"System evaluates system": method for evaluating the efficiency of IS

Dita Blazkova

Faculty of Business and Management
Brno University of Technology

blazkovad@fbm.vutbr.cz

DOI: 10.20470/jsi.v7i4.278

Abstract: *In paper I deal with the possible solution of evaluating the efficiency of information systems in companies. The large number of existing methods used to address the efficiency of information systems is dependent on the subjective responses of the user that may distort output evaluation. Therefore, I propose a method that eliminates the subjective opinion of a user as the primary data source.*

Applications, which I suggests as part of the method, collects relevant data. In this paper I describe the application in detail. This is a follow-on program on any system that runs parallel with it. The program automatically collects data for evaluation. Data include mainly time data, positions the mouse cursor, printScreens, i-grams of previous, etc. I propose a method of evaluation of the data, which identifies the degree of the friendliness of the information system to the user. Thus, the output of the method is the conclusion whether users, who work with the information system, can handle effectively work with it.

Keywords: information system, evaluation, efficiency, effectiveness, method, functionality, user interface, application

1. Introduction

Is an information system (IS) effective? This question hides countless other questions. The efficiency level is the general output of the IS evaluation process. Efficiency may be defined as such utilisation of economic sources that brings maximum level of satisfaction achievable with the given inputs and technology (Samuelson, Nordhaus, 1995, s. 768). Generally the efficiency level is the ratio between the benefits of a certain activity and the costs incurred for its implementation (Koch, Dovrtěl, Hrůza, Neničková, 2010, p. 127).

Drucker (1993) extended the term by dividing it into two: Efficiency and Effectiveness. The relationship between them was shown by him through their impact on business prosperity. Its explanation may be applied generally, though, for example to IS evaluation. This is how efficiency also acquired its definition: the relationship between the goals and their impact. (Koch et al., 2010, p. 127).

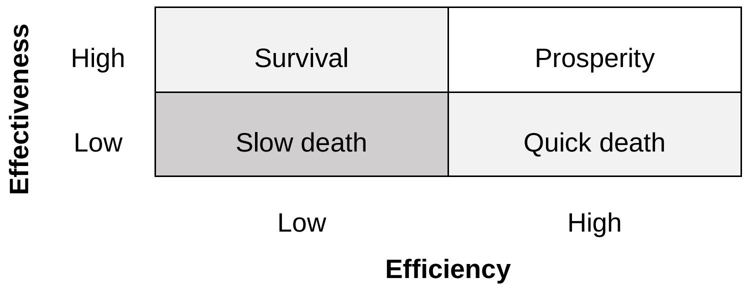


Fig. 1: Efficiency versus Effectiveness in effect on business prosperity; *source: Drucker, 1993*

Efficiency may be described as "doing things correctly", unlike Effectiveness defined as expediency, i.e. "doing correct things". These seemingly similar terms hide a big difference. Using an exaggerated example, the task to build a pool can be implemented with low Efficiency with the help of dynamite. On the other hand in the case of low Effectiveness the pool will be built on top of an inaccesible mountain (Koch). In the IS area it is for example ineffective to take a log time and a lot of labour to search for the required data and the empoloyee will face low effectiveness if the IS fails to provide the relevant information.

IS is a subset of the Business system. The purpose of IS is to provide correct information in the correct place and at the correct time. The Business system is the actual enterprising with all its related components (people, activties, goals, intentions, technical means, information etc.). The question of efficiency/effectiveness of IS addresses the issue of compatibility and compliance between IS and the business (Bruckner, Voříšek, Buchalcevová, Stanovská, Chlapek, Řepa, 2012, p. 14-17).

And why should IS be efficient? For the user to be able to use its function. The purpose of IS deployment in a cmpany follows from the joint study of Economist Intelligence Unit and IBM Global Services. In the context of the research the managers listed their priorities in IS performance – see Figure 2. (Ruščáková, 1999)

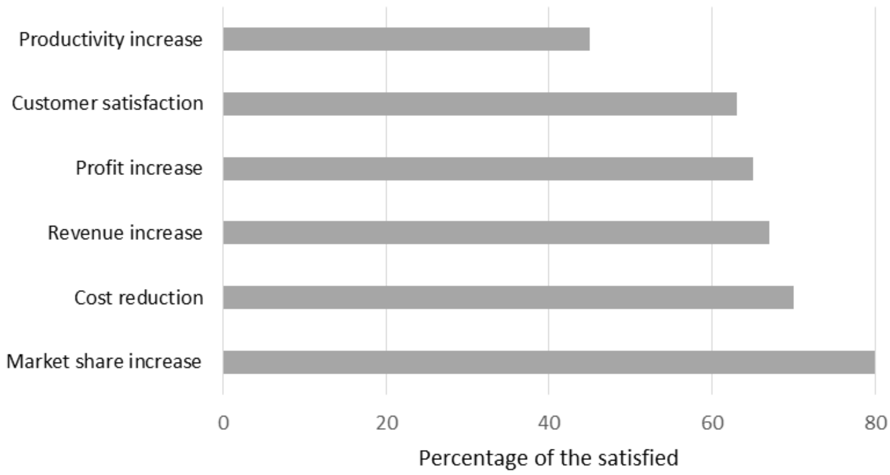


Fig. 2: Satisfaction with individual benefits of IS; *source: Molnár, 2000*

The above diagram shows the purpose of IS from the managerial point of view. IS are not only used by managers, though. Molnár (2000, p. 17) distinguishes 4 categories of IS users and defines their expectations as follows:

- owners – permanent valuation of their invested assets,
- managers – assistance in successful company management with minimum resources,
- employees – better work environment, higher socials tatus, increased feeling of affiliation with the company,
- customer – acquisition of a product/service with higher added value at affordable price.

I would personally extend employee expectations with improved morals for in the case of ineffective or user unfriendly IS the employee´s morals decrease with every use. Often an employee (especially the older generations) is unable to “make friends” with the corporate IS, which leads fo fluctuations for the employees are usually not positioned to affect change of the introduced system.

There are generally two approaches to data acquisition for Efficiency/Effectiveness evaluation. The first mines user data and inevitably bears the risk of human factor. An example may be the countless evaluation algorithms often based on the questionnaire research method, or less frequently on interview or observation. In the case of all these methods the output faces a high level of subjective bias, whether on the part of the user or on the part of the assessor.

The other approach is based on automated data collection. Even in this case non-measurable effects must be taken into account. For example if the application measures time spent by compulter work it still cannot distinguish cases of thoughtless staring at the monitor, leaving the computer for a time or study of data on the screen without mouse movement.

The existing systems evaluating Effectiveness for example include the RescueTime application, detecting active applications and creating statistics of time spent on computer by work, entertainment etc. on the basis of the application sorting. The results are processed and displayed on a web nterface where the computer time is assessed.

Another representative of the evaluation systems focuses on evaluation of web site efficiency and is called ClickTale (or others like LuckyOrange, Clixpy, MouseFlow, Mpathy, etc.). The system records click locations, mouse movements and scrolled areas on the screen. The obtained data are used for creation of "heatmaps" where the frequency of phenomena is depicted with corresponding colours. Apart from this metrics the application accumulated time and fact data on web forms, errors, links and site leaving.

The following text deals with automated data acquisition for the purpose of evaluation of IS Efficiency, basically using certain standard metrics, some extended and others added. Their identification and description is based on standard principles (Bruckner et al., 2012, p. 137), based on personal observations.

The chosen metrics do not address complex evaluation of Efficiency of IS but only part of it: The extent to which the IS is user-friendly, or how the set of users working with the IS can cope with it. The proposed method therefore only provides data and algorithms for evaluation of the metrics. The Efficiency scale (specification of the limit from which the IS begins to be efficient) is outside the scope of this article.

2. Auxiliary Application for Data Collection

For successful application of the method input data collection software is needed to monitor:

- Mouse clicks (number, frequency, location),
- Mouse movements (occurrence, direction, trajectory)
- System routes,
- Screen focus (time, frequency),
- Time in the system,
- Forms (speed of processing, frequency of repeated values),
- Detection of user behaviour characteristics,
- Repeated batches (i-grams of statuses and activities).

The application would be set to be automatic and start spontaneously after installation on the operating system booting and run on the background without disturbing the user. The application would recognize automatically when the IS window in inactive and suspend data collection. The application would not collect data outside the IS window coordinates.

The method is primarily conceived for complex offline IS, i.e. solutions in the form of a set of (heterogeneous) applications are not supported by this method variant[1]. The only thing the user needs to enter before the first use is identification of the application (IS) to be tested. After that the application would only modestlyinform about being active and continue to be displayed in the notification area only.

For easier data collection from different users it is better for the evaluation to be accessible by log in to a web interface. This assumes that the application sends the collected data to a remote server database. In the sense of performance optimisation auxiliary files would be created on the disc containing the collected data and sent off in batches (in certain intervals or on occurrence of a defined trigger: for example on closingthe underlying IS).

3. Monitored Data

3.1 Mouse Clicks

Mouse clicks are a commonly monitored user activity. This method records time, screen coordinates, screen identification (open part of the IS) and other data about every click.

These data may be used for creation of the generally known "heatmaps" (Freundlieb et al., 2014) showing in colour the most frequent click locations on the given screen (see Picture 2 on the left). After every click the screen needs to be checked for changes: expanded hidden drop-down menu,

[1] *Example of unsupported IS: Excel tables connected with Notepad documents in directory structure and E-mails in Outlook or browser.*

forwarding of the user to another window etc. every such change causes change of the screen identification or suspends click monitoring (for example after expanding a ComboBox[2] click detection will be suspended as long as these are generated inside the coordinates of the expanded element).

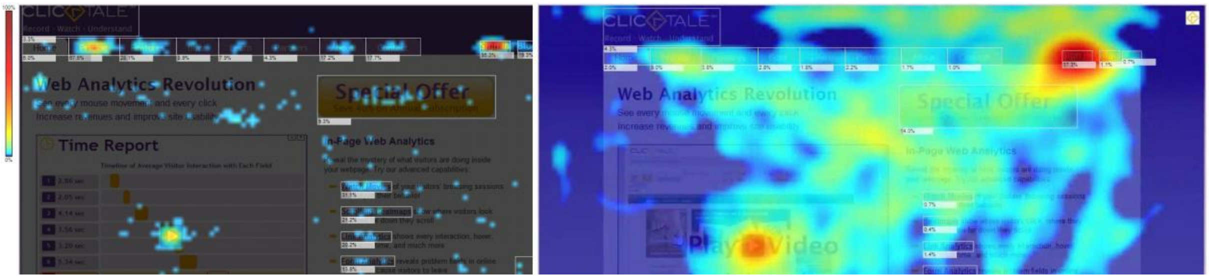
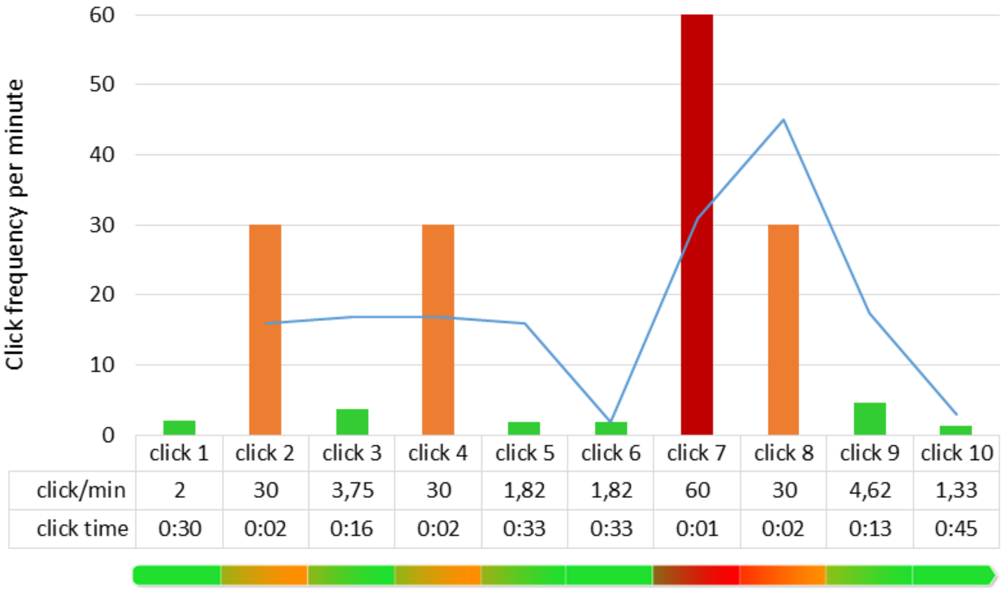


Fig. 3: Heatmap of clicks (on the left) and Heatmap of mouse movements (on the right); *source: ClickTale Ltd, 2014*

In addition to graphic display of click frequency in a point on the screen time data will be compared and a diagram of click frequency per minute will be calculated by the following formula:

$$f(t_i) = \left(\frac{t_i - t_{i-1}}{60}\right)^{-1}, i = 1, 2, \ldots, n \qquad (1)$$

where *t* is the click time and *i* the click index. The result is shown in clicks per minute. The graphic representation of the output values is a bar diagram where higher speed is displayed with dark red colour. The diagram is extended with a trend line (moving average of two values).



	click 1	click 2	click 3	click 4	click 5	click 6	click 7	click 8	click 9	click 10
click/min	2	30	3,75	30	1,82	1,82	60	30	4,62	1,33
click time	0:30	0:02	0:16	0:02	0:33	0:33	0:01	0:02	0:13	0:45

Fig. 4: Example of click frequency per minute; source: in-house *(based on sample data)*

Further information mined form the monitored data include comparison of the numbers of clicks in the individual screens and compilation of a diagram of development of numbers of clicks in the course of work with the IS (number of clicks per hour of day/night).

However, heatmaps can rarely stand on their own. (Jacko, 2009, p. 37) Therefore, other methods are necessary to use with them.

[2] *ComboBox is form control. It is called "drop-down list" too.*

3.2 Mouse Movements

Graphical representation are created like in the case of clicks (see Picture 2 on the right). (Leivaa, Huang, 2015)

For individual screens the mean speed of mouse movement between clicks or stops is calculated for all users in sum and for individual user groups. In the latter casethe speed data are sorted into three groups pursuant to the quantiles specified on the basis of the highest the and lowest speeds. The resultng speed data show separately the slowest, the quickest users and the users working in average speed.

Mouse movement on the screen is expressed as vector with speed and direction monitoring. The most frequent mouse movements are shown for parts of the screen and the given directions with 45° tolerance. The output diagram resembles schematic depictions of water or air streams.

High values of the directing trajectory are taken in this section as a critical manifestation of inefficiency. Value increase is caused by three types of user behaviour:

1) sudden slow down of the mouse movement with the movement implemented within a relatively small area around a control – usually in the spiral shape.
2) Empty click or repeated clicks with zero effect (without any element activation).
3) Slow approaching an object and second waiting for the help bubble.

These phenomena mainly occur in the case of extremely small controls, elements with vague demarcation of the functional area and elements with unclear function.

3.3 System Routes

Clicks in IS may cause, inter alia, transfer to another screen. Each of these screens forms a node in which the user spends a certain period of time. Theseroutes start on the homepage on the system strt up and end by the IS closing. Indexed sequence of these nodes can be used for findling redundant (unused) routes, the most frequently used routes and mapping of ineffective system browsing.

The method of data evaluation based on this metric is detailed in the chapter entitled Repeated Batches.

The evaluation focuses on screens, modules or windows. The structure of these phenomena in the mathematical set terminology is as follows: A = {modules}, B = {windows}, C = {screens}, with $A \subseteq B \subseteq C$.

One of the IS evaluations is conceived for the window level where the application automatically maps all tunnels in the system opening new windows, creating an oriented network diagram from them. Without connection to user behaviour in the IS the diagram calculates all circles (minimum closed sub-diagrams) and all available paths and sequences (a path means a sequence consisting of unrepeated nodes while the other sequences can include repeated nodes). The diagram shows system transparency and throughput, in addition pointing out redundant connections.

This initial diagram is then compared to one created on the basis of the user browsing the system. The output includes both percentages of use of the available routes and nodes and the number (and list) of the unused nodes, paths and circles.

3.4 Screen Focus

A screen means unique display of a part of IS. This does not always mean different system windows, but even a minor change in the same module (except for expansion of a ComboBox). For change detection there are functional principles for image recognition (such as OpenCV library).

Screen focus means measurements which parts of the screen are most frequently and most extensively displayed. The term "Screen Focus" is called "Attention Heatmap" too. (ClickTale Ltd, 2016) This statistic is just marginal, for most IS do not use scrolling at all, except for data excerpts from databases (which do not produce relevant data for this method). The vertical axis (Y) of data structuring is instead usually extended to the depth of the system (Z axis) – with new windows. Despite this some user interfaces may use scrolling (for example due to poor quality optimisations for different monitor sizes).

3.5 Time in System

Measurement of time in IS with distnction of the visited modules is often used in applications designed for employee controlling. The metric is however also useful for detection of the most frequently/the least used parts of the system and therefore for system redundance for the given user group. Values in absolulte numbers and in percent may for example suggest modules which the employer pays but does not utilise.

The evaluation is against divided to screens, modules and windows. The structure isthe same as the one describes in the System Routes chapter.

3.6 Forms

Forms are separately measured units. The measurements include speed of movement between form items and frequency of value changes within items and overall in the form.

Like in the Mouse Movement chapter where the speed of users was evaluated separately by quantile in this case the users will again be divided to three groups (quick, average and slow) and their speed will be analysed separately.

The numbers of corrections per user rather show their trend towards errors but if you calculate frequency of errors across all users and this frequency exceeds the defined limit it needs to be noted for its cause may be a poorly characterised element.

Further evaluations based on data of this metric are described in the chapter entitled Repeated Batches.

3.7 User Behaviour Characteristics Detection

Users are different. Some hesitate before every button pushing, others hastily pass through the system and often return. Users also differ in the speed of typing and mouse movement. Therefore a unique assessment approach is not appropriate. The most remarkable differences result from time data comparisons. See for example switching between form items. Some users write text chains on the keyboard and use the mouse for switching between items. Others (usually the quicker ones) use tabulators or particular key combinations for the switching.

In the chapter entitled Mouse Movement the users are divided to three groups, slow, average and quick. This classification is based on quantiles specified pursuant to the minimum and maximum value of the speed of all users taken together. The means within these groups bring more information. The same classification will be used for all time comparing metrics.

Another important characterstic feature of user behaviour is recognition of the moment when the user leaves the computer. A user may stop working, leave the current window open (in this case the application collects data as if the user was working) and leave the computer for a couple of minutes. This case is virtually undistinguishable from the group of users not moving the mouse when reading from the screen and only reading by their eye movement.to distinguish between these two groups I suggest a highly unreliable and only informative method: in my experience there are two groups of users distinguished by the way they read from the monitor.

The first group of users move the mouse along the lines like the finger along the line of a printed text. They do this inadvertently and virtually without exceptions. When these users do not click, write or move the mouse they are not by the monitor and the application for data collection can be suspended after several seconds of inactivity (15s).

The remaining users form the second group. Hence if the phenomenon or ”reading with the mouse” is not detected then the appliation may only be suspended after much longer inactivity (3min).

Technically speaking I have also come across a third and a fourth gourp of users who sometimes drop the mouse and point to the monitor withtheir fingers. One of them is a group of advanced users working in specialised graphic programs who are not likely to leave the peripheral devices at rest for a long time. The other are complete beginners who are unfortunately undetectable – on the other hand these users represent a statistic exception in the group of users connected with any professional IS (for whom this method is designed) and therefore they need not be considered.

3.8 Releated Batches

The sequences of some of the abovementioned monitored phenomena can repeat. Like releated sequences of two or three letters are searched in encoded texts with the help of frequency analysis in this case repeated i-item sequences are searched. The items can be represented by system status or user activity.

This analysis is only performed on request for evaluation. For that reason it is different from the abovementioned metrics collecting data on an ongoing basis.

In the complete sequence of items where an item may be represented for example by a visited IS screen i-grams are searched[3]. Characteristics of i-gram include:

- at least two occurrences in the sequence of all items,
- an i-gram is a sequence of items, i.e. consists of at least two items (a digram, i.e. a 2-gram),
- none of the items is repeated in the sequence (in the Diagram Theory terminology thisis not a Sequence but rather a Path).

The highest possible i-gram equals to one half of the total number of items in a sequence.

Search for i-grams also monitors item indexing (and time data) for in the context of an i-gram lower order data may occur whose elimination may in effect increase efficiency in comparison to replacement of just higher order i-grams. Regarding the rule of non-repeatability of items in i-gram even the sub-i-gram cannot occur i an i-gram more than once. Thanks to that sub-i-grams (paths that are as a whole part of another i-gram) occurring in the total sequence in the same number as their superimposed i-grams may be excluded.

Regarding the purpose of i-gram search the times of addition of the individual items in the log are not relevant. The total sequence has a given first and a variable last item. The first item (with the lowest index) will be the one activated after the first start up of the IS. The last item will be the one closing the IS. System closing will be recorded as another item in the sequence with diferentiation between closing with and without saving. Interruption of IS activity for example by computer hibernation is not taken as IS closing. When searching for i-grams these closing items are of special importance. They can only occur at the end of an i-gram. Most IS (with just a few exceptions) open with the same homepage after start up equipped with the same structure of items. Regardless the screen in which the used closed the IS before the new start is always the same. In addition a repeated batch which may be replaced for example with a sigle button may also include IS closing.

I-grams found pursuant to the abovementioned rules will be submitted to the assessor on the basis of item indexation, i.e. the list of the sub -i-grams will be placed under the relevant i-gram. The application will propose to the assessor the most convenient i-grams for elimination, or the assessor will be recommended to unify the sequence into a single node. As soon as one is selected, the total improvement score is calculated together with individual scoresfor each i-gram.

Schematic example

Step 1)

I-grams found:	
sequence:	D → A → B → C → D → E → A → B → C → D → F → A → B → D → E → A
2x 4-gram:	**A → B → C → D**
└ excluded sub-i-grams:	A → B → C, B → C → D, B → C, C → D
└ 3x sub-2-gram:	A → B
2x 3-gram:	**D → E → A**
└ excluded sub-i-grams:	D → E, E → A

[3] *Repeated sequence (batch) with i items (i-order)*

Benefits of i-gram eliminations (total score):

A → B → C → D	50%
D → E → A	37.5%
A → B	37.5%

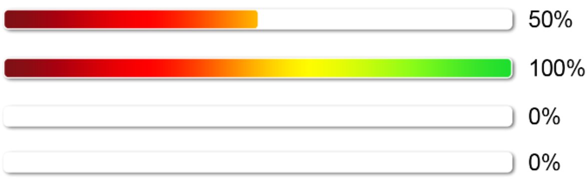
Step 2)

Selected i-gram: A → B

Step 3)

Calculated replacement score (A → B = X):

2x 4-gram:	**X → C → D**
└ 3x sub-2-gram:	X
2x 3-gram:	**D → E → A**
└ excluded sub-i-grams:	D → E, E → A



Step 4)

Recalculation:

sequence:	D → X → C → D → E → X → C → D → F → X → D → E → A
2x 3-gram:	**X → C → D**
└ excluded sub-i-grams:	X → C, C → D
2x 2-gram:	**D → E**

Practical example

Let us say that a back office staff member most frequently uses the customer management module. Thus his activity includes a found tetragram (4-gram) in the following system route: Homepage → Customer management module → Customer list → New customer addition. The system recommends sparing the system browsing and creation of either a form for new customer addition or a button redirecting this particular user group to the depth of the system where the form can be found.

The recommended button may already exist in the system without being used by this particular user group. In this case the assessor must decide whether to improve the user skills (for example by training) or the IS (by highlighting the button, change of the icon, its location or size).

If in the schematic example above the items (capitalised) represented the visited IS screens (or windows) then the assessor would be given a chance to study in detail the most recommended i-grams before selection of the i-gram for elimination. For example from the temporal point of view: The user may show a longer delay between two particular items, which may mean that the user takes information necessary for further proceeding from the screen by looking at it. It would therefore be ineffective to replace all elements in the i-gram by one step without associated measures. Whether the user takes data from the screen may also be found (with a certail level of error tolerance) with the User Behaviour Characteristic Detection metrics.

The very merger of an i-gram into a single node does not reprogram the IS. Let me remind that this method only points to system inefficiencies. Hence before the assessor decides which i-gram to select he should first find out the subject of search by the user: i-gram may just be a transfer path – with the simple solution of replacement of the path with a single button for the path shortening. Whether the end of the i-gram hides a form or a data excerpt a the subject of the search may be found by other monitored phenomena (such as mouse clicks in connection with time separation pursuant to the time data of the i-gram).

Thus selection of the most relevant i-gram as well as the subsequent adaptation of the IS both need the human factor (an informed developer).

The application provides a log of individual users. Thus repeated batches may be monitored both on the user level and on the IS level.

On the user level (with the basic data set represented by a single user data) the number of i-grams is monitored and the recommendation for merger is issued for those i-grams whose number is higher than the value calculated by the following formula:

$$f(i) = \frac{n}{i} \times 0{,}01p \quad (2)$$

Where *f(i)* is the minimum necessary number of i-grams for inclusion among the recommended (*f(i)* is rounded down to whole numbers),

n is the total numberof items in the sequence (for example uploaded screens including duplicities),

i is the multiple of the i-gram (inthe cae of pentagram *i=5*),

p is the sensitivity percentage set by the user – default setting will be 75%.

On the application level (with the basic data set represented by all user data) the number of i-grams is monitored as well but the formulafor the recommended is more plastic including the user dimension:

$$f(i_a) = \frac{n_a}{i} \times 0{,}01p;\ \ a = 1, 2, \ldots, m \quad (3)$$

$$g(i_a) \geq f(i_a);\ \ u \geq m \times 0{,}01p \quad (4)$$

$$h(i) = 0{,}01p \times \left(m^{-1} \times \sum_{a=1}^{m} \frac{n_a}{i}\right) \quad (5)$$

$$h(i) \leq u^{-1} \times \sum_{k=1}^{u} g(i_k) \quad (6)$$

where *f(i)* is the minimum necessary number of i-grams of one user for inclusion among the recommended,

n is the total numberof items in the sequence of one particular user,

i is the multiple of the i-gram,

p is the sensitivity percentage set by the user,

a is one particular user regardless whether an i-gram was found in his sequence,

k are individual users with the recommended i-gram in their sequences,

u is the total number of users with the recommended i-gram in their sequences (i.e. frequency of the i-gram above the set sensitivity percentage),

m is the number of users enrolled in the test,

$g(i_a)$, or $g(i_k)$ corresponds to the number of the found i-grams of one particular user *a*, or *k* – in connection with variable *k* this is the sum of the found i-grams whiche xceed the sensitivity level *p* per user

h(i) is the minimum limit number of i-gram to be achieved for inclusion among the recommended.

Inequations *(4)* and *(6)* condition further progress of the algorithm. Condition *(6)* thus says that the recommended i-grams may only include those whose ratio of frequency and number of users in the selection set (found i-grams) amounts at least to *p* percent of the same ratio in the basic set of i-grams (maximum possible number of i-grams in user sequences that need to be taken separately).

The improvement score (in %) is calculated by the classical principle (see formula 7). For the sake of clarity I will use the distinction between i-gram and a-gram here. I-gram means the originally measured inefficient path. A-gram is a path selected for elimination, i.e. for replacement with a single node.

$$s(i) = \left(1 - \frac{r - (q \times b)}{r}\right) \times 100 \quad (7)$$

where *r* is the number of items in an i-gram (total score: in sequence),

q is the number of items in an a-gram,

b is the number of occurrences of the a-gram in the given i-gram (total score: in sequence),

i is the multiple of the i-gram.

In the score calculation within individual i-grams *b* will always equal to one. In the total score the *r* and *b* quantities do not relate to any i-gram but to the total measured sequence of items.

The Repeated Batch algorithm (i-gram measurements) can be applied with just minor deviations to more of the monitored phenomena described in the chapters above. Thus i-gram items may include:

- clicked buttons outside forms,
- uploaded screens (or activated IS windows),
- values in individual forms (character chains entered in TextBoxes, the selected OptionBoxes, the selected values of ComboBoxes, the valules of ScrollBars, etc.).

The first two item types are very similar but the item set “uploaded screens” represents a subset of the set of “clicked buttons”, as not all buttons cause screen change.

For the third type the high occurrence of i-grams may be resolved by setting (or resetting) of the default values of certain form items, or direct option of “quick fill out” – there may be more of these.

The outputs will display for the assessor as a table with recommended i-grams sorted in the descending order according to frequency and grouped according to the superimposed i-gram. The table will include the sequence (i-gram) identification, absolute frequency, relative frequency, mean frequency per user, a link to more detailed tables including the list of i-gram frequency per user (the number of users with given i-gram frequency; i-gram frequency per user), a link to the site with graphic representations of the most frequent sequences (such as sequences of visited screens in the form of printscreens), and other potential data. For example if an IS screen is an i-gram item, then a link will be aded to a detailed breakdown of the times spent over the individual items of the given i-gram, a heatmap of clicks and a vector map of mouse movements for the individual items etc.

The assessor will be reminded directly in the excerpt about longer delays betwen i-gram items (or about the recorded “mouse reading” – see the chapter “User Behaviour Characteristics Detection”), for these may represent a serious contraindication for selection. On the application level (the basic set are data from all users) the duration times will be averaged.

4. Discussion

The proposed data collection system mines data, evaluates them and turns them into information independently of the subjective judgement of the user or the assessor. The actual evaluation and status finding does not require a specialist and may be performed by anybody familar with the tested IS. In these two aspects the proposed method substantially differs from most IS evaluation systems.

At the same time all of the abovementioned proposed techniques of data collection accumulate data for evaluation by hard metrics. They can be processed quantitatively and the objective stimulus of the decision can be traced back easily.

On the other hand, a combination of this proposed method with one of the questionnaire methods might generate a new approach to IS in the form of the difference between a subjective opinion of the user/assessor and the output evaluation of this method that might point to weak points of the proposed application (orthe questionnaire method) or form another basis for decisions about IS restructuring.

5. Conclusion

The proposed method reads user behaviour and draws conclusions regarding IS efficiency from it. The user, his activity and its motives spring from very complex processes, though, which only allow for very limited mapping on the current knowledge level and with the proposed tools. Metaphorically speaking the assessor will try to estimate the whole 100-piece jig saw from 3 pieces of the puzzle. If you do not want to use costly biometric equipment this method is the most objective affordable evaluation method. Despite this a lot of attention will need to be paid to testing and tuning after completion.

For example the users are only divided to three groups on the basis of quantiles with regard to their speed of work. The testing may however show the need for a more detailed user clustering. Another point for tuning up is the sensitivity percentage for search of the recommended i-grams. The chapter “Default Batches” sets 75% as default but statistical survey may show this limit as unrealistic requiring lowering this percentage down. The evaluation system as such will need further streamlining. The system should be designed sufficiently user-friendly and intuitive to allow the assessor to easily understand the relatively complex data and principles and draw conclusions from them. For example in the course of the whole evaluation partial informative scores will be set up for the individual sections and subsequently the total score will correspond to the level of improvement of the IS efficiency.

Following tuning the data collection application its orientation might be diversified to web IS or IS consisting of a set of heterogeneous applications and a specific evaluation method might be created for each.

Acknowledgment

This paper was supported by grant FP-S-13-2148 ‘The Aplication of ICT and Mathematical Methods in Business Management’ of the Internal Grant Agency at Brno University of Technology.

References

Bruckner, T., Voříšek, J., Buchalcevová, A., Stanovská, I., Chlapek, D., Řepa, V., 2012: *Tvorba informačních systémů: principy, metodiky, architektury*. Praha: Grada

ClickTale Ltd, 2014: *Product Tour* [online]. Available at: http://www.clicktale.com/products/tour [Accessed 2015-03-23]

ClickTale Ltd, 2016: *Heatmaps* [online]. Available at: https://www.clicktale.com/solutions/features/heatmaps/ [Accessed 2016-10-19]

Drucker, P., 1993: *Inovace a podnikavost: Praxe a principy.* 1st ed. Praha: Management Press

Freundlieb M., Gräuler M., Teuteberg F., 2014: A conceptual framework for the quality evaluation of sustainability reports, *Management Research Review*, 37(1), pp. 19 - 44

Jacko, J. A., 2009: *Human-Computer Interaction*. Berlin: Springer.

Koch, M., Dovrtěl, J., Hrůza, T., Neničková, H., 2010: *Management informačních systémů*. 3rd ed. Brno: CERM

Leivaa L. A., Huang J., 2015*: Building a better mousetrap: Compressing mouse cursor activity for web analytics*, *Information Processing & Management*, 51(2), pp. 114 - 129

Molnár, Z., 2000: *Efektivnost informačních systémů.* Praha: Grada

Ruščáková, D., 1999: V čem spočívá a jak je možné zjistit hodnotu IT. *ComputerWorld*, 17/99.

Samuelson, P. A., Nordhaus, W. D., 1995. *Ekonomie:* 2nd ed. Praha: Svoboda.

JEL Classification: M15, M21