

Дослідження мотивації користувача для управління життєвим циклом інформації

Павел Беліков

Факультет комп'ютерних наук, Санкт-Петербурзький державний університет інформаційних технологій, механіки та оптики, Росія, Санкт-Петербург, Кронверкський проспект 49, E-mail: pbelikov@gmail.com

Загальновідомо, що в даний час інформація є одним з найважливіших ресурсів будь-якої організації. У зв'язку з цим, методи, які допомагають у зниженні сукупної вартості володіння цією інформацією, стають все більш і більш актуальними.

Одним з таких методів є управління життєвим циклом інформації, який ґрунтується на концепції багаторівневої інфраструктури зберігання даних.

Життєвий цикл інформації можна визначити як послідовність тимчасових інтервалів, чия інформація проходить від створення до знищення, і кожен з цих часових інтервалів характеризується відповідними властивостями і класом пристроїв зберігання даних.

Основна властивість будь-якої інформації на кожному етапі її життєвого циклу і є вартість. Таким чином, інформація знаходиться на кожному етапі життєвого циклу до тих пір, поки його цінність розташована між визначеними верхніми і нижніми порогоми.

Існують різні методи оцінки інформації. Метод, заснований на оцінці ймовірності доступу користувачів, використовується.

Визначеними є рівняння для оцінки ймовірності мотивованого доступ користувача до даних.

Експеримент проводиться з метою зобразити залежність ймовірності доступу до інформації від кількості мотивованих користувачів.

При визначенні цінності інформації, було помічено наступну особливість: збільшення кількості вмотивованих користувачів з постійною загальною кількістю користувачів в системі приводить до збільшення ймовірності доступу до інформації, і, отже, цінності інформації.

Актуальність цієї роботи полягає у тому, що вона описує підхід до оцінки даних, які базуються виключно на математичних визначеннях і не вимагає жодного керівництва користувача для підтримки управління життєвим циклом інформації механізму. Таким чином, це дозволить знизити суб'єктивність, яка виникає при визначенні та здійсненні стратегії в області інформаційного управління життєвим циклом.

Переклад виконано Малиновською О. А., центр іноземних мов «Universal Talk», www.utalk.com.ua

User motivation research for information lifecycle management

Pavel Belikov

Department of Computer Science, Saint-Petersburg State University of Information Technologies, Mechanics and Optics, RUSSIA, Saint-Petersburg, Kronverkskiy prospekt 49, E-mail: pbelikov@gmail.com

Information is one of the most important resources of organization nowadays. One way to control data storage resources is implementation of lifecycle management concepts. This work is devoted to research of probabilistic method of information lifecycle management. Use of this method will reduce the task of determination of information lifecycle to single-criterion one. The generalized definition of information lifecycle is provided. Value of information is defined as its main property. Influence of user motivation on probability access is depicted.

Keywords – information, data, lifecycle, probability, user, class of storage devices, value, single-criterion task, motivation.

I. Introduction

The goal of this work is to research influence of user motivation on probability of user access to information in data storage systems. The research is conducted in annex to known method of storage organization - information lifecycle management.

II. Information lifecycle

Information lifecycle can be defined as a sequence of consecutive time intervals, which information passes from creation to destruction and each of these time intervals is characterized by the corresponding properties and class of storage devices.

It is logical that at each subsequent stage of lifecycle various classes of storage devices are used, and they are characterized by worse performance, lower security requirements and lower unit cost of storage per unit than on the previous stage.

According to [1], information lifecycle management is “the process of deciding the value of information”. In this case we can say that value is the main property of information.

In this work method, proposed in [2], is used. This method is based on evaluation of probability of user access to data in storage system.

To determine time of information storage, one can assume that the time within which data and information should be stored in storage system, is defined as time, during which the probability of access to this data is different from chosen threshold. As annex to certain stages of lifecycle this hypothesis is as follows: time of storage of data on certain stage of lifecycle is defined as time during which probability of access to this data is in certain interval. Hence, the conclusion is made that task of parameterization of lifecycle can be reduced (in basic environment) to a single-criterion task, which variable is criterion of access probability followed by automated valuation of information.

III. Determining the value of information

Let's assume that we use information system which consists of two sets of users (M_U^I and M_U^II) and two sets of data (M_D^A and M_D^B). Sets of users altogether compose set M_U , and sets of data compose set M_D , as shown in Fig. 1. Users in set $M_U^I \in M_U$ are motivated, meaning that they are the only ones who can access data in set $M_D^A \in M_U$.

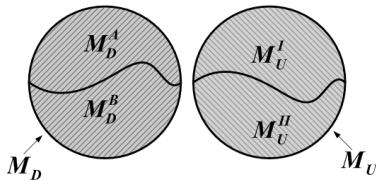


Fig. 1 Sets of data and users.

Probability P is the probability of target event. Target event is access of users in set M_U^I to data of set M_D^A . Let's assume that if the probability P is higher than upper threshold, that would mean the transfer of information to the previous stage of its lifecycle (if applicable), with the probability P less than a certain lower threshold, it is transferred to the next stage of lifecycle.

The more probable access to data of set M_D^A means that all information in system (contained in the whole set M_D) will stay on current level of its lifecycle for a longer period of time.

Probability P is expressed as follows, see Eq. (1):

$$P = P(E_k \cap E_U \cap E_{UI} \cap E_D \cap E_{DA}) \quad (1)$$

In Eq. (1) E_k is event of single access try, E_U - access of one of the users of set M_U , E_{UI} - access of one of the users of set M_U^I , E_D - single unit of data (no matter belonging to set M_D^A or M_D^B) is accessed, E_{DA} - single unit of data in set M_D^A is accessed.

According to rules of probability multiplication [3] this equation is expressed as follows, see Eq. (2):

$$P = P(E_k) \times P(E_U) \times P(E_{UI}) \times P(E_D) \times P(E_{DA}) \quad (2)$$

These probabilities in detail are expressed as follows, see Eq. (3):

$$P = \frac{1}{k} \times \frac{1}{N_U} \times \frac{N_U^I}{N_U} \times \frac{1}{N_D} \times \frac{N_D^A}{N_D} \quad (3)$$

Thus, the probability of access of user of set M_U^I to data in set M_D^A is expressed as follows, see Eq. (4):

$$P = \frac{N_U^I N_D^A}{k N_U N_D^2} \quad (4)$$

In Eq. (4) k is number of access tries, N_U is total amount of users in set M_U , N_U^I is amount of users in set M_U^I , N_D is amount of units of data in set M_D and N_D^A is amount of units of data in set M_D^A .

IV. Experiment

Using Eq. (4), experiment has been conducted. The goal of the experiment was to determine dependence of probability of user access to information against amount of motivated users.

As initial data were chosen: $k=1$, $N_D = \{100\}$, $N_D^A = \{50\}$, $N_U = \{100\}$, $N_U^I = \{10, 25, 50\}$.

Results are presented in Fig. 2.

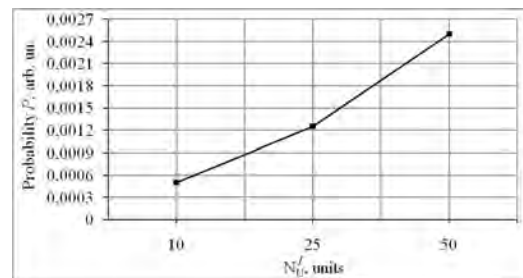


Fig. 2 Dependence of probability of access to information against amount of motivated users.

According to Fig. 2 it is obvious that having larger amount of motivated users (in set M_U^I) with constant total amount of users (set M_U) leads to higher probability of access P and therefore information, which is contained in data of set M_D^A , becomes more valuable.

Conclusion

Further development of probabilistic method of information valuation and its extension with parameters, reflecting legal requirements and methods of value prediction, will greatly simplify the task of information lifecycle management implementation and minimize the subjectivity of such implementation.

References

- [1] Al-Fedaghi S., "On Information Lifecycle Management", IEEE Computer Society, vol. 12, pp. 335-342, 2008.
- [2] Belikov P., "New method of information lifecycle management. Dynamics of information value in databases.", Proceedings of IX International scientific conference New information technologies and systems (NITaS'2010), pp. 223-229, 2010.
- [3] Bronshtein I.N., Semendyaev K.A., "Handbook of mathematics for engineers and students of technical colleges", Nauka, p. 443, 1986.