Data protection in transactional and statistical applications of databases

B. SZAFRAŃSKI, R. BAŁAZY

boleslaw.szafranski@wat.edu.pl, balazyrafal@gmail.com

Military University of Technology, Faculty of Cybernetics
Kaliskiego 2, 00-908 Warsaw, Poland

The article describes a discussion on the issue of data protection in databases. The discussion attempts to answer the question about the possibility of using a transactional database system as a system capable of data protection in a statistical database. The discussion is preceded by a reminder of the basic issues related to data protection in databases, including reminder of flow control models, access control models and the inference. The key element of the article is the analysis, based on the example of the Oracle database management system, whether data protection mechanisms in transactional databases can be effective in case of data protection in statistical databases.

Keywords: database, database system, data protection in database, flow control, inference control, statistical database, Oracle.

DOI: 10.5604/01.3001.0014.4439

1. Transactional and statistical databases − Definitions

In the PWN Encyclopedia [11], the database is defined as "a set of interconnected data, stored in the memory of computers and used by the application programs of an institution or organisation together with software enabling the definition, use and modification of that data".

In the item [1] the authors explain that "a database is in fact nothing other than a collection of data that exists over a long period of time, often for many years".

In turn, under the item [2] the database "is a set of data related to one another. By data we mean in this case known facts which can be somehow recorded".

In this article, a database is understood as a set of data, collected according to specific rules, recorded and stored using existing and available techniques on an electronic data carrier, managed by specialized software called database management system.

The database management system according to [2] is "(...) a set of programs that enable creation and operation of a database. (...) is a universal software system which enables defining, constructing, manipulating and making databases available to various applications and users".

A statistical database system is a system, in which data are collected and processed solely for statistical purposes. In such a system, access to the database is limited only to statistics from a set of acceptable statistics, i.e. those whose disclosure does not cause a loss of confidentiality of data values of individual objects (e.g. people), based on which these statistics were calculated.

A transactional database system is a system, in which data reflecting transactions taking place in a given organisation (e.g. in its management processes), is collected and processed. In such a system, access to the database is limited only by the access rights granted to its users and its purpose may be both access to statistics and data values of individual objects.

Where this will not give rise to doubt, the terms "statistical database" and "transactional database" will respectively be used interchangeably, instead of the terms "statistical database system", "transactional database system".

Different expectations in the way data are accessed make data protection rules for transactional and statistical databases significantly different.

These differences result in a different terminology and a set of problems and issues that make the need for research and work in relation to statistical databases still relevant.

It should also be noted that even a cursory review of available Database Management Systems (DBMS) leads to the conclusion that there are no solutions dedicated to statistical

applications of database technology among the DBMSs offered on the market.

The next part of the article will be entirely devoted to the analysis (based on the example of the most popular Oracle system) of protection mechanisms occurring in transactional applications of database technology in order to assess their usefulness in statistical databases.

This analysis will be preceded by a brief presentation of the classification of data protection methods developed by D. Dennig and published for the first time in the book [3].

2. Data protection in databases

The data protection methods discussed in [3] have been divided into methods concerning the protection of transactional and statistical databases, as well as general purpose methods. The first group includes flow control and access control, the second group includes inference control, the third group includes encryption control.

3. Access control

Access control is the method based on an access matrix model, called the Lampson model or a model of capability lists (authorisations).

The basic elements of the model [4] are:

- a set of active objects called subjects − S,
- a set of passive objects called objects − O,
- a set of operations − T,
- set of privileges (access rights) − P, which is the set of the triplets
  $p_{ij} = (s_i, o_j, t_n)$, where $s_i \in S$, $o_j \in O$, $t_n \in T$, $i = \{1,2,3,..\}$, $j = \{1,2,3,\ldots\}$, $n = \{1,2,3,\ldots\}$.

The set of access privileges is presented in the form of the matrix A, in which columns correspond to objects $\{o_1, o_2, \ldots, o_j\}$ from the set O, rows correspond to subjects $\{s_1, s_2, \ldots, s_i\}$ from the set S, and the element of the matrix A $[o_i, s_j]$ is the p_{ij} access privilege granted to a given subject.

The rule of authorization of any access request in this model is to check whether for a given active entity s_i in matrix A there is (exists) a privilege of access to the object specified in the request, allowing the execution of this request.

4. Flow control

Flow control refers to the rules of data flow control between the objects, distinguished in the model, used in database processing.

The Flow Control Model (FCM) is the following ordered quadruplet:

$$FCM = <O, K, F, R>,$$

where:

- the set of objects – O,
- the set of the confidentiality classes – K,
- the function F that assigns each element of the object set O the confidentiality class from the set K,
- the flow relation – R is defined on pairs of confidentiality classes; i.e. $R \subset K \times K$. The confidentiality classes k_1 and k_2 belong to the relation, i.e., $(k_1, k_2) \in R$, if and only if data from objects with confidentiality class k_1 can flow to objects with confidentiality class k_2.

The rule for authorizing data flows in the FCM model is that data flows between the objects of the set O (including the direction of the flow) are allowed if the confidentiality classes held by these objects form a pair belonging to the relationship R.

In addition to access and flow control for (transactional and statistical) databases, a complementary method is encryption, which is the process of converting data into ciphertext as a way to protect data confidentiality. Data encryption is complementary to other data protection methods and is not subject to further analysis in the article.

5. Inference control

The inference control according to [4] is to ensure that: "access to data in a statistical database is limited to making data available only in the form of statistics to those, which do not lead, directly or indirectly, to the disclosure of data relating to individual identified information objects in the database (e.g. a particular person, transaction or entity, ...)".

Inference control is the essential element of data protection in the statistical database. The problem is the lack of dedicated solutions in available database management systems that can provide application control mechanisms.

The key question is therefore whether, and how, data protection mechanisms built-in in commercially available database management systems can be used to protect the statistical databases managed by these systems.

6. Data protection mechanisms in Oracle system versus protection of statistical databases

In Oracle system[1] data protection mechanisms are classified as direct and indirect data protection mechanisms.

Direct data protection mechanisms are the mechanisms that secure access to objects storing or processing data or making data available.

Indirect data protection mechanisms complement the database protection at the level of the database system. These mechanisms include database access management, data encryption management, data masking management and database monitoring management (e.g. Oracle Database Vault, Oracle Data Masking, Oracle Audit Vault, etc.).

It should be noted that the data protection mechanisms identified and described below are designed to protect transactional databases (this is the main application area of the Oracle system) rather than statistical databases.

Analyzing the sources on the Oracle system, a set of mechanisms was identified, in a wide range of fields related primarily to modification and making data available.

7. Data protection mechanisms in the Oracle system

The characteristics of the identified mechanisms were presented in the following order: name of the producer, shortened principle of operation, group of protection methods (according to the classification [3]), description of the conducted study, assessment of the effectiveness of protection of the statistical database, justification. Wherever possible, results are documented with a presentation made in the application for communication with the database.[2]

To this end, the *Test* database was created, which was used to illustrate how data protection mechanisms can be assessed.[3]

Name of the mechanism: *Virtual Private Database (VPD)*

The principle of operation:

A means of implementing access control by limiting a database queries, undisclosed for the user and outside his control and defined in the security rules, by automatically adding additional conditions to the search result restriction clause.

Group of protection methods: access control

Method of analyzing the mechanism

In order to analyze the VPD mechanism, the table *Person* has been fed as presented in Figure 1. Next, using the language of defining the database schema, the so-called data protection[4] policy was defined (the principle of data limitation). The configuration of this policy is presented in Figure 2.

The proper operation of VPD required the definition of the database function to be created (Figure 3). For the VPD configuration defined this way, tests were conducted by calculating the values of statistics such as minimum (MIN), maximum (MAX), sum (SUM), arithmetic mean (AGV) for data from the table *Person*. These statistics have been called the reference statistics. They will be used to assess the quality of the statistics[5] received after the Oracle data protection mechanisms mentioned below are activated.

After collecting the reference statistics and activating the VPD mechanism three attempts were made[6] to calculate the statistics. The results of these attempts and the conclusions are presented below.

---

[1] Selecting the Oracle database management system is justified by the widespread use and the largest market share (as of February 9, 2020). https://www.statista.com/statistics/809750/worldwide-popularity-ranking-database-management-systems/ https://db-engines.com/en/ranking

[2] SQL Developer software version 4.1.4.21 was used.

[3] During the research, the table named Person was used, containing five columns named, in sequence, *Name* (character content), *Surname* (character content), *Acccount_Number* (numerical content), *Account_State* (numerical content), the column where the research was conducted is *Account_State* column.

[4] In the documentation, the concept of a protection policy or security policy can be met. For the purposes of the article, the author implemented the concept of the access limitation principle because the function which is the most important element of the VPD mechanism in fact always has a limiting role.

[5] The quality of statistics may be understood as the absolute value of the difference between the value of the reference statistic and the value of the statistics obtained during a survey. The lower the absolute value derived from difference, the better the quality of the statistics obtained.

[6] The number of attempts equal to three is not justified by the statements and mathematical relationships. The aim was to show that the ODR mechanism generates different data each time, which has a significant impact on the quality of the subsequent statistics (which is further shown).

ID	NAME	SURNAME	ACCOUNT_NUMBER	ACCOUNT_STATE
1	Name1	surname1	1	12000
2	Name2	surname2	2	366600
3	Name3	surname3	3	2221
4	Name4	surname4	4	225
5	Name5	surname5	5	5
6	Name6	surname6	6	22200
7	Name7	surname7	7	2256
8	Name8	surname8	8	14000
9	Name9	surname9	9	336
10	Name10	surname10	10	2266

Fig. 1. The table Person with test data used to test the VPD mechanism

```
BEGIN
 DBMS_RLS.ADD_POLICY (
  object_schema     => 'system',
  object_name       => 'persin',
  policy_name       => 'hide_rows',
  policy_function   => 'vpd_func',
  sec_relevant_cols => 'account_state');
END;
```

Fig. 2. A script in the database schema definition language that configures the VPD mechanism

```
CREATE OR REPLACE FUNCTION vpd_func (
 v_schema IN VARCHAR2,
 v_objname IN VARCHAR2)

RETURN VARCHAR2 AS
con VARCHAR2 (200);

BEGIN
 con := 'ID>2';
 RETURN (con);
END vpd_func;
```

Fig. 3. The configuration function of VPD

Test number	MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
1	5	366600

Fig. 4. Reference results of the statistics MIN, MAX obtained from the table Person

Test number	SUM(ACCOUNT_STATE)	AVG(ACCOUNT_STATE)
1	422109	42210,9

Fig. 5. Reference results of the statistics SUM, AVG for the table Person

Test number	MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
1	5	366600
2	5	366600
3	5	366600

Fig. 6. Results for the MIN, MAX statistics after activation of the VPD mechanism for the table Person (number of attempts 3)

Test number	SUM(ACCOUNT_STATE)	AVG(ACCOUNT_STATE)
1	410109	45567,6666666666666666666666666666666667
2	410109	45567,6666666666666666666666666666666667
3	410109	45567,6666666666666666666666666666666667

Fig. 7. Results for the SUM, AVG statistics after activation of the VPD mechanism for the table Person (number of attempts 3)

In accordance with the defined configuration, the VPD mechanism limited the set of answers by omitting the row with *id = 1*. The lack of this row affected the value of statistics in a predictable way, i.e. after starting the query it was not included in the calculation of statistics in subsequent attempts (1, 2, 3).

Effectiveness of protection of the statistical database: can be used to complement other methods

Justification:

The additional criteria in the query restricting clause are the way to further restrict access to the data in the table. Where the solution to protect data in a statistical database is to deliberately limit the resulting response sets, taking into account the impact on the subsequent quality of the statistics, and to accept this quality, the mechanism can be considered acceptable as complementary only.

The disadvantage of the VPD is the imposition of additional criteria on queries according to a predefined template, which can be modified intentionally only. The Oracle system does not have mechanisms to decide whether or not to run a VPD for a particular query. This would be important, for example, where the provision of certain data could raise the sensitivity of previously obtained statistics.

The name of the mechanism: *Oracle Data Redaction (ODR)*

The principle of operation:

The ODR mechanism consists in modifying the response set for a given query in such a way that the original values of selected rows or columns are replaced by substitute values generated according to defined data swapping rules.[7] Importantly, the changed data is not saved in the database, but is only an answer to the query.

Group of protection methods: access control

[7] In the documentation there are several algorithms for converting values in columns.

Method to test the mechanism

For the purpose of investigating the ODR mechanism, the test database named *Test*, containing the table *Person*, which had been fed as before (Figure 8), was built again. After feeding with data, the ODR mechanism was configured as presented in Figure 9. Then, tests were performed to calculate the statistics (MIN, MAX, SUM, AVG) called the reference statistics, for the table *Person*, and compare them with the statistics obtained after the ODR mechanism had been activated. The number of attempts was 3.[8]

	ID	NAME	SURNAME	ACCOUNT_NUMBER	ACCOUNT_STATE
1	1	Name1	surname1	1	12000
2	2	Name2	surname2	2	366600
3	3	Name3	surname3	3	2221
4	4	Name4	surname4	4	225
5	5	Name5	surname5	5	5
6	6	Name6	surname6	6	22200
7	7	Name7	surname7	7	2256
8	8	Name8	surname8	8	14000
9	9	Name9	surname9	9	336
10	10	Name10	surname10	10	2266

Fig. 8. The table Person

```
GRANT BEGIN
  DBMS_REDACT.add_policy(
    object_schema => 'system',
    object_name   => 'person',
    column_name   => 'account_state',
    policy_name   => 'redact_account_state',
    function_type => DBMS_REDACT.full,
    expression    => '1=1'
  );
END;
```

Fig. 9. A script in the database schema definition language that configures the VPD mechanism

ID	NAME	SURNAME	ACCOUNT_NUMBER	ACCOUNT_STATE
1	Name1	surname1	1	8258
2	Name2	surname2	2	139897
3	Name3	surname3	3	498
4	Name4	surname4	4	116
5	Name5	surname5	5	4
6	Name6	surname6	6	13361
7	Name7	surname7	7	1034
8	Name8	surname8	8	2699
9	Name9	surname9	9	135
10	Name10	surname10	10	1679

Fig. 10. The result of the first execution of a query retrieving all data from the table Person while the ODR mechanism is enabled

[8] The number of attempts was intended to show that each time the ODR mechanism generates different data.

ID	NAME	SURNAME	ACCOUNT_NUMBER	ACCOUNT_STATE
1	Name1	surname1	1	4940
2	Name2	surname2	2	171490
3	Name3	surname3	3	1615
4	Name4	surname4	4	139
5	Name5	surname5	5	0
6	Name6	surname6	6	12647
7	Name7	surname7	7	347
8	Name8	surname8	8	1693
9	Name9	surname9	9	232
10	Name10	surname10	10	1818

Fig. 11. The result of the second execution of a query retrieving all data from the table Person while the ODR mechanism is enabled

ID	NAME	SURNAME	ACCOUNT_NUMBER	ACCOUNT_STATE
1	Name1	surname1	1	4851
2	Name2	surname2	2	154954
3	Name3	surname3	3	1640
4	Name4	surname4	4	173
5	Name5	surname5	5	1
6	Name6	surname6	6	5126
7	Name7	surname7	7	1372
8	Name8	surname8	8	11025
9	Name9	surname9	9	261
10	Name10	surname10	10	1550

Fig. 12. Result of the third execution of a query retrieving all data from the table Person while the ODR mechanism is enabled

Test number	MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
1	5	366600

Fig. 13. The reference values of the statistics MIN, MAX for the table Person

Test number	SUM(ACCOUNT_STATE)	AVG(ACCOUNT_STATE)
1	422109	42210,9

Fig. 14. The reference values of the statistics SUM, AVG for the table Person

Test number	MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
1	4	139897
2	0	171490
3	1	154954

Fig. 15. Results of the statistics MIN, MAX for subsequent queries 1, 2, 3 when the ODR mechanism is enabled

Test number	SUM(ACCOUNT_STATE)	AVG(ACCOUNT_STATE)
1	503043	16768,1
2	194921	19492,1
3	180953	18095,3

Fig. 16. Results of the SUM, AVG statistics for subsequent queries 1, 2, 3 (when the ODR mechanism is enabled)

Analyzing the above results you can see that new values for the *account_state* column were generated by the number generator built into the ODR mechanism. Due to the fact that the ODR mechanism ensures data confidentiality when generating query responses (without changing the source data) after each attempt to execute query, the values of the *account_state* column were different than the results from the previous attempt.

The results of the reference statistics and the statistics obtained during the operation of the ODR mechanism vary considerably. This disqualifies the use of the ODR mechanism as a direct means of data protection in the statistical database.

Effectiveness of protection of the statistical database: can be used to complement other methods

Justification:

Converting the original value of an object into a random value and the resulting data protection is an ineffective way to protect statistical resources.

The problem is the uncontrolled change of data of individual information objects, which results in significant corruption of[9] statistics. This is shown above (in this case, statistics such as sum, average, minimum, maximum value proved to be useless).

The Oracle Data Redaction (ODR) mechanism also allows to swap data values in rows of numeric type[10] in a more controlled way, i.e. one of the data exchange algorithms (for rows of numeric type) is to swap a given numeric value by drawing a new number from the <*A*, *B*> range. In the configuration of this operation it can be assumed that the <*A*, *B*> range will be limited by a minimum and maximum value determined from the existing data in the rows of the modified column. This type of operation will also have a negative impact on the quality of statistics.

Another way to use ODR is to swap row values in a column by swapping data between rows[11] (e.g. a $N + 3$ row value is assigned to a N row, and a $N + 4$ row value is assigned to a $N + 3$ row and a N row value is assigned to a $N + 4$ row). After the data value conversion operation, the set of row values in the column will remain unchanged.

The advantage of this operation is low variability of statistics at a low level of detail, the disadvantage is the risk of significant variability of detailed statistics[12]. The usefulness of such statistics needs to be further discussed.

In addition, it should be noted that the swap of the original value of an object and the resulting data protection may be omitted (in the context of an inference) by performing additional queries that attempt to supplement the set of information obtained in earlier queries. Intentional modification of the response may result in a temporary delay or difficulty in obtaining information on a single statistical database object.

Where the aim is to know more precise properties of a object, it may be necessary to make further queries which could be handled by other, more effective protective mechanisms.[13]

Name of the mechanism: *Transparent Data Encryption (TDE)*

The principle of operation: Encrypting selected columns in tables based on encryption algorithms.

Group of protection methods: encryption

Effectiveness of protection of the statistical database: can be used to complement other methods

Justification: Distortion or corruption of data.

Name of the mechanism: Oracle Label Security (OLS)

[9] The article assumes that a statistic is corrupted when its result differs from the reference result by, for example, an order of magnitude or by p %.

[10] The mechanism also allows for the swap of data in other types, however, from the point of view of the article it is not of interest.

[11] This type of operation in selected cases may have negative influence on the values of statistics or even falsify them. Especially when the adjacent rows in the table refer to the criterion by which the statistics are calculated, e.g. row i refers to the salary *P1* from department *A* and row *i+1* refers to the salary *P2* from department *B*. This swap will result in *P1* and *P2* being wrongly included in Departments *A* and *B*.

[12] The statistics at a low level of detail are called e.g. average salary of all employees, average age of the patient in the outpatient care centre, etc. The statistics at a high level of detail are e.g. the average salary of employees by divisions - departments, the average age of a patient in a outpatient care centre divided by disease group, e.g. cardiovascular disease, etc.

[13] Mechanisms to prevent the recognition of the value of a single object in a statistical database.

The principle of operation:

The complementation of the table access control mechanism, implemented at the data level in a table, by adding an additional column or group of columns to the table with defined values called labels, to which OLS provides configurable access based on defined security policies.

Group of protection methods: access control

The way the mechanism is tested:

By using the OLS mechanism a new column called *label_security* was added to the table *Person*. Then each of the rows was randomly assigned a value from the set *label_value* = {*label1*, *label2*, *label3*}, thus obtaining a table presented in Figure 17.

For the table presented this way, reference statistics (min, max, avg, sum) were calculated and compared with selected statistics obtained for the system user having access to selected labels. The first case involved a user who has access to data labelled as *label*1, while the second case involved a user who has access to data labelled as *the label*2, *the label*3. The results and conclusions are set out below.

ID	NAME	SURNAME	ACCOUNT_NUMBER	ACCOUNT_STATE	LABEL_SECURITY
1	Name1	surname1	1	12000	label1
2	Name2	surname2	2	366600	label2
3	Name3	surname3	3	2221	label1
4	Name4	surname4	4	225	label1
5	Name5	surname5	5	5	label2
6	Name6	surname6	6	22200	label2
7	Name7	surname7	7	2256	label2
8	Name8	surname8	8	14000	label3
9	Name9	surname9	9	336	label3
10	Name10	surname10	10	2266	label3

Fig. 17. View of the table Person after OLS mechanism is activated

MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
5	366600

Fig. 18. The reference value MIN, MAX for the table Person after OLS is applied

SUM(ACCOUNT_STATE)	AVG(ACCOUNT_STATE)
422109	42210,9

Fig. 19. The reference value SUM, AVG for the table Person after OLS is applied

MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
225	12000

Fig. 20. The result of the statistics MIN, MAX for a user who has access to the data labelled as *the label*1

SUM(ACCOUNT_STATE)	AVG(ACCOUNT_ST...
14446	4815,33333333...

Fig. 21. The result of the statistics SUM, AVG for a user who has access to the data labelled as *the label*1

MIN(ACCOUNT_STATE)	MAX(ACCOUNT_STATE)
5	366600

Fig. 22 The result of the statistics MIN, MAX for a user who has access to the data labelled as *the label*1, *the label*2

SUM(ACCOUNT_STATE)	AVG(ACCOUNT_STATE)
405507	57929,57142857142...

Fig. 23. The result of the statistics SUM, AVG for a user who has access to the data labelled as *the label*1, *label*2

Analyzing the above results, it can be concluded that the OLS mechanism does not distort data, but only restricts access to data based on which statistics are calculated.

Effectiveness of protection of the statistical database: can be used to complement other methods

Justification:

The labelling mechanism may serve as a means to restrict access to data, being a complementary way to protect data in the statistical database, provided that it is deliberate and intentional to deny access to selected groups of data and the consequences of this action are known (primarily the impact on the quality of statistics, i.e. overstatement, understatement, misinterpretation).

The OLS mechanism does not have features that would allow for the use of denying or making selected data available in an autonomous manner. That is to say, a database system could dynamically limit access to the data based on reasons of risk that confidentiality can be lost.

8. Conclusions

The implementation of a statistical database requires much more sophisticated data protection mechanisms than is in case of transactional databases. While in transactional applications available implementations of data protection methods meet data protection needs, in the case

of statistical solutions there is a deficit of technical but also conceptual solutions.

Existing theories outline the nature of protection problems, but in many cases the available knowledge is more of a direction of research than the existence of conceptual and technical solutions that can be exploited or developed in intensive and targeted way.

It should be noted that the pursuit of new or developing existing theories and guidelines for statistical databases is as important as research and an intuitive attempt to use existing and proven solutions. The realization of such an approach is the qualification of the Oracle system in the context of the potential usefulness of the data protection mechanisms contained therein in the statistical protection of the database.

The available Oracle solutions such as VPD, OLS, ODR, TDE for statistical database protection have proven to be insufficient, although they may at least provide a basis to build effective mechanisms to protect statistical data or complement other existing ones. Their properties and operation are the result of many years of theoretical and practical work on data protection in transactional databases and, above all, the result of the experience resulting from their use.

Currently, in the case of statistical databases, data protection is essentially a set of theoretical issues which provide guidelines and proposals on how to implement potential protection mechanisms for such databases.

The problem (apart from the conceptual deficit) is the difficulty and complexity of effective implementation of data protection principles and rules in statistical databases. It is difficult to ensure the confidentiality of the data through tailor-made inference control.

One possible approach to solve this problem is to use transactional databases and their data protection mechanisms as a complementary solution.

The question remains open as to whether it is possible to directly extend the data protection mechanisms in transactional databases by effective ways of protecting data in statistical databases and whether such extension can become an adequate response to the needs to protect such resources.

9. Bibliography

[1] Galcia-Molina H., Ulman J.D., Widom J., *Systemy baz danych. Pełny wykład*, WNT, Warszawa 2006.

[2] Elmasri N., *Wprowadzenie do systemów baz danych*, Helion, Gliwice 2016.

[3] Denning D.E., *Cryptography and Data Security*, Addison-Wesley Publishing Company, California 1982.

[4] Szafrański B., "Podstawy budowy skutecznych metod ochrony statystycznych baz danych", *Wiadomości statystyczne,* 3(670), 71−85 (2017).

[5] Denning D.E., "A lattice model of secure information flow", *Communications of the ACM*, Vol. 19, No. 5 (1976).

[6] Corporation, Oracle, "Oracle Database 12 c Security".

[7] Corporation, Oracle, "Oracle Database Security Guide 12c".

[8] Bell D.E., LaPadula L.J., *Secure computer systems: Mathematical foundations and model*, Bedford: The Mitre Corp., 1973.

[9] Denning D.E., "Secure information flow in computer systems", Ph.D. Thesis, Purdue University, 1975.

[10] Oracle Corporation, "Oracle Database Online Documentation 12c".

[11] https://encyklopedia.pwn.pl/haslo/baza-danych;3875256.html.

Ochrona danych w transakcyjnych i statystycznych zastosowaniach baz danych

B. SZAFRAŃSKI, R. BAŁAZY

Artykuł prezentuje dyskusję dotyczącą problematyki ochrony danych w bazach danych zawierającą próbę odpowiedzi na pytanie o możliwości użycia transakcyjnego systemu bazodanowego jako systemu zdolnego do ochrony danych w statystycznej bazie danych. Dyskusja poprzedzona jest przypomnieniem podstawowych zagadnień związanych z ochroną danych w bazach danych, w tym modeli sterowania przepływem, sterowania dostępem oraz wnioskowaniem. Kluczowy element artykułu stanowi analiza, na przykładzie systemu zarządzania bazą danych Oracle, czy mechanizmy ochrony danych w transakcyjnych bazach danych mogą być skuteczne w przypadku ochrony danych w statystycznych bazach danych.

Słowa kluczowe: baza danych, system bazy danych, ochrona danych w bazie danych, sterowanie przepływem, sterowanie wnioskowaniem, statystyczna baza danych, Oracle.