration of action or time between steps.
- The sequence of actions performed by a user is more informative than precise action quoted out of the same series.
- Graphical methods are very effective to spot behavior patterns and density anomalies.

A practical foundation in web behavior modeling, highlighting the importance of this domain for the modern business and the impact applied to it by the users' behavior as stated in [3, 5, 6, 10]. [11] highlights the efficiency of user behavior modeling in risk assessment of information systems (web-based). Considering the emerging development of mobile devices' utilization, many researchers [12–14] are paying significant attention to the mobile sector of the web and new opportunities related to additional data sources and IoT [14–17]. Like [19–21], many authors are focused on practical challenges of behavior modeling for gamification. ML and AI are showing promising results in the classification (clustering) of vast groups of users, but some methods are not applicable for small groups of people. This fact may be considered a core disadvantage of big data in behavior analysis [21–27]. Researchers are paying particular attention to the interaction of users and computers [28] and user-to-user interaction [29] as parts of users' behavior. Finally, the holistic approach of defining the core user's characteristics, coming from their behavior, is applied in work [30].

The listed works show possible areas of improvement and the proven value of research in this field.

E. The solution proposed. For this research, we assume that the chosen strategy is a solid artifact and should not be decomposed to review step-by-step decision making and sequential influence of decisions made thru the method. In turn, this paper will examine simple cases where strategy is short and unable to be divided into significant steps. At first glance, such an approach may lose accuracy and effectiveness, but it will cover simple cases where a single behavioral act or short sequence is observed. A detailed review of complex strategies is going to be the subject of additional work.

2. Materials and methods

A. Model. Let us assume that we have a game or strategic interaction 'N'. Interacting sides are a user 'u' and a company 'c'. The formal definition of this game will look like this:

$$N = \{c, u\}, \tag{1}$$

$$S = \{S_c, S_u\}. \tag{2}$$

Consequently, a pool of users $\{u_1, u_2, u_3, \ldots, u_n\}$ will interact with one and the only 'c' thru several games $\{N_1, N_2, N_3, \ldots, N_n\}$. This generalization is fundamental to link groups of users, being a part of practical research, with a single abstract or precise game.

'c' uses a web interface to interact with any 'u' in a standard predefined way. [5] explicitly shows the impact of UI on users' decisions. Disclosed research confirms that an abstract strategy of 'c', exposed in a manner of UI, can interact with users' decision-making process. Besides that, [28] reviews the interaction of a user and a computer as a part of a user's behavior. 'c' is willing to nurture any of 'u' (preferably all) to deliver a desirable sequence of actions. Typically, such a sequence of steps is called an optimistic (positive) user scenario or a 'happy path.' Normally, we are working with a set of positive scenarios. And a strategy of 'c' allows it. As you may conclude from the example given below (Figure 1), a UI may provide the user multiple options leading to equivalently positive results. Generally, there is no practical difference between proposed channels of communication (in case if both are valid), and there is no point in contacting the same user twice even if he typed in both contact fields.

Alternatively, any strategy of 'u', mistreating a proposed UI, is a negative user scenario. A strategy of will 'c' include validation and tips aimed to block a user from negative scenarios or guide to positive.

Considering all these facts, we may conclude that 'c' demonstrates a solid strategy able to be recognized by any of 'u' and any of 'u' may interact with 'c' using different techniques.

Since *clear business objectives guide* 'c', we observe a strategic interaction, having not only a strategy but an aim. And his range of strategies S_c, is represented by a single strategy

$$S_c = r. \tag{3}$$

Where r can be defined as an array $r = [A_1, A_2, \ldots, A_n]$, representing a sequence of mandatory actions $(A_1, A_2, \ldots, A_n)$, expected from any of 'u'. In shorter form $r[]$. Such (or similar) representation is widely used in Markov models [7] to describe the sequence of users' actions.

In turn, 'u' is a user applying a set of strategies S_u. As was already mentioned, these strategies may refer to positive as well as negative scenarios. Each strategy considers as an artifact is one of the user's characteristics. As was noted by [8], a sequence of the user's actions is very informative to describe his behavior.

Having any defined $r[]$, we can compare it with actions performed by 'u'. If

$$r[] = s[]_u, \tag{4}$$

then S_u is a strategy describing a positive scenario and vice versa. As far as 'c' will not limit a time spent by 'u', a user may deliver needed actions using various time ranges. From this perspective, executing the same behavioral act using different amounts of time could be considered various strategies. [7] and [8] show the effective use of time utilized by users for a certain step to amend the model of their behavior. Let us take $i \in \mathbb{N}$ as an amount of time used to perform strategy s_u^i. In this case, expression (4) will appear as:

$$r[]^{\infty} = s[]_u^i = s[]_u^{i+1}. \tag{5}$$

After adding time limits usually used in UI, the expression (4) will appear as:

$$t_{max} > i > t_{min} \Rightarrow s[]_u^i = r[]_{t_{max}}^{t_{min}}. \tag{6}$$

Such a form may also help to describe the time needed to be spent on specific actions.

In addition, 'c' may not limit the number of actions performed in between mandatory actions (reading of other pages, correcting fields already completed, etc.). In this case, any $s[]_u^i$, representing a sequence of actions including mandatory steps, will be considered as a positive scenario (strategy). And expression (4) will transform into:

$$r[] \subseteq s[]_u. \tag{7}$$

However, the expression (4) will remain valid as a description of cases when the precise sequence is strongly demanded.

It is important to note that more complex combinations of steps and requirements of 'c' for their execution are out of the scope of this work and will be reviewed separately. In other words, we consider $r[]$ or any $s[]_u$ as linear and out of internal dependencies.

The final definition of strategy set including all aspects can be represented as:

$$S = \left\{ r[]_{t_{max}}^{t_{min}}, s[]_u^i \middle| i \in \mathbb{N} \right\}. \tag{8}$$

As users may and will probably have a limited amount of time, we need to review the expression $i \in \mathbb{N}$. Let us assume that there is $p \in \mathbb{N}$, defining efforts the user can apply due to physical or any other limitations. In this case,

$$S = \left\{ s_u^i \middle| \substack{i \in \mathbb{N} \\ i \leqslant p} \right\} = P. \tag{9}$$

From this, we can conclude that the set of strategies earlier defined as (8) can be represented as:

$$S = \left\{ r[]_{t_{max}}^{t_{min}}, s[]_u^i \middle| i \in P \right\}. \tag{10}$$

It is naturally that 'c', as well as 'u, will utilize some resources to perform their strategies. To take these resources into account, let's define them as:

$$V = \{v_c, v_u\}. \tag{11}$$

Where V are values of a game for all players, v_c is a value of the game for player 'c', which may be found as marketing, operations, loyalties or any other direct or indirect countable expenses applied by 'c' to make 'u' deliver needed actions. In some cases, v_c can be defined by the actions of 'u' (f.e., processing expanses appearing only if 'u' applies registration). In turn, v_u, defines resources of 'u', utilized to perform strategy he uses. If r does not demand any material input from 'u', and no notable resource utilization, we can use $v_u = f(i)$ to identify the value of time used by 'u', as the only resource utilized. This is very important as many web interactions demand only time for certain actions. Despite the simplicity of this expression, it helps us to operate with v_u without knowing the precise value of time for a 'u'. For example, comparing two strategies $s[]_u^i$ and $s[]_u^{i'}$ we may state that

$$i > i' \Rightarrow v_u^i > v_u^{i'}. \tag{12}$$

Benefits of 'c' and 'u' are also should be taken into consideration, as they define their motivation. Let us define it as:

$$B = \{b_c, b_u\}. \tag{13}$$

In such expression, b_c can be given as precise profit delivered by 'u' to 'c' or it may identify a fact of conversion. In all cases, b_c and v_c are two values connecting this model with P&L used in modern business.

B. Outcomes. For positive scenarios, defined in expressions (4)–(7), a game outcome for 'c' will have an expression:

$$r[]_{t_{max}}^{t_{min}} \subseteq s[]_u^i \Rightarrow U_c\left(r[], \forall\left\{s[]_u^i \middle| t_{max} > i > t_{min}\right\}\right) = b_c - v_c. \tag{14}$$

In turn, the similar expression for negative cases will look like this:

$$\forall \begin{matrix} r[]_{t_{max}}^{t_{min}} \nsubseteq s[]_u^i \\ t_{max} < i \\ i < t_{min} \end{matrix} \Rightarrow U_c\left(r[], \forall s[]_u^i\right) = -v_c. \tag{15}$$

It is essential to highlight that being rational 'c' will not apply an unprofitable strategy. Coming from this introductory statement, we need to raise significant expression:

$$b_c - v_c > 0. \tag{16}$$

This statement will be used widely in our further conclusions related to validation of different strategies able to be applied by 'c'. We may call (16) the formal definition of player's 'c' condition to take part in a game. Although (16) is logical and, as will be further shown, can be used as a condition, in practical use, it should be extended to cover all outcomes of a pool. This requirement appears from the fact that overall profitability of 'c' will come from a number of positive and negatives attempts of multiple 'u'. Therefore, (16) should appear as:

$$Cb_c - Hv_c > 0. \tag{17}$$

Where C is a conversion rate or an expected conversion rate, given in a number of positive outcomes per H attempts (hosts).

It is interesting that being rational, 'u' will act similar to 'c' (approach to gain profit), so we can raise a statement equivalent to (16) regarding any of 'u':

$$b_u - v_u > 0. \tag{18}$$

However, as 'u' most likely will act independently, there is now a form of (18) expanding to all pools. Any correlated actions of multiple 'u' are generally possible but out of the scope of this research. If there are no additional resources demanded from 'u', but time, we can use earlier defined p to modify (18) to:

$$b_u - f(p) \geqslant 0. \tag{19}$$

This expression shows that being rational 'u' will not spend more time than an amount having a similar or lower value than a benefit, accepted from 'c' as a result of interaction in a positive scenario.

A b_u has no general definition. A value, delivered to 'u', as a consequence of positive interaction with 'c', should be defined in each particular case. Alternatively, it may be considered an abstract value; as for further conclusions, the (19) is convenient enough.

The outcomes of a game for 'u' are equivalent to (14), (15), but assuming that there are no additional resources demanded 'u', but time, we can use p to modify these statements respectively:

$$r[]^{t_{min}}_{t_{max}} \subseteq s[]^i_u \Rightarrow U_u\left(r[], \forall\left\{s[]^i_u \middle| \substack{t_{max}>i>t_{min}\\ p\geqslant i}\right\}\right) = b_u - f(i), \tag{20}$$

$$\forall \begin{array}{c} r[]^{t_{min}}_{t_{max}} \not\subseteq s[]^i_u \\ t_{max} < i \\ i < t_{min} \end{array} \Rightarrow U_u\left(r[], \forall s[]^i_u\right) = -f(i). \tag{21}$$

C. Dominating strategies. Let us assume that there is a strategy $s[A]^i_u$, where A is a simple observation during which 'u' can accept demanded actions ($r[]$). 'u' is able to perform this observation despite any strategy of 'c'. This observation lasts l amount of time. Certainly, $r[] \not\subseteq s[A]^l_u$, so (21) will collapse to:

$$r[] \not\subseteq s[A]^l_u \Rightarrow U_u\left(\forall r[], s[A]^l_u\right) = -f(l). \tag{22}$$

Let us also assume that A is a minimal holistic action 'u' able to complete

$$\forall s[]^i_u \neq 0 \Rightarrow l < \forall i. \tag{23}$$

In this case, being rational 'u' will not apply any other strategy than $s[A]^i_u$, if he understands that demanded actions are unable to be completed within p amount of time. Alternatively, in described circumstances, any other strategy is strongly dominated. The formal expression of domination is following:

$$\begin{array}{c} \forall s[]^i_u \neq 0 \\ t_{min} > p \\ p \geqslant l \end{array} \Rightarrow U_u\left(r[]^{t_{min}}_{t_{max}}, s[A]^l_u\right) > U_u\left(r[]^{t_{min}}_{t_{max}}, \forall s[]^i_u\right). \tag{24}$$

It should be admitted that if $p < l$, it will obviously fail to meet the requirements of 'c' as well as complete $s[A]^i_u$. Usually, such behavior refers to accidental attempts or low interest from a taken 'u'. However, in some cases, it can mean that the UI is unclear and confuses users. It is crucial to highlight this outcome. It explicitly shows that the given model can produce conclusions based on the idea of users' rationality and aim-based actions, demanding no test groups. Other similar outcomes will be shown further or elaborated by examining the model, amended with the observed system's parameters.

In the case of 'u' concludes that he can complete all demanded actions within p amount of time; being rational, he will complete the needed sequence. Considering (19), we may define:

$$\begin{array}{c} \forall s[]^i_u \neq 0 \\ r[]^{t_{min}}_{t_{max}} \subseteq s[]^i_u \\ p > t_{min} \end{array} \Rightarrow U_u\left(r[]^{t_{min}}_{t_{max}}, \forall\left\{s[]^i_u \middle| p \geqslant i\right\}\right) > U_u\left(r[]^{t_{min}}_{t_{max}}, \forall\left\{s[]^j_u \middle| t_{min} > j\right\}\right). \tag{25}$$

From reviewing (24) and (25) may appear like 'u' always defines he will complete $r[]$, or he will dismiss the session immediately. However, in many cases, it goes alternatively. For example, a user may deliver strategy $s[A, A_1, A_2]^p_u$, meanwhile $r[A, A_1, A_2, A_3]$ is demanded and obviously $t_{min} > p$. As we know from (24), being rational 'u' should have been applied strategy $s[A]^l_u$. Despite cases when 'u' is really irrational, such situation indicates that during action A 'u' made a wrong assumption about $r[]$ and t_{min} specifically. In fact, the accuracy of this assumption depends on the user's ability to evaluate his performance and the efforts needed to complete $r[]$. However, as shown further, sometimes it indicates that the UI is unclear and provides a pore impression of actions required.

Let us assume that b_u is a constant value appearing before strategic interaction and forming user's aim to interact with 'c'. From this perspective, (20) will have one and the only variable — $f(i)$. Based on this and (12), we can state any strategy $s[]_u^i$ will strongly dominate strategy $s[]_u^{i+1}$, if they are leading to the same result. This statement can be represented as:

$$\begin{matrix} \forall s[]_u^i \neq 0 \\ r[]_{t_{max}}^{t_{min}} \subseteq s[]_u^i \\ p > t_{min} \end{matrix} \Rightarrow U_u\left(r[]_{t_{max}}^{t_{min}}, s[]_u^i\right) > U_u\left(r[]_{t_{max}}^{t_{min}}, s[]_u^{i+1}\right). \tag{26}$$

Besides applying the same strategy within a different time, users may use different strategies, covering $r[]$. This statement was described in (7). Practically it will case the utilization of various amounts of time in most of the cases so that it will be covered by (26). However, it may happen that two strategies, both covering $r[]$, will take the same amount of time. In this case, (12) is not applicable, and we can confirm only weak dominance:

$$\begin{matrix} \forall s[]_u^i \neq 0 \\ p > t_{min} \end{matrix} \Rightarrow U_u\left(r[]_{t_{max}}^{t_{min}}, \left\{s[]_u^i \middle| r[]_{t_{max}}^{t_{min}} = s[]_u^i\right\}\right) \geqslant U_u\left(r[]_{t_{max}}^{t_{min}}, \forall\left\{s[]_u^i \middle| r[]_{t_{max}}^{t_{min}} \subseteq s[]_u^i\right\}\right). \tag{27}$$

(26) and (27) are essential for understanding the user's outcome maximization.

Consequently, 'u' is rational, so he will try to apply the most straightforward strategy, taking minimal time.

$$Argmax_u = U_u\left(r[]_{t_{max}}^{t_{min}}, \left\{s[]_u^{t_{min}} \middle| \begin{matrix} r[]_{t_{max}}^{t_{min}} = s[]_u^{t_{min}} \\ p > t_{min} \end{matrix}\right\}\right) = b_u - f(t_{min}). \tag{28}$$

Earlier, we defined $r[]$ as a predefined strategy. However, let's assume that 'c' runs set of 100 games $\{N_1, N_2, N_3, \ldots, N_{100}\}$ with multiple users $\{u_1, u_2, u_3, \ldots, u_{100}\}$ to define an overall outcome of a pool (17). After evaluating pool outcomes 'c' is able to apply $r[]$ for a new collection of users. Being rational and being aware of (24) and (25) 'c' will try to maximize his outcome in the next pool.

The very first assumption is that 'c' being rational will apply strategy $r'[]$ giving him the best conversion rate will improve the profit part of the equation (17). Any strategy giving a lower conversion rate is strongly dominated. The formal definition of domination is represented below:

$$\begin{matrix} r[]_{t_{max}}^{t_{min}} \cup r'[]_{t_{max}}^{t_{min}} \subseteq s[]_u^i \\ H = H' \\ C > C' \end{matrix} \Rightarrow U_c\left(r[]_{t_{max}}^{t_{min}}, \forall s[]_u^i\right) > U_c\left(r'[]_{t_{max}}^{t_{min}}, \forall s[]_u^i\right). \tag{29}$$

To gain a deeper understanding of actions 'c' applies let's improve (17) to the more detailed view. Let us assume that 'c' applied a strategy $r[A, A_1, \ldots, A_n]$. We mentioned earlier that practically any action of 'u' demands spend from 'c' . So, we can take for our model not a single v_c, but a panel of values $v_c^A, v_c^{A_1}, \ldots, v_c^{A_n}$ representing spends of 'c' taking place for each particular action of any 'u'. It gets more sense in the context of technologies like Lambda (one of the Amazon services). However, careful research in this direction will show many cases where we can define costs for any action performed. For the matter of this paper, it is enough to understand that it is a meaningful outcome of IT systems' operation. Also, let's replace H with values $h^A, h^{A_1}, \ldots, h^{A_n}$, representing a number of times each action took place thru H scenarios. Consequently, h^A will be numerically equal to H, as a very first action taking place in any $s[]_u^i$. In addition, as we earlier considered conversion rate as given in a number of positive outcomes per H attempts (hosts), we can take C as equal to h^{A_n}. Considering listed assumptions, (17) will transform in:

$$h^{A_n} b_c - \sum \left[v_c^A h^A, v_c^{A_1} h^{A_1}, \ldots, v_c^{A_n} h^{A_n}\right] > 0. \tag{30}$$

As (29) focuses only on positive outcomes and its' impact on the profitability of 'c', we need to examine the polynomial replacing the generalized value v_c. This is obvious that it shows the structure of spends 'c' will have relatively to strategies applied by all 'u'. We can elaborate here two critical outcomes. Firstly, 'c' will apply strategies nurturing 'u'to perform precisely a minimal set of needed actions. Any other strategy, bringing the same conversion rate, will be strongly dominated, as shown here:

$$U_c\left(r[]^{t_{min}}_{t_{max}}, \forall\left\{s[]^i_u \middle| r[]^{t_{min}}_{t_{max}} = s[]^i_u\right\}\right) > U_c\left(r[]^{t_{min}}_{t_{max}}, \forall\left\{s[]^i_u \middle| r[]^{t_{min}}_{t_{max}} \subseteq s[]^i_u\right\}\right). \tag{31}$$

It is interesting to admit, that based on (27) and (31), any strategy defined by the expression $r[]^{t_{min}}_{t_{max}} \subseteq s[]^i_u$ is strongly dominated by both sides.

Secondly, 'c' will try to nurture 'u' to perform only $s[A]^l_u$ in case of negative scenarios. The formal expression is given below.

$$\begin{matrix} r[] \nsubseteq \forall \overline{s}[]^i_u \\ s[A]^l_u \subseteq \forall \overline{s}[]^i_u \end{matrix} \Rightarrow U_c\left(r[]^{t_{min}}_{t_{max}}, s[A]^l_u\right) > U_c\left(r[]^{t_{min}}_{t_{max}}, \forall \overline{s}[]^i_u\right). \tag{32}$$

It should be highlighted that in case of any combination of factors giving strong dominance in cases (29), (31), (32), we should apply to actual values of equation (30) and should not expect any dominance by default. However, (30) is explicit enough to understand domination between two strategies in the practical study clearly.

3. Results

3.1. Indication of model inputs

As shown by many pieces of research [2, 5–8], the modern technology stack proposes various tools, enabling science and industry professionals to collect accurate data about users' actions. Considering this fact, we will not focus on tools and methods of data collection. The use of such a model will demand records describing users' strategies and the time utilized. In more complex cases, any other resource should be tracked to amend a model. Let us look at Figure 1. We can assume the following actions as part of $r[]$ and possible acts of 'u':

Expected actions per UI given in Figure 1 are given below.

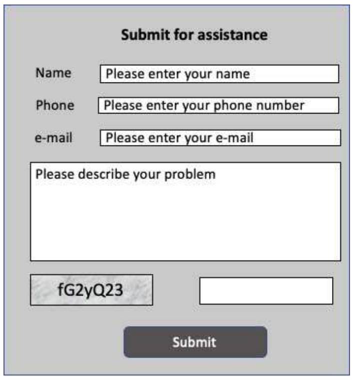

Fig. 1. Mockup of subscription form (private-owned IT service company).

Table 1.

Index	UI Element	Description
A	all	User inspects form to decide does he wants to submit for assistance and if this procedure is affordable for him now
$\mathbf{A_1}$	'Name' field	A user enters his name.
$\mathbf{A_2}$	'Phone' field	A user enters his phone.
$\mathbf{A_{2'}}$	'E-mail' field	A user enters his e-mail
$\mathbf{A_3}$	Text field	A user enters a description of his problem.
$\mathbf{A_4}$	'Submit' button	User hits 'Submit' button
$\mathbf{A_x}$	all	User quits submission form

Analyzing the given user interface, we can conclude that 'c' applies two equivalent strategies $r[A, A_1, A_2, A_3, A_4]^{t_{min}}_{t_{max}}$ and $r'[A, A_1, A_2, A_3, A_4]^{t_{min}}_{t_{max}}$. t_{min} can be defined by examining user experience, while t_{max} can be taken as session expiration time. After collecting timestamps (data from Google Analytics) of actions applied by more than 2000 'u' we can analyze their strategies to find common patterns. First of all, we need to exclude $s[A]^l_u$, strategies to work with a pool of actually interested in further interaction, but it worth to admit that users, applied $s[A]^l_u$, and reasons of their choice is a subject for a separate analysis. To easily recognize commons in actions of multiple 'u' let's inspect the frequency distribution of i for $s[]^i_u$, excluding $s[A]^l_u$, as was mentioned before. However, let's review positive ($s[A, A_1, A_2, A_3, A_4]^i_u$ or $s[A, A_1, A_{2'}, A_3, A_4]^i_u$) and negative $\overline{s}[A, \ldots, A_x]^i_u$ cases as separate charts. It is easy to recognize that a positive chart behaves as expected. We can see normal distribution from 89th to 117th seconds of the session (Figure 2).

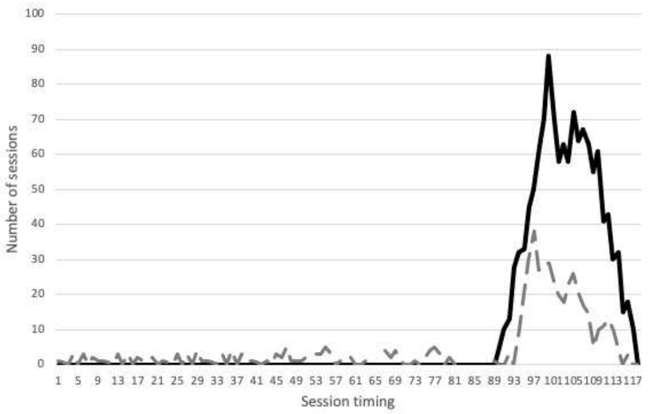

Fig. 2. Frequency distribution of i for $s[]^i_u$. Solid line: ($s[A, A_1, A_2, A_3, A_4]^i_u$ or $s[A, A_1, A_{2'}, A_3, A_4]^i_u$), dashed line: $\overline{s}[A, \ldots, A_x]^i_u$.

The negative cases' chart discloses potential opportunities for improvement. Besides slight noise available along with all charts, we can see a group of results close to the distribution of positive cases. Further analysis showed that two similar strategies cover nearly all cases: ($\overline{s}[A, A_1, A_2, A_x]^i_u$, $\overline{s}[A, A_1, A_{2'}, A_x]^i_u$). In other words, around 3% of users failed to deliverA_3. At the same time, from (24), we can assume that such strategies are strongly dominated by $s[A]^l_u$. As was already mentioned, such behavior can be a result of unclear UI.

3.2. Identifying the economic result

The recommendation, given to 'c' in this case was to exclude A_3 from the demanded sequence of actions. To understand the impact of such improvement on profit and loss, let's review the consequent changes.

The majority of outliers, applying strategies previously $\left(\overline{s}[A, A_1, A_2, A_x]^i_u, \overline{s}[A, A_1, A_{2'}, A_x]^i_u\right)$ will switch to positive scenarios, as excluded step A_3, was the root cause of their confusion. In other words, having no action A_3, making $p < t_{min}$, 'u' will be able to apply a dominant strategy, leading to mutual profit. In addition, we can assume the existence of a group of users, using the strategy $[A]^l_u$ previously, because demanded step A_3 was making $p < t_{min}$. Such users are going to switch to positive scenarios as well.

From the loss perspective, as UI becomes clearer, users, applying negative scenarios, will tend to apply $s[A]^l_u$ only. As was previously shown, it leads to the elimination of operational spending. By modifying impression (30) to an evaluation of this improvement, we can come to an equation, explicitly showing the economic result of improvement:

$$\Delta = ({h'}^{A_4} - h^{A_4})b_c - \sum \begin{bmatrix} v_c^{A_1}({h'}^{A_1} - h^{A_1}) \\ v_c^{A_2}({h'}^{A_2} - h^{A_2}) \\ v_c^{A_{2'}}({h'}^{A_{2'}} - h^{A_{2'}}) \\ v_c^{A_4}({h'}^{A_4} - h^{A_4}) \end{bmatrix}. \tag{33}$$

Consequently,

$$\Delta > 0 \Rightarrow U_c\left(r'[]\,{}^{t_{min}}_{t_{max}}\right) > U_c\left(r[]\,{}^{t_{min}}_{t_{max}}\right). \tag{34}$$

4. Discussion

It is worth highlighting one interesting observation, explicitly showing the elegancy of the model proposed. Let us assume that there is

$$\sigma = \frac{h^{A_n}}{h^A} \tag{35}$$

representing the probability of A_n step execution, meaning the successful completion of $\forall s[]^i_u$, meeting requirements of $r[A, A_1, \ldots, A_n]$. Consequently, we can state that

$$1 - \sigma = \frac{h^A - h^{A_n}}{h^A}. \tag{36}$$

Where $1 - \sigma$ represents the probability of $\forall s[]^i_u$, failing to meet requirements of $r[A, A_1, \ldots, A_n]$. As was proven in (24), (25), (26), and (27), being rational, 'u' will apply either $s[A]^l_u$ or $s[]^i_u = r[]$ and there is no clear dominance between these two strategies, but only the choice of 'u'. Other words, we can say that 'u' will apply mixed strategy

$$S_u\left\{(1-\sigma)s[A]^l_u; \sigma s[]^i_u | s[]^i_u = r[]\right\}. \tag{37}$$

As being rational, he will use $s[A]^l_u$ as one and the only strategy, not meeting requirements of $r[A, A_1, \ldots, A_n]$ with a probability $1-\sigma$; and $s[]^i_u = r[]$ as one and only strategy, meeting requirements of $r[A, A_1, \ldots, A_n]$ with a probability σ. Consequently, the average outcome of 'c' can be represented as:

$$\overline{U}_c = \sigma U_c(s[]^i_u) + (1-\sigma)U_c(s[A]^l_u). \tag{38}$$

Or using (14) and (15):

$$\overline{U}_c = \sigma\left(b_c - \sum\left[v^A_c, v^{A_1}_c, \ldots, v^{A_n}_c\right]\right) - (1-\sigma)v^A_c. \tag{39}$$

Further, using (35) and (36), we can get:

$$\overline{U}_c = \frac{h^{A_n}}{h^A}\left(b_c - \sum\left[v^A_c, v^{A_1}_c, \ldots, v^{A_n}_c\right]\right) - \frac{h^A - h^{A_n}}{h^A}v^A_c. \tag{40}$$

By multiplying both sides of an equation by h^A we are coming to the equivalent of expression (30). It brings us to the significant and exciting conclusion: the profitability requirement, expressed thru conversion rate, (17) and the outcome of 'c' from a mixed strategy of 'u' (37) are actually the same expression.

In more general understanding, it means that the economic result of interaction with users is a straight outcome of strategies they are applying. And from (24), (25), (26), and (27), we know that their choice is led by a benefit they are trying to get and a value of a strategy applied. All these enable us to conclude the high potential of Game Theory in user behavior modeling, despite multiple assumptions and limitations.

5. Conclusions

The proposed model delivers a sufficient modeling mechanism able to provide a versatile description of users' behavior. Even though it can operate with frequency distributions to model a group behavior, it can work with data of one or very few users. Developing an image of users' behavior from the strategy aimed to gain precise goals focuses on individuals' motivation rather than statistic impressions. At the same time, meaning different user's behavior can be compared with a clear identification of consequences. The given example shows the ability to elaborate assumptions and predict the results of their implementations. The tool kit to evaluate practical results is given.

As the model is based on abstract actions and integrates with P&L calculations in a domain-agnostic way, it can be used in various applications. At the same time, the whole concept of 'domination' inherited from the classic Game Theory provides explicit assumptions on users' behavior. As domination is always based on the superior outcomes delivered by a particular strategy, the model brings convenience in reasoning for elaborated assumptions.

Last but not the list, is the ability to work with user data given from different sources. Focused on behavior rather than on personality, the model can work with anonymized data. These two points allow simple integration with existing systems and businesses.

Considering all listed benefits, as well as requirements raised by prof. Kobsa [1] and the disadvantages of known statistic methods, we can conclude that the given model can find wide use in modern business. However, as it covers simple strategies (sequences of actions with no decision-making inside), the need to study more complex interactions is obvious.

[1] Kobsa A. Generic user modeling systems. User Modeling and User-Adapted Interaction. **11** (1–2), 49–63 (2001).

[2] Zhang M., Wang Y., Chai J. Review of User Behavior Analysis Based on Big Data: Method and Application. International Conference on Advances in Mechanical Engineering and Industrial Informatics (AMEII 2015). 99–103 (2015).

[3] Hassan M. T., Junejo K. N., Karim A. Bayesian Inference for Web Surfer Behavior Prediction. Lahore: Dept. of Computer Science, Lahore University of Management Sciences (2007).

[4] Borges J., Levene M. Data Mining of User Navigation Patterns. International Workshop on Web Usage Analysis and User Profiling. 92–112 (1999).

[5] Cottam J. A., Blaha L. M. Bias by default? A means for a priori interface measurement. Cognitive Biases in Visualisations. 46–58 (2017).

[6] Wang G., Zhang X., Tang S., Zheng H., Zhao B. Y. Unsupervised Clickstream Clustering for User Behavior Analysis. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 225–236 (2016).

[7] Petrovskiy M. A data mining approach to learning probabilistic user behavior models from the database access log. Proceedings of the First International Conference on Software and Data Technologies – Volume 2: ICSOFT, 73–78 (2006).

[8] Beutel A. User Behavior Modeling with Large-Scale Graph Analysis (Ph.D. paper). Pittsburgh, PA: Computer Science Department, School of Computer Science, Carnegie Mellon University (2016).

[9] Wall E., Arcalgud A., Gupta K., Jo A. A Markov Model of Users' Interactive Behavior in Scatterplots. 2019 IEEE Visualization Conference. 81–85 (2019).

[10] Kumbarovska V., Mitrievski P. Behavioral-based modeling and analysis of Navigation Patterns across Information Networks. Journal of Emerging Research and Solutions in ICT. **1** (2). 60–74 (2016).

[11] Canali D., Bilge L., Balzarotti D. On the effectiveness of risk prediction based on users browsing behavior. Proceedings of the 9th ACM symposium on Informatics, computer and communications security. 171–182 (2014).

[12] Cabafero L., Hervas R., Gonzňlez L, Fontecha J., Mondéjar T., Bravo J. Characterization of mobile-device tasks by their associated cognitive load through EEG data processing. Future Generation Computer Systems. **113**, 380–390 (2020).

[13] Ellavarason E., Guest R., Deravi F. Evaluation of the stability of swipe gesture authentication across usage scenarios of a mobile device. Eurasip Journal on Information Security. **2020**, Article number: 4 (2020).

[14] Sharma K., Giannakos M., Dillenbourg P. Eye-tracking and artificial intelligence to enhance motivation and learning. Smart Learning Environments. **7**, Article number: 13 (2020).

[15] Sultan K, Ali H., Ahmad A., Zhang Z. Call details record analysis: A spatiotemporal exploration toward mobile traffic classification and optimization. Information. **10** (6), 192 (2019).

[16] Stylios I., Kokolakis S., Thanou O., Chatzis S. Behavioral biometrics & continuous user authentication on mobile devices: A survey. Information Fusion. **66**, 76–99 (2021).

[17] Lee H., Upright C,, Eliuk S., Kobsa A. Personalized visual recognition via wearables: A first step toward personal perception enhancement. Personal Assistants: Emerging Computational Technologies. 95–112 (2018).

[18] Gerina F., Massa S. M., Moi F., Reforgiato Recupero D., Riboni D. Recognition of cooking activities through air quality sensor data for supporting food journaling. Human-Centric Computing and Information Sciences. **10**, Article number: 27 (2020).

[19] Papoutsoglou M., Kapitsaki G., Angelis L. Modeling the effect of the badges gamification mechanism on personality traits of Stack Overflow users. Simulation Modelling Practice and Theory. **105**, 102157 (2020).

[20] Smiderle R., Rigo S. J, Marques L. B,, Pesanha de Miranda Coelho J. A., Jaques P. A. impact of gamification on students' learning, engagement, and behavior-based traits. Smart Learning Environments. **7**, Article number: 3 (2020).

[21] Bovermann K., Bastiaens T. Towards a motivation design? Connecting gamification user types and online learning activities. Research and Practise in Technology Enhanced Learning. **15**, Article number: 1 (2020).

[22] Aizerman M. A., Braverman E. M., Rozonoer L. L. Method of Potential Functions in the Theory of Learning Machines. Nauka, Moscow (1970), (in Russian).

[23] Labayen V., Magaiia E., Moratй D., Izal M. Online classification of user activities using machine learning on network traffic. Computer Networks. **181**, 107557 (2020).

[24] Kobsa A. What is explained by AI models? Artificial intelligence. 174–189, (2018).

[25] Wassouf W. N., Alkhatib R., Salloum K. Predictive analytics using big data for increased customer loyalty: Syriatel Telecom Company case study. Journal of Big Data. **7**, Article number: 29 (2020).

[26] Yu H., Sun L., Zhang F. A Robust Bayesian Probabilistic Matrix Factorization Model for Collaborative Filtering Recommender Systems Based on User Anomaly Rating Behavior Detection. KSII Transactions on Internet and Information Systems. **13** (9), 4684–4705 (2019).

[27] Vuong T., Saastamoinen M., Jacucci G., Ruotsalo T. Understanding user behavior in naturalistic information search tasks. Journal of the Association for Information Science and Technology. **70** (11), 1248–1261 (2019).

[28] Wu T., Zheng K., Wu C., Wang X. User identification using real environmental human-computer interaction behavior. KSII Transactions on Internet and Information Systems. **13** (6), 3055–3073 (2019).

[29] He S., Zheng X., Zeng D. D. Modeling user behavior with competitive interactions. Information and Management. **56** (4), 463–475 (2019).

[30] Wang M., Wang G., Zhang Y., Li Z. A high-reliability multi-faceted reputation evaluation mechanism for online service. IEEE Transactions on Services Computing. **12** (6), 836–850 (2019).

Прогнозування економічних результатів удосконалення бізнес-логіки моделювання сценаріїв користувача із використанням теорії ігор

Кузьмін О. Є., Станасюк Н. С., Берднік Д. А., Гайдучок О. В.

Національний університет "Львівська політехніка", вул. С. Бандери, 12, 79013, Львів, Україна

Запропоновано новий підхід до моделювання поведінки користувачів на основі теорії ігор. Початкова інтенсивність, застосована стратегія, отриманий прибуток та ресурси, що використовуються як невід'ємні атрибути поведінки користувачів – все це було враховано при розробленні нового підходу. Метод охоплює різні аспекти мотивації та раціональних дій користувачів, а не лише статистичний образ набору даних. Крім того, дана модель тісно пов'язана з параметрами прибутку та збитку, оперуючи прибутками та використаними ресурсами як частинами вхідних даних моделі. Запропонована модель може забезпечити ефективне моделювання, спрямоване на перевірку економічних результатів існуючих інтерфейсів та прогнозування результатів нових.

Ключові слова: *поведінка користувачів, теорія ігор, стратегія, позитивні та негативні сценарії, економічний результат.*