Analysing Buyers' Burstiness in E-Business: Parameter Estimation and Practical Applications

Andreas Ahrens and Jelena Zaščerinska
Hochschule Wismar, University of Technology, Business and Design Philipp-Müller-Straße 14, 23966 Wismar, Germany

Keywords: E-Business Process, Buyers' Burstiness, Gap Processes, Binary Customer Behaviour.

Abstract: Optimization of e-business process assists in earning profits in e-business. For the success of the optimization of e-business process, a simulation model based on gap processes for the analysis of buyers' burstiness in e-business process has been recently developed. However, the model has to be validated in terms of input parameter values and distributions. The research question is as follows: What are practically relevant input parameter values and distributions of the model based on gap processes for the analysis of buyers' burstiness in e-business process? The aim of the research is to validate the simulation model based on gap processes for the analysis of buyers' burstiness in e-business process in terms of input parameter values and distributions underpinning elaboration of a new research question on the model validity. The meaning of the key concepts of validation, model validation and model validation approach is studied. The results of the present research show that the simulation model for analysis of buyers' burstiness e-business process in terms of input parameter values and distributions is valid. The novel contribution of the paper is revealed in the newly created research question on the proposed model for evaluation of buyers' burstiness in e-business process. Directions of further research are formulated.

1 INTRODUCTION

Optimization of e-business process assists in earning profits in e-business. Optimization of e-business process implies choices about quantity of goods to be delivered, number of the staff to be employed as highlighted in (Ahrens et al., 2015), goods' pricing, goods discounts, computer software to be installed, networking between a business company and its customers to be established, etc. Additionally, such a result of business process as purchase and/or sale of a good or service indicates the output of this process.

For the success of the optimization of e-business process, a simulation model based on gap processes for the analysis of buyers' burstiness in e-business process has been recently developed (Ahrens and Zaščerinska, 2016). Existing models do not take into account the context of e-business process. E-business process proceeds under certain conditions. One of the conditions is bursty processes that are quite common in our daily live. Table 1 demonstrates the phenomenon of burstiness in a range of scientific fields (Ahrens and Zaščerinska, 2016).

Beginning in 1960 Gilbert presented the first model in telecommunications which emphasized that bit errors occurred in bundles or, in other words, bursts (Gilbert, 1960; Elliott, 1963). Since then, the issues of a general procedure to evaluate the performance or, in other words, e-business process in the present research, as well as a basic set of parameters or, in other words, criteria, are still relevant today.

In business including e-business, burstiness of workload is traditionally analyzed (Heinrich, 2014). However, the paradigm has changed from an input based business process or, in other words, burstiness of workload to an outcome based process or, in other words, burstiness of buyers (Ahrens et al., 2015). The shift from analysis of burstiness of workload to evaluation of burstiness of buyers allows increasing the efficiency of e-business process and, consequently, e-business profit.

In e-Business, buyers' burstiness has to be highlighted as such a condition. Buyers' burstiness in e-business reflects the real environment in e-business. By phenomenon's burstiness, intervals of high-activity alternating with long low-activity periods are meant.

A new model for analyzing buyers burstiness in e-business process was presented in (Ahrens and Zaščerinska, 2016). For the design of a mathematical

Ahrens, A. and Zaščerinska, J.
Analysing Buyers' Burstiness in E-Business: Parameter Estimation and Practical Applications.
DOI: 10.5220/0006407800710077
In *Proceedings of the 14th International Joint Conference on e-Business and Telecommunications (ICETE 2017) - Volume 2: ICE-B*, pages 71-77
ISBN: 978-989-758-257-8

Table 1: Burstiness in different scientific fields.

Scientific field	Phenomenon of burstiness
Telecommunications	Burstiness of bit-errors in data transmission
Economics	Burstiness of crises
Natural sciences	Burstiness of disasters or earthquakes
Logistics	Burstiness of traffic
Social media	Burstiness of hot topic, keyword or event
Business	Burstiness of workload
E-Business	Burstiness of buyers

model for evaluation of buyers' burstiness in e-business process, the synergy between e-business and telecommunications is promoted as the phenomenon of customers in the e-business process as well as bit-errors in data transmission appear to be of a similar nature, namely, the bursty nature. It should be noted that the present research is not limited to only two scientific disciplines, namely e-business and telecommunication, but is based on a number of scientific disciplines such as business, social media, logistics, literature, etc. Such mathematical models that consider the bursty nature of bit-errors in data transmission have been successfully implemented in telecommunications for optimizing data communication protocols and will be adopted in this work to the buyers' burstiness in e-business process. The proposed model for analyzing buyers burstiness in e-business process is able to take the buyers' concentration into account since the buyers' concentration cannot be considered when only analyzing the buyers' probability (Ahrens and Zaščerinska, 2016). Thus, this model is novel as it offers two parameters, namely, the buyers' probability and the buyers' concentration, for analyzing buyers burstiness in e-business process.

A comparison of the model of evaluation of burstiness of hot topic, keyword, etc. in social media shown by the group of Japanese researchers (Kotozaki et al., 2015) with the model for evaluation of buyers' burstiness in e-business process is reflected in Table 2. The comparative analysis of Table 2 reveals that Kleinberg's burst detection algorithm, which is based on a queuing theory, is built on a sequence of phenomena while gap distribution function is featured by sequential independence of gaps between two buyers. The comparative analysis assists in concluding that e-business process is characterized by sequential independence of gaps between two buyers.

However, the model has to be validated in terms of input parameter values and distributions as the model was already evaluated by experts (Ahrens et al., 2016b). The novelty of this contribution is that the model validity is checked by analyzing the visitor/buyer relationship in two different e-shops. Whereas the first e-shop is focused on selling kitchen furniture, the second e-shop is specialized in selling everyday products.

The research question is as follows: What are practically relevant input parameter values and distributions of the model based on gap processes for the analysis of buyers' burstiness in e-business process? The aim of the research is to validate the simulation model based on gap processes for the analysis of buyers' burstiness in e-business process in terms of input parameter values and distributions underpinning elaboration of a new research question on the model validity. The meaning of the key concepts of validation, model validation and model validation approach is studied. Moreover, the analysis demonstrates how the key concepts are related to the idea of e-business process.

The remaining part of this paper is organized as follows: Section 2 introduces buyers' burstiness in e-business process. Here, a mathematical model for evaluation of buyers' burstiness in e-business process via gap processes is presented. In Section 3 the conceptual framework on modal validation is introduced. The associated results of an empirical study will be discussed in Section 4. Finally, some concluding remarks are provided in Section 5.

2 MODEL DESCRIPTION

The present part of the paper reveals the simulation model based on gap processes for the analysis of buyers' burstiness in e-business process with the focus on the description of buyers' burstiness in e-business process and estimation of parameters of buyers' burstiness.

Table 2: Comparison of models for evaluation of burstiness in social media and e-business process.

Model's Element	Social Media	E-Business Process
Feature	Sequence of batched georeferenced documents	Sequential independence of gaps between two buyers
Approach	Kleinberg's burst detection algorithm based on queuing theory	Gap distribution function

2.1 Buyers' Burstiness in E-Business Process

By binary customer behaviour, *to buy, or not to buy* is meant as depicted in Figure 1 (Ahrens et al., 2015; Ahrens and Zaščerinska, 2016). It should be noted that, in the present contribution, the terms customer and buyer are used synonymously. Within the binary decision (*to buy, or not to buy*) paradigm, the e-business process such as selling or buying is a success if it finishes with a deal such as a sale or a purchase (Ahrens et al., 2016c). Gap in the present contribution means the buying process which ends without a purchase (Ahrens et al., 2016c). Buyers' burstiness is a feature of e-business process. By buyers' burstiness, intervals of buyers' high-activity alternating with long low-activity periods within a fat-tailed inter-event time distribution is meant (adopted from (Karsai et al., 2012)).

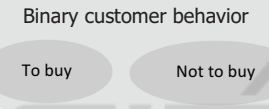

Figure 1: Elements of customers' binary option.

An efficient modelling of buyers' burstiness in e-business processes requires that the characteristic variables are considered with a given precision. Models have to take both – the buyers' probability p_e as well as the concentration of the buyers – into account as pointed in Table 3.

Table 3: Characteristics of buyers' burstiness.

Characteristics	
Buyers' probability	Buyers' concentration
p_e	$(1-\alpha)$

Further on, the levels of buyers' burstiness are summarized in Table 4.

Models based on gap processes allow a realistic evaluation of buyers' burstiness in e-business process. Fig. 2 illustrates an e-business process between buyers described by gaps. In (Ahrens et al., 2015; Ahrens and Zaščerinska, 2016), the e-business process was defined by a buyers-gap distribution function $u(k)$ defining the probability of a gap larger than k, i. e.

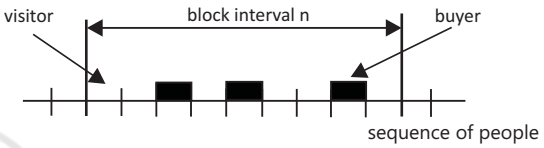

Figure 2: Buyers' gap for describing binary customer behavior.

$$u(k) = P(X \geq k) \quad . \tag{1}$$

For the buyers-gap distribution function $u(k)$ the following expression was identified

$$u(k) = ((k+1)^{\alpha} - k^{\alpha}) \cdot e^{-\beta \cdot k} \quad 0 \leq k \leq \infty \tag{2}$$

with

$$\lim_{k \to \infty} e^{-\beta \cdot k} = 0 \qquad \beta > 0 \tag{3}$$

and

$$\beta \approx {p_e}^{1/\alpha} \quad . \tag{4}$$

The e-business process, i. e. the buyers' characteristics, is modeled by two parameters, namely buyers' probability p_e and the buyers' concentration $(1-\alpha)$. Re-writing of $u(k)$ leads to the buyers-gap density function $v(k)$, i. e.

$$v(k) = P(X = k) \quad , \tag{5}$$

which describes the probability of a gap X equal to k. Using (1), buyers-gap density function $v(k)$ can be calculated as follows

$$\begin{aligned} u(k) &= v(k) + v(k+1) + v(k+2) + \cdots \\ u(k+1) &= v(k+1) + v(k+2) + \cdots \quad . \end{aligned}$$

Finally, by calculating the difference between $u(k)$ and $u(k+1)$ the buyers-gap density function $v(k) = P(X = k)$ can be obtained

$$v(k) = u(k) - u(k+1) \quad . \tag{6}$$

Table 4: Levels of buyers' burstiness $(1-\alpha)$.

L1	L2	L3	L4	L5
very low	low	average	high	very high
$0.00-0.10$	$0.11-0.20$	$0.21-0.30$	$0.31-0.40$	>0.41

The stochastic nature of the e-business process is defined by the buyers-gap density function $v(k)$ or the buyers-gap distribution function $u(k)$, which depends on the buyers' probability p_e and the buyers' concentration $(1-\alpha)$. Figure 3 illustrates the buyers' burstiness within a sequence of e-shop visitors, which can be obtained by analyzing (1) for given parameters of p_e and $(1-\alpha)$.

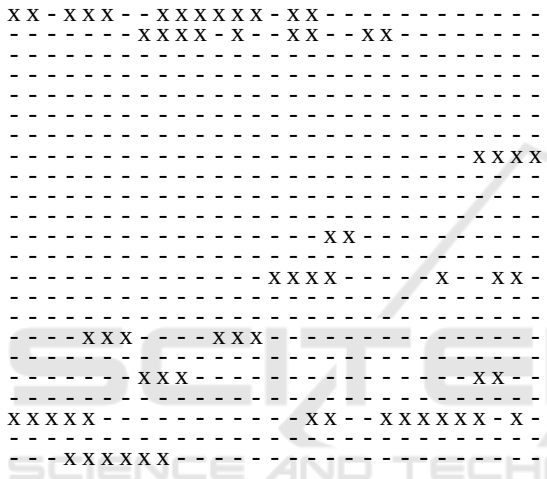

Figure 3: Buyers' burstiness (represented by "x") within a sequence of visitors (represented by "-").

2.2 Buyers' Burstiness Parameter Estimation

Next to the generation of data streams which describe the bursty nature of buyers (see also Fig. 3), it is also essential to be able to analyze a given set of data for the buyers' probability p_e and buyers' concentration $(1-\alpha)$. In the following section the estimation of p_e and $(1-\alpha)$ are shown based on a given set of collected data.

By analysing the probability $\mathbf{P}(E)$ of a pattern E within an interval of n visitors, information about the distribution of the buyers can be obtained. Assuming that the gaps between buyers are statistically independent the probability $\mathbf{P}(E)$ of a pattern E in an interval of n visitors with e buyers at the positions $n_1, n_2, \cdots, n_e$ can be obtained

$$\mathbf{P}(E) = p_e \cdot u(n_1 - 1) \cdot u(n - n_e) \cdot \prod_{v=2}^{e} v(n_v - n_{v-1} - 1) \ . \quad (7)$$

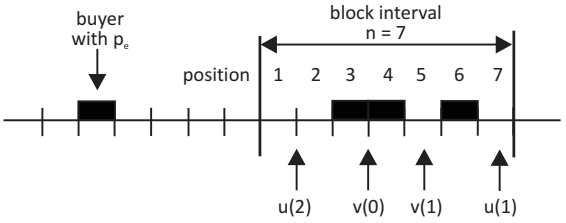

Figure 4: Calculation of the probability $\mathbf{P}(E)$ of a pattern E within an interval of $n = 7$ visitor with $e = 3$ buyers at the positions $n_1 = 3$, $n_2 = 4$ and $n_3 = 6$.

Figure 4 illustrates the calculation of the probability $\mathbf{P}(E)$ of a pattern E within an interval of $n = 7$ visitors. Here, the $e = 3$ buyers are at the positions $n_1 = 3$, $n_2 = 4$ and $n_3 = 6$. The probability $\mathbf{P}(E)$ of such a pattern E within an interval of $n = 7$ visitors is given by

$$\mathbf{P}(E) = p_e \cdot u(2) \cdot v(0) \cdot v(1) \cdot u(1) \ . \quad (8)$$

For $n = 2$ one gets:

$$\mathbf{P}(E) = p_e \cdot v(0) \quad (9)$$

By analysing a captured data stream it is possible to determine the values $\mathbf{P}(E)$ and p_e for $n = 2$. The probability $\mathbf{P}(E)$ can be obtained as follows

$$\mathbf{P}(E) = \frac{\mathbf{E}\{\text{number of neighbouring buyers}\}}{\mathbf{E}\{\text{number of visitors}\}} \quad (10)$$

with the parameter $\mathbf{E}\{\cdot\}$ denoting the expectation functional. The number of neighbouring buyers are counted when after a buyer immediately the next buyer appears, i. e. the distance k between two buyers is $k = 0$ (neighbouring buyers). The buyers' probability p_e is given by

$$p_e = \frac{\mathbf{E}\{\text{number of buyers}\}}{\mathbf{E}\{\text{number of visitors}\}} \ . \quad (11)$$

Combining (9), (10) and (11) we get the probability $\mathbf{P}(E)$ for $n = 2$:

$$\mathbf{P}(E) = \frac{\mathbf{E}\{\text{number of neighbouring buyers}\}}{\mathbf{E}\{\text{number of buyers}\}} \ . \quad (12)$$

Taking the buyers-gap distribution function $u(k)$ into consideration

$$u(k) = ((k+1)^{\alpha} - k^{\alpha}) \cdot e^{-\beta \cdot k} \quad 0 \le k \le \infty \quad (13)$$

the buyers' concentration $(1-\alpha)$ can be calculated. By analyzing the buyers-gap density function $v(k) =$

$P(X = k)$, which describes the probability of a gap X equals to k we get

$$v(k) = u(k) - u(k+1) \ . \tag{14}$$

Analyzing the parameter $v(0)$

$$v(0) = u(0) - u(1) \tag{15}$$

the buyers' concentration $(1-\alpha)$ can be estimated. Analyzing

$$u(0) = 1 \tag{16}$$

and

$$u(1) = (2^{\alpha} - 1)\,\mathrm{e}^{-\beta} \ . \tag{17}$$

with the assumption

$$\mathrm{e}^{-\beta} \approx 1 \quad \text{for} \quad \beta \ll 1 \ , \tag{18}$$

we get

$$\frac{\mathbf{P}(E)}{p_{\mathrm{e}}} = v(0) \approx 2 - 2^{\alpha} \ . \tag{19}$$

From this equation, the buyers' concentration $(1-\alpha)$ is given as

$$(1-\alpha) \approx 1 - \log_2 [2 - v(0)] \ . \tag{20}$$

Practically, analysing a captured sequence for the buyers' concentration $(1-\alpha)$ requires the calculation of the parameter $v(0)$. Here, the equation

$$v(0) = 1 - u(1) \tag{21}$$

has to be analysed for a given data sequence. The parameter $u(1)$ describes the probability for buyers' gaps larger than 1, i. e.

$$u(1) = P(X \geq 1) \ . \tag{22}$$

Let us assume that a given sequence contains 9807 buyers. After analysing the buyers' gaps it was found out that 8247 gaps with $k \geq 1$ are within the sequence. Following, we have 1560 situations where after a buyer in the distance of $k = 0$ another buyer appears. Here we get

$$v(0) = 1 - \frac{8247}{9807} = 0{,}159 \ . \tag{23}$$

Finally, in this sequence we can expect a buyers' concentration $(1-\alpha)$ of

$$(1-\alpha) \approx 1 - \log_2 [2 - 0{,}159] = 0{,}1195 \ , \tag{24}$$

concluding, that the calculated buyers' concentration $(1-\alpha)$ shows a low level of burstiness.

3 CONCEPTUAL FRAMEWORK

No model can be accepted unless it has passed the tests of validation, since the procedure of validation is vital to ascertain the credibility of the model (Martis, 2006).

Validation is the task of demonstrating that the model is a reasonable representation of the actual system: that it reproduces system behaviour with enough fidelity to satisfy analysis objectives (Govindarajan, 2014).

A model is usually developed to analyse a particular problem and may therefore represent different parts of the system at different levels of abstraction (Govindarajan, 2014). As a result, the model may have different levels of validity for different parts of the system across the full spectrum of system behavior (Govindarajan, 2014). For most models there are three separate aspects which should be considered during model validation (Govindarajan, 2014):

- Assumptions,
- Input parameter values and distributions and
- Output values and conclusions.

The assumptions of the simulation model based on gap processes for the analysis of buyers' burstiness in e-business process were validated (Ahrens et al., 2016a). Experts positively evaluated the presented simulation model (Ahrens et al., 2016b). Consequently, the present contribution will concentrate on such a model validation as input parameter values and distributions. Real system measurements will be carried out of three approaches to model validation (Govindarajan, 2014):

- Expert intuition,
- Real system measurements and
- Theoretical results/analysis.

It should be noted that any combination of the three approaches may be applied as appropriate to the different aspects of a particular model (Govindarajan, 2014).

Real system measurements mean comparison with a real system that is the most reliable and preferred way to validate a simulation model (Govindarajan, 2014). In practice, however, this is often infeasible because the measurements would be too expensive to carry out as assumptions, input values, output values, workloads, configurations and system behaviour should all be compared with those observed in the real world (Govindarajan, 2014).

4 PRACTICAL APPLICATIONS

For data analysis the visitor/buyer relationship was captured in two e-shops as highlighted in Table 5.

Table 5: E-Shop's characteristics for analysing visitor and buyer relationship.

E-Shop Number	Shops Specialization
1	selling kitchen furniture
2	selling everyday products

E-shop 1 is specialized in selling kitchen furniture. Here a low level of buyers' concentration $(1-\alpha)$ as well as a low buyers' probability p_e is expected. Whereas e-shop 2 sells everyday products, here people visit with a high wish to buy. That is why a higher level of buyers' concentration and buyers' probability is assumed. The data are captured within 10 minutes intervals. A "x" appeared when at least one buyer was registered within the 10 minutes interval as revealed in Fig. 5 and 6. The data were collected for two month.

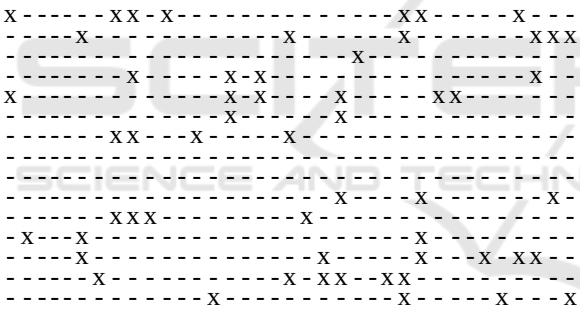

Figure 5: Buyers' distribution (in parts) within E-Shop 1 (buyer represented by "x" and visitor represented by "-").

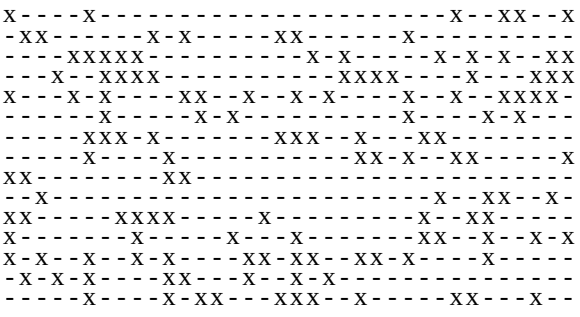

Figure 6: Buyers' distribution (in parts) within E-Shop 2 (buyer represented by "x" and visitor represented by "-").

With the estimated values for the buyers' probability and buyers' concentration it is now possible to calculate the averaged gap-length between two buyers. The averaged gap-length can be calculated as follows

$$\mathbf{E}(k) = \sum_{k=0}^{\infty} k \cdot v(k) = \sum_{k=0}^{\infty} u(k) - 1 \qquad (25)$$

and depends on the buyers' probability and buyers' concentration. The buyers' probability is given by

$$\mathbf{E}(k) + 1 = \frac{1}{p_e} \ . \qquad (26)$$

Data analysis presented in Tab. 6 allows estimating the levels of buyers' probability and buyers' concentration for the investigated e-shops based on Tab. 4. Tab. 7 presents the levels of the buyers' concentration for the investigated e-shops: In e-shop 1 the buyers' concentration refers to Level 1 or the very low level of buyers' burstiness and in e-shop 2 the buyers' concentration indicates level 2 or the low level of buyers' burstiness.

Table 6: Estimated levels of buyers' probability and buyers' concentration for the investigated E-Shops.

E-Shop	Buyers' probability p_e	Buyers' concentration $(1-\alpha)$
1	0,091	0,039
2	0,252	0,187

Table 7: Estimated levels of buyers' concentration for the investigated e-shops.

Shop	Levels of buyers' concentration
1	L1
2	L2

Summarizing content analysis (Mayring, 2004) of the data reveals that the buyers' burstiness in

- E-Shop 1 (the furniture shop) is of the very low level and
- E-Shop 2 (the everyday product shop) is of the low level.

5 CONCLUSIONS

The practical applications of the simulation model based on gap processes for the analysis of buyers' burstiness in e-business process allow drawing a conclusion that the model is valid.

The empirical findings of the research allow drawing the conclusions on the levels of buyers' probability and buyers' concentration for the two investigated e-shops.

The following research question has been formulated: Is the simulation model based on gap processes

for the analysis of buyers' burstiness in e-business process valid in terms of output values?

The present research has limitations. The interconnections between *e-business process, binary customer behaviour, the buyers' burstiness* and *gap processes* have been set. The study is also limited by

- Considering only one aspect of model validation such as input parameter values and distributions and
- Examining the model validation in terms of input parameter values and distributions only through one approach, namely real system measurements.

Another limitation is the empirical study based on two cases only. Therein, the results of the study cannot be representative for the whole area. Nevertheless, the results of the research, namely practical application of the mathematical model for evaluation of buyers' burstiness in e-business process based on gap processes and its results, may be used as a basis of analysis of buyers' burstiness in e-business process. If the results of other cases had been available for analysis, different results could have been attained. There is a possibility to continue the study.

Further research tends to facilitate the practical applications of the validated simulation model for evaluation of buyers' burstiness in e-business process. Despite that initial validation attempts will concentrate on the output of the model, and only if that validation suggests a problem more detailed validation will be undertaken (Govindarajan, 2014), the search for relevant methods, tools and techniques for further model validation is proposed. Consideration of output values is planned. Use of the combination of the three approaches is intended for application in order to validate the simulation model (Govindarajan, 2014):

- Expert intuition,
- Real system measurements and
- Theoretical results/analysis.

A comparative research on validation of simulation models for evaluation of burstiness in other scientific fields could be carried out, too.

REFERENCES

Ahrens, A., Purvinis, O., Zaščerinska, J., and Andreeva, N. (2015). Gap Processes for Modelling Binary Customer Behavior. In *8th International Conference on Engineering and Business Education*, pages 8–13, Ostfold University College, Fredrikstad (Norway).

Ahrens, A., Purvinis, O., Zaščerinska, J., and Andreeva, N. (2016a). Education for Health Society: Indicators of Burstiness in Research. *Education in a Changing Society*, 1:28–40.

Ahrens, A., Purvinis, O., Zaščerinska, J., and Andreeva, N. (2016b). Experts' Evaluation of the Model for Analysis of Buyers' Burstiness in E-Business. In *4th Annual International Scientific-Practical Application Conference Trends in Science and Higher Education Studies Under Conditions of Globalization.*, Panevezys (Lithuania). (accepted for publication).

Ahrens, A. and Zaščerinska, J. (2016). Gap Processes for Analysing Buyers Burstiness in E-Business Process. In *International Conference on e-Business (ICE-B)*, Lisbon (Portugal).

Ahrens, A., Zaščerinska, J., Purvinis, O., and Andreeva, N. (2016c). A Model for Simulation of Binary Customer Behaviour in a Bursty Business Process. In *International Forum on Innovative Technologies and Management for Sustainability.*, Panevezys (Lithuania).

Elliott, E. O. (1963). Estimates of Error Rates for Codes on Burst-Noise Channels. *Bell System Technical Journal*, 42(5):1977–1997.

Gilbert, E. N. (1960). Capacity of a Burst-Noise Channel. *Bell System Technical Journal*, 39:1253–1265.

Govindarajan, M. (2014). *Decision Making Methods.* IGI Global, New York, wang, j.: encyclopedia of business analytics and optimization edition.

Heinrich, R. (2014). *Aligning Business Processes and Information Systems: New Approaches to Continuous Quality Engineering.* Springer, New York, 1 edition.

Karsai, M., Kaski, K., Barabasi, A., and Kertesz, J. (2012). Universal Features of correlated Bursty Behavior. *Scientific Reports*.

Kotozaki, K., Tamura, K., and Kitakami, H. (2015). Identifying Local Burstiness in a Sequence of Batched Georeferenced Documents. *International Journal of Electronic Commerce Studies*, 6(2):269–288.

Martis, M. S. (2006). Validation of Simulation Based Models: A Theoretical Outlook. *Electronic Journal of Business Research Methods*, 4(1):39–46.

Mayring, P. (2004). Qualitative Content Analysis. In *A Companion to Qualitative Research*, pages 266–269. SAGE, Glasgow (UK).