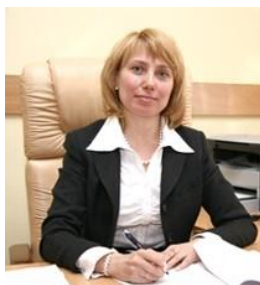


УДК 339.13

## ВИЗУАЛИЗАЦИЯ ПУТИ ОНЛАЙН-ПОКУПАТЕЛЯ СРЕДСТВАМИ ПРОЦЕССНОЙ АНАЛИТИКИ В ЯЗЫКЕ R



**В.А. Пархименко**  
Заведующий кафедрой экономики БГУИР, кандидат экономических наук, доцент



**Е.Н. Живицкая**  
Проректор по учебной работе, кандидат технических наук, доцент



**С.В. Наркевич**  
Преподаватель кафедры экономики, магистр экономических наук

Белорусский государственный университет информатики и радиоэлектроники,  
Республика Беларусь  
Email: parkhimenko@bsuir.by, jivitskaya@bsuir.by, s.v.narkevich@yandex.ru

### **В.А. Пархименко**

Окончил БГУИР по специальности "Экономика и управление народным хозяйством" (специализация "Экономическая информатика"). Работал на предприятиях ГПО "Белэнерго". С 2010 заведующий кафедрой экономики БГУИР. Научные интересы: Data Mining в маркетинге, математическое моделирование маркетинга, маркетинг программных продуктов и ИТ-услуг. Имеет более 100 научных публикаций.

### **Е.Н. Живицкая**

Окончила Минский радиотехнический институт по специальности "Автоматизированные системы управления", аспирантуру БГУИР, магистратуру Международной высшей школы бизнеса и менеджмента технологий Белорусского государственного университета по специальности «Бизнес-администрирование» (МВА), специальный факультет иностранных языков по переподготовке кадров Минского государственного лингвистического университета. Автор более 160 научных работ, из них 4 монографии, 1 учебное пособие с грифом министерства образования РБ, 10 учебно-методических пособий. Под руководством Живицкой Е.Н. подготовлены и защищены одна кандидатская и 13 магистерских диссертаций.

### **С.В. Наркевич**

Окончил БНТУ (специалист) и БГУИР (магистр наук). Работал инженером-экономистом в ГП «НИПТИС», экономистом группы ценообразования и договорной работы и отдела бюджетирования в ОАО "Керамин". С 2003 года преподаватель кафедры экономики БГУИР.

**Аннотация.** В статье рассматривается подход к визуализации пути онлайн-покупателя (customer journey) средствами процессной аналитики (process mining) набора пакетов VupaR в языке R.

**Ключевые слова:** визуализация, онлайн, покупатель, аналитика, язык R.

Построение так называемого пути покупателя (customer journey map) – традиционная задача в рамках маркетингового анализа. Ее сущность заключается в определении тех шагов, которые проходит покупатель (пользователь, потребитель) в процессе взаимодействия с компанией – от первоначального узнавания о существовании компании до повторных покупок (лояльности) или ухода к конкурентам.

У пути пользователя могут быть выделены несколько измерений («слоев») [1]. По мнению авторов, важно различать

- стадии покупательского поведения (реакции) с точки зрения теории (осознание потребности, поиск информации, оценка вариантов, приобретение, повторная покупка, формирование лояльности и т.д.);
- конкретные действия потребителя (например, просмотр баннерной рекламы и переход от нее к посадочной веб-страницы определенного продукта);
- точки контакта, т.е. «многочисленные и разнообразные ситуации, места и интерфейсы соприкосновения клиента с компанией» [2];
- обслуживающие подразделения, т.е. структурные звенья компании, которые обеспечивают взаимодействие компании с потребителем в той или иной точки контакта в ходе того или иного действия потребителя, находящегося на той или иной стадии покупательской реакции.

Обычно путь покупателя визуализируют в форме таблицы или графика, описывающих цепочки типовых шагов покупателя и, возможно, все остальные упоминаемые выше «слои». Эту визуализацию называют «картой» (customer journey map).

При этом существуют разные подходы к построению такой карты. Один из первых и наиболее естественных – ее построение пути потребителя силами экспертов, вообще без использования информационных и иных технологий. Понимание потребителя до сих пор остается во многом во власти интуиции и опыта эксперта, который формализует те шаги и состояния, через которые проходит клиент.

Однако в случае онлайн-магазинов (с их большим объемом посетителей и транзакций) интуиция и опыт, вероятно, существенно уступают тем возможностям, которые дают современные вычислительные мощности и технологии интеллектуального анализа данных.

В этом контексте классическим инструментом анализа пути потребителя является, по мнению авторов, отчет о потоке пользователей (Users Flow report) в Google Analytics [3] или другие отчеты о потоках (Behavior Flow, Events Flow, Goal Flow) [4].



Рисунок 1. Пример потока пользователей (Users Flow) в Google Analytics [5]

В статье авторы показывают, что одно из возможных и, насколько мы осведомлены, еще не используемых решений в задаче построения customer journey map – это применении методологии процессной аналитики.

Процессная аналитика (process mining) подразумевает анализ реализации бизнес-процессов на основе автоматического анализа журнала событий («логов»), т.е. анализа совокупности фактически зарегистрированных событий [6].

В рамках такой аналитики классически решаются три важных подзадачи:

– обнаружение (Discovery): строится модель бизнес-процессов, исходя из журналов событий;

– проверка соответствия (Conformance checking): анализируются расхождения между бизнес-процессами «на бумаге» и тем, как они на самом деле реализуются;

– оптимизация (Performance Mining): бизнес-процессы совершенствуются с использованием информации, полученной в ходе предыдущей задачи.

Существует довольно большое множество алгоритмов процессной аналитики, хотя базовым (эталонным) является  $\alpha$ -алгоритм ( $\alpha$ -algorithm), предложенный van der Aalst, Weijters и Maruster в 2003 году [7].  $\alpha$ -алгоритм позволяет построить сети Петри на основе данных из журналов событий («логов»), т.е. совокупности событий с зарегистрированным временем их возникновения.

Известными программными средствами для процессной аналитики, по мнению Wil van der Aalst, являются следующие (таблица 1).

В данной статье, как указывалось выше, предлагается использовать методологию процессной аналитики для построения пути онлайн-потребителя. В качестве технического обеспечения процессной аналитики предлагается использовать набор специализированных пакетов VupaR для языка R.

Пакет VupaR был создан в 2016 году исследовательской группой в области экономической информатики (Business Informatics research group) в Университете города Хассельт (Бельгия). В настоящее время его поддержка осуществляется Gert Janssenswillen [9].

Таблица 1.

Основные программные средства для процессной аналитики [8]

Наименование	Производитель	Вебсайт
ProM ProM Lite	Open Source	www.promtools.org
RapidProM	Open Source	www.rapidprom.org
Celonis Process Mining	Celonis GmbH	www.celonis.de
Disco	Fluxicon	www.fluxicon.com
Interstage Business Process Manager Analytics	Fujitsu Ltd	www.fujitsu.com
Minit	Gradient ECM	www.minitlabs.com
myInvenio	Cognitive Technology	www.my-invenio.com
Perceptive Process Mining	Lexmark	www.lexmark.com
QPR ProcessAnalyzer	QPR	www.qpr.com
Rialto Process	Exeura	www.exeura.eu
SNP Business Process Analysis	SNP AG	www.snp-bpa.com
webMethods Process Performance Manager	Software AG	www.softwareag.com

VupaR включает в себя ряд пакетов (edeaR, eventdataR, petrinetR, processmapR, processmonitR, xesreadR и др.), каждый из которых обеспечивает различные аспекты процессной аналитики. Обобщая, скажем, что набор пакетов VupaR позволяет трансформировать исходные данные в требуемый формат журнала событий («лога») и осуществить ряд методов визуализации и анализа бизнес-процессов.

Рассмотрим апробацию предлагаемого подхода к визуализации пути онлайн-потребителя средствами VupaR.

Отметим, что реализация данного подхода в нашем случае потребовало также использования функционала таких пакетов, как jsonlite, stringr, tidyverse, lubridate.

Исходными данными для анализа послужила информация, предоставленная на условиях анонимности интернет-магазином одной из стран СНГ.

Представленные данные (файлы в формате json) содержат традиционную для интернет-магазинов информацию о Посетителях (Visitors), Визитах (Visits), Событиях (Events), Товарных категориях (Categories), Товарах (Offers), Просмотрах страниц (View\_page), Просмотрах товаров (View\_product) и т.д.

В качестве объекта анализа был Визит, т.е. посещение интернет-магазина конкретным посетителем в конкретный промежуток времени, в течение которого этот пользователь просматривает разные веб-страницы, включая страницы с описанием конкретного товара, и по итогу совершает или не совершает покупку. Именно в рамках одного визита и строился путь покупателя, и обобщения всех возможных путей позволило осуществить анализ и сделать необходимые визуализации.

Для того, чтобы методы набора пакетов VupaR могли работать корректно, исходные данные должны соответствовать определенной модели (рисунок 2) и содержать информацию о действии/событии (activity), времени события/события (timestamp), ресурсе, задействованном в событии (resource) и т.п.

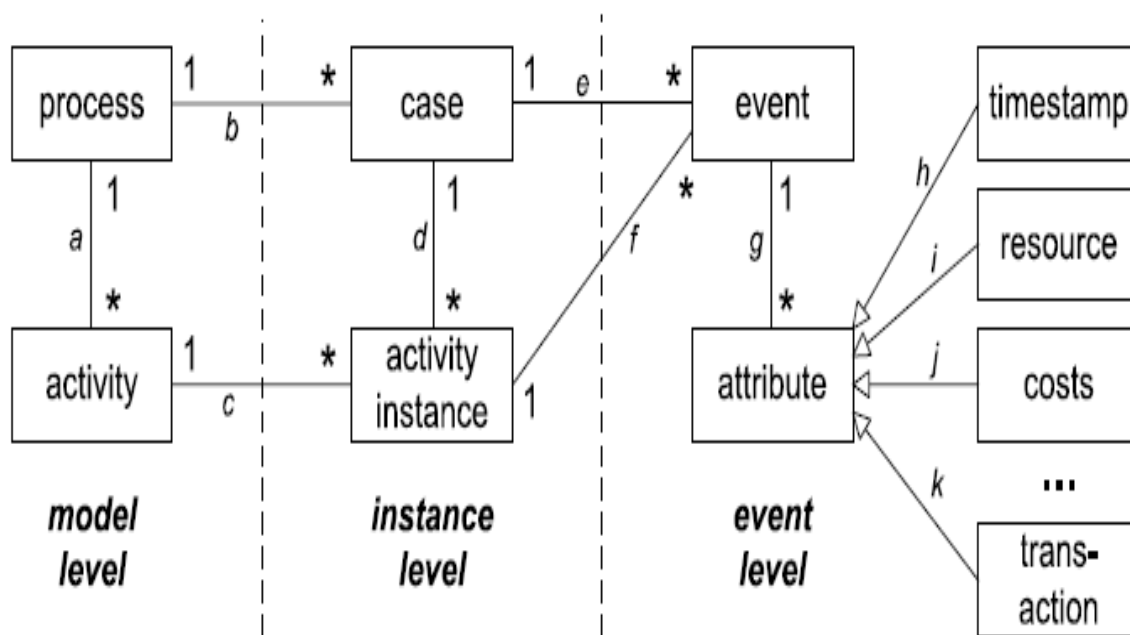
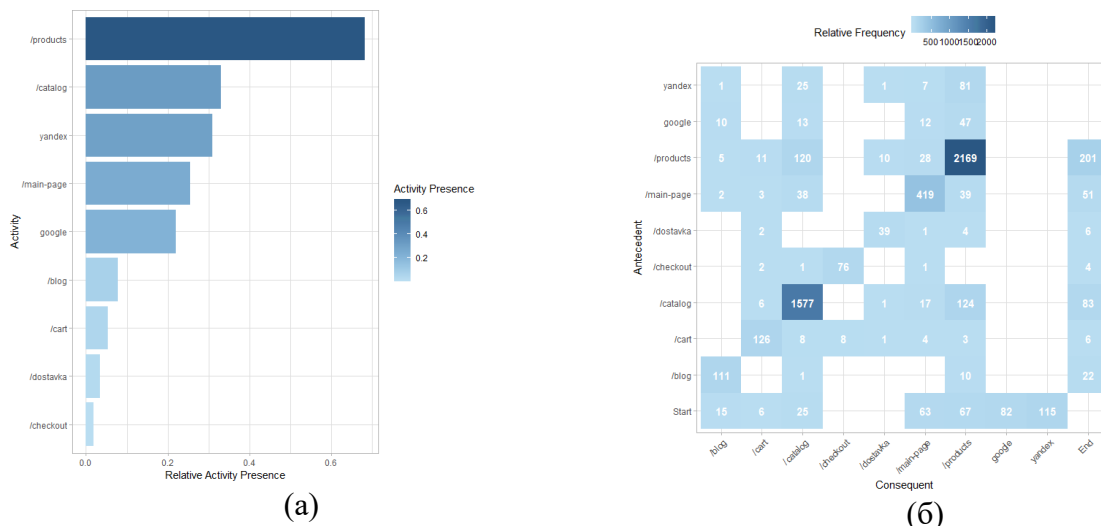


Рисунок 2. Модель данных о событии в рамках VupaR [9]

Таким образом, в ходе рассматриваемой в данной статье апробации предложенного подхода к построению пути покупателя Визит рассматривался как Case, а Просмотры страниц в ходе одного визита рассматривались как Activity в терминологии VupaR. Также источник прихода потребителя в интернет-магазин (Google, Yandex, сайты партнеров и т.п.) рассматривался как самая первая Activity (для этого время события – Timestamp – был принудительно определен за счет вычитания одной секунды от просмотра самой первой страницы на сайте в ходе визита).

После трансформации исходных данных в данные требуемой структуры с помощью функции eventlog() был создан журнал событий.

Использование функций `activity_presence()` и `precedence_matrix()` позволило осуществить дескриптивный анализ и визуализацию его результатов (рисунок 3) с точки зрения основных действий посетителя интернет-магазина (т.е. просмотра конкретных страниц и/или перехода на сайт из внешних источников) и частоты переходов от одного события/действия к другому (матрица предшествования).



Примечание. При визуализации осуществлена фильтрация по частоте событий/действий на уровне 90%.

Рисунок 3. Результаты дескриптивного анализа: (а) – основные действия онлайн-потребителя; (б) – и матрица предшествования

Визуализация непосредственно пути покупателя (посетителя) была осуществлена с помощью функции `process_map()` с фильтрацией по разным уровням частоты событий/действий и устойчивых цепочек действий/событий (traces):

```
my_log %>%
  filter_activity_frequency(percentage = 0.9) %>%
  filter_trace_frequency(percentage = 0.8) %>%
  process_map()
```

Результаты визуализации приведены на рисунке 4.

Особый интерес представляет визуализация пути посетителей интернет-магазина, которые завершили свой визит покупкой [10]. На рисунке 5 представлены сети Петри, построенные посредством функции `process_map()` по посетителям, отобранным по условию, что покупка сделана:

```
my_log %>%
  filter(Purchase_done == 1) %>%
  process_map()
```



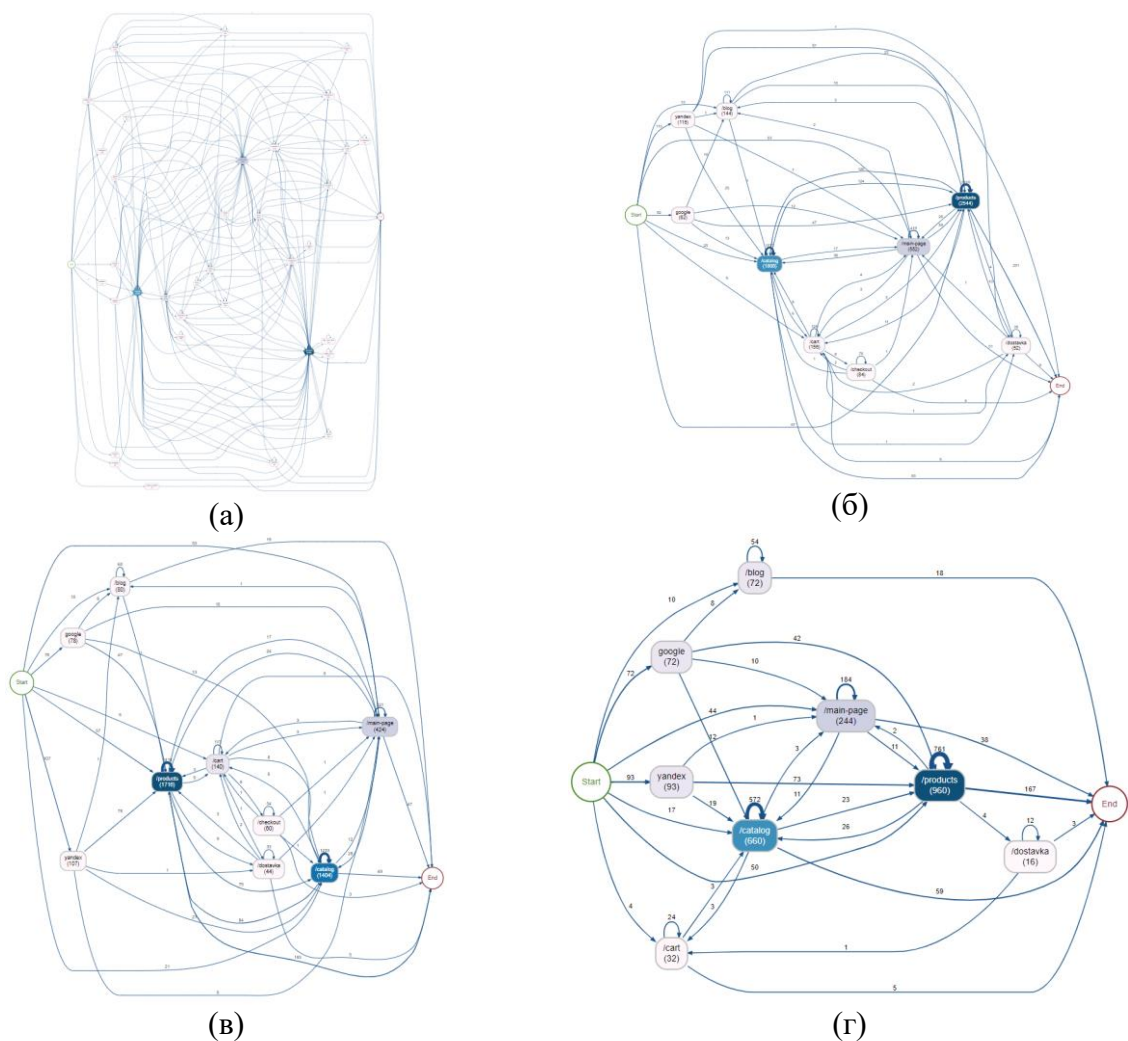


Рисунок 4. Результаты визуализации пути онлайн-потребителя по разным уровням частоты событий/действий и устойчивых цепочек действий/событий: (а) – 100%/100% , (б) – 90%/90%, (в) – 90%/80%, (г) – 90%/70%

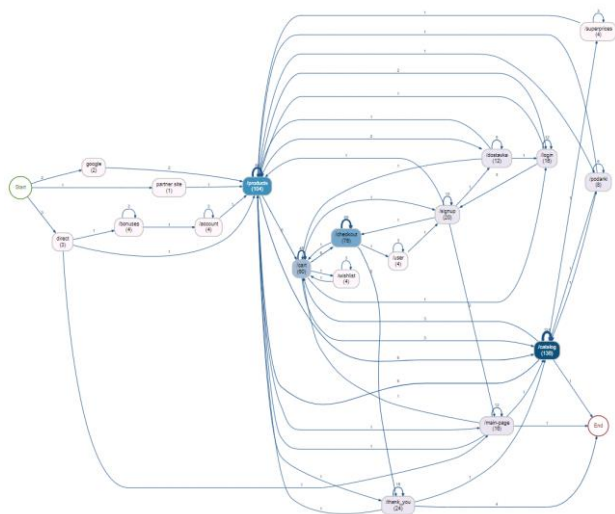


Рисунок 5. Путь онлайн-потребителей, совершивших покупку

Наконец, еще одной возможностью визуализировать отдельные аспекты пути посетителя интернет-магазина является построение графиков во временном разрезе (рисунок 6).

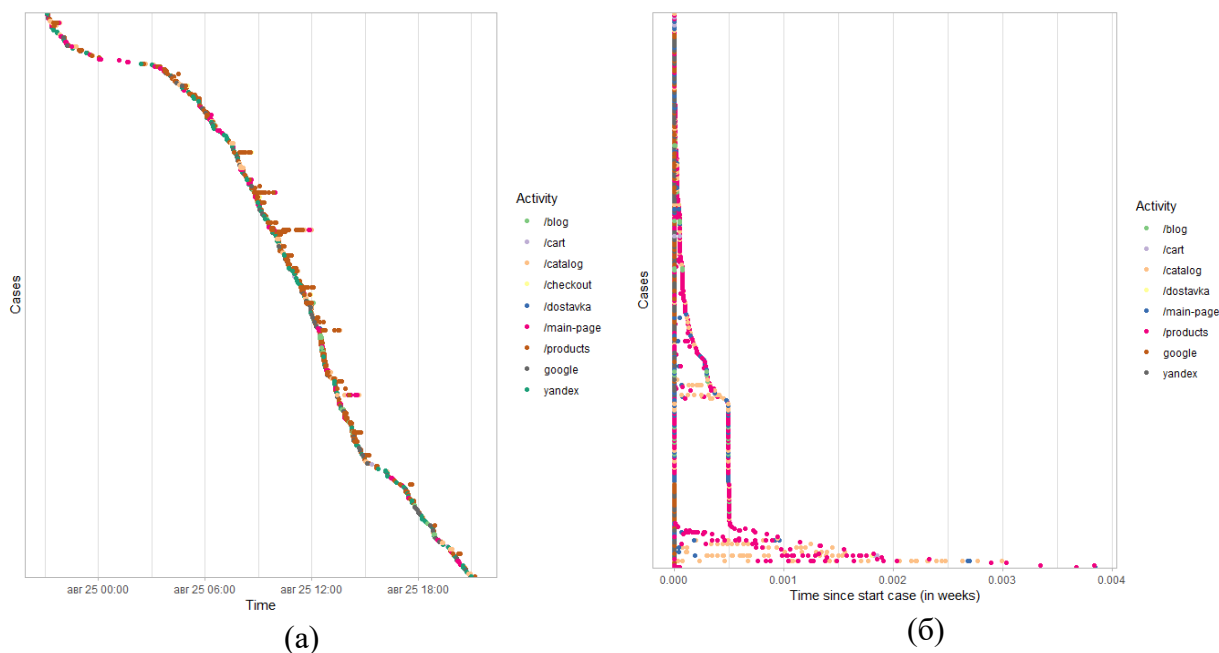


Рисунок 6. Визуализация путей всего множества потребителей во временном разрезе: (а) – абсолютные значения, (б) – относительные (длительность)

В качестве заключения стоит отметить, что предложенный подход позволяет достаточно просто построить различные визуализации разных аспектов пути посетителя (покупателя) интернет-магазина. В данной статье рассмотрены лишь некоторые основные. В то же время открытым остается вопрос о существенной эффективности предложенного подхода по отношению к существующим инструментам, однако авторы полагают, что его основным достоинством является гибкость и адаптивность за счет широкого использования возможностей языка R.

### Литература

- [1] Crowe, D. How to Map a Customer Journey in Ecommerce: The Data Behind Consumer Psychology and Experience // Analytics. Shopify [Электронный ресурс]. – Режим доступа: <https://www.shopify.com/enterprise/customer-journey-map-ecommerce>. – Дата доступа: 04.02.2019.
- [2] Манн, Игорь. Точки контакта. Простые идеи для улучшения вашего маркетинга / Игорь Манн, Дмитрий Турусин. — 4-е изд. — М. : Манн, Иванов и Фербер, 2015. — 156 с.
- [3] Analyze your data with Users Flow // Analytics Help [Электронный ресурс]. – Режим доступа: <https://support.google.com/analytics/answer/1713056?hl=en>. – Дата доступа: 04.02.2019.
- [4] About the flow visualization reports // Analytics Help [Электронный ресурс]. – Режим доступа: <https://support.google.com/analytics/answer/2519986>. – Дата доступа: 04.02.2019.
- [5] Isenberg, E. A Beginner's Guide to Google Analytics Users Flow // TechWyse Internet Marketing [Электронный ресурс]. – Режим доступа: <https://www.techwyse.com/blog/website-analytics/a-beginners-guide-to-google-analytics-users-flow/>. – Дата доступа: 04.02.2019.
- [6] W.M.P. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. – Springer Verlag, 2011. – 352 p.
- [7] van der Aalst, W.M.P., Weijters, A.J.M.M., Maruster, L. Workflow Mining: Discovering process models from event logs // IEEE Transactions on Knowledge and Data Engineering. –2003. – vol 16. – P. 1128 – 1142 [Электронный ресурс]. – Режим доступа: <http://www.wis.win.tue.nl/~wvdaalst/publications/p245.pdf>. – Дата доступа: 04.02.2019.

[8] van der Aalst, W.M.P. How to get started with process mining? // RWTH Aachen University [Электронный ресурс]. – Режим доступа: <http://www.padsweb.rwth-aachen.de/wvdaalst/courses/howtogetstarted-withprocessmining.pdf>. – Дата доступа: 04.02.2019.

[9] Business Process Analysis in R [Электронный ресурс]. – Режим доступа: <https://www.bupar.net/index.html>. – Дата доступа: 04.02.2019.

[10] Parkhimenka, U., Tatur, M., Zhvakina, A. "Heuristic approach to online purchase prediction based on internet store visitors classification using data mining methods." Information and Digital Technologies (IDT), 2017 International Conference on. IEEE, 2017.

## **VISUALIZATION OF ONLINE CUSTOMER JOURNEY USING PROCESS MINING METHODOLOGY IN R LANGUAGE**

**V.A. PARKHIMENKO**

*Head of the Department of Economics of BSUIR, Candidate of Economic Sciences, Associate Professor*

**E.N. ZHIVITSKAYA**

*Vice Rector for Academic Affairs, Candidate of Technical Sciences, Associate Professor*

**S.V. NARKEVICH**

*Lecturer of the Department of Economics, Master of Economic sciences*

*Belarusian State University of Informatics & Radioelectronics, Belarus*

*Email: parkhimenko@bsuir.by, jivitskaya@bsuir.by, s.v.narkevich@yandex.ru*

**Abstract.** The article describes an approach to visualization of online customer journey using process mining methodology via BupaR suite in R language.

**Keywords:** visualization, online, customer, analytics, R language.