

М.Н. Дубовенко, В.М. Белов, А.С. Булгакова, В.В. Зосимов, Т.М. Гонтарь

Информационная технология анализа психологической информации из социальных сетей средствами индуктивного моделирования

Приведены результаты применения комбинированного итерационно-комбинаторного алгоритма индуктивного моделирования для анализа психологической информации. Определены критерии информативности показателей для исследования Интернет-зависимости. Получена новая формула расчета по диагностическому вопроснику Янг, позволяющая более точно проводить первичную оценку уровня Интернет-зависимости у пользователей социальных сетей.

The results of the combined iterative-combinatorial algorithm of the inductive simulation for the analysis of psychological data are presented. The criteria of informative indicators for the study of Internet addiction defined. A new formula for calculating the Young's diagnostic questionnaire is obtained, which allows to realize more accurate initial assessment of the level of Internet Addiction among users of social networks.

Наведено результати застосування комбінованого ітераційно-комбінаторного алгоритму індуктивного моделювання для аналізу психологічної інформації. Визначено критерії інформативності показників для дослідження Інтернет-залежності. Отримано нову формулу розрахунку за діагностичним опитувальником Янг, що дозволяє більш точно проводити первинну оцінку рівня Інтернет-залежності серед користувачів соціальних мереж.

Введение. Современное состояние проблемы Интернет-зависимости пребывает в достаточно неоднозначной ситуации, поскольку до сих пор нет окончательно признанных критериев этой зависимости. До настоящего времени для диагностирования Интернет-зависимости чаще других используют тест американского психиатра Янг (*K. Young*) [1], в том числе и в исследованиях в русскоязычных странах [2].

С появлением социальных сетей начали проводить исследования с целью изучения Интернет-зависимости. Предварительные результаты таких исследований приводят *Griffits* и *Kuss* [3], которые показывают, что многие особенности *Web 2.0*, способствуют появлению зависимости от социальных сетей. Исходя из этого, можно предположить, что Интернет-зависимость однозначно связана с уровнем развития Интернет-технологий, которые по своей природе динамично изменяются и, без сомнения, Интернет претерпел значительные изменения со времени своего возникновения. Отсюда – дискуссия: насколько значимы заложенные в тесте Янг вопросы, относительно определенных показателей и критериев Интернет-зависимости, адекватно используемых для последующего определения степени такой зависимости [4].

Рассмотрим структуру теста на Интернет-зависимость Янг. Изначально автором теста Янг для исследования Интернет-зависимости были адаптированы критерии определения игромании по Международной классификации болезней (*DSM-IV*). Затем ею была предложена шкала Интернет-зависимости, состоящая из 20 вопросов, варианты ответов на которые имеют оценку от одного до пяти баллов [1, 5]. Вопросы теста приведены в таблице. Минимальное количество полученных после тестирования баллов – 20, а максимальное – 100. При этом уровень Интернет-зависимости Y рассчитывался по формуле:

$$Y = \sum_{i=1}^{20} w_i \quad (1),$$

где i – номер вопроса, w_i – количество баллов по i -му вопросу.

Далее значение Y сопоставляется с интервалами I , по которым принимается заключение о следующих значениях оценки диапазона изменений:

$I_1 : Y \in [20; 49]$ – норма;

$I_2 : Y \in [50; 79]$ – Интернет вызывает некоторые проблемы;

$I_3 : Y \in [80;100]$ – Интернет вызывает значительные проблемы (диагностируется как Интернет-зависимость) [1].

Имеются исследования, где авторы [6] используют модифицированный тест Янг с учетом только шести критериев, среди которых:

- I Значимость Интернета.
- II Чрезмерное использование Интернета.
- III Пренебрежение работой ради Интернета.
- IV Предвкушение выхода в Интернет.
- V Отсутствие контроля при использовании Интернета.

VI Пренебрежение социальной жизнью ради Интернета.

Проведенная верификация этого модифицированного теста корреляционными методами показала, что показатели, критерии и сам тест достаточно валидны [6]. Однако отсутствие весовых коэффициентов при w_i в формуле (1) придает всем показателям равную значимость, что в нынешних условиях развития и доступности Интернета ощутимо изменилось. По отношению к началу использования теста произошли значительные технологические, экономические и социальные изменения, что естественно сказалось на условиях потребителей Интернет-услуг.

Постановка задачи

Возникает задача анализа показателей и критериев, используемых в тесте для определения Интернет-зависимости с целью более строгого определения информативных показателей и, в связи с этим, модификации формулы расчета уровня Интернет-зависимости.

Методы исследования

Для решения указанной задачи следует установить взаимосвязь между информативными показателями в формуле расчета уровня Интернет-зависимости. Для этого была разработана информационная технология, состоящая из трех этапов:

- 1) интеграции программного приложения в социальные сети;
- 2) тестирования пользователей социальных сетей на Интернет-зависимость;
- 3) анализа данных и моделирования.

Схема технологии приведена на рис. 1.

Этап 1 – интеграция программного приложения в социальные сети детально описана в работах [7–9].

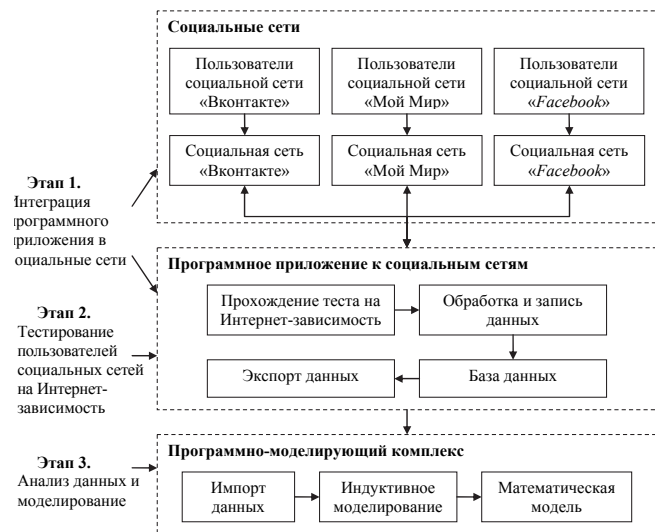


Рис. 1. Схема информационной технологии анализа психологической информации из социальных сетей средствами индуктивного моделирования

На **этапе 2** для получения данных использован авторский программный комплекс, интегрированный в социальные сети «ВКонтакте», «Мой Мир», «Facebook». Пользователям социальных сетей было предложено анонимно пройти автоматизированный тест на Интернет-зависимость. Результаты тестирования обрабатывались, фиксировались в базу данных программного комплекса и были экспортированы в программно-моделирующий комплекс для анализа.

На **этапе 3** проведен анализ средствами индуктивной самоорганизации моделей по экспериментальным данным или «индуктивного моделирования», которое использует индуктивный подход: от конкретных данных наблюдений к общей модели. Итерационные алгоритмы [10–12] решают задачи построения моделей по выборке $W = (Xy)$ данных n наблюдений за m входными и одной выходной переменными [10–12].

В качестве алгоритма индуктивного моделирования в нашей задаче использован комбинированный итерационно-комбинаторный алгоритм (КИКА), при котором возможны пары как из промежуточных аргументов, так и из про-

межуточных и начальных, с применением комбинаторной оптимизации частных моделей [10] (рис. 2).

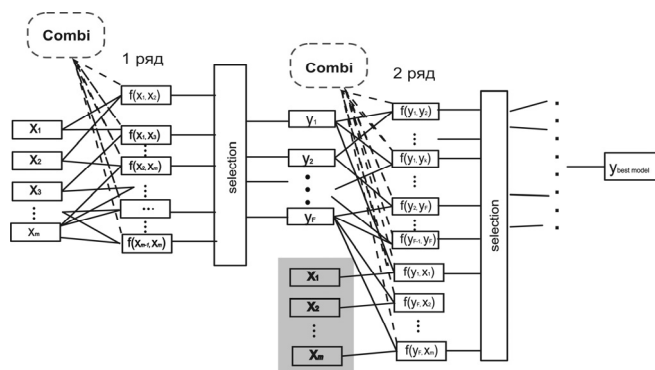


Рис. 2. Схема работы комбинированного итерационно-комбинаторного метода группового учета аргументов

Рассмотренный алгоритм ищет оптимальную модель решения задачи оптимизации:

$$f^* = \arg \min_{f \in \Phi} CR(y, f(X, \hat{\theta}_f)), \quad (2)$$

где CR – критерий регулярности:

$$AR_{B|A} = \|y_B - \hat{y}_{B|A}\|^2 = \|y_B - X_B \hat{\theta}_A\|^2 \quad (3)$$

основан на разбитии выборки W на две части A и B , где $\hat{\theta}_A$ – оценка параметров на подвыборке A с помощью метода наименьших квадратов. Искомая оптимальная модель представляет собой нашу формулу расчета уровня Интернет-зависимости.

В качестве экспериментальных данных для определения зависимости между показателями были взяты их значения (20 показателей модифицированного теста Янг), полученные от пользователей в ходе тестирования. Данные прошли проверку на достоверность [8]. Авторами получен информационный массив данных тестирования в течение шести месяцев от 3626 пользователей. Таким образом, выборка содержит 20 переменных и 3626 точек наблюдений ($n_A = 2050$, $n_B = 1576$).

Точность модели определяется по формуле коэффициента детерминации:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} 100\%, \quad (4)$$

где \bar{y} – среднее значение, \hat{y}_i – выход модели.

Построение модели проводилось средствами программно-моделирующего комплекса, реализующего работу итерационных алгоритмов. Программное обеспечение позволяет работать с различными наборами данных, которые могут храниться и редактироваться в формате *Excel* и текстовом редакторе Блокнот. Программа имеет удобный интерфейс для работы с данными, на основе которых строятся модели любой сложности и структуры с применением различных выборок. Лучшие модели представляются системой для графического и качественного анализа, они хранятся в базе данных вместе с промежуточными расчетами и результатами экспериментов для дальнейшего применения.

Программно-моделирующий комплекс позволяет строить модели сложных систем с использованием итерационных алгоритмов метода группового учета аргументов, определять влияние того или иного показателя на целевой фактор, а также анализировать и отбирать информативные показатели [13].

Обсуждение результатов

Итак, получена новая формула расчета оценки Интернет-зависимости с использованием модифицированного теста Янг:

$$Y' = 2,1 + 1,18x_1 + 1,12x_3 + 1,06x_4 + 1,45x_5 + 1,14x_6 + 1,05x_8 + 1,56x_9 + 1,42x_{10} + 1,3x_{12} + 1,16x_{13} + 1,56x_{14} + 1,13x_{15} + 1,39x_{16} + 1,16x_{17} + 1,34x_{18}. \quad (5)$$

Новое минимальное значение Y' составляет 21,12, максимальное 97,2. Интервалы, на которых строятся заключения, соответственно смещены к таким значениям:

$$\begin{cases} I'_1 : Y \in [21,12; 48,7], \\ I'_2 : Y \in [48,71; 77,23], \\ I'_3 : Y \in [77,24; 97,2]. \end{cases} \quad (6)$$

Проведено сравнение расчетов Y и Y' . Адекватность новой формулы определяется попаданием Y и Y' в один и тот же интервал I . Ус-

тановлено, что в 95 процентах случаев ($R^2 = 95\%$) значения Y и Y' попадают в один и тот же интервал.

Среди 20 исследуемых вопросов неинформативными оказались пять. В таблице приведено сравнение результатов исследования на информативность показателей тестов, проведенное авторами и другими исследователями [6]. Серым цветом выделены вопросы, которые по результатам работы оказались незначительными, что позволило исключить их из диагностического вопросника на Интернет-зависимость. Кроме этого, наши расчеты привели к тому, что из шкалы Интернет-зависимости Янг в русскоязычном секторе остались только критерии: I, II, III, V, VI.

Структура теста на Интернет-зависимость

№	Вопросы						
		I	II	III	IV	V	VI
1	2	3	4	5	6	7	8
19	Как часто Вы предпочитаете провести больше времени в онлайн-нежели прогулку с кем-то?	+					
13	Как часто ругаетесь, кричите или иным образом выражаете свое недовольство, когда кто-то беспокоит Вас, когда Вы в онлайн?	+					
12	Как часто Вы ощущаете, что жизнь без Интернета скучна и пуста?	+					
15	Как часто Вы предвкушаете или фантазируете в оффлайне, чем займетесь как только окажетесь онлайн?	+					
10	Как часто Вы блокируете беспокоящие мысли о вашей реальной жизни успокаивающими мыслями об Интернете?	+					
2	Как часто Вы пренебрегаете домашними делами ради времени в онлайн?	+					
14	Как часто Вы пренебрегаете сном, засиживаясь в Интернете допоздна?	+					
20	Как часто в оффлайне Вы испытываете депрессию, подавленность или нервозность, которые проходят, когда Вы возвращаетесь в онлайн?	+					
1	Как часто Вы замечаете, что проводите в онлайн больше времени, чем изначально предполагали?	+					
18	Как часто Вы пытаетесь скрыть количество времени, которое провели в онлайн?	+					
6	Как часто Ваши успехи в работе или учебе страдают из-за количества времени, проведенного в онлайн?			+			
8	Как часто Ваша производительность труда снижается из-за Интернета?			+			

Продолжение таблицы

1	2	3	4	5	6	7	8
9	Как часто Вы занимаете оборонительную позицию и скрываете, когда кто-то спрашивает, чем вы занимаетесь в онлайн?			+			
11	Как часто Вы предвкушаете очередной выход в онлайн?				+		
7	Как часто Вы проверяете электронную почту прежде чем сделать что-то более важное?				+		
17	Как часто Вы терпите поражение в попытках сократить время, которое проводите в онлайн?					+	
5	Как часто Вы испытываете раздражение, когда окружающие интересуются количеством времени, которое проводите в онлайн?					+	
16	Как часто Вы говорите себе: «Еще пару минут», находясь в онлайн?					+	
4	Как часто Вы заводите новые онлайн-знакомства с интернет-пользователями?						+
3	Как часто Вы предпочитаете пребывание в онлайн интимному общению с партнером?						+

Таким образом, специальный математический анализ показал, что наиболее информативными критериями Интернет-зависимости, согласно полученным результатам, есть:

- высокая значимость Интернета для пользователя;
- чрезмерное использование Интернета;
- пренебрежение работой из-за Интернета;
- отсутствие контроля в использовании Интернета;
- пренебрежение социальной жизнью из-за Интернета.

Заключение. Применение комбинированного итерационно-комбинаторного алгоритма индуктивного моделирования для анализа психологической информации позволяет достичь высокой точности.

Получена новая формула расчета по диагностическому вопроснику Янг, в которой взвешен каждый критерий и исключены неинформативные показатели, что позволяет более точно проводить первичную оценку уровня Интернет-зависимости у пользователей социальных сетей.

Окончание на стр. 92

Использование информационной технологии анализа психологической информации из социальных сетей средствами индуктивного моделирования дало возможность отобрать наиболее информативные показатели и критерии для анализа Интернет-зависимости, создать оригинальный вопросник. Эта технология может быть применена и для других подобных задач.

1. *Netaddiction.com*. Internet Addiction Test. – http://netaddiction.com/index.php?option=com_bfquiz&view=onepage&catid=46&Itemid=106
2. Лоскутова В.А. Интернет-зависимость как форма нехимических аддиктивных расстройств // Автореф. диссертации. – <http://user.lvs.ru/vita/avtoreferat.htm>
3. Kuss D., Griffiths M. Online Social Networking and Addiction—A Review of the Psychological Literature // *Int. J. Environ. Res. Public Health*. – 2011. – N 8(9). – P. 3528-3552. – <http://www.mdpi.com/1660-4601/8/9/3528/>
4. Nicki A. Dowling, Kelly L. Quirk. Screening for Internet Dependence: Do the Proposed Diagnostic Criteria Differentiate Normal from Dependent Internet Use? // *CyberPsychology & Behavior*. – Febr. 2009. – N 12(1). – P. 21–27.
5. *Internet Addiction Scale (Young's Diagnostics Questionnaire)*. – http://www.knowmo.ca/capacity/addictionmeasures/addictionmeasureslist/10-12-31/Internet_Addiction_Scale_Young%E2%80%99s_Diagnostic_Questionnaire.aspx
6. Widyanto L., McMurran M. The Psychometric Properties of The Internet Addiction Test // *CyberPsychology & Behavior*. – Febr. 2004. – N 7(4). – P. 443–450.
7. Белов В.М., Дубовенко М.Н. К проблеме Интернет-зависимости // *Кибернетика и вычислительная техника*. – 2010. – 161. – С. 53–60.
8. Дубовенко М.Н., Белов В.М. Концептуальный алгоритм классификации психологических проблем пользователей на основе приложений в социальных сетях // Там же. – 2011. – 165. – С. 3–15.
9. Дубовенко М.Н., Белов В.М. Некоторые аспекты технологии разработки социальных приложений, адаптированных к нескольким социальным сетям в задачах исследования зависимости от Интернета // *УСиМ*. – 2012. – № 3. – С. 39–44.
10. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным. – М.: Радио и связь, 1986. – 118 с.
11. Степашко В.С., Булгакова О.С., Зосимов В.В. Гібридні алгоритми самоорганізації моделей для прогнозування складних процесів // *Індуктивне моделювання складних систем: 36. наук. пр.* – Київ: МННЦ ІТiС, 2010. – 2. – С. 75–87.
12. Булгакова О.С., Степашко В.С. Порівняльний аналіз ефективності ітераційних алгоритмів МГУА за допомогою обчислювальних експериментів // *Вісн. ЧДТУ*. – 2011. – № 1. – С. 41–44.
13. Булгакова О.С. Програмний комплекс моделювання складних систем на основі ітераційних алгоритмів МГУА з можливістю мережевого доступу // *Матеріали 14 міжнар. наук. конф. «Системний аналіз та інформаційні технології» SAIT'2012*. – С. 176–177.

Поступила 18.07.2012

Тел. для справок: +380 44 503-9544 (Київ)

E-mail: dep150@ukr.net

© М.Н. Дубовенко, В.М. Белов, А.С. Булгакова,
В.В. Зосимов, Т.М. Гонтарь, 2012